



UNIVERSITA' DELLA CALABRIA

Dipartimento di Ingegneria Meccanica, Energetica e Gestionale

Scuola di Dottorato "Pitagora" in Scienze Ingegneristiche

Dottorato di Ricerca in Ingegneria Meccanica

CICLO

XXVIII

TITOLO TESI

*Integrazione di tecniche ottiche e acustiche di imaging 3D
in ambiente subacqueo.*

Settore Scientifico Disciplinare ING-IND 15

Direttore:

Ch.mo Prof. Sergio Rizzuti

Coordinatore:

Ch.mo Prof. Leonardo Pagnotta

Supervisore:

Ch.mo Prof. Fabio Bruno

Dottorando: Dott. Antonio Lagudi

Sommario

Introduzione.....	1
Obiettivi della tesi.....	4
Struttura della tesi.....	5
1 Stato dell'arte.....	6
1.1 Sensori e metodologie per l' <i>imaging</i> 3D subacqueo.....	6
1.1.1 Metodologie di <i>sensing</i>	6
1.1.2 Tipologie di sensori.....	10
1.2 <i>Data Fusion</i>	35
1.2.1 <i>Multisensor data fusion</i>	36
1.2.2 Integrazione di sistemi 3D ottico - acustici.....	40
2 Metodologia per la calibrazione di una camera ottico - acustica 3D.....	52
2.2 Configurazione del sistema	53
2.2.1 Descrizione sottosistema ottico	53
2.2.2 Descrizione sottosistema acustico	56
2.2.3 Layout del sistema.....	58
2.3 Analisi su differenti approcci alla calibrazione estrinseca del sistema.....	60
2.4 Metodologia di calibrazione	67
2.4.1 Canale di <i>sensing</i> acustico	68
2.4.2 Canale di <i>sensing</i> ottico.....	70
2.4.3 Registrazione ottico – acustica	79
2.4.4 Elaborazione statistica.....	80
2.5 Set-up sperimentale	81
2.6 Sperimentazione.....	83
2.6.1 Acquisizioni immagini.....	83
2.6.2 <i>Image Enhancement</i>	85
2.6.3 Ricostruzione ottica 3D.....	86
2.6.4 Registrazione ottico – acustica	86
2.6.5 Elaborazione statistica.....	87
2.7 Risultati.....	91
3 Prototipo camera 3D ottico - acustica.....	94
3.1 Descrizione del sistema	95
3.1.1 Architettura <i>hardware</i>	95

3.1.2	Architettura <i>software</i>	101
4	Metodologia per la modellazione 3D di siti archeologici sommersi	112
4.1	Metodologia di rilievo, ricostruzione 3D e geo-localizzazione di siti archeologici sommersi	113
4.1.1	Ispezione e segmentazione sito archeologico	118
4.1.2	Acquisizione ed elaborazione dati ottici.....	125
4.2	Ricostruzione ottico- acustica sito archeologico Capo Colonna.....	126
4.2.1	Acquisizione ed elaborazione dati zona 2: fondale.	126
4.2.2	Acquisizione ed elaborazione dati zona 1: “relitto Punta Scifo D”	130
4.2.3	Acquisizione ed elaborazione dati acustici.....	135
4.2.4	Data fusion ottico – acustico	145
	Conclusioni finali.....	151
	Riferimenti.....	153

Introduzione

L'osservazione e l'esplorazione del regno subacqueo rappresentano argomenti di grande interesse fin da tempi antichi. Ciò ha consentito, nel corso degli anni, di ottenere enormi progressi tecnologici in svariati settori legati a queste attività: sviluppi negli apparati di immersione, nei sensori ottici e acustici, nella realizzazione di veicoli comandati a distanza (*Remotely Operated Vehicles, ROV*) o veicoli subacquei autonomi (*Autonomous Underwater Vehicles, AUV*) sono, senza dubbio, tra gli esempi più notevoli di tali progressi.

Sebbene la meta di viaggiare nelle profondità oceaniche sia inseguita da secoli, è comunque piuttosto recente il tentativo di porre enfasi sull'ottenimento e sull'interpretazione dell'informazione al di là di quella resa disponibile dal sistema visivo umano; ciò è stato reso possibile sia dalle nuove tecniche e tecnologie oggi impiegate, nonché da una migliore comprensione dei parametri fisici relativi all'ambiente subacqueo: metodi convenzionali di acquisizione immagini per mezzo di fotocamere ad alta risoluzione, metodi di scansione laser, metodi elettromagnetici, o tecniche acustiche di acquisizione, sono solo alcune delle tecnologie attualmente impiegate con profitto in applicazioni marine.

Dal punto di vista storico, i progressi nel campo dell'*imaging* subacqueo sono stati caratterizzati da una lenta evoluzione, ragion per cui sono state necessarie decine di anni per raggiungere l'attuale stato dell'arte. Ciò è stato in parte causato dai limiti derivanti da una non piena comprensione del fenomeno della trasmissione delle radiazioni attraverso mezzi ottici come l'aria e l'acqua di mare, limiti nell'*hardware* ottico e nell'elettronica dei sensori, restrizioni nel trasferimento dei dati, nella registrazione e nell'elaborazione dei segnali. Progressi continui in queste aree hanno effettivamente limitato la riduzione del contrasto e del raggio d'azione risultanti da effetti fisici intensi della propagazione luminosa nell'acqua marina; osservazioni simili possono anche essere fatte per la propagazione del suono e quindi per l'*imaging* acustico.

Una delle tecnologie che negli ultimi anni sta emergendo prepotentemente in questo ambito è rappresentata dall'*imaging* 3D. Questa può essere definita come l'intero

processo che a partire dall'acquisizione di dati metrici dell'ambiente circostante, ottenuti attraverso l'impiego di una varietà differente di tecniche e strumenti, si conclude con la visualizzazione tridimensionale della scena osservata. In questo senso, il termine assume un significato molto più ampio di quello che comunemente gli viene attribuito, cioè la sola ricostruzione 3D della scena. Quest'ultima, in realtà, rappresenta solo il primo passo dell'intero processo, il quale include anche le successive fasi di modellazione e visualizzazione.

Nel campo delle applicazioni subacquee i sistemi di *imaging* 3D sono sempre più impiegati sia in ambito commerciale, come nelle perforazioni *off-shore* o nei moderni rilevamenti bentonici, sia in applicazioni scientifiche, quali il monitoraggio di flora e fauna marina [1-3] o in archeologia subacquea [4]. In particolare, quest'ultimo settore, ha avuto enormi benefici dall'introduzione di tecniche e tecnologie d'*imaging* 3D nelle tradizionali metodologie impiegate. Difatti, la fase di rilievo tridimensionale è divenuta di fondamentale importanza in quanto, oltre a fornire una mappatura dettagliata ed esaustiva del sito archeologico in fase di documentazione, risulta necessaria per condurre scavi ed esplorazioni in quanto consente di acquisire una certa sicurezza e quindi evitare procedure errate che potrebbero danneggiare il patrimonio archeologico stesso.

Tutte queste attività hanno come filo conduttore la necessità d'impiegare tecniche e tecnologie in grado di creare modelli 3D di elevata accuratezza e precisione, non trascurando al contempo aspetti fondamentali quali il costo, la facilità d'utilizzo, il livello di automazione dell'intero processo.

Indipendentemente dalla particolare soluzione adottata, l'ambiente subacqueo impone dei vincoli severi che rendono il soddisfacimento dei requisiti sopra esposti un compito arduo; ciò a causa delle molteplici problematiche derivanti dalle proprietà dell'acqua quale mezzo trasmissivo e dalle difficili condizioni operative. *Scattering* ed assorbimento [5] sono i fenomeni che maggiormente degradano le prestazioni dei sistemi ottici di *imaging* 3D a causa di un decadimento del contrasto nelle immagini e dell'attenuazione dell'intensità della luce, con conseguente perdita di dettagli ed alterazione dei colori. Tali fenomeni sono ben noti, e diversi sono i lavori presenti in letteratura riguardanti l'impatto dell'acqua sulla propagazione luminosa e sulla resa cromatica [6-8]. D'altro canto, l'uso di sensori acustici di *imaging* 3D mitiga

notevolmente l'impatto che tali fenomeni hanno sui sistemi d'acquisizione; ciò, comunque, al costo di ottenere immagini 3D a risoluzione molto minore se comparata a quella tipica dei sistemi ottici e ad una più difficile interpretazione dei dati.

In questo contesto, l'integrazione di sensori ottici e acustici e la conseguente fusione dei dati da essi acquisiti sta emergendo, negli ultimi anni, come una tecnica estremamente promettente nel campo dell'*imaging* subacqueo. La natura complementare di queste due tipologie di sensori, che rende la loro integrazione particolarmente interessante, origina allo stesso tempo diverse problematiche riguardanti la fusione dei dati da essi acquisiti. In generale, sonar e dispositivi ottici forniscono tipologie d'informazione completamente differenti e dunque principi di elaborazione diversi sono applicati per manipolare ed estrarre informazioni da queste due sorgenti di dati. Il principale problema nella fusione dei dati acquisiti da entrambi i sensori è che essi operano a risoluzioni altamente differenti. Un sistema ottico è in grado di fornire immagini ad alta risoluzione, ma il suo *range* è limitato dall'attenuazione della luce in acqua, mentre i sensori acustici possono operare in *range* di visibilità più ampi e fornire informazioni 3D anche in condizioni di acque torbide, a spese però di una più bassa risoluzione ed una più difficile elaborazione ed interpretazione dei dati. Ad oggi, solo pochi laboratori in tutto il mondo hanno prodotto lavori pionieristici sulla fusione di dati ottico - acustici in ambiente sottomarino. Si vedano, ad esempio, i lavori riportati in [9, 10], in cui gli autori descrivono l'uso di tecniche che consentono la sovrapposizione di fotomosaici su mappe batimetriche 3D del fondale. In questo caso, le importanti informazioni qualitative provenienti dalle immagini ottiche sono fuse con la mappa digitale del terreno, la cui definizione geometrica è ottenuta attraverso acquisizioni sonar. Rilievi ottici e acustici possono anche essere integrati utilizzando sistemi ottici a luce strutturata con sonar ad alta frequenza, come riportato in [11]. Questo approccio è molto robusto e accurato in condizioni di scarsa visibilità, ma non riporta informazioni qualitative della scena acquisita. Secondo quanto riportato in [12], la fusione dei dati ottici e acustici può essere realizzata utilizzando due strategie, o "livelli" differenti. Un primo approccio, definito *data level fusion*, semplicemente prende l'informazione 3D estratta da ogni modalità in maniera indipendente e l'allinea utilizzando una trasformazione affine calcolata o attraverso la selezione manuale di coppie di punti 3D

da entrambe le rappresentazioni o attraverso l'applicazione di algoritmi automatici. L'altro approccio è una strategia più complessa, definita *signal level fusion*, che utilizza i dati grezzi provenienti da entrambi i sensori per produrre una singola e coerente rappresentazione 3D della scena acquisita. Già da questi pochi esempi si evince come tale problematica viene affrontata con approcci altamente differenti, evidenziando in tal modo come le tecniche di fusione dipendono fortemente dai vincoli imposti dalla particolare configurazione di sistema adottata (tipologia di sensori impiegati, struttura del modello dati, specifica applicazione per cui tale sistema è stato concepito).

Comunque, nonostante le enormi difficoltà nel combinare le tecniche di *imaging* 3D ottiche e acustiche, le molteplici innovazioni tecnologiche ed i progressi, soprattutto nel campo dei sensori acustici, hanno gradualmente permesso la generazione di dati di buona qualità ed elevata risoluzione in grado di essere integrati con maggiore profitto e dunque la possibilità di poter implementare nuove tecniche di fusione e di integrazione di sensori di *imaging* 3D in ambiente subacqueo.

Obiettivi della tesi

Il principale obiettivo di questa tesi è lo studio di diverse metodologie per l'integrazione d'informazione ottica e acustica al fine di ottenere una rappresentazione tridimensionale della scena subacquea. A tal fine, questo obiettivo generale è stato suddiviso nei seguenti obiettivi specifici:

- studio dei sensori e delle metodologie di *imaging* 3D maggiormente impiegate in applicazioni subacquee, analizzando principali caratteristiche e prestazioni;
- studio dello state dell'arte sulle principali problematiche inerenti la fusione di dati provenienti da sorgenti differenti ed in particolar modo i lavori riguardanti la fusione di dati acquisiti per mezzo di sensori d'*imaging* 3D ottici e acustici in ambiente subacqueo;
- progettazione, test e validazione di una metodologia di calibrazione per l'integrazione di dati 3D acquisiti attraverso una camera ottica stereoscopica ed una camera acustica 3D;
- progettazione, secondo un approccio di tipo *signal level fusion*, di strategie per la fusione e la rappresentazione dell'informazione ottenuta attraverso un

prototipo di sistema composto da una camera ottico-acustica per l'acquisizione in *real-time* di immagini 3D in ambiente subacqueo in condizioni di torbidità variabile e da un *software* per la visualizzazione in realtà aumentata (AR) dei dati acquisiti;

- progettazione e test di una metodologia per il rilievo, la modellazione 3D e la geo-localizzazione di reperti e siti archeologici sommersi impiegando tecniche ottiche e acustiche secondo un approccio di tipo *data level fusion*.

Struttura della tesi

La presente tesi è organizzata nel modo seguente:

Il Capitolo 1 ha lo scopo di studiare lo stato dell'arte concernente l'integrazione di sensori *d'imaging* 3D ottici e acustici. A tal fine, è presentata una panoramica dei principali sensori ottici e acustici comunemente impiegati in applicazioni subacquee. Successivamente, dopo aver introdotto i concetti fondamentali delle tecniche di *data fusion* e le principali problematiche inerenti tale disciplina è fornita una panoramica dei principali lavori che integrano tecniche di *imaging 3D* ottico e acustico in ambiente subacqueo.

Il Capitolo 2 descrive la progettazione, l'implementazione ed i test riguardanti una metodologia di calibrazione specificatamente realizzata al fine di poter integrare una camera ottica stereoscopica e una camera acustica 3D.

Il Capitolo 3 presenta un prototipo di camera ottico – acustica subacquea progettata al fine di generare, in *real-time*, una rappresentazione tridimensionale della scena osservata. Viene illustrata l'architettura *hardware* e *software* del sistema, così come le funzionalità e la logica di visualizzazione dei dati 3D ottico – acustici.

Infine, il Capitolo 4, è dedicato alla descrizione e successiva applicazione di una metodologia per il rilievo, la modellazione 3D e la geo – localizzazione di siti archeologici sommersi che utilizza dati ottici e acustici secondo un approccio di tipo *data level fusion*.

1.1 Sensori e metodologie per l'*imaging* 3D subacqueo

La tecnologia odierna comprende un vasto assortimento di *hardware* e *software* di trattamento di immagini rapidi, economici e specializzati, oltre che di fotocamere di elevata qualità e non troppo costose. Sistemi laser vengono impiegati per operazioni di misura quantitative 2D e 3D, insieme con elaborazioni Sonar utilizzando veloci microprocessori, *field-programmable gate arrays* (FPGA, dispositivo a semiconduttori contenente componenti logici e interconnessioni entrambi programmabili) o DSP (*Digital Signal Processor*). Algoritmi prima troppo dispendiosi dal punto di vista computazionale sono adesso realizzabili e, dato il continuo migliorare dell'*hardware*, diventano possibili applicazioni sempre nuove.

A tal riguardo, in questa sezione, viene presentata una panoramica delle principali metodologie e dei sensori utilizzati nell'ambito dell'*imaging* 3D subacqueo. Verranno dapprima descritte le diverse metodologie di *sensing* per l'acquisizione di dati 3D e successivamente le differenti tipologie di soluzioni *hardware* esistenti.

1.1.1 Metodologie di *sensing*

I sensori di *imaging* 3D possono essere classificati in tre principali categorie, dipendentemente dalla metodologia di misura: Triangolazione, Tempo di volo, Modulazione.

Come mostrato in Fig. 1.1, un determinato sensore può appartenere a più di una classe, il che significa come esso, al fine di ottenere dati 3D, utilizzi differenti metodologie o una loro combinazione. Un'ulteriore classificazione, tradizionalmente adottata in letteratura scientifica, suddivide i sensori in base al loro comportamento energetico: avremo allora dispositivi attivi o passivi.

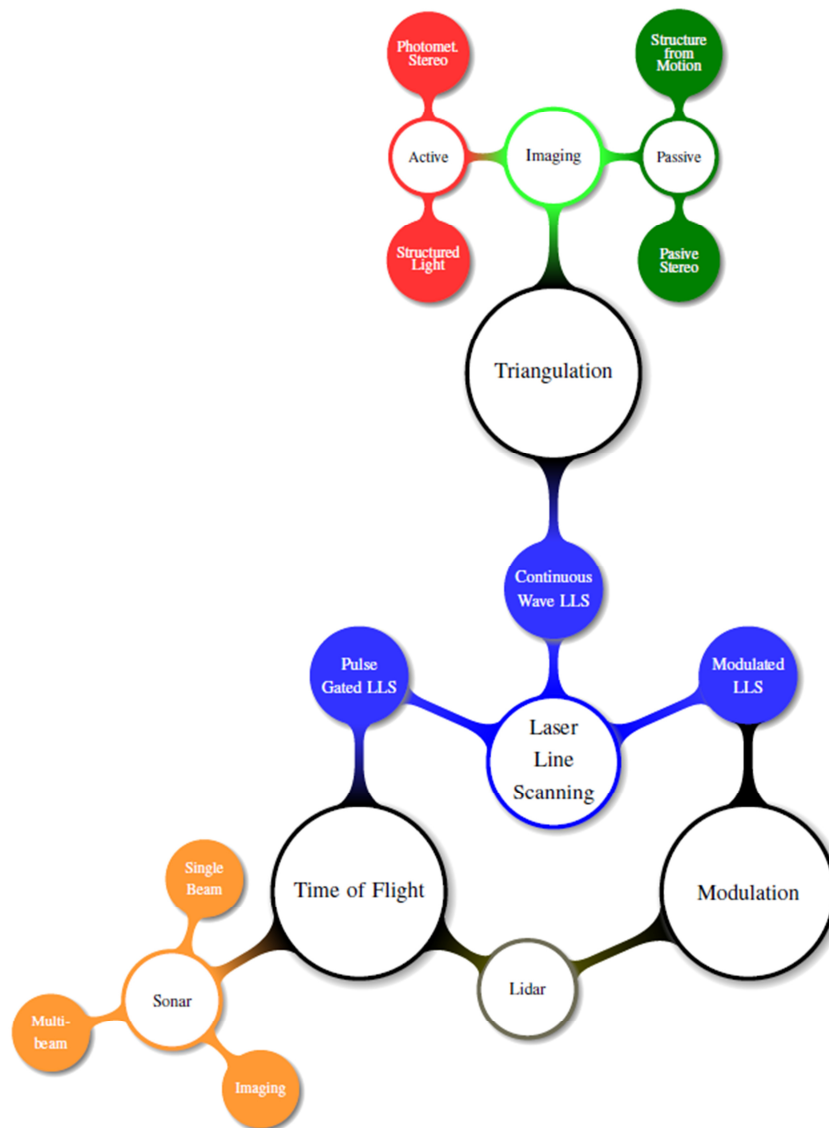


Fig. 1.1: classificazione dei sensori di *imaging* 3D.

2.1.1.1 Tempo di volo

Il **tempo di volo** (spesso indicato con **TOF**, dall'inglese ***Time Of Flight***) indica la misura del tempo impiegato da un oggetto, una particella o un'onda (acustica, elettromagnetica o di altro tipo) per percorrere una certa distanza in un determinato mezzo. Dalla conoscenza della velocità del segnale nel mezzo di propagazione è dunque possibile ricavare misure di distanza. Le metodologie basate sulla discriminazione temporale consentono di ottenere misurazioni a grande distanza, esempio tipico è rappresentato dai sistemi Sonar, anche se l'accuratezza delle misure può essere influenzata da variazioni della temperatura dell'acqua, della sua salinità e dalla pressione, in quanto questi parametri incidono sulla velocità del suono.

Per brevi distanze, un piccolo ritardo nella temporizzazione può causare un elevato errore di misura. Inoltre, alcuni sensori richiedono una distanza minima di lavoro che dipende dalla loro geometria.

Sonar, LIDAR e *Pulse Gated Laser Line Scanning* (PG-LLS) sono alcuni esempi di sensori *hardware* che utilizzano tale tecnica per acquisire dati 3D.

2.1.1.2 Triangolazione

I metodi di triangolazione si basano sulla misura della distanza da due o più dispositivi (siano essi sensori o ricevitori) di una *feature* comune o un target di cui sono noti alcuni parametri. Ad esempio, due camere ottiche possono ottenere la profondità della scena attraverso l'acquisizione di due immagini prese da due posizioni leggermente differenti. Un algoritmo di corrispondenza stereo cerca di identificare punti corrispondenti (*features*) nelle due immagini, relative allo stesso punto nella scena, risolvendo il cosiddetto "problema della corrispondenza". Una volta identificate ed opportunamente filtrate, le *features* rimanenti vengono proiettate nello spazio come raggi provenienti dalle due camere. Il triangolo formato dalla *feature* nello spazio tridimensionale e le due camere è alla base della triangolazione.

La limitazione maggiore dei sistemi basati sulla triangolazione è il campo di vista (FOV) sia dell'emettitore che del ricevitore. Le loro prestazioni tendono a degradare all'aumentare della distanza dal target. Infine, maggiore è la separazione tra le camere (*baseline*) migliore è la risoluzione in profondità, a condizione che vi sia una sovrapposizione tra i campi di vista.

Esistono diversi sensori che determinano l'informazione 3D per triangolazione: tra i sistemi di *imaging* attivi possiamo annoverare quelli a luce strutturata, i Laser strip e quelli stereo fotometrici (PhS). Esempi di sistemi passivi sono quelli basati sulla tecnica Structure From Motion o di Stereo Visione mentre tra quelli a scansione laser lineare si ha il Continuous Wave Laser Line Scanning (CW-LLS).

2.1.1.3 Modulazione

Mentre l'approccio nel dominio del tempo utilizza l'ampiezza e il tempo per discriminare i segnali diffusi provenienti dal target, i sistemi che operano nel dominio

della frequenza eseguono questa operazione utilizzando le differenze in ampiezza e in fase di un segnale modulato. I segnali diffusi subiscono un processo di *scattering* multiplo con una conseguente dispersione temporale dell'impulso trasmesso. Tale fenomeno consente solo alle componenti a bassa frequenza di poter essere trasmesse in maniera efficiente, mentre quelle ad alta frequenza sono perdute.

In ambiente subacqueo, le principali tecniche di modulazione/demodulazione coerente non possono essere applicate a segnali alle frequenze ottiche sia a causa dell'elevata dispersione sia per la dipendenza dei coefficienti di assorbimento e di *scattering* dalla lunghezza d'onda del segnale ottico stesso. Esiste un minimo per questi coefficienti nella porzione blu – verde dello spettro elettromagnetico. Tali fenomeni rendono la modulazione in intensità di una portante laser l'unica tecnica di modulazione per l'*imaging* 3D subacqueo. Tale metodo basa il suo funzionamento su una modulazione in ampiezza di una portante laser attraverso un segnale a frequenza dell'ordine del GigaHertz. Per tale ragione, i sistemi necessitano di sensori molto sensibili ed accurate scale temporali. Tipicamente i ricevitori sono dei Photomultiplier Tubes (PMT) o matrici di fotodiodi avalanche (APD). Questi sensori vengono attivati per un determinato periodo temporale (*time window*) e i fotoni incidenti in questo intervallo vengono acquisiti, trasformati in segnali elettrici ed integrati. Infine, dopo la fase di demodulazione, l'informazione 3D è ottenuta come differenza di fase rispetto al segnale trasmesso.

2.1.1.4 Sensori attivi e passivi

In generale, così come accade per i sistemi terrestri, i sensori di *imaging* 3D comunemente impiegati in ambito subacqueo possono essere classificati come attivi o passivi, dipendentemente dalla modalità di utilizzo della sorgente luminosa.

I sistemi passivi utilizzano la luce al fine di acquisire immagini di oggetti che sono stati illuminati da una certa sorgente diversa da quella associata al sistema di acquisizione. Alcuni esempi sono l'*imaging* di oggetti mediante luce solare o altre fonti di illuminazione come la bioluminescenza. L'*imaging* passivo è particolarmente richiesto per operazioni "nascoste" come le ispezioni di oggetti effettuata dalla Marina Militare o per osservazioni della fauna marina. I dispositivi attivi fanno invece uso di una sorgente di luce; un esempio semplice è costituito da un sistema a fotocamera

subacquea che utilizzi illuminazione artificiale ad impulsi o continua. Tali sistemi offrono vantaggi sostanziali per l'acquisizione subacquea, dal momento che la luce incidente può essere collimata in raggi molto ristretti e può essere monocromatica (laser), con la possibilità di durata molto breve (sistemi di illuminazione a impulsi, laser pulsati). Inoltre, i sistemi attivi permettono solitamente un'acquisizione a distanze maggiori e con maggior contrasto rispetto ai sistemi passivi che usano la luce del sole. In più, con laser a impulsi brevi, possono operare perfino a livelli moderati di luce circostante. L'*imaging* subacqueo attivo è però molto spesso limitato a causa della radiazione che subisce il *backscatter* dovuto alle particelle presenti nel volume d'acqua illuminato, poiché la radiazione retrodiffusa si sovrappone alla luce "utile" proveniente dagli oggetti per riflessione.

1.1.2 Tipologie di sensori

2.1.1.5 Sonar

Il termine SONAR nasce come acronimo dell'espressione inglese *SOund NAVigation and Ranging*. Vi sono due principali categorie di sonar: attivi e passivi.

I sistemi sonar passivi di solito dispongono di grandi database di "firme sonore". Un sistema computerizzato utilizza questi database per identificare classi di navi, azioni (cioè velocità di una nave, tipologia di arma rilasciata) o particolari tipologie di imbarcazioni. Questi sensori non sono evidentemente utilizzati per applicazioni di *imaging* 3D, quindi verranno tralasciati in questo studio.

Nei principali sistemi acustici 3D attivi, la scena è "insonificata" attraverso un impulso acustico e i segnali di *backscatter* sono acquisiti utilizzando *array* bidimensionali di sensori. Questi sistemi sono classificati in tre principali categorie, dipendentemente dalla metodologia di elaborazione del segnale: *beamforming systems*, *holographic systems*, *acoustic lens – based systems*.

1.1.2.1.1 Beamforming systems

Un *beamforming system* riceve i segnali di *backscatter* utilizzando un *array* 2D di sensori e li elabora così che l'eco proveniente dalla direzione di puntamento sia

amplificato a differenza di quelli che provengono da altre direzioni che vengono invece completamente attenuati. In tal modo si ottengono i dati 3D per ogni direzione di puntamento mentre un'immagine 3D completa viene costruita ripetendo questo processo variando l'angolo di puntamento in maniera simile a quanto avviene per un sistema di scansione raster. In Fig. 1.2 è illustrato un tipico *beam pattern* per questa tipologia di sistemi.

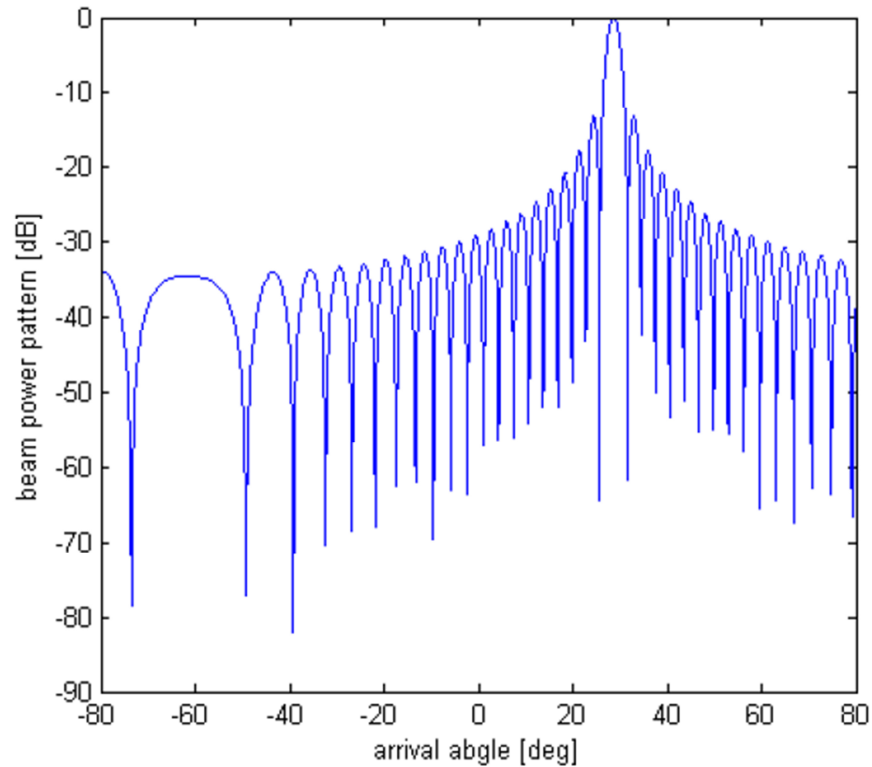


Fig. 1.2: *beam pattern* di un *array* di 64 elementi con spaziatura di 1.5 mm, direzione $\theta = 30^\circ$, frequenza = 300 KHz.

Nell'ipotesi in cui i sensori siano equispaziati di una quantità d e centrati nell'origine del sistema di riferimento, il *beam pattern* può essere scritto in forma chiusa come [13]:

$$BP_{BMF}(\omega, \beta, \theta) = \frac{\sin[\omega Nd(\sin\beta - \sin\theta)/2c]}{\sin[\omega Nd(\sin\beta - \sin\theta)/2c]} \quad (1.1)$$

dove N è il numero di sensori; θ la direzione di puntamento; d la distanza tra i sensori; $\omega = 2\pi f$ la velocità angolare; f la frequenza; c la velocità del suono nel mezzo; $\beta \in [-\varphi_{max}, \varphi_{max}]$ e $2\varphi_{max}$ è la copertura angola verticale del sonar.

Il *beam pattern* rappresenta un importante strumento per valutare le *performance* del sistema attraverso due parametri chiave: “forma del lobo principale” e “livello dei lobi secondari”. La larghezza del lobo principale rappresenta la risoluzione angolare del sistema, mentre il livello dei lobi secondari influisce sulla quantità di rumore presente sui dati finali (artefatti sulle immagini 3D generate). L’equazione (1.1) è valida in caso di un *array* lineare di sensori, comunque non sufficiente a produrre un’immagine 3D della scena acquisita. In effetti, al fine di poter generare un’immagine acustica 3D è necessario poter indirizzare il fascio nello spazio tridimensionale. Supponendo di avere un *array* quadrato di sensori equispaziati, il *beam pattern* nello spazio 3D può essere espresso come:

$$BP_{BMF}(\omega, \beta_\alpha, \beta_e, \theta_\alpha, \theta_e) = BP_{BMF}(\omega, \beta_\alpha, \theta_\alpha)BP_{BMF}(\omega, \beta_e, \theta_e) \quad (1.2)$$

dove θ_α e θ_e rappresentano gli angoli di puntamento in azimut ed elevazione rispettivamente. La figura 1.3 mostra il *beam pattern* nello spazio 3D di un *array* di trasduttori con direzione di puntamento $\theta_\alpha = 30^\circ$ $\theta_e = 0^\circ$.

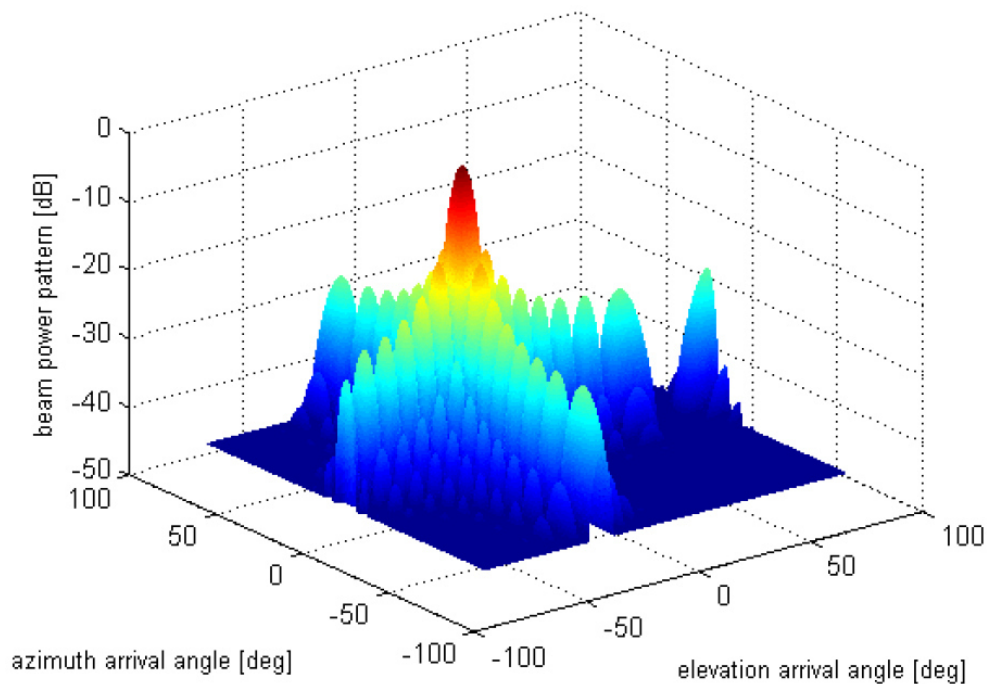


Fig. 1.3: *beam pattern* di un *array* bidimensionale di 15 x 15 elementi con spaziatura $\lambda/2$, direzione di puntamento $\theta_\alpha = 30^\circ$ $\theta_e = 0^\circ$, frequenza = 300 KHz.

1.1.2.1.2 Holographic systems

Esistono due tipologie di algoritmi utilizzati nell'ambito dei sistemi olografici: il primo, impiegato per sistemi a banda stretta, si basa su un'approssimazione di Fourier/Fresnel mentre il secondo usa la tecnica SVD (*Singular Value Decomposition*) al fine di stimare la riflettività acustica della scena per sistemi a banda larga [13].

Nei sistemi olografici a banda stretta l'algoritmo "sfasa" i segnali acquisiti dai sensori 2D e applica una trasformata di Fourier discreta al fine di passare dall'indice $l = 1 \dots N$ dei sensori all'indice delle frequenze spaziali k . Il *beam pattern* che si ottiene applicando tale algoritmo può essere espresso in forma chiusa come [13]:

$$BP_{HOL}(\omega, \beta, k) = \frac{\sin\left[\frac{\omega Nd}{2c}\left(\sin\beta - \frac{2\pi kc}{\omega Nd}\right)\right]}{\sin\left[\frac{\omega d}{2c}\left(\sin\beta - \frac{2\pi kc}{\omega Nd}\right)\right]} \quad (1.3)$$

dove N è il numero di sensori; k è l'indice delle frequenze spaziali; d è la distanza tra due sensori adiacenti; $\omega = 2\pi f$ la velocità angolare; f la frequenza; c la velocità del suono nel mezzo; $\beta \in [-\varphi_{max}, \varphi_{max}]$ e $2\varphi_{max}$ è la copertura angola verticale del sonar.

E' evidente come l'equazione (1.3) sia del tutto identica alla (1.1) riguardante il *beam pattern* di un *beamforming system*. A tal riguardo, la relazione che lega l'indice k alla direzione di puntamento θ si può esprimere come:

$$\sin\theta = \frac{2\pi kc}{\omega Nd} \quad (1.4)$$

dall'equazione (1.4) si evince allora come questa particolare tipologia di sistema soffra delle stesse problematiche presenti nei *beamforming system*, cioè il problema dei lobi secondari, problemi di *aliasing* per $d > \lambda/2$ e una non perfetta messa a fuoco. Il *beam pattern* risulta identico a quello mostrato in Fig. 1.3.

Nella seconda tipologia di sistemi olografici il problema è riformulato in termini di equazioni matriciali. Il campo di pressione $s(\omega, p)$ è un vettore $N \times 1$, dove N rappresenta il numero di sensori piazzati in p_l con $l = 1 \dots N$; il vettore $c(\omega, r)$ di

dimensione $M' \times 1$, contiene invece il valore di riflettività di ogni cella nello spazio immagine (Fig. 1.4).

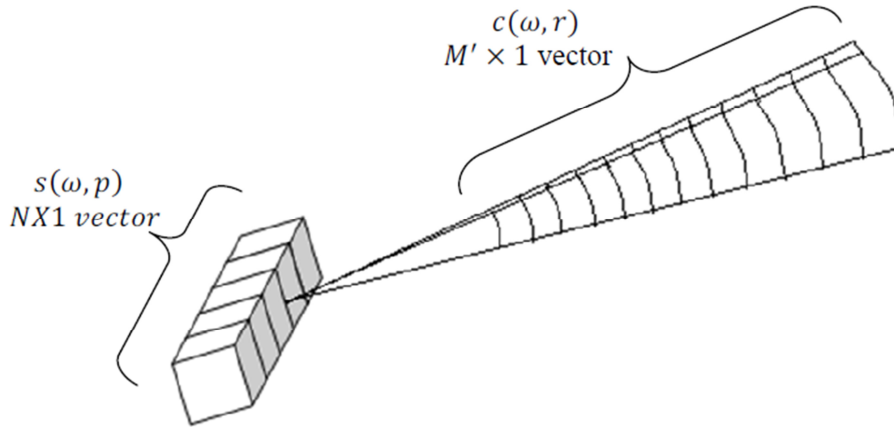


Fig. 1.4: rappresentazione delle matrici nella metodologia olografica per un singolo fascio. La figura rappresenta un *array* lineare. Nei sonar 3D l'*array* è bidimensionale.

La relazione tra queste matrici può essere espressa come:

$$s(\omega, p) = U(\omega, p, r)c(\omega, r) \quad (1.5)$$

dove $U(\omega, p, r)$ è una matrice di trasferimento di dimensione $N \times M'$.

Nell'equazione (1.5), l'intero problema di *imaging* è trasformato in una stima del vettore $c(\omega, r)$ data la matrice di trasformazione $U(\omega, p, r)$ e il vettore delle misure $s(\omega, p)$. La tecnica SVD è utilizzata in questo caso al fine di determinare la matrice pseudo – inversa di $U(\omega, p, r)$, così che il vettore stimato $\tilde{c}(\omega, r)$ si possa esprimere come:

$$\tilde{c}(\omega, r) = U^H(U \cdot U^H)^+s \quad (1.6)$$

dove H rappresenta la complessa coniugata e trasposta mentre $^+$ la pseudo – inversa della matrice.

L'equazione (1.6) di solito è scritta per ogni *layer* dello spazio immagine il quale risulta essere suddiviso in molti *layer* sferici concentrici al fine di minimizzare la dimensione delle matrici calcolate.

1.1.2.1.3 Acoustic lens systems

I sistemi a lenti acustiche lavorano in maniera simile a quelli ottici. Di base, questi sistemi usano lenti acustiche al fine di focalizzare i segnali di eco su una retina 2D di sensori. Ogni sensore riceve un segnale che rappresenta l'eco proveniente da una direzione ben definita. La misura del tempo di volo di questo segnali consente di costruire immagini 3D piuttosto che 2D come avviene per i sistemi ottici.

Lo sfasamento che nei sistemi olografici è realizzato via *software*, in questo caso avviene in maniera naturale quando i segnali passano attraverso le lenti. Nei sistemi a lenti acustiche il campo d'acquisizione è limitato dall'apertura delle lenti, così che esso può essere calcolato utilizzando la seguente equazione:

$$S'(\omega, x) = \begin{cases} S(\omega, x)e^{j\omega x^2/2CF} & -A/2 \leq x \leq A/2 \\ 0 & \text{altrove} \end{cases} \quad (1.7)$$

dove F è la lunghezza focale delle lenti ed A la loro apertura.

Il *beam pattern* dei sistemi a lenti acustiche si può esprimere in forma chiusa come:

$$BP_{lens}(\omega, \beta, \theta) = 2 \frac{\sin[\omega A(\sin\beta + \sin\theta)/2c]}{\omega A(\sin\beta + \sin\theta)/2c} \quad (1.8)$$

Il *beam pattern* per questa tipologia di sistemi risulta molto simile a quello dei sistemi *beamforming*; tuttavia, i sistemi a lenti acustiche soffrono di alcune problematiche dovute alla natura fisica delle lenti stesse: la regolazione della distanza di messo a fuoco tale che essa sia prossima a quella di diffusione richiede una complessa disposizione meccanica del sistema ed inoltre lo sfasamento prodotto dalle lenti è applicabile solo nella zona di approssimazione di Fresnel.

1.1.2.1.4 Problemi nei sistemi acustici di *imaging* 3D con array planari

Lo sviluppo di sistemi acustici di *imaging* 3D attraverso l'utilizzo di *array* planari è ostacolato principalmente da due fattori: prima di tutto, il costo di realizzazione dell'*array*, che si compone di un numero elevato di trasduttori acustici; in secondo luogo, il carico computazionale necessario ad elaborare i segnali ottenuti dai sensori acustici stessi.

Nella progettazione dell'*array* di sensori bisogna inoltre considerare il problema dell'*aliasing* il quale impone un vincolo, espresso in termini di lunghezza d'onda, sulla distanza massima tra due sensori adiacenti ($\lambda/2$). D'altra parte, al fine di ottenere una risoluzione elevata, è necessario realizzare un *array* con un'ampia estensione spaziale. Da quanto detto emerge che per soddisfare i requisiti sopra esposti l'*array* planare sarà composto da migliaia di sensori acustici. Nei sistemi reali, il problema della progettazione dell'*array* viene affrontato come un problema di ottimizzazione al fine di ridurre il numero di sensori acustici e di trovare le posizioni e i pesi ottimali; gli obiettivi del problema di ottimizzazione sono quelli di abbassare il livello dei lobi laterali, di ridurre la larghezza del lobo principale e di evitare il problema dei *grating lobe* nel *beam pattern*. Questa procedura riduce il costo della matrice planare e tronca l'onere computazionale richiesto per produrre la mappa 3D.

L'*imaging* 3D subacqueo eredita molte delle problematiche dell'*imaging* 2D, cioè *speckle noise*, riflessioni speculari, coerenza spaziale e temporale del mezzo trasmissivo. Lo *speckle noise* è causato dall'interferenza dei segnali di eco provenienti da superfici rugose. Esso può essere notevolmente attenuato grazie all'uso di filtri adattativi [14] o attraverso l'impiego di tecniche statistiche [15]. D'altra parte, la riflessione speculare si manifesta principalmente nell'*imaging* di oggetti fatti dall'uomo, poiché questi tendono ad avere superfici lisce che portano alla generazione di immagini a basso dettaglio e con un alto contrasto. Infine, le proprietà spaziali e temporali del mezzo tendono a distorcere le immagini prodotte nel caso di scansioni ad ampio *range*.

1.1.2.1.5 T Configuration

Molti sistemi sonar sfruttano la dualità di puntamento dell'approccio *beamforming* sia in trasmissione che in ricezione al fine di costruire un'immagine 3D attraverso una scansione combinata della scena. In questi sistemi, un *array* di trasmissione verticale illumina una sottile *slice* in elevazione, mentre un *array* di ricezione orizzontale acquisisce i segnali provenienti da differenti angoli di azimut attraverso ogni *slice* emessa. Questa configurazione è chiamata "*T Configuration*" o "*Mill's Cross Configuration*"; il suo principale vantaggio risiede nel basso numero di trasduttori e canali di trasmissione/ricezione il che si traduce in un sistema più economico e

semplice al costo comunque di avere tempi di scansione 3D molto più lunghi se comparati a quelli ottenibili con un sistema formato da un *array* 2D di sensori. Ad esempio, per ottenere 128 x 128 punti ad una distanza di 100 m il sistema ha bisogno di emettere 128 impulsi successivi con un ritardo minimo di 0.13 s impiegando, dunque, 17.06 s per una scansione completa della scena 3D.

1.1.2.1.6 Concentrator lens

Questa tipologia di sistemi può essere considerata come la versione meccanica degli *Acoustic lens systems*. In questi sistemi, i dati 3D subacquei possono essere generati a partire da un sonar 2D equipaggiato con queste lenti in modo da formare *pencil beams* attraverso i quali effettuare un'intera scansione 3D del campo di vista del sistema ruotando il sonar di 45° gradi. Il sistema di lenti riduce la larghezza di banda verticale dai 14° gradi di *default* fino a 3°/1° grado, dipendentemente dalla particolare soluzione adottata. Un esempio di questa tipologia di sonar è rappresentato dal sistema DIDSON [16]. Al fine di compensare i leggeri movimenti del sistema, questa metodologia richiede l'utilizzo di sensori GPS e di movimento per la misura degli angoli di *heading*, *pitching* e *rolling*. Inoltre, poiché si utilizza un sistema meccanico di rotazione, i tempi di scansione dell'intero campo di vista risultano maggiori se comparati rispetto a tutti gli altri sistemi di acquisizione acustica 3D.

1.1.2.1.7 Sistemi batimetrici

I *Multibeam echosounder* (MBES) sono sistemi tipicamente impiegati al fine di generare mappe batimetriche del fondale. La "spazzata" del sonar è prodotta da centinaia di *narrow beams* adiacenti secondo una configurazione a "ventaglio" (Fig. 1.5). La mappa batimetrica 3D impiegando tale tipologia di sistema è generata come risultato del movimento della piattaforma sulla quale il sonar è installato.

Vi sono due modalità attraverso le quali poter elaborare i dati acquisiti dal MBES al fine di generare mappe 3D del fondale. Nel primo approccio i punti 3D acquisiti vengono filtrati in modo da eliminare i valori anomali presenti, mentre la superficie generata viene resa più liscia attraverso la generazione di una griglia ottenuta a partire dai dati filtrati. In altre parole, l'interpolazione è eseguita sui punti 3D filtrati con lo scopo di

produrre un DTN (*Digital Terrain Model*) dell'area acquisita. Differenti algoritmi appartenenti a questa modalità di elaborazione sono stati proposti in letteratura, alcuni basati sull'uso di filtri passa banda per la rimozione dei valori anomali e su strutture dati di tipo *wavelets* o *quad-tree* per la generazione delle superfici [17].

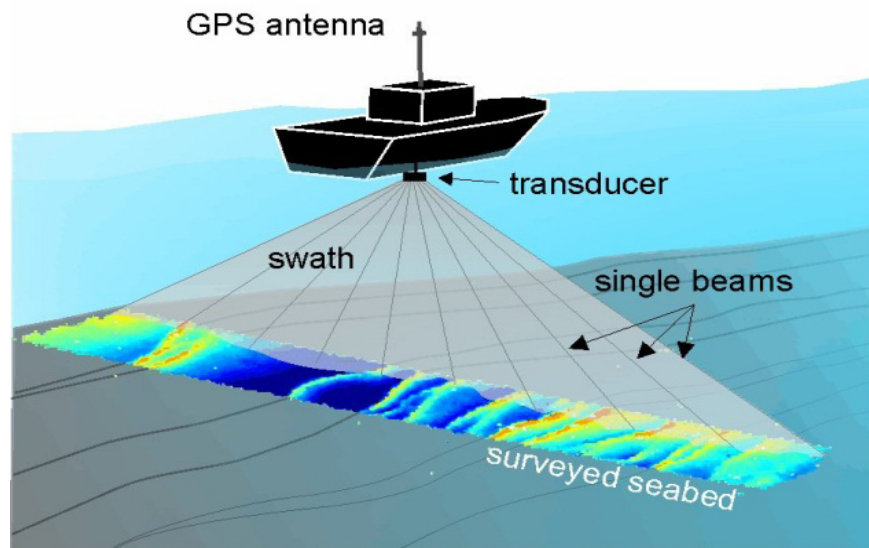


Fig. 1.5: mappa batimetrica 3D generata utilizzando un *Multibeam Echosounder*.

Il secondo approccio modella il fondale marino utilizzando una TIN (*Triangulated Irregular Network*) [18]. In questa metodologia i dati 3D vengono filtrati come in precedenza ma usa tali dati direttamente senza l'uso di nessuna griglia. L'algoritmo TIN genera modelli 3D con migliori dettagli rispetto al caso precedente al costo di avere comunque una implementazione più complessa [19].

1.1.2.1.8 Sistemi interferometrici

I sonar interferometrici utilizzano misure di fase per determinare l'angolo di arrivo del segnale di eco. Gli elementi riceventi, posti a distanza nota, determinano l'angolo di arrivo θ calcolando l'*offset* di fase dell'eco di ritorno. Questi sistemi non basano il loro funzionamento sulla metodologia di *beamforming*, ma piuttosto su una misura accurata della fase del segnale ricevuto su ogni singolo elemento, generando in tal modo mappe batimetriche combinando le misure dell'angolo di arrivo θ con misure di distanza. Questa tipologia di sonar è stata utilizzata per decenni anche se problemi di rumore, sia interno che ambientale hanno fortemente inciso sul loro sviluppo. Ulteriori

problemi che affliggono tali sistemi riguardano la necessità di discriminare tra molteplici angoli di arrivo (differenti segnali di eco provenienti da diverse direzioni) e l'effetto dello *shifting* del *footprint* (cioè la ricezione simultanea su più elementi riceventi di segnali di eco prodotti da zone vicine di fondale) [20]. Comunque, i sistemi interferometrici sono stati notevolmente migliorati negli ultimi anni grazie all'introduzione di nuove tecnologie nel campo elettronico e nella realizzazione dei trasduttori, nonché negli algoritmi di calcolo della fase dei segnali di eco.

2.1.1.6 LIDAR

I sistemi LIDAR (*Light Detection And Ranging*) sono stati inizialmente sviluppati per mappatura idrografica e sorveglianza. Questi sistemi sono stati in uso per diverse decadi ed hanno permesso il miglioramento della tecnologia laser ad alta potenza. Lo stato dell'arte dei LIDAR idrografici commercialmente disponibili in commercio consentono misure di profondità dalla superficie marina a circa 50 m con una precisione di 0.25 m [21], con una potenza tipica di 2 kW (Fig. 1.6).

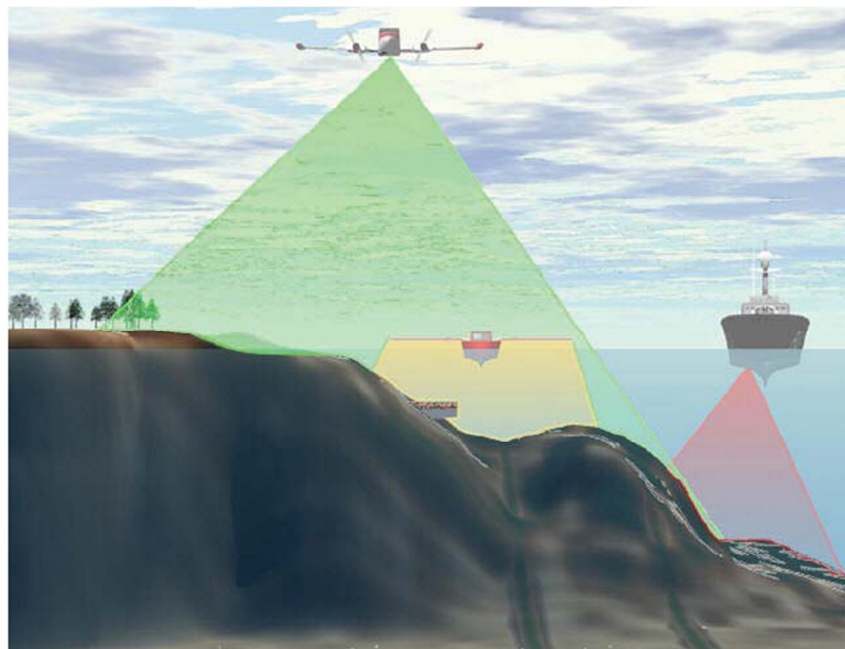


Fig. 1.6: Telerilevamento mediante LIDAR.

I sistemi *range gated* possono essere usati all'interfaccia aria-mare. Il grande vantaggio dei sistemi impiegati in aria è che hanno la capacità di accoppiare la velocità di

ispezione di elicotteri o piccoli aerei con la velocità di *imaging* della luce; ciò potrebbe fornire un modo potenzialmente molto rapido per l'analisi di oggetti subacquei, ma che richiederebbe comunque che il sistema lavorasse attraverso l'acqua marina. Dal momento che ciò comporterebbe lo studio di fenomeni variabili sia col tempo che con lo spazio, dei quali sono note solo considerazioni statistiche, il problema non si presenta di facile soluzione [22].

Il *Magic Lantern System* della "Kaman Aerospace" (Tucson, Arizona) è stato usato con successo nella Guerra del Golfo per rilevare la presenza di mine [23]. Un altro set di sistemi è stato impiegato per l'indagine di pesci epipelagici [24]. Variazioni di questi sistemi vanno dal semplice *hardware* di scansione, con PMT *gated*, a dispositivi che sono piuttosto simili a quelli *range gated*. I risultati, al momento, hanno presentato immagini *fuzzy* che potrebbero esser adattate alla stima delle dimensioni e della densità di pesci o altri oggetti; acquisire immagini dei dettagli di oggetti sommersi è un problema più complesso.

Nel tipo più semplice di sistemi *range gated*, viene usato un breve impulso di luce per creare una lastra di illuminazione che viene quindi riflesso dagli oggetti di interesse. Attivando la fotocamera ad un istante preciso, valutato in base al *range* di interesse, si ottiene un determinato ritardo del cammino ottico e relative informazioni prive di *backscatter*. Tali sistemi risultano essere soggetti al fenomeno del *blur* a causa del *forward scattering*, ma sono comunque pienamente in grado di fornire buone immagini a *range* estesi. Un sistema che è stato posto sotto sviluppo dai ricercatori canadesi usa un approccio *range gated* in coppia con acquisizione mediata per formare immagini *range gated*. Il sistema LUCIE (*Laser Underwater Camera Image Enhancer*) impiega un laser che emette brevi impulsi di luce che sono quindi riflessi dai *target* ed acquisiti da una fotocamera ICCD. Fotocamera e laser sono sincronizzati in modo tale che la prima si attivi dopo uno specifico ritardo nel percorso di propagazione della luce. La luce, il cui tempo per effettuare il proprio percorso circolare viene confrontato con quello della generazione del laser e quello di accensione del sistema di acquisizione, viene percepita e di conseguenza è possibile registrare un'immagine relativamente priva di *backscatter*. Un aspetto ingegnoso del sistema LUCIE è che viene utilizzata una frequenza di ripetizione molto elevata di 2000 impulsi al secondo, ciò permette un'acquisizione mediata *hardware* della luce raccolta, catturando immagini ripetute e

compensando il basso rapporto segnale-rumore dei sistemi precedenti che usavano una sola immagine.

Un prototipo più nuovo, attualmente sotto sviluppo, userà un laser più potente, accrescerà il campo di acquisizione e fornirà perciò immagini di maggiore chiarezza in condizioni più torbide.

2.1.1.7 Time Resolving Imaging

Uno sviluppo recente nell'acquisizione *time resolved* (a separazione temporale) è un sistema che utilizza un raggio conico di luce laser a impulsi in coppia con *hardware* a ricezione rapida che permette di impiegare la tecnica di *time resolution* dell'illuminazione riflessa.

Sviluppato da Arete Associates, lo *Streak Tube Imaging Lidar* (STIL) ha dimostrato di recente capacità di *imaging* 3D ad elevata risoluzione per l'identificazione elettro-ottica (EOID) di oggetti subacquei. Come un sistema LIDAR, lo STIL misura il tempo di volo; una nuova caratteristica è invece costituita dalla capacità di misurare anche l'ampiezza del segnale che ha subito *backscatter*. Questi dati possono essere elaborati per formare sia un'immagine di contrasto che un'immagine di *range* (un'immagine di *range* differisce da una convenzionale immagine a gradazioni di grigio per il fatto che a ogni pixel non è associato un livello di intensità luminosa, bensì una misura della distanza tra il sensore e un punto della superficie di un oggetto presente nella scena osservata. In Fig. 1.7 è rappresentato schematicamente il funzionamento tipico dello STIL. Di tali sensori esistono molti tipi diversi, che si trovano di sovente utilizzati in combinazione), al fine di fornire una piena rappresentazione 3D della scena subacquea. Per l'EOID, gli obiettivi di progetto sono una risoluzione laterale di 1,25 cm, e 2,5 cm di risoluzione di *range*.

Lo STIL è un sistema di *imaging* attivo che usa un trasmettitore di laser a impulsi ed un ricevitore *streak tube* per applicare il *time resolving* alla luce retrodiffusa. Il raggio laser è diffuso in una dimensione usando una lente cilindrica per formare un raggio conico. La luce sottoposta a *backscatter* viene acquisita da una lente convenzionale posizionata su una fenditura posta di fronte al fotocatodo dello *streak tube*, ed è sottoposta a *time/range resolving* tramite deflessione elettrostatica all'interno dello *streak tube*, generando un'immagine di *range-azimuth* ad ogni impulso laser.

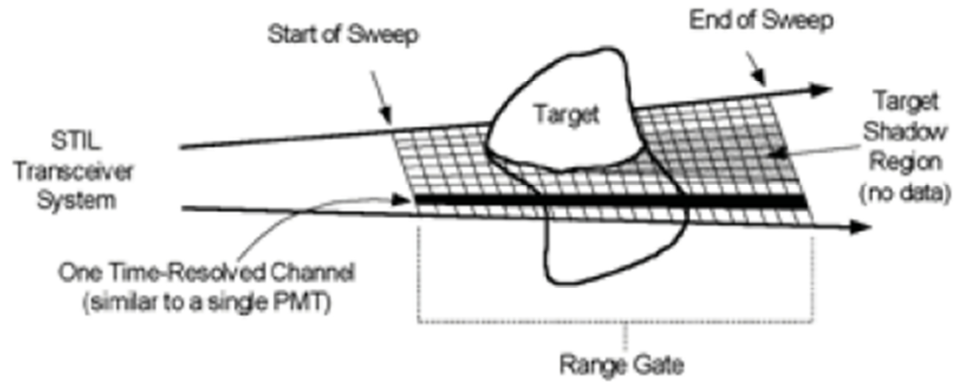


Fig. 1.7: rappresentazione schematica della modalità di funzionamento tipica dello STIL.

Orientando il raggio conico perpendicolarmente alla traccia del veicolo, viene campionata la dimensione trasversale intervenendo sulla frequenza di ripetizione dell'impulso (PRF) del laser alla velocità di avanzamento del veicolo, spazzando quindi il volume oceanico in modo *pushbroom*. Questa implementazione priva di scanner garantisce piena copertura del volume senza bisogno di ricorrere a componenti mobili.

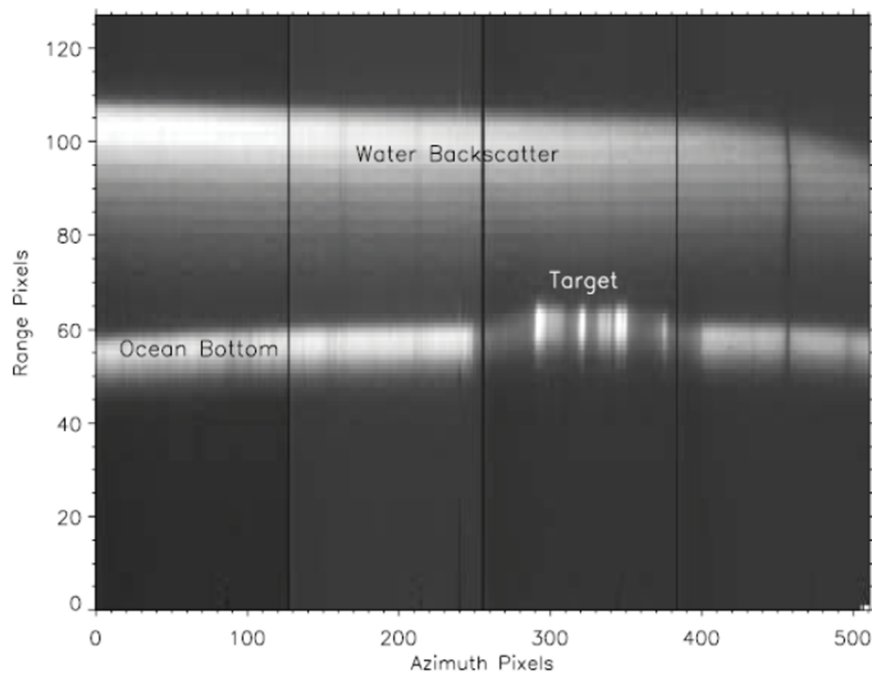


Fig. 1.8: Immagine *range-azimuth* ricavata da un'unica acquisizione (a sinistra), e profilo del range, rappresentante il decadimento esponenziale del *backscatter* dell'acqua, informazioni sul range ottenute dal tempo di volo e sul contrasto ricavate dall'ampiezza dell'informazione di ritorno dal fondale.

La Fig. 1.8 è un'immagine *range-azimuth* derivante da una singola acquisizione che illustra il decadimento esponenziale del *backscatter* dell'acqua e l'informazione di ritorno dal fondale, ed evidenzia come il *backscatter* sia separato temporalmente da quest'ultimo, causandone l'esclusione. Questo preciso campionamento temporale rende inoltre il sensore immune alla luce solare circostante. L'informazione di ritorno include sia dati sul tempo di volo che forniscono a loro volta una misura quantitativa dell'altezza dell'oggetto posto sul fondo (in questo caso un tronco di cono), sia il livello radiometrico, che è proporzionale alla riflettività dell'oggetto posto sul fondo. Ogni *shot* laser determina quindi distanza e contrasto del fondale, per ogni pixel posto lungo la direzione trasversale al moto. Mediante l'allineamento dell'informazione contenuta da ogni pixel con la direzione longitudinale al moto, è possibile ottenere immagini di *range* e di contrasto della scena sommersa (Fig. 1.9).

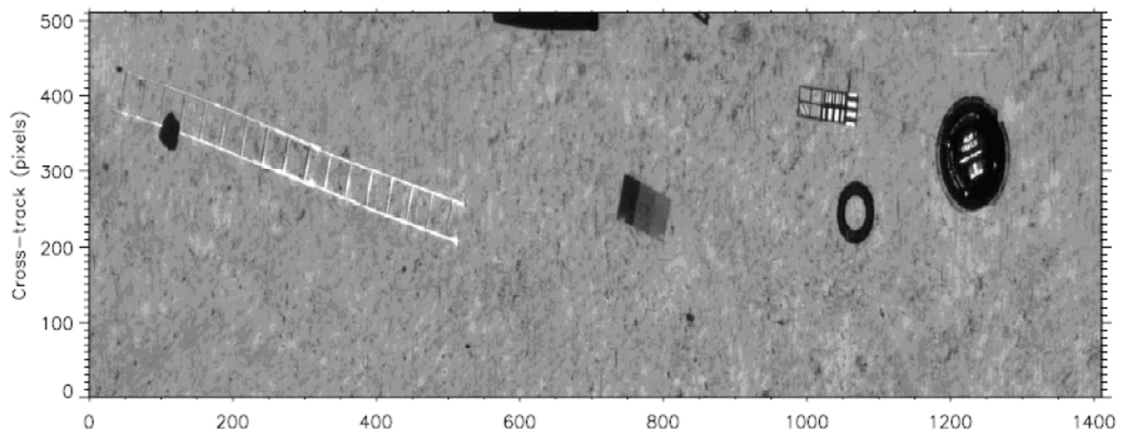


Fig. 1.9: Immagine risultante da più acquisizioni.

Sotto la sponsorizzazione della ricerca dell'*Office of Naval*, la tecnologia STIL si è evoluta dal suo status di semplice prova di laboratorio di un concetto: la prima generazione del sistema fu sviluppata nel 1995 usando un sistema *streak tube* disponibile in commercio per dimostrare, in laboratorio, la capacità di *imaging 3D*; un sistema di seconda generazione venne poi realizzato nel 1998 per una configurazione di impiego subacqueo.

La terza generazione rappresenta una configurazione compatta, adatta alla sorveglianza di ambienti oceanici di elevata profondità. In collaborazione con la *Coastal System Station* della Marina Militare, questo sistema integrato fu testato

durante l'estate 2000 nel Golfo del Messico; esso fu schierato sia da una nave che da un elicottero, a velocità superiori ai 25 nodi.

Sotto le condizioni di acqua relativamente chiara riscontrate durante l'esperimento, la qualità delle immagini risultava buona, poterono essere identificati in maniera chiara oggetti posti in profondità. Le prestazioni in condizioni di acqua torbida subivano invece, come ipotizzato, un degrado significativo. Questa è l'unica tecnica che garantisce un'immagine 3D regolare di target sommersi e che fornisce la profondità reale. Si deve inoltre considerare la possibilità di utilizzare un PMT al posto dello *streak tube*; quest'ultimo è però preferibile in quanto fornisce una serie di vantaggi, tra i quali la possibilità di utilizzare centinaia o migliaia di canali *time-resolved* funzionanti contemporaneamente, il che permette di raccogliere molte più informazioni per ogni impulso laser in confronto al PMT, che è un dispositivo dotabile di un unico o comunque di pochi canali.

Inoltre, in termini di intensità, uno *streak tube* garantisce un *range* dinamico molto più ampio rispetto ad un PMT. Per ciò che riguarda la risoluzione, tenendo conto del fatto che essa varia anche in funzione delle condizioni dell'acqua (torbidità e livello di *scattering*), sotto l'ipotesi di perfetta limpidezza si potrebbe raggiungere il valore di 0,5 in (12,7 mm), con dispositivo agente ad una distanza dal fondale marino pari a 30 piedi, ossia poco più di 9 metri.

I laser impiegati sono quelli a Nd:YAG a 532 nm, molto vicini al picco della curva di trasmissione per l'acqua oceanica; la potenza è dell'ordine dei 5 mJ per impulso, la cui durata è normalmente di 5 ns. Attualmente il dispositivo è impiegato nell'ispezione dei fondali marini finalizzata alla ricerca di mine presenti sul fondale marino, essendo particolarmente adatto a tale scopo grazie all'elevata risoluzione; in questo campo, infatti, altri sistemi hanno dimostrato di essere molto meno indicati, procurando una serie di falsi allarmi. Le mine giacenti sul fondale risultano essere una seria minaccia per la navigazione: esse stanno passivamente alla ricerca della traccia acustica o magnetica di una nave passante e non causano danneggiamento tramite l'esplosione. Infatti, quando scoppiano, generano una bolla d'aria che si innalza e solleva la nave dall'acqua facendo in modo che si frantumi sotto l'azione del proprio peso. Altre applicazioni sono la batimetria in volo, che consiste nella misurazione della profondità del mare o di altri corsi d'acqua; ciò è particolarmente utile per il monitoraggio di

regioni costiere. Mentre una barca dotata di un sistema sonar non può coprire un'area vasta fornendo risoluzione elevata, un sistema aereo per la batimetria garantisce una copertura molto più ampia in modo molto rapido ed economico. Un altro campo è la mappatura terrestre; l'aspetto più interessante per chi si interessa di questo tipo di indagini è che non si ottiene solo una mappa della superficie, ma anche un'immagine dell'area, il che consente, mediante un confronto tra le due, di poter interpretare parti della mappa come ad esempio dossi.

2.1.1.8 Sistemi a scansione laser per linee

1.1.2.1.9 Scanning sincronizzato

Nel caso della scansione sincrona (*synchronous scanning*), un approccio noto come sistema a Scansione Laser per Linee (*Laser Line Scan, LLS*) è sotto sviluppo da poco più di dieci anni, principalmente da Brian Coles di Raytheon. Il sistema consiste in un raggio laser CW ad elevata collimazione che viene fatto passare attraverso uno specchio in rotazione, in modo tale che possa sondare il fondale marino in direzione perpendicolare rispetto al moto della piattaforma. Un set di rilevatori ottici effettua acquisizione sincrona in modo tale che venga catturata solo luce proveniente da una piccola area del fondale marino che per la maggior parte non ha subito *scattering*.

Uno sviluppo recente di questo sistema è una versione a colori che permette fino a quattro lunghezze d'onda diverse. Ogni ricevitore consiste di un gruppo ottico rotante di input, uno ad apertura controllabile, un PMT, un preamplificatore, di elettronica di processazione di segnali ed un convertitore analogico-digitale.

I quattro gruppi ottici di input ad uno di output sono montati su un albero guida (*drive shaft*) per assicurare la sincronizzazione meccanica dello spot laser sul fondale.

I gruppi ottici di input/output impiegano specchi a quattro facce, dando origine a quattro linee di scansione a 90° per ogni rotazione dell'albero. Le immagini ottenute per linee di scansione sono formate a partire dalla porzione centrale a 70° di ogni linea di scansione mediante digitalizzazione dell'output elettrico da ogni ricevitore ad un valore di 12 bit ad un numero di pixel selezionabile dall'utente (512, 1024, 2048, o 4096) per linea di scansione. Immagini bidimensionali vengono formate attraverso il moto della piattaforma, assicurandosi che le linee di scansione successive siano

effettivamente spostate l'una dall'altra. La dimensione dei pixel prodotti da tale sensore dipende dall'altitudine dello stesso, dalla velocità del rimorchio, dalla frequenza di scansione e dal numero di pixel per linea di scansione. La dimensione del pixel in direzione longitudinale è, in centimetri, $771 * \text{velocità (nodi)/rpm}$, dove rpm rappresenta il numero di giri al minuto dell'albero guida.

La velocità di scansione è definibile dall'utente fino ad un limite di 4000 rpm. La dimensione del pixel in centimetri, nella direzione trasversale, è di $122 * \text{Altezza (m) / Npixel}$, dove Npixel (512, 1024, 2048, o 4096) è il numero di pixel per ogni linea di scansione a 70°. Per esempio, dimensioni del pixel di 1 centimetro in entrambe le direzioni si ottengono ad un'altitudine di 8,5 m e ad una velocità del rimorchio di 5,2 nodi nel momento in cui lo scanner lavora ad un numero di giri di 4000 rpm e lungo una singola linea di scansione sono presenti 1024 pixel.

È interessante notare che questo sensore può assumere diverse configurazioni al fine di acquisire tipi di dati differenti: nella sua configurazione *Fluorescence Imaging Laser Line Scan* (FILLS), si utilizza un laser a ioni Argon a 488 nm, ed i quattro ricevitori sono equipaggiati con filtri di interferenza centrati rispettivamente sui valori (per esempio) di 680 nm, 488 nm, 515 nm, e 570 nm: in questa configurazione, il canale 2 acquisisce immagini attraverso luce diffusa elasticamente, mentre i canali 1, 3, e 4 creano immagini rispettivamente con luce rossa, verde e gialla. La sorgente di questa luce rossa, verde e gialla è in maniera predominante fluorescenza naturale stimolata dal laser a 488 nm (comunque, a causa della loro vicinanza spettrale, vi è una piccola fuoriuscita della luce diffusa elasticamente a 488 nm nell'immagine del canale verde). In una seconda configurazione, può essere adottato un laser gassoso ad Argon/Krypton, che produce simultaneamente output a 647 nm (rosso), 515 nm (verde) e 488 nm (blu). Quando tre dei ricevitori sono dotati di filtri di *matching*, è possibile ottenere i dati richiesti a creare immagini RGB. In questa configurazione non sono installati filtri nel canale rimanente. I sensori LLS riducono gli effetti dannosi del *backscatter* ed il *forward scatter*, nonché la sfocatura e la luminescenza, realizzando immagini attraverso: *a)* uno spot laser molto ridotto e *b)* sistema di ricezione a scansione sincrona basata su piccoli spot di ricezione. Come illustrato in Fig. 1.10, sul fondale, tali spot possiedono una forma approssimativamente rettangolare; la loro larghezza trasversale è tipicamente di 10 milliradiani o meno.

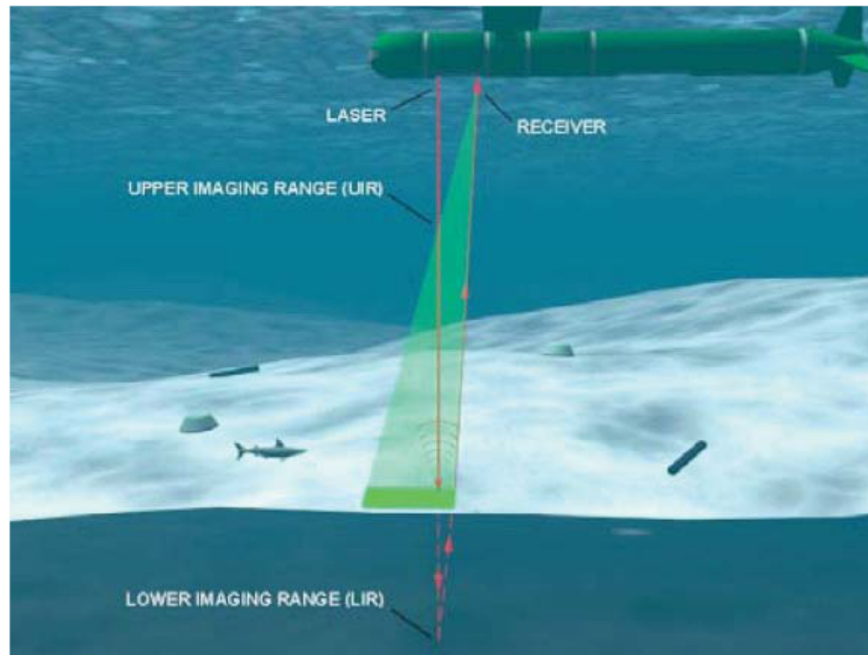


Fig. 1.10: Esempio di sistema LLS installato a bordo di un veicolo subacqueo (sinistra). Per ogni ricevitore, l'*upper imaging range* (UIR) è la distanza più vicina alla quale il campo di acquisizione dello scanner interseca il laser, mentre il *lower imaging range* (LIR) è la distanza più lontana alla quale si verifica l'intersezione tra il prolungamento del campo di acquisizione dello scanner ed il prolungamento del raggio laser.

La profondità di campo, controllabile dall'utente, ne determina invece la lunghezza longitudinale. La profondità del campo è controllata dai fattori di *upper imaging range* (UIR) e *lower imaging range* (LIR) del sensore, e viene settata in maniera tale da bilanciare l'altitudine del sensore. Da un punto di vista fisico, l'UIR ed il LIR di ogni ricevitore sono controllati attraverso componenti di apertura posti frontalmente ad ogni PMT. Per questo sensore, l'UIR minimo è di circa 15 piedi. Nella pratica, comunque, i limiti inferiore e superiore del campo di acquisizione sono stabilite dalle condizioni di purezza dell'acqua.

1.1.2.1.10 Luce strutturata: *light stripe scanning*

L'acquisizione subacquea ottica è un mezzo chiave che permette numerose applicazioni oceanografiche. Sicuramente una delle sfide principali poste dall'acquisizione ottica nelle suddette applicazioni è la forte degradazione della qualità dell'immagine per *scattering* dovuto alle impurità e agli organismi presenti nell'acqua. Inoltre, le sorgenti luminose naturali come la luce solare si attenuano completamente prima di raggiungere profondità significative. Per questo motivo, è stata sviluppata una

grande varietà di tecniche di illuminazione attiva (strutturata) da impiegare proprio in queste applicazioni, a partire dall'uso di sottili raggi laser e sorgenti confocali, fino a tecniche più sofisticate *time-gated* e *synchronization gated*.

Tali metodi migliorano la visibilità attraverso la riduzione degli effetti ottici dello *scattering*. In ogni caso, solo pochissime tecniche analizzano in maniera esplicita l'influenza dello *scattering* sull'aspetto delle scene sottoposte ad illuminazione attiva. L'illuminazione strutturata è inoltre sfruttata ampiamente nella *computer vision* per la ricostruzione 3D di scenari. Nella maggior parte dei casi, si assume comunque implicitamente che la luce non sia né diffusa né assorbita dal mezzo in cui la scena e la sorgente sono immerse (come avviene in aria pura), perciò, non è semplice tener conto degli effetti dello *scattering* nel momento in cui si applicano i metodi di luce strutturata in ambiente subacqueo.

Al fine di ottenere sia un miglioramento nella visibilità mediante l'uso di luce strutturata e di analizzare le immagini acquisite per recuperare le proprietà della scena e del mezzo, è necessario dare risposta a tre domande: innanzitutto, quali sono gli effetti dello *scattering* risultanti dall'interazione tra luce strutturata ed il mezzo e la scena; in secondo luogo, come superare tali effetti al fine di ottenere i risultati per cui i metodi basati sulla luce strutturata sono stati studiati in ambito tradizionale, nella *computer vision*; infine, se esistono informazioni aggiuntive da estrarre dagli effetti dello *scattering* che, mediante tecniche di *computer vision* generiche, non potrebbero essere rilevate. In riferimento alla tecnica della scansione a strisce di luce (*light stripe scanning*) è stato costruito un modello analitico di formazione dell'immagine [25] basato sulla fisica dello *scattering* singolo; questo modello descrive le interazioni della luce strutturata con il mezzo e la scena.

Mediante l'uso di questi o di strisce di luce, è possibile sviluppare un semplice algoritmo per rilevare in maniera affidabile gli oggetti ed ottenere così una ricostruzione 3D della scena, anche in presenza di forte *scattering*. Si deve considerare che, l'interazione tra la luce strutturata ed il mezzo consente di stimare le proprietà del mezzo stesso; tale risultato può essere sfruttato allo scopo di rimuovere gli effetti dello *scattering* e valutare quindi l'aspetto della scena come se questa fosse immersa in aria pura. Per verificare il metodo attraverso esperimenti reali, è stato costruito [25] un setup consistente in un serbatoio in vetro riempito con un mezzo dispersivo (latte

diluito), un proiettore DLP ed una fotocamera posta all'esterno del serbatoio, come rappresentato in Fig. 1.11.



Fig. 1.11: Setup per la *light stripe scanning*

Si noti che la calibrazione di questo setup richiede di gestire la rifrazione della luce alle interfacce vetro-aria; questa procedura di calibrazione è simile, nel principio, a [26] e non richiede né una calibrazione geometrica esplicita della fotocamera e del proiettore o la conoscenza di particolari punti di rifrazione o degli indici di rifrazione dei mezzi. I risultati mostrano elevata precisione in confronto alla realtà inerente alla terraferma (2% - 6% di errore RMS (*Root Mean Square*) relativo nella ricostruzione 3D su 6 diverse concentrazioni dei mezzi) che corrispondono a quelli ottenibili solitamente negli scenari subacquei. Sebbene l'attenzione sia stata focalizzata sulla scansione a strisce di luce, i risultati possono essere utilizzati per ampliare diverse altre tecniche di *imaging* come *grid coding* [27] *gated* [28] e a polarizzazione attiva [29, 30].

2.1.1.9 Sistemi fotogrammetrici

Con il termine fotogrammetria si intendono tutte quelle procedure che utilizzano immagini fotografiche di un oggetto per ricavarne la posizione, la forma e le

dimensioni. Effettuare il rilievo di un oggetto significa ricavarne la posizione spaziale di tutti i punti di interesse (Fig. 1.12).

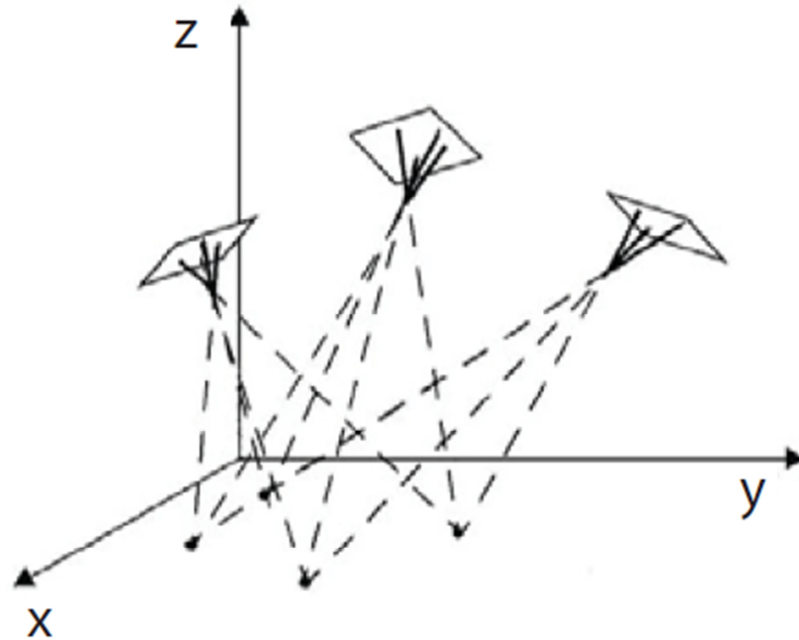


Fig. 1.12: Principio base della fotogrammetria, cioè osservazione di un oggetto da diversi punti di vista.

Mediante la fotogrammetria questa operazione viene fatta, in gran parte, non direttamente sull'oggetto, ma utilizzando le prospettive fotografate centrali mediante osservazione stereoscopica; essa è stata sviluppata molto prima della stessa invenzione della fotografia, e, sebbene sia nata per il rilievo architettonico, si è sviluppata principalmente per il rilevamento topografico del territorio, ed è stata applicata in gran parte come fotogrammetria aerea.

Malgrado l'elevato costo e la complessità degli apparecchi usati per svolgere le operazioni fotogrammetriche, come gli *stereorestitutori* (strumento atto a restituire le coordinate geometriche tridimensionali *stereo* da una coppia di prese fotogrammetriche), abbiano in passato ostacolato la diffusione dei metodi fotogrammetrici, oggi la fotogrammetria rappresenta uno strumento di acquisizione di dati metrici e tematici tra i più affidabili, rapidi ed economici; è il metodo più qualificato per la definizione spaziale degli elementi del territorio e per la misura delle loro più minute modificazioni.

Inoltre, le nuove modalità fotogrammetriche consentono descrizioni grafiche e numeriche più pertinenti, e assicurano la possibilità di istituire periodici e sistematici controlli ai fini della salvaguardia e conservazione di beni naturali, culturali, ambientali, e così via. In particolare, la recente disponibilità e diffusione di computer in grado di manipolare grosse quantità di dati, unitamente ai progressi tecnologici nella *computer graphics*, hanno permesso di eseguire in ambiente digitale, a costi inferiori ed in modo più semplice e rapido, il rilievo fotogrammetrico e di produrre elaborati numerici facilmente utilizzabili nell'ambito dei GIS (*Geographical Information System*). Questa evoluzione ha portato ad applicare le metodologie della fotogrammetria in settori dove questa raramente era utilizzata in passato (ingegneria, geologia, ambiente, ecc.). Solo negli ultimi anni, quindi, la fotogrammetria digitale ha conosciuto un vero sviluppo sia nell'ambito della ricerca che delle applicazioni; essa si configura come una tecnica innovativa che si basa sui principi della preesistente fotogrammetria analitica e si avvale in maniera sostanziale degli apporti delle tecnologie più avanzate nell'ambito dell'informatica, dell'elettronica e della *computer vision*.

La tendenza moderna della fotogrammetria vede ormai consolidato il passaggio a tecniche digitali, anche se molte delle procedure e degli algoritmi tipici della fotogrammetria tradizionale costituiscono ancora il punto di partenza. La fotogrammetria digitale è infatti, a tutti gli effetti, erede di quella analitica: da quest'ultima mutua la maggior parte degli algoritmi, da quello dell'orientamento interno alla *triangolazione aerea*, con l'unica differenza che tali algoritmi vengono applicati su immagini ricavate non da fotogrammi convenzionali, ma digitali; ovviamente, la transizione dalla fotogrammetria analitica a quella digitale richiede la modifica e l'estensione dei modelli esistenti e lo sviluppo di nuovi concetti. Come spesso accade per molte nuove tecnologie, anche la fotogrammetria digitale ha iniziato la sua storia risolvendo più velocemente e a costi più contenuti alcuni problemi che, in fotogrammetria analitica, richiedevano grandi sforzi economici e operativi. Per rendersi conto di questo fatto, è sufficiente pensare alle notevoli semplificazioni dei restitutori digitali, ormai costituiti da un software gestibile da un comune PC, alla possibilità di giungere all'individuazione automatica di coppie di punti omologhi nell'esecuzione di alcune fasi di orientamento. In più, la sempre maggiore comprensione delle logiche e delle potenzialità della fotogrammetria digitale, particolarmente nel

corso dell'ultimo decennio, ha consentito di delineare nuovi prodotti che la tecnica analitica non consentiva di concepire.

Per comprendere i nuovi sviluppi della fotogrammetria nel passaggio dall'analitico al digitale si devono analizzare differenze e analogie: la principale differenza si ritrova nella diversa natura del dato primario. In fotogrammetria tradizionale si utilizzano immagini fotografiche che, mentre nella fotogrammetria digitale il dato primario è costituito da immagini digitali derivanti dall'acquisizione diretta, con fotocamere digitali, o indiretta, con scanner, a partire dai fotogrammi tradizionali.

In fotogrammetria analitica l'operatore è tenuto a scegliere i punti necessari alle varie operazioni di orientamento e restituzione, a individuare punti omologhi e quindi a misurarne le coordinate dei punti *fiduciali*. In fotogrammetria digitale i punti possono essere ancora scelti dall'operatore, ma l'individuazione dei punti omologhi può essere automatica e le misure sono già state eseguite durante la fase di acquisizione delle immagini digitali. Tale differenza comporta una semplificazione degli strumenti: l'uso di fotogrammi tradizionali richiede infatti la presenza di uno strumento sofisticato di misura delle coordinate immagine dei punti omologhi e di un dispositivo che consenta la visione stereoscopica artificiale, mentre l'uso di immagini digitali può prescindere da entrambi; il restitutore digitale è costituito semplicemente da un PC e dal software.

I progressi tecnologici attuali forniscono innumerevoli mezzi per lo sviluppo di questa tecnica: ad esempio, per ciò che riguarda l'acquisizione dei dati, sono in continua evoluzione gli scanner per la digitalizzazione delle immagini tradizionali, le camere digitali e i vari sensori in grado di fornire immagini digitali di alta qualità e con un elevato contenuto di informazione geometrica. Il problema dell'immagazzinamento dei dati e il trattamento delle immagini in forma digitale è stato superato con la disponibilità di calcolatori con capacità di memoria e velocità di calcolo sempre maggiori. Si potrebbe concludere che la fotogrammetria digitale è ancora fotogrammetria analitica, poiché da essa eredita tutto il patrimonio di algoritmi, da quello dell'orientamento interno alla triangolazione aerea, con la differenza che tali algoritmi vengono applicati su dati di misura effettuati non su fotogrammi convenzionali, ma su immagini digitali, con uno strumento che si basa sulla tecnologia del computer, anziché quella tradizionale ottico-meccanica. Il sistema fotogrammetrico digitale (*hardware e software*) per la gestione delle fasi di presa e restituzione) è

formato da due unità logicamente distinte: l'unità di acquisizione e l'unità di restituzione.

L'unità di acquisizione fornisce l'immagine digitale in modo che, ad ogni pixel, possa essere immediatamente associata la coppia di coordinate riferite al sistema *fiduciale*, già corrette da errori sistematici (distorsioni ottiche, errori geometrici e radiometrici dello scanner, ecc.). L'unità di restituzione gestisce le immagini digitali provenienti dal sistema di acquisizione, consentendo di eseguire tutte le operazioni di tipo fotogrammetrico, come orientamenti, acquisizione dati, restituzione, ecc. In Fig. 1.13 è mostrata una ricostruzione di un modello 3D, per mezzo di stereo fotogrammetria digitale. Infine, bisogna ricordare che di recente la stereofotogrammetria viene impiegata insieme a tecniche basate sulla luce strutturata [31].

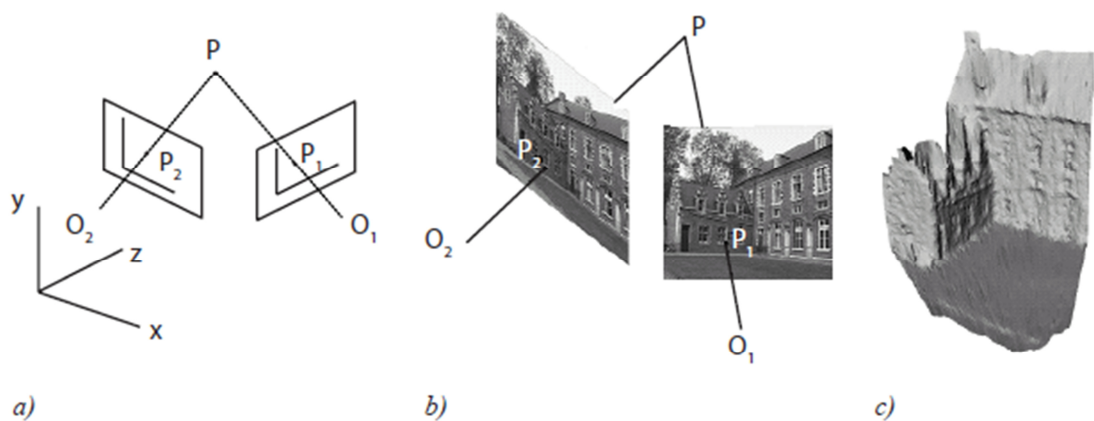


Fig. 1.13: a) Caso generale della *stereo-fotogrammetria*; b) immagini acquisite; c) modello 3D.

1.1.2.1.11 Fotogrammetria digitale subacquea

La ricerca in atto fa uso di sistemi video subacquei e tutt'ora di sistemi fotogrammetrici per ottenere, di un oggetto, misure quantitative come dimensioni, area, volume, forma, densità e movimento, così come le distanze da un punto o un piano di un oggetto o superficie.

Metodi automatizzati richiedono invece, il riconoscimento e la segmentazione di molteplici spot e raggi laser o, in alternativa, potrebbero usare le informazioni di una fotocamera e di una lente in combinazione con algoritmi che scartino regioni fuori fuoco. Dal momento che questi sistemi offrono un'aggiunta a basso costo e basso impatto a quello video standard, è facile registrare un'applicazione sempre crescente

di questa tecnologia nel campo della gestione delle risorse marine, in particolare la documentazione, il monitoraggio e la stima di lungo termine dell'estensione geografica di habitat critici. Molte di queste applicazioni richiedono specializzazione unica (ossia trattamenti di immagini sequenziali automatizzati basati sulla distanza percorsa, correzione dell'immagine per rollio e beccheggio del veicolo, inclusione di metadati dai sensori del sistema, ecc.) che non è possibile trovare in sistemi disponibili in commercio come "C-Map Systems Video Ruler", "Tritech ISS-Soft", e "Eos Systems PhotoModeler Pro" [32]. Per esempio, i ricercatori al *CSIRO Marine Research* (CMR), in Australia, usano un sistema con doppia video camera con quattro laser di riferimento per ognuna, equipaggiato per un piccolo ROV per studiare le dinamiche di ecosistemi bentici. In questa applicazione, viene usato un tool aggiuntivo Optimas™ di elaborazione di immagini sviluppato alla MBARI, chiamato *Laser Measure*, al fine di estrarre misure quantitative dalla collezione di immagini riferita al laser [21]. Nuovi algoritmi sono stati attualmente aggiunti all'insieme delle funzionalità, perché determinino la lunghezza focale dello zoom mediante la posizione di un singolo laser. Simile a questo è un sistema a tre laser di riferimento, *3-Beam*, sviluppato per i ricercatori del *Washington State Department of Fish and Wildlife* al fine di valutare la densità di popolazione di pesci che vivono in prossimità del fondale [33].

Questo sistema consente l'elaborazione automatica di immagini in sequenza sulla base del tempo, della distanza percorsa o per campionamento casuale. Le immagini sono collezionate ad un angolo di inclinazione valutato tenendo conto del rollio e del beccheggio del veicolo, e i dati circa la posizione del veicolo sono conservati per ogni immagine, nel caso sia necessario il georeferenzamento.

In un viaggio nel maggio 2005 dei ricercatori del *Canadian Department of Fishers* (DOF), i video ed i dati osservati furono raccolti a bordo del sommergibile "Nuytco's Aquarius" per confrontare le stime di densità di popolazione della specie ittica di pesce *Carpiodes cyprinus* su un'area di ispezione ottenute con metodi *transect* e *light striping*. Un quarto laser, parallelo ma fuori asse, fu aggiunto al sistema *3-Beam* per derivare la misura della rugosità locale della superficie bentica per dare assistenza nella classificazione degli habitat [21].

Altre tecniche per ottenere misurazioni di rugosità o di rilievi sono stati sviluppate recentemente; esse impiegano schemi di raggi laser [34, 35]. Due sistemi basati su

cinque laser di riferimento montati sui “DOE Phantom XTL ROV”, che sono anche usati per ispezioni del fondale marino, sono l’“Autonomous Benthic Image Scaling System” (ABISS) dell’Università di Plymouth [36] ed il sistema *5-Beam* del “NOAA Southwest Fisheries Science Center” [37].

In situazioni in cui la scansione a strisce di luce è un processo troppo lungo per essere considerato pratico (ad esempio nel caso di scene dinamiche), la stereofotogrammetria [38] risulta essere una valida alternativa. Tradizionalmente, essa si configura come una tecnica per la ricostruzione di una scena (normale alla superficie e albedo) a partire da un numero ridotto di immagini della scena stessa acquisite mediante direzioni di illuminazione differenti. In letteratura esistono diverse varianti di tale tecnica [38, 39], ma nessuna di queste risulta essere efficiente in un mezzo di trasmissione dispersivo [25]. La scelta più semplice del problema, assume che le superfici siano lambertiane (disperdono la stessa intensità lungo tutte le direzioni) le sorgenti distanti, le interreflessioni trascurabili e la fotocamera ortografica. In assenza di *scattering*, è noto che tre immagini di una scena illuminata da direzioni diverse note sono sufficienti per determinare univocamente le normali alla superficie e gli albedo [29]. Verrà perciò in un primo momento determinato il numero di sorgenti necessarie in presenza di *scattering*, quindi mostrato come le proprietà della scena possano essere recuperate dalle immagini corrispondenti.

1.2 Data Fusion

Col termine *data fusion* ci si riferisce al processo d’integrazione multipla di dati (cioè fusione di diverse tipologie di dati) riguardanti lo stesso oggetto reale, in una rappresentazione consistente, accurata ed utile.

In particolare, si definisce *multisensor data fusion* una tecnologia capace di combinare informazioni di diverse sorgenti (strumenti e dispositivi) allo scopo di generare un’unica rappresentazione [41]. A tale scopo, è necessario studiare metodi efficienti per attuare in maniera automatica o semi-automatica la trasformazione delle differenti tipologie di informazione in modo tale da fornire una rappresentazione efficace, migliore e più completa della misura che si intende effettuare. Il processo di *data fusion* richiede una ricerca multidisciplinare che prende in prestito idee e metodi da

molti altri campi come quello del *signal processing*, dell'*information theory*, della *statistical estimation* e dell'intelligenza artificiale. Le tecniche di *data fusion* sono ampiamente utilizzate in vari settori quali il *network* di sensori, la robotica, il video e l'*image processing*, giusto per citarne alcuni.

In letteratura esistono varie "concettualizzazioni" del processo di *fusion* di dati. La più comune e nota concettualizzazione di sistemi di fusione è rappresentata dal modello JDL (Joint Directors Laboratories) [42] la cui classificazione è basata in funzione della tipologia di dati in *input* e sui relativi *output* prodotti ed utilizzata, principalmente, in ambito militare. Il modello JDL originale considera il processo di fusione in quattro crescenti livelli di astrazione: oggetto, situazione, impatto e processo di rifinitura. Questa formalizzazione è focalizzata sui dati (*input* ed *output*) piuttosto che sul *processing*. Un'alternativa è rappresentata dal *Framework* di Dasarathy [43], che modella il sistema di fusione da un punto di vista dell'ingegneria del *software*, cioè come un flusso dati caratterizzato da *input/output* così come da "funzionalità" (processi). Una diversa concettualizzazione si trova nel lavoro presentato da Goodman [44] che è basata sulla nozione di *random sets*. Gli aspetti distintivi di questo *framework* sono l'abilità nel combinare l'incertezza nelle decisioni con le decisioni stesse, presentando uno schema generico di rappresentazione d'incertezza. Uno dei più recenti e astratti *framework* di fusione è proposto da Kokar [45]. Questa formalizzazione è basata sulla "teoria delle categorie" ed è vincolata ad essere sufficientemente generale da includere tutti i tipi di fusione: *data fusion*, *feature fusion*, *decision fusion* e fusione di informazioni relazionali. Essa è da considerare come il primo step verso lo sviluppo di una teoria formale di fusione. La maggiore novità di questo lavoro è l'abilità di esprimere tutti gli aspetti di multi-sorgenti del processo d'informazione, cioè sia dati che *processing*. Inoltre, consente di ottenere combinazioni consistenti di algoritmi, con performance misurabili e dimostrabili, aprendo la strada ad applicazioni di metodi formali standardizzati ed automatici ai sistemi di fusione.

1.2.1 *Multisensor data fusion*

Esistono un certo numero di problematiche che fanno del *data fusion* un *challenge task*. La maggior parte di questi problemi aumentano con la quantità di dati da

integrare, con le imperfezioni e le diverse tecnologie dei sensori, nonché con la natura dell'ambiente di applicazione. Analizziamone le principali:

- *Imperfezione dati*: i dati forniti dai sensori sono sempre affetti da qualche livello di imprecisione come l'incertezza delle misure. Gli algoritmi di *data fusion* devono quindi essere capaci di esprimere tali imperfezioni efficacemente e sfruttare la ridondanza dei dati per ridurre i loro effetti.
- *Outliers e dati spuri*: le incertezze nei sensori aumentano non solo con l'imprecisione ed il rumore nelle misurazioni, ma sono anche causate dalle ambiguità e dall'inconsistenza presente nell'ambiente e dall'incapacità di distinguerli tra loro [46]. Gli algoritmi di *data fusion* devono essere dunque capaci di sfruttare i dati ridondanti per alleviare tali effetti.
- *Dati conflittuali*: l'integrazione di questi dati può essere problematica specialmente quando il sistema di fusione è basato su *evidential belief reasoning* e *Dempester's rule of combination* [47]. Per evitare di produrre risultati counter intuitive, gli algoritmi di *data fusion* devono trattare dati altamente conflittuali con particolare attenzione.
- *Modalità dati*: le reti di sensori possono acquisire dati qualitativamente simili (omogenei) o differenti (eterogenei) come audio, immagini, misure tattili di un fenomeno. Entrambi i casi devono poter essere gestiti efficacemente da uno schema di *data fusion*.
- *Correlazione dati*: questa problematica è particolarmente importante e comune nell'implementazione di fusioni distribuite, ad esempio nelle reti di sensori wireless.
- *Allineamento/registrazione dati*: i dati devono essere trasformati da ogni sistema di riferimento locale del sensore ad uno comune, prima di poter effettuare la fusione. Tale problema d'allineamento è spesso riferito col termine "registrazione dei sensori" e trattato mediante calibrazione degli errori indotti dai nodi individuali di sensori. La registrazione dei dati è di importanza cruciale nello sviluppo con successo di un sistema di fusione.
- *Associazione dati*: problemi di *tracking multi target* introducono maggiore complessità nei sistemi di fusione, se confrontati con quelli a singolo *target*. Una di

queste riguarda il problema dell'associazione dei dati, la quale può realizzarsi in due modalità: associazione misura-traccia o traccia-traccia. La prima si riferisce al problema di identificare da quale *target*, se presente, proviene ogni misura, mentre la seconda tratta la distinzione e la combinazione di tracce che stimano lo stato reale del *target* stesso [48].

- *Processing framework*: il processo di *data fusion* può essere eseguito in maniera centralizzata o decentralizzata. Quest'ultima è usualmente preferibile per le reti di sensori *wireless* poiché consente, ad ogni nodo di sensori, di elaborare "localmente" i dati acquisiti. Ciò è molto più efficiente se si considerano le risorse di comunicazione richieste da un approccio centralizzato, in cui ogni misura deve essere inviata al nodo centrale prima di essere elaborata e fusa.
- *Operational timing*: l'area di copertura dei sensori può comprendere un vasto ambiente, composto da molteplici aspetti che variano in modalità differente. Dunque, anche in caso di sensori omogenei, la frequenza operativa dei sensori può essere differente. Un metodo di *data fusion* ben progettato deve includere scale multi temporali con l'obiettivo di trattare efficacemente tale variazione temporale nei dati. Nei sistemi di fusione distribuiti, parti differenti di dati possono seguire differenti strade prima di giungere al centro di fusione, con conseguente possibilità di arrivi fuori sequenza. Questo problema necessita di un'opportuna gestione, specialmente nelle applicazioni *real-time*, al fine di evitare potenziali decadimenti delle performance.
- *Static vs dynamic phenomena*: i fenomeni sotto osservazione possono essere tempo invarianti o variabili in funzione del tempo. In quest'ultimo caso può essere necessario, per gli algoritmi di *data fusion*, incorporare nel processo di fusione una tabella temporale delle misure [49]. In particolare, l'aggiornamento dei nuovi dati, cioè quanto velocemente le sorgenti di dati catturano le variazioni e le memorizzano, giocano un ruolo vitale nella validità dei risultati della fusione. Ad esempio in qualche recente lavoro [50], gli autori hanno eseguito un'analisi probabilistica della storia temporale delle misure aggiornate, al fine di assicurare l'immediatezza dei dati di *input* e per migliorare l'efficienza del processo di fusione.

- *Data dimensionality*: i dati di misura acquisiti possono essere pre-processati sia localmente, in ogni nodo sensore, o globalmente, nel centro di fusione, per essere successivamente compressi in dati di dimensione minore, assumendo che sia tollerato un certo livello di perdita d'informazione a valle del processo di compressione stesso. Questo fase di pre-processing è utile poiché consente di risparmiare sia sulla larghezza di banda del canale di comunicazione che sulla potenza richiesta per la trasmissione dei dati, in caso di pre-processing locale [51], o limitare l'onere computazione del nodo centrale di fusione, in caso di pre-processing globale [52].

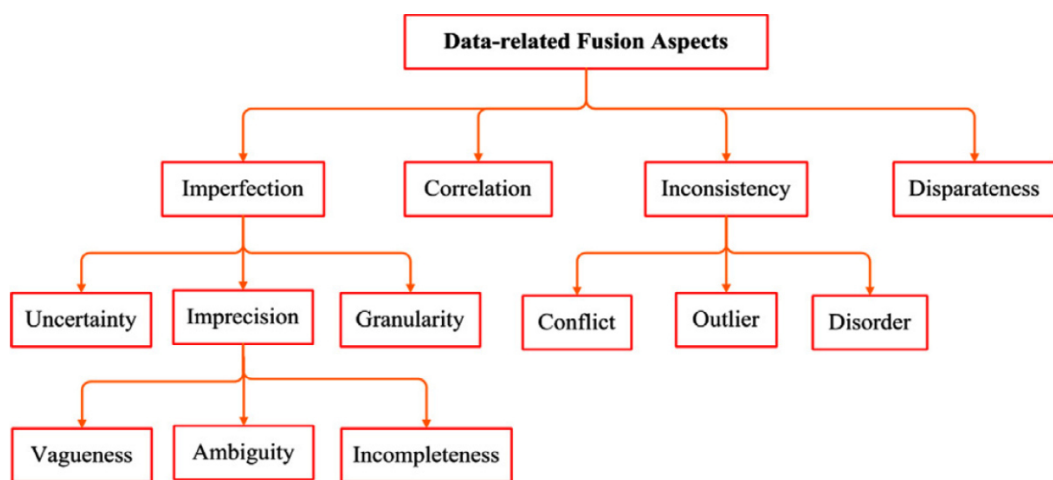


Figura 1.14: Tassonomia delle metodologie di data fusion: differenti algoritmi possono essere classificati in funzione dei quattro problemi dei dati di input (*imperfection, correlation, inconsistency, disparateness*) [41].

Sebbene molti di questi problemi siano stati identificati ed investigati in maniera approfondita, nessun algoritmo di *data fusion* è in grado di risolvere contemporaneamente tutte queste problematiche. In effetti, le differenti metodologie presenti in letteratura, si focalizzano sulla risoluzione di un sottoinsieme di tali tematiche, determinata in funzione della particolare applicazione considerata.

Nel caso in esame, in cui l'obiettivo è quello di integrare in maniera efficiente i dati ottici e acustici acquisiti dalla camera progettata, sarà affrontato principalmente, in base alla classificazione riportata, il problema dell' Allineamento/Registrazione di dati.

In Fig. 1.14 è mostrato uno schema della tassonomia utilizzata per classificare i diversi algoritmi di *data fusion* presenti in letteratura scientifica, basata sulle diverse problematiche a cui possono essere soggetti i dati di input di un sistema di *data fusion*. Essi, in generale, possono essere "imperfetti", "correlati", "inconsistenti" e/o in

disparate “forme/modalità”. Ognuna di queste quattro categorie principali può essere ulteriormente suddivisa in sotto categorie che affrontano problemi più specifici.

Per quanto concerne l’imperfezione dei dati, vengono considerati tre aspetti principali: “incertezza”, “imprecisione” e “granularità”. Un dato è definito “incerto” quando il grado di confidenza associato all’informazione presente sul dato stesso è minore di 1. Un dato “impreciso” è uno che si riferisce a molti oggetti, piuttosto che ad un solo. La “granularità” dei dati si riferisce invece all’abilità di distinguere tra gli oggetti che sono descritti dai dati stessi, essendo questi dipendenti dal set di attributi fornito. Inoltre, per quanto concerne “l’imprecisione”, essa si può manifestare come “ambiguità”, “indeterminatezza” o “incompletezza” dei dati. Continuando ad analizzare la tassonomia di Fig. 1.14, si hanno i dati “correlati” (cioè dipendenti) che, come già descritto, rappresentano un problema nei sistemi di data fusion e che dunque richiedono una trattazione opportuna. Troviamo poi “l’inconsistenza” nei dati di input, classificati ulteriormente in dati “conflittuali”, “spuri” o “fuori sequenza”. Infine, i dati possono essere classificati in funzione delle diverse modalità attraverso le quali essi sono forniti: avremo, dunque, una o più modalità di fornitura (dati generati da diversi sensori fisici o da operatori umani).

1.2.2 Integrazione di sistemi 3D ottico - acustici

L’integrazione di sensori ottici ed acustici, realizzata al fine di superare i limiti di visibilità presentati dai sistemi ottici di acquisizione 3D subacquea (causati da torbidità ed insufficiente illuminazione), costituisce un campo di ricerca emergente, sviluppatosi solo da pochissimi anni e pertanto ancora pochi sono i lavori presenti in letteratura. La natura complementare di queste due tipologie di sensori, che rende la loro integrazione particolarmente interessante, origina allo stesso tempo diverse problematiche concernenti la fusione dei dati da essi acquisiti. In generale, sonar e camere ottiche forniscono tipologie d’informazione completamente differenti (dati puntuali di *range* e immagini 2D, rispettivamente) e dunque principi di elaborazione diversi sono applicati per manipolare ed estrarre informazioni da queste due sorgenti di dati.

Comunque, il problema principale nella fusione dei dati acquisiti da entrambi i sensori è che essi operano a risoluzioni altamente differenti. Una camera ottica è in grado di

fornire immagini ad alta risoluzione, ma il suo range è limitato dall'attenuazione della luce in acqua, mentre i sensori acustici possono operare in range di visibilità più ampi e fornire informazioni 3D anche in condizioni di acque torbide, a spese però di una più bassa risoluzione ed una più difficile interpretazione dei dati

Uno dei primi lavori riguardanti la fusione di dati provenienti da acquisizioni ottiche ed acustiche, realizzata al fine di ottenere una rappresentazione più accurata del fondale marino è stato presentato da Singh [9]. Il metodo proposto combina informazioni d'intensità ad alta risoluzione ottenute da fotomosaici con mappe microbatimetriche di tipo acustico elaborate da un sonar di tipo *pencilbeam* operante alla frequenza di 675 kHz (Fig. 1.15).

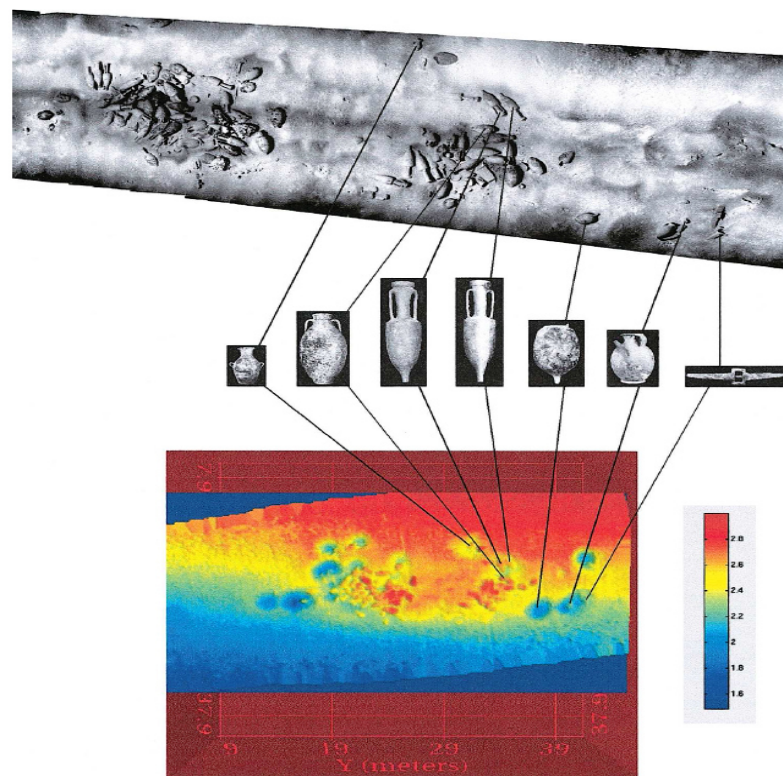


Fig. 1.15: immagine del mosaico deformato e della mappa batimetrica di un relitto [9].

La mappa microbatimetrica a bassa risoluzione è interpolata al fine di poter essere adattata alla risoluzione del fotomosaico. Punti corrispondenti vengono allora manualmente identificati attraverso le due rappresentazioni e utilizzando un *warp* ad elementi finiti il fotomosaico viene deformato nella nuova *mesh* definita attraverso i punti corrispondenti. Infine, questo fotomosaico deformato viene utilizzato come *texture* sovrapponendolo alla mappa batimetrica. Al fine di valutare quantitativamente

l'errore, i punti corrispondenti sono utilizzati per ottenere una trasformazione affine tra il fotomosaico e la mappa batimetrica legando le due rappresentazioni in un sistema di riferimento comune. I residui associati forniscono una misura della distorsione geometrica del fotomosaico, la quale è mostrata essere piccola.

Gli stessi autori hanno descritto in [12] un approccio per la fusione di dati subacquei 3D ottenuti dalla combinazione di immagini ottenute attraverso tecniche di *Structure from Motion* (SFM) e dati batimetrici acquisiti utilizzando un sonar subacqueo di tipo *pencilbeam*. L'algoritmo SFM, generalmente, stima simultaneamente sia la struttura del terreno che la traiettoria del veicolo. Comunque, questo significa che, poiché deve essere mantenuta la consistenza tra entrambe le stime, l'accuratezza della struttura può essere influenzata da errori nella stima della traiettoria tracciata dal veicolo, e viceversa, la precisione della traiettoria può essere influenzata da errori di ricostruzione della scena. Nel loro articolo, gli autori esaminano l'idea di utilizzare entrambi i sensori al fine di forzare l'algoritmo SFM a convergere ad una soluzione che meglio si adatta ai dati acquisiti dal sensore acustico. Un primo approccio, definito dagli autori come *data level fusion*, semplicemente prende l'informazione 3D estratta da ogni modalità in maniera indipendente e l'allinea utilizzando una trasformazione affine calcolata attraverso la selezione manuale di punti da entrambe le rappresentazioni. Una volta allineate, le due rappresentazioni possono essere combinate e confrontate. L'altro approccio presentato è una fusione a "livello di segnale", che utilizza i dati grezzi provenienti da entrambi i sensori per produrre una singola e coerente rappresentazione 3D. Un algoritmo SFM iterativo *multi-resolution* è utilizzato così che in ogni iterazione l'informazione 3D proveniente dalle altre modalità (nel caso specifico dai dati di navigazione provenienti da un sistema LBL) sia combinata. Come risultato, la soluzione SFM per la struttura del terreno diviene consistente con i dati di navigazione e la superficie ricostruita viene migliorata notevolmente grazie ai risultati dell'algoritmo SFM.

Nel lavoro di William [53] è presentato un algoritmo capace di modellare una scogliera naturale utilizzando dati acustici ed ottici acquisiti da un robot autonomo. Essi hanno sviluppato un'applicazione del metodo SLAM per il *tracking* di un veicolo operante in un ambiente non strutturato utilizzando le caratteristiche naturali del terreno. A tal fine essi utilizzano un sonar a scansione *pencilbeam* (Tritech SeaKing Imaging Sonar)

montato direttamente sopra la fotocamera del veicolo in modo che esso scansioni direttamente il fondale sottostante generando un profilo della superficie sopra la quale il veicolo si muove. I dati ricevuti dal sonar sono elaborati per identificare gli echo principali utilizzati al fine di inizializzare nuove *features* nella mappa SLAM. Facendo uso della traslazione tra i sistemi di riferimento solidali con i due sensori, il *footprint* generato dal sonar è proiettato sul piano immagine ed è utilizzato per identificare *features* ad alto contrasto nell'immagine dell'area illuminata dal sonar. Una volta inizializzate, le *features* sono tracciate utilizzando le osservazioni ottenute dalla camera.

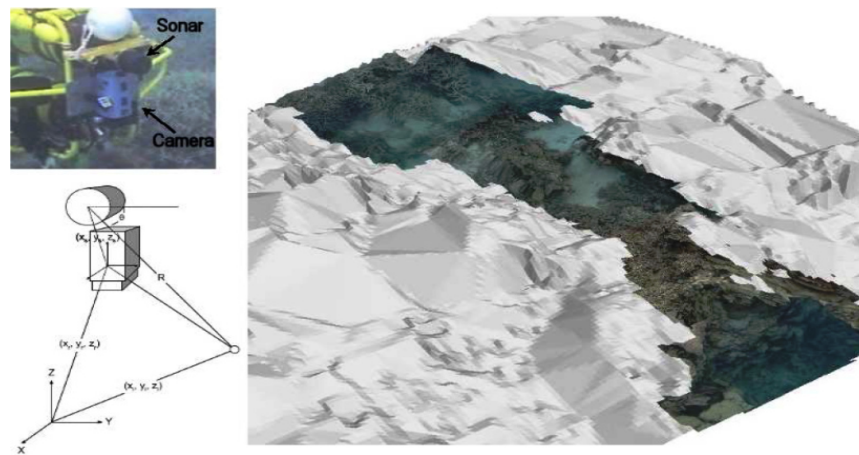


Fig. 1.16: configurazione del sistema con pencilbeam sonar e camera (sinistra). Modello del terreno risultante costruito proiettando la texture delle immagini ottiche su un modello superficiale ottenuto dai dati acustici (destra).

Le posizioni delle *features* rispetto al sistema di riferimento solidale con la camera sono ottenute dall'algoritmo SLAM che calcola la stima dello stato a posteriori in accordo all'equazione di aggiornamento del filtro di Kalman standard mantenendo quindi traccia della posizione del veicolo. In questo modo, le stime fornite dallo SLAM sulla posizione del veicolo sono ottenute senza l'utilizzo di un qualsiasi sistema di posizionamento *off-board* e legate solamente alle osservazioni della struttura del terreno utilizzando il sistema acustico e visivo. Le posizioni risultanti del veicolo sono usate per allineare la scansione del sonar generando dunque una ricostruzione 3D del terreno della scogliera sottostante. Alla fine, *la texture* ottenuta dalle immagini ottiche è proiettata sulla superficie risultante ottenendo una realistica rappresentazione della scogliera che risulta visivamente consistente (Fig. 1.16).

Nei lavori presentati da Fusiello [54, 55] è proposta l'integrazione di sensori ottici ed acustici al fine di migliorare la comprensione della scena subacquea ed assistere i piloti dei ROV. L'integrazione di un video e di dati 3D ottenuti da una camera acustica ad alta risoluzione (Echoscope) è ottenuta geometricamente registrando questi dati rispetto ad un modello noto della scena osservata (Fig. 1.17).

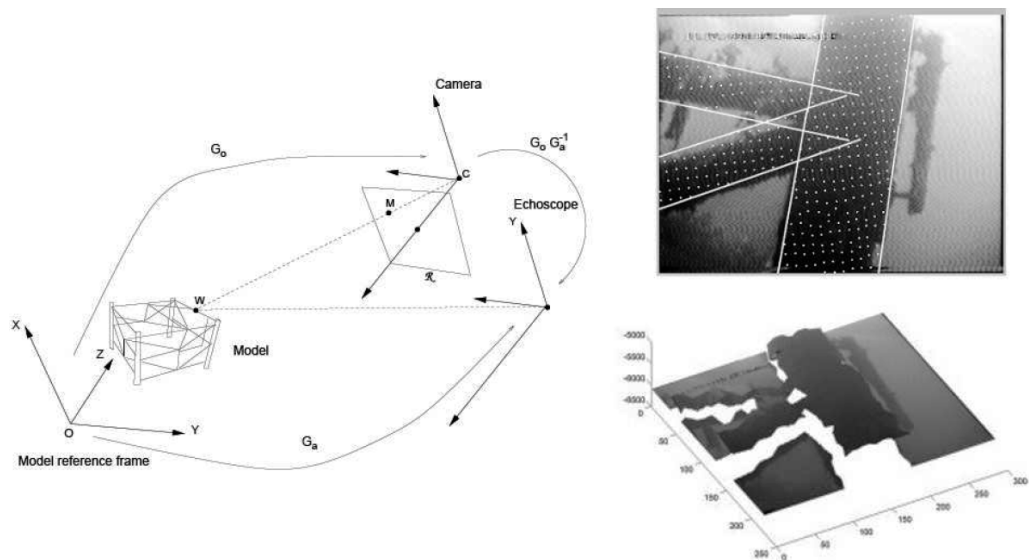


Fig. 1.17: setup di calibrazione con modello, sonar Echoscope e camera (sinistra). Immagine ottica con i punti 3D acustici proiettati (punti bianchi) e modello proiettato sovrapposto su di essi (in alto a destra). Ricostruzione 3D con l'immagine reale (texture) mappata (in basso a destra).

In [10] la catena di elaborazione delle due modalità d'acquisizione è descritta per una particolare applicazione, avente come oggetto una piattaforma petrolifera. Dopo alcune fasi di pre-processing, i dati acustici vengono registrati rispetto ad un modello noto utilizzando l'algoritmo ICP (*Iterative Closest Point*) [56] mentre l'allineamento ottico è eseguito utilizzando un algoritmo sviluppato da Lowe [57] che ricerca la posa della camera che porta al miglior *matching* tra ogni segmento e la proiezione dei segmenti del modello sull'immagine. Una volta che le pose di entrambi i sensori sono calcolate rispetto all'oggetto fisso osservato, la posa relativa della camera ottica e di quella acustica può essere stimata.

In questo modo l'approccio non richiede né una procedura di calibrazione utilizzando un oggetto costruito appositamente né il bisogno di identificare corrispondenze ottico acustiche (che sono infatti molto difficili da stabilire a causa della risoluzione e del rumore del dispositivo acustico). Comunque, poiché la scena stessa è utilizzata come

target di calibrazione, deve essere disponibile un modello CAD dello scenario osservato. Questa calibrazione geometrica porta alla registrazione della mappa di profondità a bassa risoluzione con le immagini ottiche. In aggiunta, modelli di oggetti possono essere sovrapposti alle immagini acquisite generando dunque un rappresentazione in realtà aumentata della scena subacquea.

Recentemente, Negahdaripour [58, 59] ha proposto un nuovo paradigma di ricostruzione 3D basato sullo sviluppo di una camera ottico-acustica in configurazione stereo. Lo scopo è quello di utilizzare le viste sovrapposte di un sistema di *imaging* stereo ottico-acustico composto da una camera ottica ed un sonar 2D DIDSON al fine di migliorare la precisione della ricostruzione 3D in condizioni di alta torbidità. La geometria epipolare è fondamentale per la ricostruzione 3D da viste stereo calibrate poiché essa semplifica il problema di stabilire le corrispondenze tra coppie di punti nelle due immagini limitando il compito ad una ricerca monodimensionale lungo le linee epipolari. Il vincolo epipolare che stabilisce le corrispondenze tra le proiezioni degli stessi punti della scena nelle due immagini (sonar e video) è stabilito in [60]. A causa dei modelli di proiezione asimmetrici in un sistema ottico-acustico, la geometria epipolare differisce nelle due viste. Le equazioni derivate mostrano che per ogni punto nell'immagine acustica, l'accoppiamento nell'immagine ottica si lega ad una sezione conica. E' anche mostrato come l'accoppiamento di ogni punto sull'immagine ottica si lega su particolari curve trigonometriche. Vale la pena sottolineare che la geometria epipolare fornisce le basi per la navigazione autonoma basata su mappe. In questo senso i vincoli geometrici possono essere utilizzati al fine d'integrare le acquisizioni ottiche ed acustiche del *target* nel risolvere il problema della registrazione immagine - modello per l'auto localizzazione.

Nel suo lavoro Negahdaripour ha anche affrontato il problema della calibrazione del sistema stereo ottico-acustico [61]. Una calibrazione intrinseca del sonar DIDSON è stata proposta per stimare e rettificare la distorsione delle lenti acustiche utilizzando immagini di una griglia planare nota, utilizzata al fine di fornire un'elevata riflettanza acustica. La calibrazione estrinseca della camera ottica e di quella acustica è stata stabilita utilizzando una griglia simile la quale presenta rilevanti *features* ottico-acustiche che sono state manualmente associate (Fig. 1.18). Dunque, la posa relativa di entrambe le camere è stata stimata attraverso un algoritmo di ottimizzazione che

minimizza le distanze 3D tra le ricostruzioni delle corrispondenze dalle proiezioni ottiche ed acustiche. Quindi, una volta che la calibrazione è stata effettuata, la ricostruzione 3D di un punto, dato il suo *matching* ottico acustico, può essere calcolata in forma chiusa. Comunque, se le corrispondenze non rappresentano accoppiamenti perfetti, il rumore porta a stime differenti. Al fine di superare queste problematiche, gli autori propongono [60] di utilizzare un metodo non lineare per la ricerca di una soluzione ottima nel senso di massima verosimiglianza, minimizzando l'errore geometrico. I risultati mostrano come un sistema di *imaging* stereo ottico-acustico supera un tradizionale sistema stereo ottico all'aumentare della distanza dal *target* e/o in condizioni di torbidità.

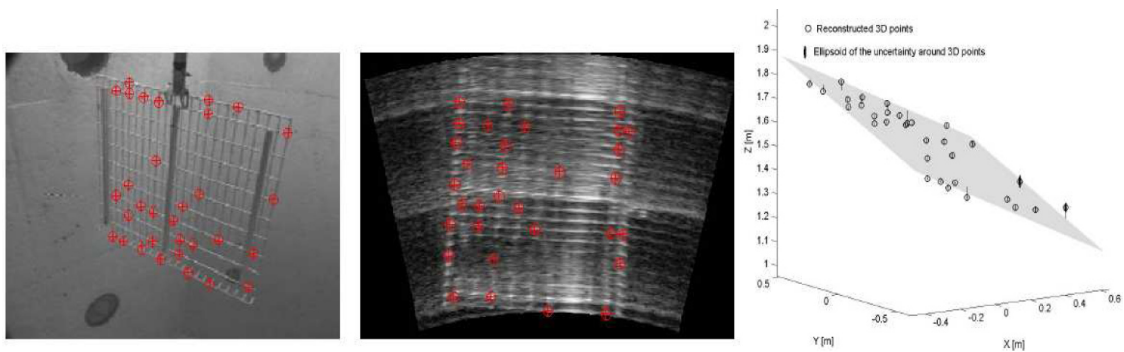


Fig. 1.18: matching *points* nelle viste ottiche ed acustiche usate per la ricostruzione 3D (sinistra, centro) e proiezioni dei punti 3D ricostruiti (destra).

In [62] gli autori propongono un metodo basato sulla corrispondenza e registrazione di *features* al fine di generare ricostruzioni 3D da distinte immagini di distanza della stessa scena. Immagini di distanza a varie risoluzioni e da differenti viste sono accoppiate stimando la trasformazione di corpo rigido tra *set* di punti. Il *set* di corrispondenze utilizzate per determinare questa trasformazione è ottenuto utilizzando *spin images* [63] che possono fornire *features* geometriche in una nuvola di punti 3D che si dimostrano essere robuste e invarianti rispetto alla risoluzione. Una volta che queste *features* sono calcolate, i due *set* sono accoppiati al fine di trovare la più alta correlazione, mentre la trasformazione che lega i punti corrispondenti è determinata utilizzando l'algoritmo di Horn [64] ed il metodo RANSAC [65]. Infine, la registrazione è rifinita utilizzando l'algoritmo ICP allo scopo di ottenere mesh che siano visivamente consistenti. Sebbene i risultati sono mostrati solamente per dati 3D

generati da un sonar *multibeam*, la tecnica mira ad essere applicata in futuro a dati 3D provenienti da diversi sistemi di acquisizione. In questo senso gli autori sottolineano come questo possa essere di particolare interesse al fine di generare ed accoppiare mesh 3D generate da sistemi stereo ottici e sistemi sonar per produrre ricostruzioni più complete.

Una nuova metodologia per il rilevamento 3D subacqueo ad alta risoluzione di siti sommersi è stata proposta nel progetto ROV 3D [66]. In particolare, l'obiettivo del progetto è stato lo sviluppo di processi e strumenti automatici di rilievo 3D subacqueo attraverso l'impiego di sensori ottici ed acustici. Al fine di ottenere un modello 3D di grandi scene complesse, a partire da queste due tipologie di acquisizioni, è stata proposta una struttura numerica ad alta risoluzione in grado di generare differenti tipologie di output (SIG 3D, MNT, mosaico, etc.). In pratica, uno scanner acustico 3D è utilizzato al fine di scansionare una grande scena, mentre un sensore ottico acquisisce, mediante restituzione fotogrammetrica, differenti aree della scena stessa. Un opportuno strumento software effettua poi la registrazione automatica di entrambe le sorgenti di dati, mentre altri algoritmi, sviluppati nel progetto, riconoscono e modellano gli oggetti di interesse. Questi dati permettono di ottenere una rappresentazione simbolica degli oggetti e della loro geometria in un preciso facsimile virtuale.

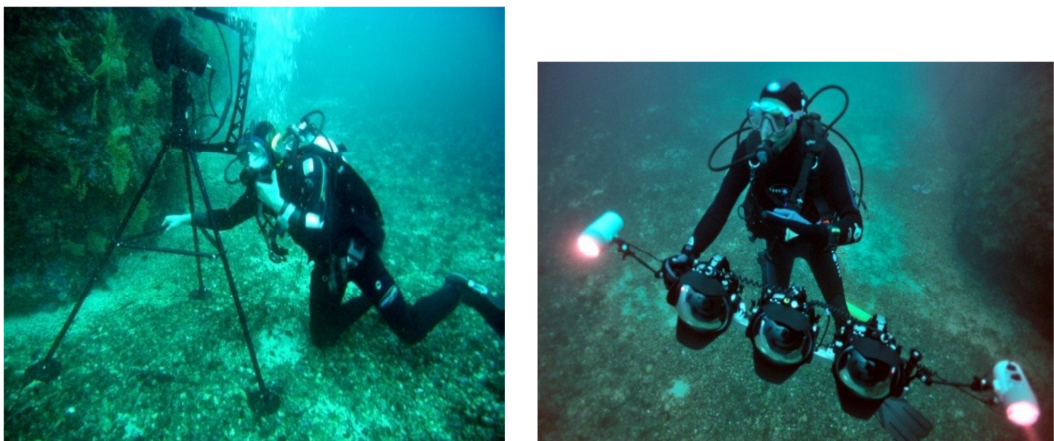


Fig. 1.19: camera acustica Blue View BV5000 *in situ*, di fronte alla grotta "Impérial du Terre" nei pressi di Marsiglia, ad una profondità di circa 30 m (sinistra). Rilevamento fotogrammetrico con l'approccio Menci (utilizzo di 3 camere sincronizzate) (destra).

La sperimentazione della fusione dei dati ottici ed acustici è stata effettuata usando una camera acustica ad alta frequenza e precisione (Blue View BV5000) ed un sistema fotogrammetrico a 3 camere reflex sincronizzate (sistema MENCI), in una grotta subacquea nei pressi di Marsiglia, ad una profondità di circa 30 m (Fig. 1.19).

Lo scanner 3D acustico attivo Blue View BV5000 fornisce una nuvola di punti 3D ad alta risoluzione del sito. A differenza dei sistemi batimetrici convenzionali che considerano una sola coordinata d'elevazione Z, generando in tal modo un profilo della scena scansionata, il sonar 3D può acquisire punti multipli d'elevazione Z per una data coppia di coordinate (X, Y). Lo scanner è montato su un supporto imperniato lungo 2 assi (orizzontale e verticale) permettendo la rotazione dall'alto verso il basso, da destra a sinistra. Con un'apertura acustica di 45° il sistema è capace di misurare un'area di 45° o 360° su una superficie sferica comprendendo l'intero ambiente circostante nel raggio di 30 m. Ogni scansione o insieme di scansioni vengono eseguite in posizione fissa e con l'ausilio di un tripode. Per ottenere una nuvola di punti 3D sufficientemente densa devono essere effettuate diverse acquisizioni da posizioni differenti, e successivamente, i risultati ottenuti devono essere combinati tra loro. Le nuvole di punti 3D acquisite da posizioni fisse, vengono registrate mediante algoritmo ICP (*Iterative Closest Point*). Il passaggio successivo è quello di posizionare l'intera nuvola di punti 3D definitiva rispetto ad un sistema di riferimento assoluto, utilizzando a tal fine punti di riferimento le cui coordinate possono essere determinate attraverso sistemi di posizionamento acustico USBL (*Ultra Short Baseline*). Due sommozzatori sono necessari per le acquisizioni con lo scanner acustico mentre altri due per l'acquisizione tramite il sistema MENCI. Quest'ultimo è composto da una parte hardware costituita da 3 camere reflex sincronizzate e da una *software house* il cui compito è quello di, attraverso opportuni algoritmi, determinare la "migliore tripletta" di immagini acquisite e da queste calcolare la mappa di profondità al fine di ottenere la nuvola di punti 3D della scena.

In Fig. 1.20 sono mostrati i risultati ottenuti dall'applicazione della metodologia proposta nella ricostruzione 3D della grotta subacquea "Impérial du Terre" presso Marsiglia, oggetto della sperimentazione.

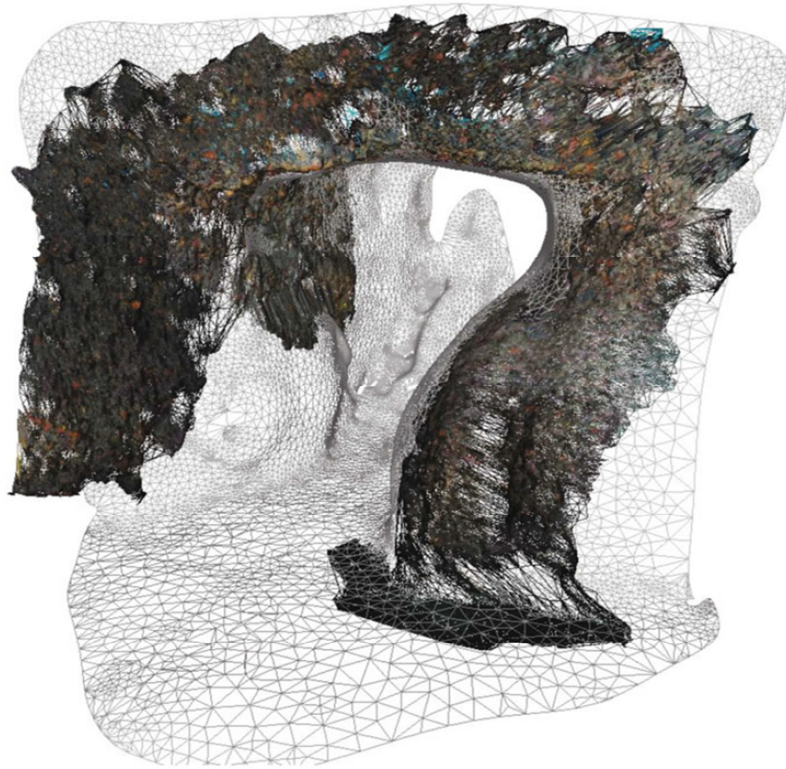


Fig. 1.20: ricostruzione 3D della grotta Impérial du Terre presso Marsiglia oggetto della sperimentazione.

La progettazione di un sistema ottico acustico per la ricostruzione 3D di scene subacquee è stato proposto nei lavori presentati da N. Hurtòs presso i laboratori VICOROB di Girona (Spagna) [67]. Il sistema ottico-acustico proposto è composto da una camera ottica accoppiata ad un sonar *multibeam* installati, attraverso un supporto rigido, a veicoli subacquei (ROV) al fine di acquisire immagini e profili di fondali marini. Il sensore acustico è impiegato allo scopo di acquisire informazioni di distanza dal fondale, mentre la camera è utilizzata per raccogliere *feature* come colore o *texture*. Per combinare le informazioni di entrambi i sensori la configurazione dei dispositivi è stata progettata in modo tale che il fascio *multibeam* intersechi l'area di proiezione dell'immagine della camera ottica. I due sistemi di sensori sono stati modellati geometricamente utilizzando il modello *standard* di camera pinhole per la componente ottica, mentre per il sonar è stato utilizzato un modello semplificato *multibeam*, considerando un numero di fasci egualmente distribuiti lungo l'apertura totale del sensore stesso. Questi modelli sono stati parametrizzati utilizzando valori reali ottenuti dalla strumentazione effettivamente installata su un AUV presente nei laboratori del VICOROB, cioè una camera Tritech Super SeaSpy e un sonar *multibeam*

DeltaT. La mappatura fra profili acustici ed immagini ottiche è stata stabilita, in un ambiente virtuale di simulazione, applicando una matrice di trasformazione rigida che relaziona i sistemi di riferimento solidali con i due sensori. In particolare, l'ambiente di simulazione creato permette di definire una traiettoria su un terreno virtuale ed acquisire immagini sintetiche di tipo ottico, così come dati *multibeam*. Data la traiettoria delle pose della camera sul terreno, il sistema di coordinate del sonar multibeam è piazzato su ognuna delle posizioni corrispondenti rispetto alla camera ottica stessa. In tal modo, i dati ottici possono essere adeguatamente riproiettati sulla mappa batimetrica ottenendo una ricostruzione 3D del terreno che comprende sia informazione acustica che visiva (Fig. 1.21).

Un approccio di questo tipo, in realtà, non avrebbe buone performance in condizioni reali poiché l'accuratezza dei dati dei tipici sistemi di navigazione supera di gran lunga la precisione intrinseca del sonar. Tuttavia, un ambiente di simulazione così realizzato consente di validare i risultati ottenuti dalle metodologie di calibrazione impiegate per i sistemi ottico acustici progettati quando la posizione relativa tra i due sensori è ben nota.

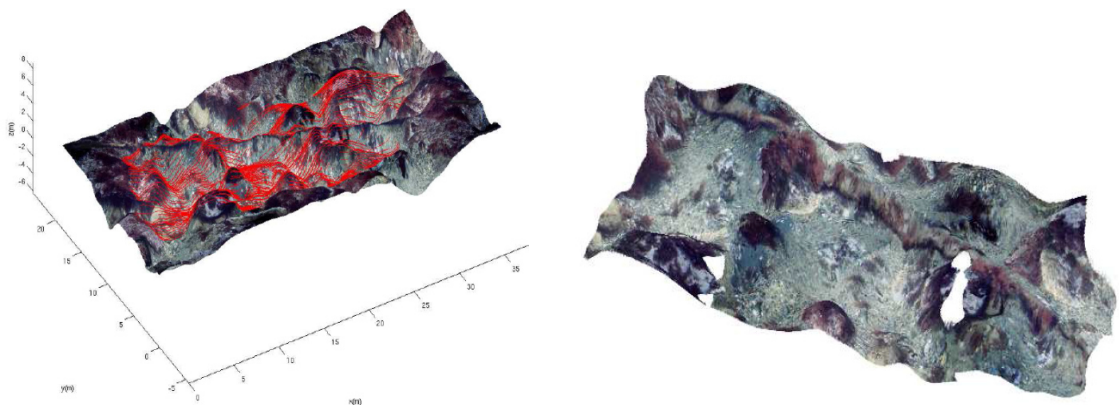


Fig. 1.21: ricostruzione 3D nell'ambiente di simulazione.

Per quanto concerne quest'ultimo aspetto, allo scopo di ottenere la posizione relativa e l'orientazione dei sistemi di coordinate del sonar multibeam e della camera ottica è stata proposta, per il sistema in esame, una metodologia di calibrazione estrinseca. Essa non si basa sull'esplicita corrispondenza ottico acustica, poiché risulta difficile stabilirla in maniera perfetta a causa della differente risoluzione presentata dai due sensori e dal rumore presente sui dati acustici. Assumendo valido il modello

semplificato per il sonar *multibeam*, il problema della calibrazione estrinseca del sistema ottico acustico proposto è stato comparato alla calibrazione di un sistema camera laser, ovvero come alla calibrazione di una camera ottica ed un sistema laser a lama di luce. Per la risoluzione di questa problematica è stata adottata la procedura di calibrazione presentata da Zhang [68] applicata al fine di calibrare una camera ottica ed un sistema laser invisibile. La componente di rumore (supposta gaussiana e uniforme) è stata considerata nel modello di simulazione ottenendo delle efficaci performance.

Metodologia per la calibrazione di una camera ottico - acustica 3D

Lo scopo di combinare tecniche di acquisizione ottiche con tecniche acustiche è principalmente quello di migliorare l'*imaging* 3D subacqueo in condizioni di scarsa visibilità. Nelle situazioni in cui vi è una luce ambientale minima (a causa di elevate profondità o elevata torbidità), l'acquisizione della nuvola di punti 3D, relativa al *target* da ricostruire, si può avvalere sia della più alta risoluzione e del maggior campo di vista delle tecnologie ottiche che del contributo delle onde sonore nell'acquisizione di quei punti che si trovano nelle "zone d'ombra".

In linea con i lavori presenti in letteratura scientifica riguardanti questo specifico filone di ricerca, la presente tesi di dottorato descrive i primi passi verso la realizzazione di una camera ottico - acustica in grado di fornire, in *real-time*, una rappresentazione tridimensionale dell'intera scena osservata. Il sistema è composto da una camera ottica stereoscopica e da una camera acustica 3D, entrambi in grado di fornire, in *real-time*, dati tridimensionali della scena acquisita. Le potenzialità offerte da una tale tipologia di sistema sono innumerevoli, basti pensare ad un suo possibile impiego come supporto all'operatore di un *Remotely Operated Vehicle* (ROV) subacqueo in fase di navigazione in acque torbide o durante operazioni di esplorazione e manutenzione realizzate con l'ausilio di manipolatori subacquei.

Dallo studio condotto sullo stato dell'arte presentato nel capitolo 1, è emerso come tutti i lavori riguardanti l'integrazione di diverse tipologie di sonar (3D *acoustic camera*, *single beam sounder*, *multibeam*) e camere ottiche presentano un approccio al *sensor fusion* che, in accordo alla classificazione stabilita in [69], è di tipo *mapping-oriented*, cioè una fusione a livello di *feature*. Ciò implica che il modello dati provenienti dai vari sensori sia descritto attraverso relazioni geometriche (orientazioni e posizioni) mentre la loro fusione sia eseguita facendo uso di corrispondenze geometriche e registrazione di nuvole di punti 3D. Un problema cruciale di queste tecniche di *sensor fusion*

riguarda l'allineamento dei dati (o calibrazione estrinseca del sistema) che consente un *mapping* tra i dati provenienti dai diversi sensori e la determinazione della trasformazione rigida che lega tra loro i sistemi di riferimento solidali con i sensori stessi.

In questo capitolo verrà presentata una metodologia di calibrazione per la determinazione della matrice di trasformazione rigida che lega i sistemi di riferimento solidali con i sensori ottici e acustici, al fine di allineare automaticamente le immagini 3D ottiche e acustiche. In generale, le metodologie di calibrazione sono altamente dipendenti dal *layout* e dai sensori che costituiscono il sistema, in special modo dalla tipologia e dalla struttura dei dati che essi forniscono. Risulta allora necessario, prima di andare a presentare il protocollo di calibrazione implementato e i risultati della campagna sperimentale condotta, andare ad analizzare l'architettura generale del sistema e le caratteristiche salienti dei sensori che lo compongono. Dunque, nella sezione 2.1 saranno analizzati in dettaglio sia le camere reali (sottosistema ottico) che il sonar utilizzato (sottosistema acustico), al fine di determinare le loro geometrie e le principali caratteristiche utili nelle successive fasi sperimentali. Nella sezione 2.2 è stata condotta un'analisi sui possibili approcci alla calibrazione del sistema per presentare in seguito, in maniera dettagliata, le vari fasi che realizzano il protocollo di calibrazione effettivamente implementato. Infine, il capitolo si conclude con la descrizione delle diverse attività ed i risultati ottenuti da una prima sperimentazione in vasca della metodologia proposta.

2.2 Configurazione del sistema

2.2.1 Descrizione sottosistema ottico

Per quanto concerne la componente ottica del sistema è stato sviluppato, presso i laboratori del DIMEG dell'Università della Calabria un primo prototipo di camera stereo ottica costituita da 2 videocamere digitali ultra compatte *Point Grey Flea 2* con pacchetto ottico *Pentax C30811TH* (Fig. 2.1), le cui specifiche tecniche sono riportate di seguito:



Fig. 2.1: camera Point Grey Flea 2 (sinistra) e pacchetto ottico PENTAX C30811TH (destra).

Camera (Point Grey Flea 2):

- *Risoluzione:* 0.8 MPx (1032 x 776);
- *Frame Rate:* 30 fps;
- *Interfaccia PC:* Firewire 800;
- *Sensore:* CCD Sony ICX204 da 1/3".

Ottiche (PENTAX C30811TH):

- *Focale:* 8.5 mm;
- *Regolazione apertura e messa a fuoco:* manuale;
- *Campo di vista (FOV):* 37,7° (orizzontale);
- *Montaggio:* CS-Mount.

La loro gestione è stata effettuata con l'ausilio del software *FlyCapture* (SDK fornito direttamente dal costruttore), in quanto la sua interfaccia consente di settare con facilità modalità di funzionamento delle camere e principali parametri di acquisizione, ad eccezione della messa a fuoco, in quanto manuale.

Poiché le camere non sono state concepite per un loro uso in ambiente subacqueo, si è proceduto alla progettazione ed alla realizzazione di due case subacquee (scafandri) che le potessero contenere (Fig. 2.2). La loro progettazione ha richiesto un'opportuna analisi degli ingombri; infatti è necessario alloggiare correttamente le camere in maniera da garantire che l'asse ottico si trovi esattamente al centro dello scafandro in modo tale da evitare problemi legati a deformazioni o alla eccessiva riduzione del campo di vista. Un'altra necessità è stata anche quella di contenerne il peso; a tal proposito si sono effettuate simulazioni per verificare la tenuta stagna della struttura

in alluminio e della parte frontale in policarbonato, ottimizzando gli spessori e la quantità di materiale da utilizzare. Per quanto concerne la progettazione dell'oblò frontale si è deciso di realizzarne, per semplicità costruttiva, uno piatto che, rispetto a quello sferico, porta ad una riduzione dell'angolo di vista dell'acquisizione, per effetto della rifrazione (Fig. 2.3). In Tab. 1 sono riportati i valori dell'angolo di vista per il pacchetto ottico utilizzato, sia in aria che in acqua, in cui è possibile osservare una riduzione stimata di circa il 25%.

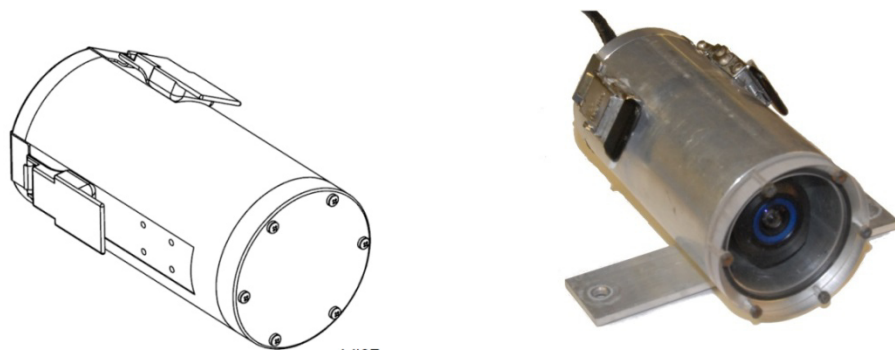


Fig. 2.2: case subacqueo per l'alloggiamento delle camere Flea2.

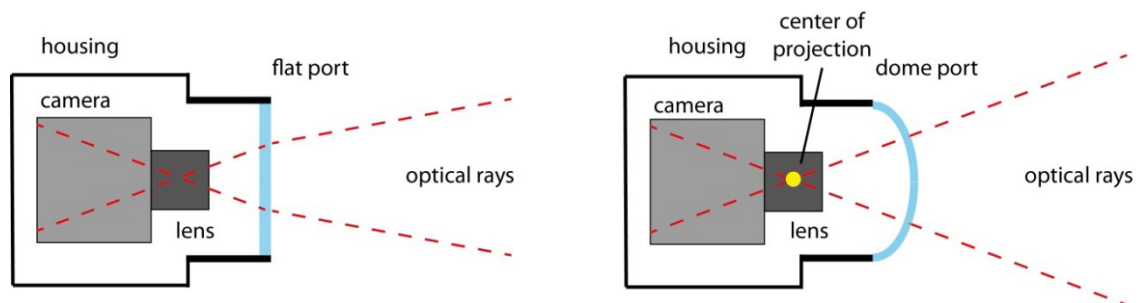


Fig. 2.3: deviazione dei raggi luminosi per una oblò piatto ed uno sferico, a causa della rifrazione.

Angolo di vista (aria)	Angolo di vista (acqua)
H : 37,7°	H: 29°
V: 32,3°	V: 24,9°

Tab.1: Angoli di vista in acqua ed in aria adottando un'ottica da 8.5 mm.

Gli scafandri sono stati testati nei laboratori del DIMEG, in una camera idrostatica a circa 4 bar per verificarne la tenuta a tali valori di pressione (Fig. 2.4). La camera idrostatica utilizzata, riempita di acqua limpida, consente di aumentare la pressione

interna fino a 10 bar attraverso una pompa di prova a mano. Questa ha permesso quindi di verificare resistenza e tenuta degli scafandri in cui vengono alloggiati le due camere.



Fig. 2.4: camera idrostatica impiegata per le prove di tenuta degli scafandri.

2.2.2 Descrizione sottosistema acustico

Per la realizzazione del sottosistema acustico è stata utilizzata una camera acustica *Coda Echoscope*, dell'azienda *CodaOctopus*. Illuminando l'intero volume di visualizzazione con un singolo *ping* (Fig. 2.5), il sonar utilizza la tecnologia *phased array* per elaborare 16.000 *beams* contemporaneamente, consentendo in tal modo di ottenere un'immagine acustica 3D in tempo reale dell'intera scena osservata.

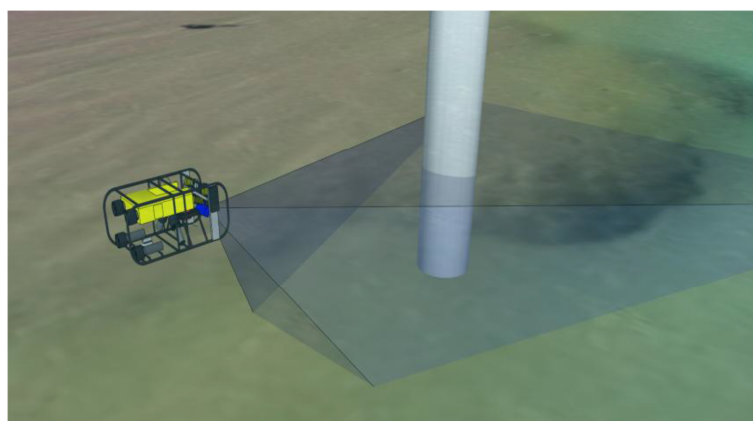


Fig. 2.5: volume di visualizzazione della camera acustica *Echoscope* (tipica installazione su ROV).

Di seguito vengono riportate le principali specifiche tecniche:

- *Frequenze di lavoro*: 375 KHz e 610 KHz (scelta operata via *software*);
- *Numero di beams*: 128 x 128 (16.384 totali);
- *Massimo range*: 120 m (375 KHz) / 80 m (610 KHz);
- *Minimo range*: 1 m;
- *Risoluzione spaziale*: 3 cm;
- *Ping rate*: fino a 12 Hz;
- *Copertura angolare*: 50° x 50°, 50° x 25°, 25° x 50°, 25° x 25° (selezionabile via *software*);
- *Spaziatura tra i beams*: 0.39° o 0.19° (dipendentemente dalla copertura angolare);
- *Dimensioni*: 380 x 300 x 160 mm (esclusi i connettori);
- *Peso*: 22 Kg (in aria) / 12 Kg (in acqua);
- *Consumo di potenza*: 3 – 6 A a 24 Vdc;
- *Profondità massima di lavoro*: 600 m (opzione disponibile fino a 3000 m);
- *Interfaccia testa sonar – unità di controllo (DIU)*: singolo cavo per dati, controllo e alimentazione;
- *Interfaccia unità di controllo (DIU)-Workstation*: Ethernet 10Mbps per ricezione dati / RS232 per trasmissione e ricezione comandi.

La camera acustica 3D *Echoscope* si compone di tre sottosistemi separati. La componente principale è rappresentata dalla testa del sonar; in essa sono alloggiati il proiettore acustico per la trasmissione del segnale a 375 KHz (o 610 KHz) e l'*array* di ricezione dell'energia acustica riflessa, composto da una matrice di 48 x 48 idrofoni. Completano il sottosistema la circuiteria elettronica interna che realizza il *beamforming* e la presenza di eventuali sensori inerziali (Fig. 2.6, sinistra).

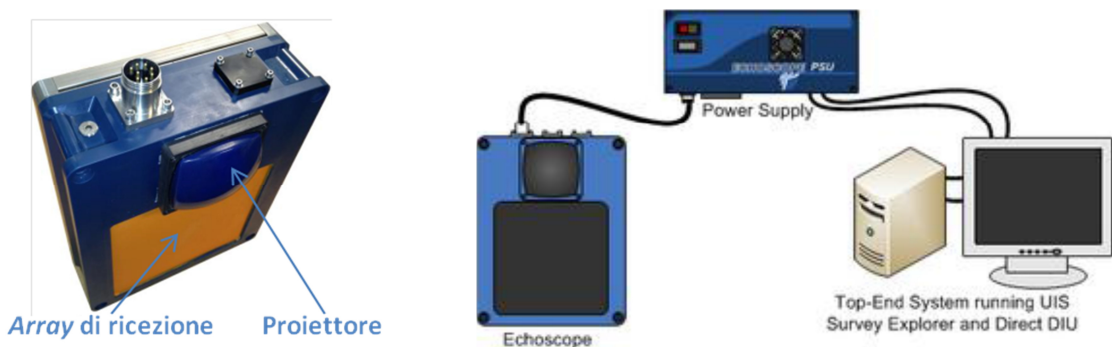


Fig. 2.6: testa del Sonar *Coda Echoscope* (sinistra), tipica configurazione del sistema per la gestione in remoto (destra).

I restanti sottosistemi sono rappresentati dal modulo di alimentazione *Power Supply unit* (PSU) e dalla *workstation* in cui vi sono installati i *software* di gestione dell'intero sistema (Fig. 2.6, destra), cioè il *Data integration Unit* (DIU), modulo che gestisce sia il flusso dati proveniente dalla testa del sonar che l'invio e la ricezione dei comandi di controllo dello stesso e il *software UIS Survey Explorer* per la visualizzazione della nuvola acustica 3D della scena osservata (Fig. 2.7).

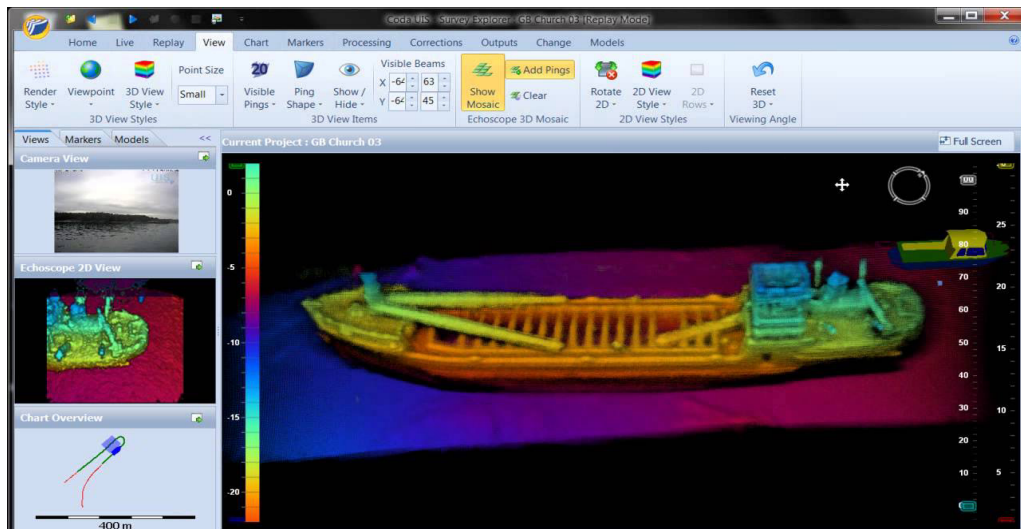


Fig. 2.7: esempio interfaccia software di visualizzazione *Coda UIS Survey Explorer*.

2.2.3 Layout del sistema

Al fine di assicurare un supporto meccanico al sistema è stata progettata una struttura costituita da profilati in alluminio ITEM, in cui sono stati alloggiati gli scafandri delle due camere ottiche e la testa della camera acustica *Echoscope*, quest'ultima connessa attraverso una piastra posteriore. L'intera struttura è stata poi fissata ad una staffa di supporto a cui è stato connesso un palo telescopico necessario alla movimentazione dell'intero sistema in fase di sperimentazione (Fig. 2.8).

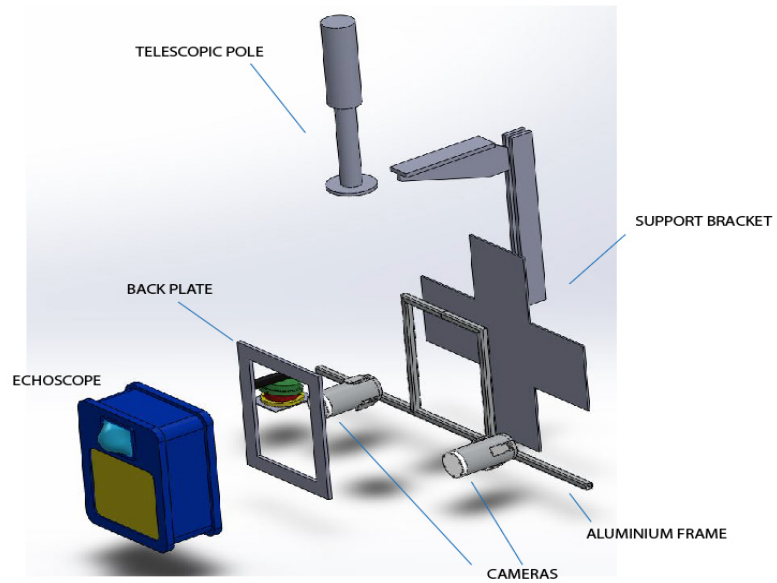
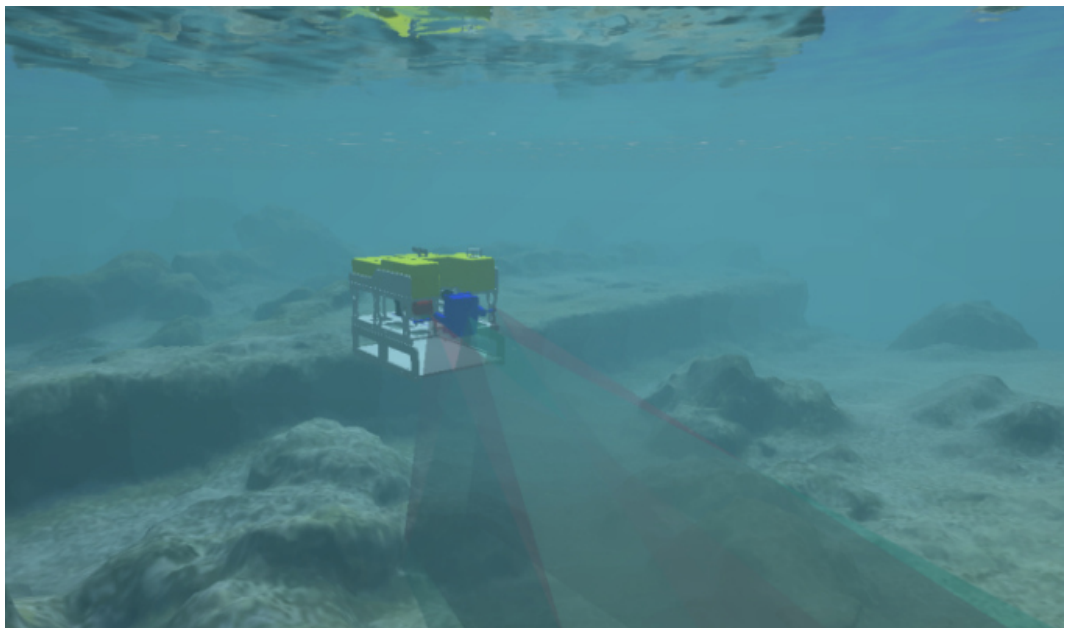


Fig. 2.8: modello 3D virtuale del sistema con relativa struttura di supporto.

La posizione relativa del sottosistema stereo ottico rispetto a quello acustico è stata definita in fase progettuale con l'obiettivo di garantire la massima sovrapposizione dei campi di vista sia orizzontale che verticale (Fig. 2.9).



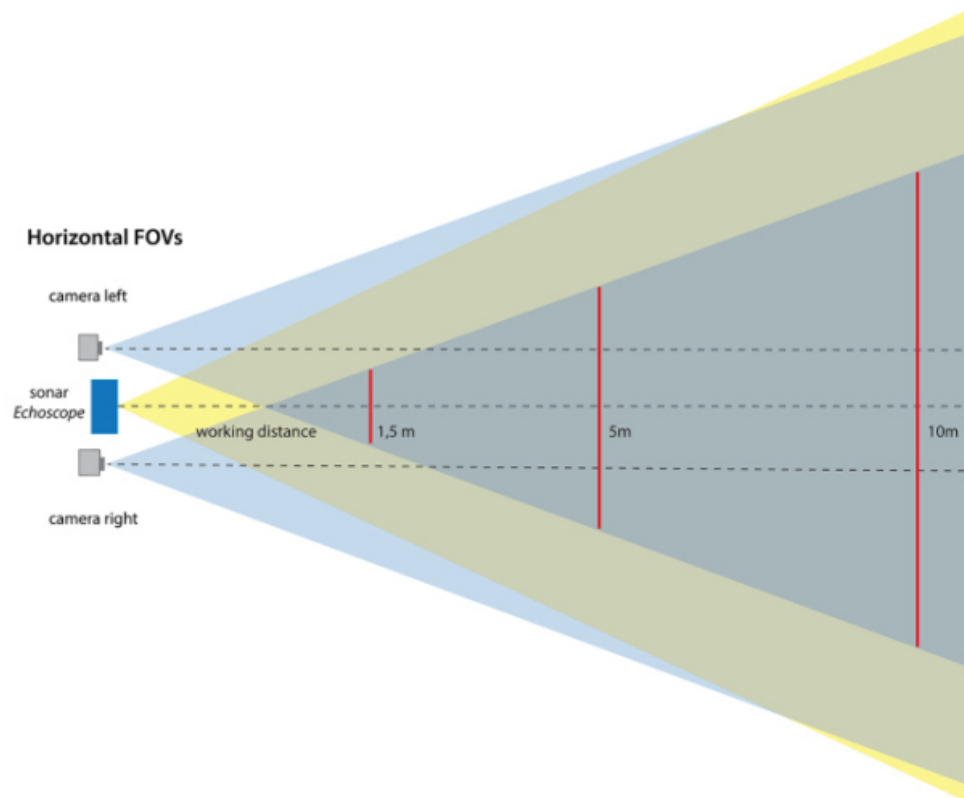


Fig.2.9: simulazione in ambiente virtuale della configurazione di sistema proposto (in alto). Grafico per lo studio della sovrapposizione degli angoli di vista di camere ottiche e sonar *Echoscope* (in basso).

2.3 Analisi su differenti approcci alla calibrazione estrinseca del sistema.

Data una struttura composta da due camere ottiche in configurazione stereo e da una camera acustica 3D, secondo il *layout* presentato nel paragrafo 2.1.3, la calibrazione estrinseca del sistema consiste nella determinazione della matrice di trasformazione, fissa ma incognita, che lega tra loro il sistema di riferimento solidale con il sottosistema stereo ottico con quello solidale con il sottosistema acustico, ottenendo in tal modo la posa (posizione ed orientamento) relativa tra i due sensori.

Indipendentemente dalla particolare soluzione adottata, la metodologia di calibrazione si comporrà, in generale, di un insieme di moduli concepiti al fine di elaborare i dati forniti in *output* dai sensori con cui è equipaggiato il sistema, cioè immagini ottiche 2D fornite dalle due camere digitali ed immagini acustiche costituite da nuvole di punti 3D, quest'ultime fornite dalla camera acustica *Echoscope*. Queste immagini risultano

essere né registrate né sovrapposte poiché i due sottosistemi presentano punti e campi di vista altamente differenti.

In linea di principio, per la calibrazione della camera ottico-acustica, risulta necessario fabbricare un opportuno oggetto in grado di dar vita e mettere in risalto *feature* distinte sia nell'immagine acustica che in quella ottica. Questo, in realtà, risulta difficile da ottenere principalmente a causa delle differenti risoluzioni spaziali presentate dai due sottosistemi. Nel nostro approccio si è scelto di utilizzare un classico pannello di calibrazione a scacchiera i cui bordi sono stati rivestiti di materiale *pluriball* il quale, grazie alla sua elevata riflettività acustica, consente di discriminare in maniera più accurata i punti 3D appartenenti effettivamente al pannello di calibrazione rispetto a quelli provenienti dai diversi oggetti che vanno a comporre la scena in cui il pannello è posizionato.

Poiché in questa fase il nostro obiettivo è quello di analizzare possibili soluzioni alternative per la determinazione della matrice di trasformazione rigida che lega tra loro le terne cartesiane solidali con i sottosistemi ottico ed acustico, supponiamo che quest'ultimi siano già stati calibrati in maniera indipendente. In particolare, per quanto concerne il sottosistema ottico, siano stati determinati, attraverso una metodologia di calibrazione che descriveremo in dettaglio nei paragrafi successivi, sia i parametri intrinseci delle singole camere ottiche (focale, punto principale, distorsioni ottiche e *skew*) sia quelli estrinseci della coppia stereo (posizione relativa fra i due sistemi di riferimento delle camere). A valle di questo processo di calibrazione, risulta possibile, in linea di principio, conoscere le coordinate cartesiane di un qualsiasi punto della scena acquisita dal sottosistema ottico rispetto al suo sistema di riferimento (tipicamente solidale con la camera sinistra della coppia stereo).

Dunque, da un punto di vista matematico, supponendo che ad un punto $p_o = [x_o, y_o, z_o]^T$ del sistema di riferimento ottico corrisponda un punto $p_a = [x_a, y_a, z_a]^T$ del sistema di riferimento acustico, la trasformazione rigida tra il sistema di riferimento solidale col sottosistema acustico rispetto a quello ottico si può esprimere come:

$$p_o = \mathbf{R}p_a + t \quad (2.1)$$

dove \mathbf{R} è la matrice ortonormale 3 x 3 che rappresenta l'orientamento della camera acustica rispetto a quella stereo ottica, mentre t è un vettore tridimensionale corrispondente alla loro posizione relativa.

Equivalentemente, indicando con \widetilde{p}_o e \widetilde{p}_a le coordinate omogenee dei punti p_o e p_a rispettivamente, la relazione precedente si può esprimere come:

$$\widetilde{p}_o = {}^o \begin{bmatrix} \mathbf{R} & t \\ 0 & 1 \end{bmatrix}_a \widetilde{p}_a \quad (2.2)$$

In cui la matrice 4 x 4 ${}^o \begin{bmatrix} \mathbf{R} & t \\ 0 & 1 \end{bmatrix}_a = {}^o \mathbf{T}_a$ rappresenta (in coordinate omogenee) la trasformazione rigida che lega il sistema di riferimento solidale con la camera acustica con quello solidale con la camera stereo ottica.

Alla luce di quanto detto, lo scopo che ci si prefigge è quindi quello di sviluppare una metodologia in grado di determinare i parametri estrinseci \mathbf{R} e t che definiscono rispettivamente l'orientamento e la posizione del sottosistema acustico rispetto a quello ottico.

Un primo semplice approccio alla risoluzione del problema può essere formulato a partire dal *setup* mostrato in Fig. 2.10:

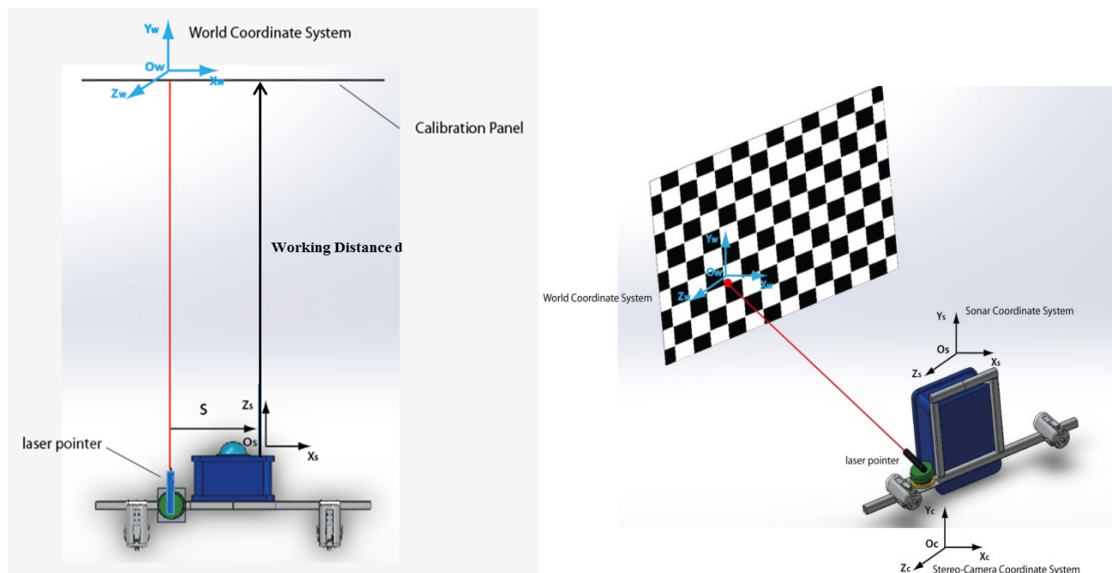


Fig. 2.10: *Setup* per la calibrazione del sistema con ausilio di puntatore laser.

Questa configurazione, come si può osservare, prevede il posizionamento del pannello di calibrazione di fronte al sistema ottico - acustico, ad una certa distanza d di lavoro, e l'impiego di un puntatore laser installato sulla struttura di supporto del prototipo, a distanza s rispetto all'origine del sistema di riferimento solidale con la camera acustica, utilizzato al fine di proiettare uno *spotlight* sul pannello di calibrazione stesso. Vincolando la posa del pannello di calibrazione in modo tale che esso risulti perfettamente ortogonale alla direzione di propagazione del fascio laser, la trasformazione rigida tra i sottosistemi ottico ed acustico può essere stimata a partire dalla conoscenza della geometria del problema. Infatti, mentre lo *spotlight* risulta "invisibile" alla camera acustica, esso può essere rilevato dal sottosistema stereo ottico e di conseguenza le sue coordinate essere determinate rispetto alla terna cartesiana solidale con lo stesso. Poiché il vincolo posto sulla posa del pannello di calibrazione equivale a considerare una matrice di orientamento \mathbf{R} unitaria, l'unica incognita del problema risulta essere il vettore di traslazione t , le cui componenti possono essere comunque stimate a partire dalla conoscenza delle distanze d ed s .

In effetti, indicando con d_s la misura della distanza di lavoro rispetto al sistema di riferimento acustico, con $p_l = [x_l, y_l, z_l]^T$ le coordinate dello *spotlight* rispetto al sistema di riferimento stereo ottico e supponendo che il puntatore laser giaccia sul piano di equazione $y_s = 0$ (secondo il sistema di riferimento acustico), avremo:

$$x_t = x_l + s; \quad y_t = y_l; \quad z_t = z_l - d_s; \quad (2.3)$$

Sebbene in linea di principio risulti di estrema facilità, la soluzione proposta, in realtà, presenta diversi aspetti critici legati principalmente alla precisione con cui può essere stabilita la posizione della terna cartesiana solidale con il sottosistema acustico (da cui dipende sia la misura s che la validità o meno del vincolo posto sulla coordinata y_t del vettore di traslazione) e della misura d_s della distanza di lavoro fornita dalla camera acustica stessa. Infine, un certo grado di approssimazione si avrà sull'ipotesi di unitarietà della matrice di rotazione \mathbf{R} , dipendentemente dal grado di precisione con cui è possibile garantire la condizione di ortogonalità tra pannello di calibrazione e direzione di propagazione del fascio laser. Nonostante questa prima ipotesi implementativa debba essere scartata a causa delle potenziali sorgenti di errore

individuata, il *setup* iniziale può essere comunque utilizzato come punto di partenza per la sintesi di una soluzione alternativa alla precedente, in grado di mitigare, se non eliminare del tutto, le diverse criticità riscontrate. Consideriamo dunque la configurazione presentata in Fig. 2.11.

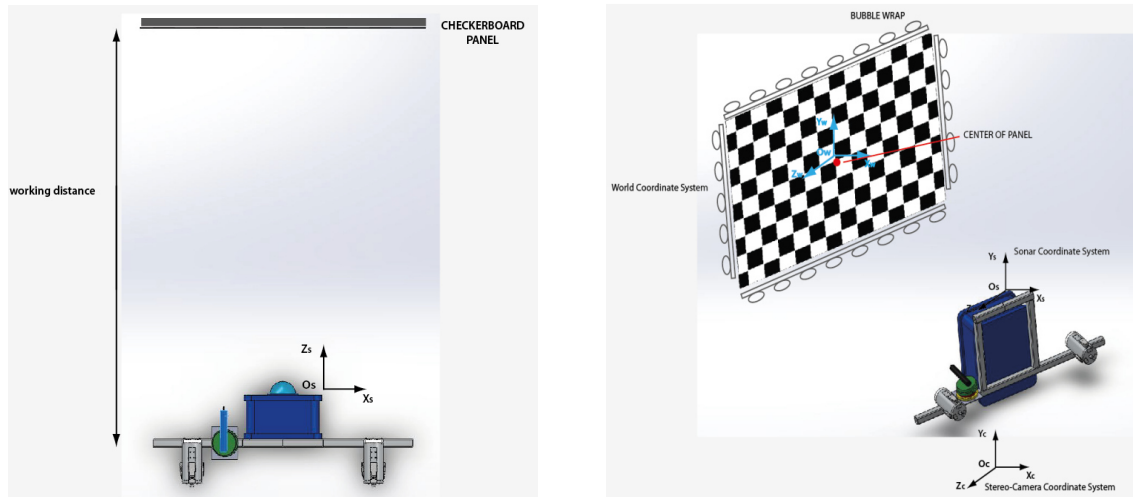


Fig. 2.11: *Setup* per la calibrazione del sistema con stima baricentro geometrico del pannello.

Anche qui il pannello di calibrazione è posto di fronte al sistema ottico-acustico, ad una certa distanza d di lavoro, ma a differenza della soluzione precedente nessun vincolo è posto circa la sua posa all'interno della scena. Inoltre, in questo approccio, non è previsto l'impiego di nessun puntatore laser. La stima dei parametri estrinseci R e t , in questo caso, si basa sulla determinazione delle coordinate cartesiane del baricentro geometrico del pannello di calibrazione rispetto sia al sistema di riferimento stereo ottico che a quello acustico e su opportune elaborazioni condotte sulla nuvola acustica 3D. Per quanto concerne il primo aspetto, cioè la determinazione delle coordinate cartesiane del baricentro geometrico del pannello di calibrazione rispetto ai sistemi di riferimento stereo ottico ed acustico, la soluzione individuata prevede una tipologia di elaborazione altamente differente per i due sottosistemi, a causa della diversa natura dei dati forniti in *output* dai due sensori. Infatti, mentre per il sottosistema stereo ottico le coordinate del baricentro del pannello di calibrazione risultano immediatamente disponibili, grazie alla sua procedura di calibrazione, più critico risulta invece essere il calcolo di tali coordinate rispetto al sistema di riferimento della camera acustica, principalmente a causa della rumorosità insita in questa tipologia di dati e per

la particolare scelta dei materiali utilizzati nella realizzazione del pannello di calibrazione. Come già detto, esso è costituito da una scacchiera realizzata in materiale DIBOND (materiale composito formato da due fogli di alluminio con un anima di materiale plastico) in modo tale da poterne permettere l'immersione in acqua, circondata da una cornice d'alluminio rivestita di *pluriball*. Poiché il materiale utilizzato è caratterizzato da una scarsissima riflettività acustica, la scacchiera risulta completamente "trasparente" durante l'acquisizione e, di conseguenza, la nuvola acustica 3D fornita dal sonar *Echoscope* presenterà punti appartenenti, in linea di principio, esclusivamente alla cornice che circonda la scacchiera stessa. Risulta allora evidente come le coordinate del baricentro del pannello di calibrazione rispetto alla terna solidale con il sottosistema acustico possano essere stimate, a partire dalla nuvola acustica 3D, solamente in maniera "indiretta". Poiché forma e dimensione del pannello sono comunque parametri noti, la soluzione proposta prevede di stimare le coordinate del baricentro geometrico come intersezione delle diagonali o delle mediane che possono essere tracciate a partire dai punti acustici 3D rappresentanti la cornice di *pluriball* che circonda la scacchiera. Dunque, dalla conoscenza delle coordinate del baricentro geometrico del pannello di calibrazione rispetto ai due sistemi di riferimento risulta immediato determinare il vettore di traslazione t .

Infatti, indicando con $p_{b,c} = (x_{b,c}, y_{b,c}, z_{b,c})$ e $p_{b,s} = (x_{b,s}, y_{b,s}, z_{b,s})$ le coordinate del baricentro geometrico, rispettivamente nel sistema di riferimento stereo ottico ed acustico, si avrà:

$$x_t = (x_{b,s} - x_{b,c}); \quad y_t = (y_{b,s} - y_{b,c}); \quad z_t = (z_{b,s} - z_{b,c}). \quad (2.4)$$

L'equazione (2.4) è valida nell'ipotesi che i due sistemi di riferimento abbiano lo stesso orientamento. In questo caso, dunque, poiché nessun vincolo è stato posto circa l'orientamento del pannello di calibrazione rispetto alla posizione del sistema ottico acustico, oltre al vettore di traslazione t , sarà allora necessario determinare preliminarmente la matrice di rotazione \mathbf{R} , non essendo più valida qui l'ipotesi di unitarietà. Tuttavia, anche per quanto concerne questa fase della metodologia di calibrazione, la soluzione proposta si basa su un approccio al problema di tipo geometrico. In effetti, poiché tutti i punti appartenenti al pannello di calibrazione sono

vincolati al piano su cui giace il pannello stesso, la matrice \mathbf{R} può essere determinata “allineando” i piani che possono essere estrapolati dai dati provenienti dai due sottosistemi. Mentre per la camera stereo ottica la sua procedura di calibrazione consente di definire con estrema precisione tale piano (è sufficiente, in tal caso, calcolare le componenti di due vettori ortogonali che giacciono sul piano stesso a partire dalla conoscenza delle coordinate dei punti ottici 3D, con un errore comparabile con la dimensione dei *pixel* che formano l’immagine ottica stessa), per quanto riguarda il sottosistema acustico, l’elevata rumorosità dei dati e la presenza di numerosi valori anomali non consentono un approccio così immediato. Questa problematica, comunque, può essere affrontata applicando opportune tecniche di *filtering* alla nuvola acustica 3D (al fine di ridurre la rumorosità presente sui dati ed eliminare la presenza di eventuali valori anomali), ed applicando successivamente un qualsiasi algoritmo di “*fitting* planare”, ottenendo in tal modo il piano che “meglio approssima” (in termini di errore quadratico medio) quello cercato.

Analizzando questa seconda metodologia per la sintesi di una metodologia di calibrazione per la camera ottico acustica proposta, si può notare come, molte delle criticità individuate nel primo approccio siano state qui superate. Sebbene la soluzione sia nettamente più robusta della precedente, la tecnica presenta, tuttavia, alcuni aspetti delicati che necessitano di un’investigazione più accurata. La problematica maggiore riguarda senza dubbio l’accuratezza dei risultati ottenuti dalle fasi di *filtering* e di *fitting* planare della nuvola acustica 3D. Infatti, a causa della scarsa risoluzione spaziale della camera acustica e dell’elevata rumorosità presente sui dati da essa forniti, un dato errore si avrà sia sulla scelta dei punti 3D dai quali tracciare le diagonali o le mediane attraverso le quali, successivamente, determinare il valore delle coordinate del baricentro geometrico del pannello rispetto al sistema di riferimento acustico, sia sulla ricerca del piano di *fitting* che meglio approssima l’orientamento della nuvola acustica 3D nello spazio. Infine, bisogna sottolineare che, a differenza delle metodologie presenti in letteratura, la procedura proposta risulta totalmente manuale.

2.4 Metodologia di calibrazione

Pur presentando alcuni aspetti di criticità che richiedono un'analisi più approfondita, in particolare un'accurata caratterizzazione dell'errore che si ha sia nella fase di *filtering* che in quella di *fitting* planare, la seconda soluzione proposta nella sezione 2.2 rappresenta, senza alcun dubbio, una solida base di partenza per la sintesi del protocollo di calibrazione cercato.

Alla luce dell'analisi condotta in precedenza, in questa sezione andremo ad analizzare in dettaglio i vari *step* che costituiscono il protocollo di calibrazione implementato, come illustrato in Fig. 2.12, in cui è mostrata la *pipeline* di elaborazione a partire dai dati provenienti dai sensori ottici ed acustici.

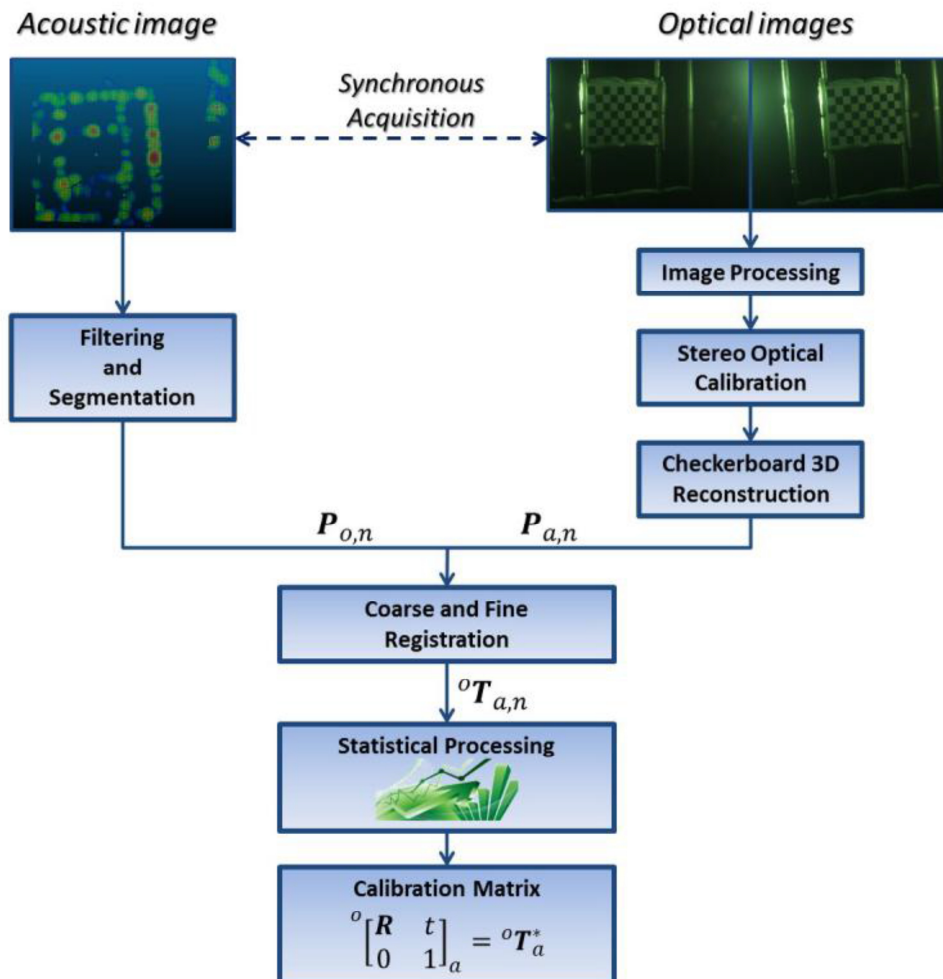


Fig. 2.12: *pipeline* di elaborazione dei dati ottici ed acustici.

La metodologia si compone di due *thread* distinti, legati ai due canali di *sensing*, che si vanno ad integrare negli ultimi stadi di elaborazione. A partire da un'acquisizione sincrona delle n pose assunte dal pannello di calibrazione durante la fase iniziale del processo, l'intera metodologia ha come obiettivo quello di ottenere n coppie di nuvole di punti 3D, dove la n -esima coppia è formata dalla nuvola di punti 3D ottica $\mathbf{P}_{0,n}$ e da quella acustica $\mathbf{P}_{a,n}$, dalle quali è possibile determinare, applicando algoritmi di *coarse* e *fine registration*, n stime ${}^o\mathbf{T}_{a,n}$ della matrice di trasformazione rigida cercata. Infine, la matrice di trasformazione finale ${}^o\mathbf{T}_a^*$ è ottenuta elaborando statisticamente il *dataset* composto dalle n matrici di trasformazione ${}^o\mathbf{T}_{a,n}$ ottenute a valle della precedente fase di registrazione.

2.4.1 Canale di *sensing* acustico

In questa sezione si descriverà la procedura di elaborazione eseguita sui dati tridimensionali ottenuti dalla camera acustica *Echoscope*. Preliminarmente, comunque, sarà condotta una breve analisi sul suo principio di funzionamento al fine di comprendere in maniera più accurata la modalità attraverso la quale si formano e vengono rappresentate le immagini acustiche tridimensionali fornite dal sistema stesso.

Illuminando la scena attraverso un impulso acustico ad alta frequenza, l'*array* bidimensionale di trasduttori dell'unità di ricezione della camera acustica *Echoscope* raccoglie i segnali di *backscattering* provenienti dagli oggetti presenti sulla scena stessa. L'intero *set* di segnali *raw* viene dunque processato in modo tale che essi "risultino in fase", formando così un insieme di segnali opportunamente elaborati (chiamati *beam signals*) i cui profili dipendono dagli *echo* provenienti da direzioni prefissate, attenuando allo stesso tempo quelli provenienti da direzioni diverse. La distanza di un punto 3D può allora essere misurata rilevando l'istante temporale in cui si ha il massimo picco di energia del segnale (Fig. 2.13, sinistra).

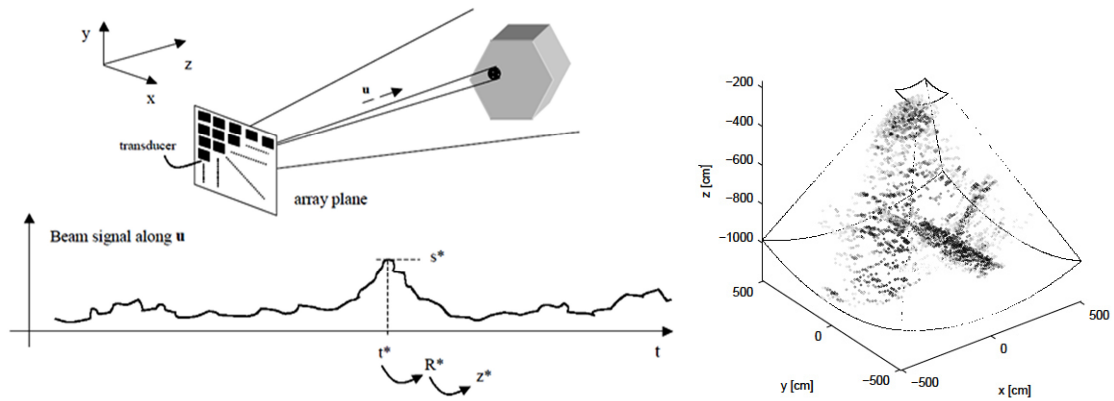


Fig. 2.13: principio di funzionamento sonar *Echoscope* (sinistra), rappresentazione tramite *voxel* immagine acustica (destra).

L'immagine 3D fornita dalla camera acustica è formata da una nuvola di 128 x 128 punti ordinati in accordo ad un sistema di riferimento polare, in modo tale che punti adiacenti corrispondono a *beam signal* adiacenti. Inoltre, l'intensità del picco massimo può essere utilizzata al fine di generare un'ulteriore immagine, registrata con la prima, rappresentante l'affidabilità delle misure 3D associate, dunque, in generale, più alta è l'intensità, più affidabile è la misura fornita.

Il sonar 3D *Echoscope* normalmente memorizza, per ogni singolo *beam*, le misure di angoli e distanza, insieme ad una matrice di movimento che descrive la sua posizione rispetto ad un sistema di riferimento locale. Questa rappresentazione è utilizzata poiché essa richiede molto meno spazio di memoria piuttosto che memorizzare una coordinata 3D per ogni *beam*, ed inoltre consente di elaborare ulteriormente i dati al fine di determinare più accuratamente la posizione 3D stessa, introducendo delle correzioni basate su un profilo di velocità dell'onda acustica. I dati, dunque, opportunamente corretti, vengono infine trasformati in coordinate cartesiane 3D in modo tale da poter essere visualizzati a schermo. Per quanto concerne quest'ultimo aspetto, la camera acustica *Echoscope* rappresenta i dati come punti nello spazio tridimensionale (*voxels*, Fig. 2.13, destra). In questa tipologia di rappresentazione, che senza dubbio è la più utile per l'interpretazione umana, i livelli di grigio, o equivalentemente la particolare colorazione dei *voxels*, è determinata dalla corrispondente intensità acustica, mentre i limiti del volume di visualizzazione vengono rappresentati attraverso delle linee solide.

2.4.1.1 Filtraggio e segmentazione

L'immagine acustica fornita dal sonar 3D può essere corrotta sia da false riflessioni causate dai lobi secondari dell'*array* di ricezione che dal rumore presente nella fase d'acquisizione dei segnali di *backscattering*, quest'ultimo modellato come *speckle noise*. I lobi secondari sono responsabili della sfocatura dell'oggetto, mentre il rumore *speckle* genera risposte nulle o molto basse all'interno dell'oggetto stesso [70]. Sebbene l'*Echoscope* esegue direttamente un'elaborazione preliminare di basso livello, risulta comunque necessario effettuare opportune operazioni sulla nuvola acustica 3D, al fine di "ripulire" le immagini ottenute dalla presenza di rumore o di eventuali valori anomali.

Da quanto affermato risulta evidente come le operazioni di *filtering* (per la riduzione del rumore presente sui dati ed eliminazione di eventuali valori anomali) e di *segmentation* (per discriminare, nella scena osservata, tra oggetti e *background*) rappresentano fasi preliminari fondamentali per l'esecuzione di un qualsiasi algoritmo di fusione applicato a questa particolare tipologia di dati. Per quanto concerne il problema della sintesi di un'efficiente metodologia di *filtering* e *segmentation* della nuvola acustica 3D risulta possibile, in linea di principio, "sfruttare" le diverse informazioni fornite dalla camera acustica stessa (range, intensità echo, matrice di connessione). A partire da queste differenti tipologie d'informazione è possibile implementare differenti algoritmi di *filtering* e *segmentation*, più o meno automatici, con prestazioni alquanto differenti tra loro [71]. La soluzione adottata in questa prima implementazione del protocollo di calibrazione in oggetto, prevede, comunque, una procedura di *filtering* e *segmentation* completamente manuale, realizzata utilizzando il software *open source* CloudCompare [72], posticipando ad eventuali future implementazioni una soluzione più o meno automatica.

2.4.2 Canale di *sensing* ottico

Obiettivo di questo paragrafo è la descrizione delle varie fasi che, a partire dall'acquisizione di una coppia di immagini da parte del sottosistema stereo ottico, consentono di ottenere una ricostruzione 3D della scena osservata.

Lo studio delle diverse tecniche e degli algoritmi per la ricostruzione 3D attraverso un'acquisizione stereo ottica rappresenta un argomento già ampiamente trattato in letteratura scientifica [73-76]. Riferendosi, dunque, a tali documenti per una trattazione più approfondita, in questa sezione, a partire da una panoramica generale sulle fasi principali che caratterizzano tale processo, andremo ad analizzare in dettaglio solo gli aspetti che effettivamente contraddistinguono i moduli che compongono il canale di *sensing* ottico del protocollo di calibrazione della camera ottico acustica proposta.

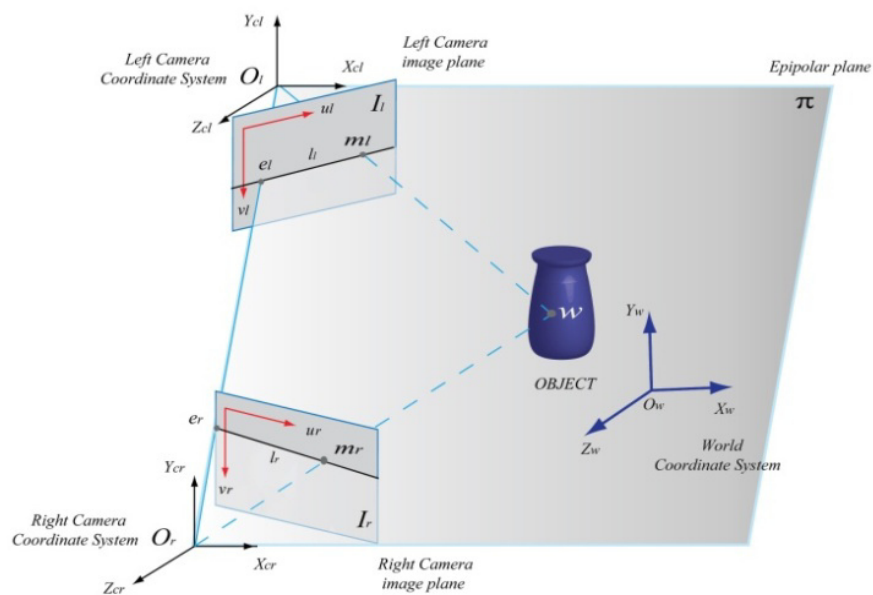


Fig. 2.14: Ricostruzione di un punto 3D di un'oggetto visto da due sistemi di riferimento diversi (due camere), in base al principio di triangolazione

In generale, al fine di ottenere le coordinate 3D di un punto di una scena ripresa da due camere, è necessario poter individuare nelle due immagini (destra e sinistra) lo stesso punto e conoscerne la posizione relativa nello spazio, come rappresentato schematicamente in Fig. 2.14, ciò in base al principio di triangolazione. Attraverso la calibrazione del sottosistema stereo ottico è possibile stimare sia la posizione relativa tra i sistemi di riferimento delle due camere sia le distorsioni ottiche dovute allo scostamento del modello di camera ideale (modello *pinhole*) da quello reale, in cui sono considerati gli effetti delle distorsioni prodotte dal pacchetto ottico.

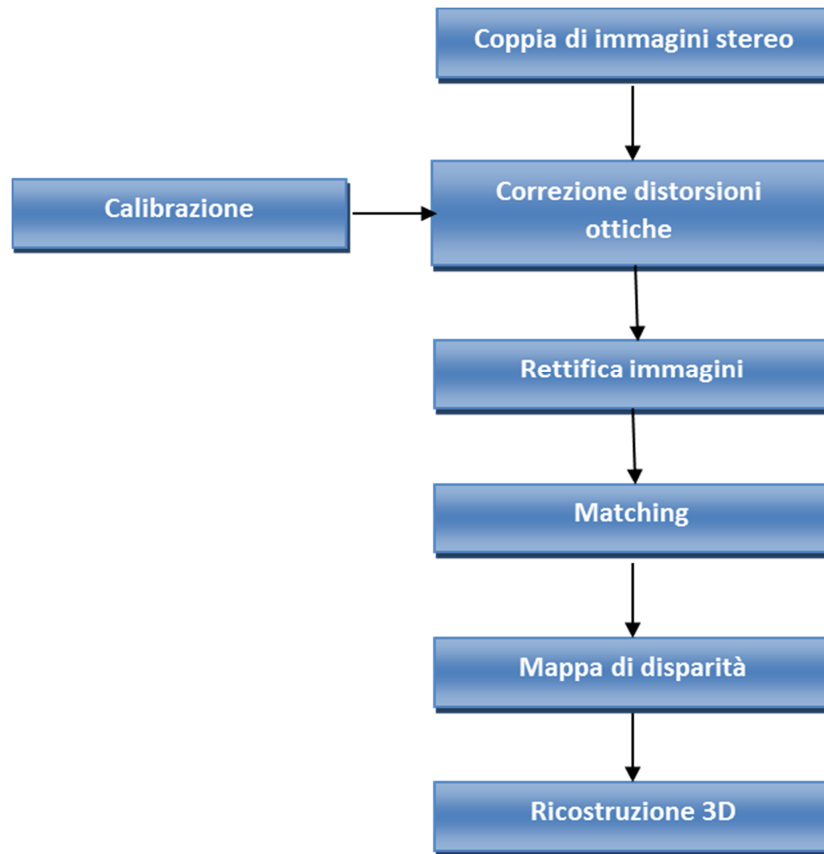


Fig. 2.15: *Flow-chart* delle fasi principali della ricostruzione 3D a partire da una coppia di immagini stereo.

In Fig. 2.15 è presentato un diagramma in cui sono rappresentate le fasi principali che contraddistinguono il processo di ricostruzione 3D. A partire dall'acquisizione di una coppia d'immagini stereo, dopo la correzione delle distorsioni, ottenuta a valle del processo di calibrazione del sistema, è possibile effettuare la *Rettifica*, ovvero è possibile ottenere una configurazione equivalente a quella *standard* (rette epipolari coniugate orizzontali e collineari, come mostrato in Fig. 2.14), applicando a ciascuna delle due immagini una trasformazione (omografia) detta, appunto, *Rettifica*. Col termine configurazione *standard* (detta anche caso normale), si intende il caso in cui gli assi ottici delle due camere sono paralleli, distanti di una certa quantità, detta *baseline*. Nella fase di *Matching* viene invece stabilita la corrispondenza lungo le linee epipolari fra punti omologhi delle due immagini, consentendo in tal modo di ottenere, in base alla *baseline* e, quindi, ai dati ottenuti dalla calibrazione, la *Mappa di disparità*. Infine, la *Ricostruzione 3D* della scena è ottenuta come nuvola di punti 3D di cui è stata calcolata la corrispondenza in entrambe le camere.

2.4.2.1 Elaborazioni immagini ottiche

A causa dell'attenuazione (dovuta ai fenomeni di assorbimento e di *scattering*) che la luce subisce nella sua propagazione in acqua, le immagini subacquee sono generalmente affette da problemi quali basso contrasto, illuminazione non uniforme, sfocatura, limitato *range* di visibilità, attenuazione dei colori (in particolare modo del canale rosso) e da rumore. Tutti queste problematiche incidono fortemente sulla qualità dei risultati ottenuti nell'applicazione delle diverse tecniche di ricostruzione 3D, ed in particolar modo nelle fasi di calibrazione del sistema stereo ottico e di applicazione degli algoritmi di *Matching*. Riduzione di visibilità e alterazione dei colori limitano, dunque, il campo d'applicazione di queste tecniche di *computer vision*, rendendo necessario l'applicazione di opportuni algoritmi in grado di migliorare la qualità delle immagini acquisite in ambiente subacqueo.

In linea generale, il *pre-processing* d'immagini subacquee può essere effettuato attraverso due diversi approcci: utilizzando "tecniche di restauro digitale" o "metodi di miglioramento delle immagini" [77,78].

Le tecniche di restauro digitale mirano a recuperare un'immagine degradata utilizzando un "modello matematico della degradazione" e della "formazione dell'immagine originale"; si tratta, essenzialmente, di un problema d'inversione. Questi metodi, anche se rigorosi, richiedono comunque la conoscenza di molti parametri del modello impiegato (come attenuazione e coefficienti di *scattering* caratterizzanti la torbidità dell'acqua) che sono solo scarsamente noti attraverso tabelle ed inoltre possono essere estremamente variabili. Un altro importante parametro richiesto è la stima della profondità di un dato oggetto nella scena.

Le tecniche di miglioramento dell'immagine utilizzano, al contrario, criteri "soggettivi qualitativi" per produrre un'immagine che visivamente risulti di qualità superiore, senza l'ausilio di alcun modello fisico per la formazione dell'immagine. Questa tecniche, non richiedendo una conoscenza a priori dell'ambiente, risultano essere più semplici e veloci rispetto ai metodi precedenti e di più largo impiego.

Una delle tecniche più diffuse per l'elaborazione delle immagini subacquee si basa sull'algoritmo ACE (*Automatic Colour Equalization*) [79] la cui implementazione è ispirata al sistema visivo umano il quale è in grado di adattarsi a condizioni d'illuminazione molto variabili ed estrarre informazioni visive dall'ambiente subacqueo

[80]. Nella prima parte del metodo ACE i dati cromatici, cioè i pixel, vengono elaborati e corretti secondo le informazioni contenute nell'immagine stessa. Successivamente, i colori dell'immagine ottenuta vengono ripristinati e migliorati.

In [81] è proposto un algoritmo automatico di miglioramento che non richiede alcun parametro di correzione. Ciascuna sorgente di errore viene corretta in maniera sequenziale. Nel primo passo si rimuove l'effetto moiré, viene poi applicato un filtro omomorfo o frequenziale per uniformare la luminosità e per aumentare il contrasto. Per ridurre il rumore di acquisizione l'algoritmo utilizza un filtro *wavelet denoising* seguito da un filtro anisotropo. Il processo viene applicato sulla componente Y dello spazio colore YCbCr, allo scopo di ottimizzare il tempo di calcolo. Infine, nella soluzione proposta in [82] si utilizzano, per migliorare le immagini subacquee, algoritmi *slide stretching* sia nello spazio RGB che HIS.

Come affermato in precedenza, il miglioramento delle immagini subacquee rappresenta una tecnica fondamentale per migliorare le prestazioni di algoritmi di estrazione di *features*. In [83] sono stati investigati gli effetti di diversi metodi di *pre-processing* sulle prestazioni del SURF detector [84] ed è stato proposto un nuovo metodo, denominato IACE (*Image Adaptive Contrast Enhancement*). Per quanto concerne il protocollo di calibrazione proposto, la soluzione adottata si basa su una nuova metodologia di filtraggio sviluppata presso il DIMEG dell'Università della Calabria, la quale rappresenta un'innovazione rispetto allo stato dell'arte noto in letteratura. Questo metodo si basa sulla tecnica dell'*Histogram Stretching* e su una procedura di ritocco manuale del colore eseguita su un campione di immagini. L'algoritmo è stato implementato utilizzando delle *batch action* al fine di estrapolare la massima quantità di informazioni da una serie di immagini affette da rumore.

Allo scopo, è stata scelta un'immagine rappresentativa dell'ambiente subacqueo per eseguire ritocco ed equalizzazione, applicando una procedura in quattro fasi, come riportato di seguito:

1. *Histogram Stretching* preliminare per migliorare il contrasto;
2. miscelazione dei canali di colore per bilanciare le informazioni mancanti sul canale rosso;

3. creazione di una serie di livelli di regolazione sull'immagine campione per il miglioramento della saturazione di alcuni colori mancanti e per le maschere di contrasto, bilanciamento colore e di compensazione;
4. esportazione delle statistiche dell'immagine campione sulle altre immagini, con diverse percentuali di miscelazione.

2.4.2.2 Calibrazione stereo ottica

In letteratura si trovano diversi algoritmi e metodi per la calibrazione di un dispositivo ottico, generalmente basati su modelli proiettivi. La calibrazione di un dispositivo ottico subacqueo deve tenere conto dell'effetto di rifrazione all'interfaccia (piatta o emisferica) aria-acrilico e acrilico-acqua, che si manifesta quando una camera è alloggiata all'interno di un *case*. Gli effetti della rifrazione possono essere considerati attraverso la modellazione esplicita dei cammini ottici utilizzando il *Ray tracing* [84] o incorporando implicitamente l'effetto della rifrazione in alcuni parametri di calibrazione della camera stessa. Nel primo caso, le camere sono tarate prima in aria e successivamente calibrate in acqua al fine di derivare la geometria delle interfacce di rifrazione, poiché la componente principale di entrambi gli effetti (rifrazione e di distorsione dell'immagine) presentano unicamente una componente radiale. Sebbene l'assunto sugli indici di rifrazione sia per la porta acrilica a cupola sia per il percorso dell'aria all'interno del *case* è affidabile, le variazioni di temperatura, pressione e salinità dell'acqua provocano piccole variazioni dell'indice di rifrazione dei mezzi che non possono essere evitati. L'approccio implicito, in alternativa, incorpora gli effetti di rifrazione delle componenti ottiche e delle interfacce di rifrazione nei parametri stessi di calibrazione della camera.

In questo lavoro, considerando la semplicità d'applicazione e l'affidabilità dei risultati che si ottengono, si è scelto di adottare una procedura basata su un approccio implicito. Parametri intrinseci ed estrinseci delle camere che costituiscono il sottosistema stereo ottico del prototipo di camera ottica acustica proposto sono ottenuti relazionando le coordinate di n punti noti w_i situati in un campione di calibrazione (nel caso specifico, il pannello di calibrazione a scacchiera), con le coordinate corrispondenti $m_i = [u_i; v_i]$ nel piano immagine. Come riportato in [76] da tale procedura si ottiene un set di $2n$ equazioni omogenee lineari che può essere

risolto col metodo dei minimi quadrati. In teoria, sei punti non complanari sono sufficienti per determinare l'incognita \mathbf{P} (cioè la *Perspective Projection Matrix*) che rappresenta il modello geometrico della camera, ma in pratica sono necessari più punti per compensare gli errori di misurazione dovuti, ad esempio, ad eventuali disturbi dell'immagine. Questo metodo è conosciuto come "Trasformazione Lineare Diretta", ed ha il pregio di minimizzare l'errore algebrico facendo ricorso ad un modello di camera semplificato, per contro, i risultati sono meno accurati. Per ottenere una maggiore precisione si può fare ricorso ad un metodo non lineare. Tale tipo di approccio riduce l'errore geometrico, portando alla formulazione di una funzione oggettiva non lineare che rappresenta la distanza tra i punti \mathbf{m}_i dell'immagine e la proiezione dei punti \mathbf{w}_i nell'immagine stessa: la cosiddetta ri-proiezione dell'errore. Il passo successivo consiste nel calcolare i parametri estrinseci del sistema (calibrazione stereo), correlando ogni singola immagine prodotta dalle camere con il sistema di coordinate globali. Lo scopo della calibrazione stereo è quello di trovare la relazione tra il sistema di coordinate globale e i due sistemi di coordinate solidali con le due camere:

$$X_{cl} = M_l X_w \tag{2.4}$$

$$X_{cr} = M_r X_w$$

dove $\mathbf{M}_l = [\mathbf{R}_{cl}; \mathbf{t}_{cl}]$ è la matrice di trasformazione tra le coordinate della camera sinistra e le coordinate del sistema globale, $\mathbf{M}_r = [\mathbf{R}_{cr}; \mathbf{t}_{cr}]$ è la matrice di trasformazione tra le coordinate della camera destra e le coordinate del sistema globale, e \mathbf{x}_{cl} , \mathbf{x}_{cr} , \mathbf{x}_w rappresentano le coordinate del punto \mathbf{w} dell'oggetto rispettivamente nella camera sinistra, destra e nel sistema globale. Le matrici di trasformazione \mathbf{M}_l ed \mathbf{M}_r rappresentano le rotazioni (\mathbf{R}_{cl} ; \mathbf{R}_{cr}) e le traslazioni (\mathbf{t}_{cl} ; \mathbf{t}_{cr}) tra il sistema di coordinate globali e i sistemi di coordinate delle camere. I parametri estrinseci possono essere ottenuti infine attraverso lo stesso processo iterativo utilizzato per stimare i parametri intrinseci di ogni camera. Tale processo si basa sulla corrispondenza dei punti del campione di calibrazione nella coppia di immagini e consiste in un processo di *problem solving* d'ottimizzazione generale, volta alla minimizzazione degli errori di ri-proiezione

tra i piani delle due camere. Dalle sperimentazioni effettuate in vasca presso i laboratori di Meccanica dell'Unical, e dal *know-how* acquisito nel corso di diverse sperimentazioni effettuate e pubblicate su diverse riviste scientifiche, è stato dimostrato che la calibrazione effettuata in acqua applicando il metodo proposto in [86] può essere usato, con profitto, per calibrare un sistema stereo ottico al fine di ottenere una ricostruzione 3D della scena acquisita. Alla luce di queste considerazioni, si è dunque deciso di applicare il metodo di Bouguet per la calibrazione del sistema stereo ottico, ed adottare quindi un metodo implicito per il calcolo della focale delle ottiche utilizzate, senza modellare matematicamente l'effetto della rifrazione all'interfaccia aria-acrilico (oblò) – acqua. Per la calibrazione di Bouguet del sottosistema stereo ottico della camera ottico acustica in oggetto, è stato utilizzato il *Camera Calibration Toolbox* di *Matlab*, che necessita dell'intervento manuale dell'operatore per selezionare i quattro spigoli esterni al fine d'individuare il piano della scacchiera di calibrazione.

2.4.2.3 Ricostruzione 3D della scacchiera

Date due immagini prese da due posizioni leggermente differenti, un algoritmo di corrispondenza stereo cerca di identificare i punti corrispondenti nelle due immagini, relative allo stesso punto nella scena, risolvendo il cosiddetto problema della corrispondenza. Per semplificare la ricerca dei punti coniugati, le immagini vengono generalmente corrette (rettifica epipolare) sistemando l'attrezzatura stereo in una particolare configurazione in cui entrambi gli epipoli sono posti all'infinito e le linee epipolari formano un fascio di linee parallele nelle immagini di destra e sinistra. In questo caso ogni coppia di punti corrispondenti giacerà sulla stessa riga dell'immagine (*scan-line*) cosicché il problema della corrispondenza sarà ridotto ad una ricerca monodimensionale lungo ogni linea epipolare.

La disparità, che codifica la profondità della scena, rappresenta la distanza tra le coordinate x di una coppia di punti corrispondenti nelle immagini rettificate destra e sinistra. È importante sottolineare che, poiché la scena è acquisita da due differenti punti di vista, è possibile incorrere in false corrispondenze dovute a distorsione radiometrica o distorsione prospettica.

Nell'approccio stereo passivo la ricerca di punti omologhi (problema della corrispondenza) avviene per mezzo di cosiddetti algoritmi di *matching* stereo, basati su misure di similarità dei pixel relativi allo stesso punto della scena nella coppia di immagini (in linea di principio, un punto della scena apparirà simile nelle immagini prodotte dalla coppia di camere stereo). Un approccio di questo tipo funziona bene con oggetti caratterizzati da una *texture* consistente, mentre si incorre in errori ed in aree non ricostruite in presenza di superfici di colore uniforme.

Vi sono due principali metodi per la corrispondenza automatica: *pattern* d'intensità immagine (*Area Based Matching*, ABM), ad esempio finestre di valori di grigio nell'intorno del punto d'interesse, ed estrazione di *features* (*Features Based Matching*, FBM), per esempio bordi e linee di separazione tra superfici con colorazione differente. In genere il metodo FBM viene integrato a quello ABM fornendo risultati sensibilmente migliori rispetto all'utilizzo della sola tecnica ABM (in molti *software* l'unica a essere implementata) soprattutto in zone in cui sono presenti molte discontinuità.

I recenti sviluppi nel campo della ricostruzione 3D testimoniano le grandi potenzialità del metodo fotogrammetrico per derivare informazioni 3D dettagliate e precise alle diverse scale, paragonabili a quelle ottenute mediante sensori attivi. Il problema cardine della stereo visione oltre al calcolo delle corrispondenze è quello del calcolo della disparità. Una coppia di punti coniugata è costituita da due punti in due immagini diverse che sono proiezione dello stesso punto nella scena reale tridimensionale. La disparità non è altro che la differenza tra due punti coniugati, supponendo di sovrapporre le due immagini. La mappa di disparità può essere rappresentata da un'immagine in scala di grigi che mostra in corrispondenza di ciascun punto dell'immagine di riferimento il valore di disparità associato a quel punto. Per favorirne la visualizzazione, i valori di disparità vengono mappati per mezzo di un opportuno fattore di scala nel *range* (0, 255).

Nella metodologia proposta la ricostruzione è stata realizzata utilizzando le librerie openCV [87] e Libelas [88]. In particolare, la libreria openCV è stata utilizzata per la rettifica della coppia stereo mentre la libreria Libelas per l'implementazione dell'algoritmo di *matching stereo* e la generazione della mappa di disparità. Infine, la nuvola di punti 3D è stata determinata come prodotto tra la matrice di ri-proiezione ottenuta a valle del processo di calibrazione e la mappa di disparità stessa.

2.4.3 Registrazione ottico - acustica

In questo paragrafo descriveremo la fase della metodologia di calibrazione che permette di ottenere le n stime ${}^0\mathbf{T}_{a,n}$ della matrice di trasformazione rigida incognita che lega il sistema di riferimento solidale con la camera acustica con quello solidale con la camera stereo ottica. Considerando la n -esima coppia $\mathbf{P}_{0,n}\mathbf{P}_{a,n}$ di nuvole di punti 3D ottico ed acustiche, la ${}^0\mathbf{T}_{a,n}$ associata è determinata come composizione di trasformazioni ottenute attraverso algoritmi di *Coarse* e *Fine Registration* [89]. Assumendo, dunque, come riferimento il sistema solidale con la camera stereo ottica, nella fase di *Coarse Registration*, sono state eseguite due operazioni:

1. Calcolo della ${}^1\mathbf{T}_{a,n}$ mediante l'orientamento manuale del sistema di riferimento solidale con la camera acustica in modo tale che l'asse Z rappresenti la profondità della scena, coerentemente con quanto avviene per il sistema ottico; questo passaggio si rende necessario poiché i dati dell'Echoscope vengono rappresentati in un sistema di riferimento in cui la profondità della scena è espressa lungo l'asse Y.
2. Allineamento della coppia di nuvole di punti 3D $\mathbf{P}_{0,n}\mathbf{P}_{a,n}$ attraverso un algoritmo automatico implementato in CloudCompare che, a partire da una stima del baricentro delle due nuvole di punti 3D, determina il vettore di traslazione t che li lega (supponendo unitaria la matrice di rotazione \mathbf{R}).

Come risultato di tale operazione si ottiene la matrice di trasformazione ${}^2\mathbf{T}_{1,n}$.

Per quanto concerne la fase di Fine registration, allo scopo di affinare i risultati ottenuti dall'algoritmo di Coarse registration, è stata applicata, alla coppia $\mathbf{P}_{0,n}\mathbf{P}_{a,n}$ allineata nella fase precedente, un'implementazione in CloudCompare del noto algoritmo di registrazione ICP (*Iterative Closest Point*) ottenendo la matrice di trasformazione ${}^2\mathbf{T}_{1,n}$.

A valle delle operazioni precedenti, dunque, la matrice di trasformazione rigida ${}^0\mathbf{T}_{a,n}$ cercata è ottenuta come (Fig. 2.16):

$${}^0\mathbf{T}_{a,n} = {}^0\mathbf{T}_{2,n} \times {}^2\mathbf{T}_{1,n} \times {}^1\mathbf{T}_{a,n} \quad (2.5)$$

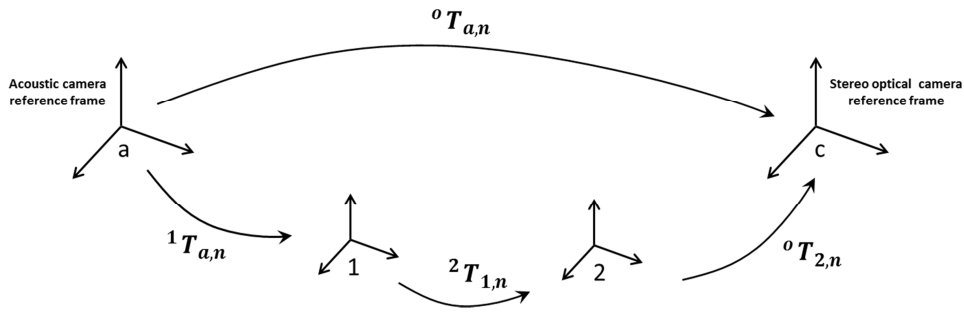


Fig. 2.16: registrazione ottico – acustica.

2.4.4 Elaborazione statistica

Differenti metodologie di elaborazione statistica possono essere applicate a partire dal *dataset* costituito dalle ${}^oT_{a,n}$ stime della matrice di trasformazione rigida cercata al fine di individuare la migliore approssimazione, definita come ${}^oT_a^*$.

Soluzioni più semplici possono consistere nell'applicazione di operazioni di media considerando tutti gli elementi ${}^oT_{a,n}$ del *dataset* o di un suo sottoinsieme determinato, ad esempio, sulla base del soddisfacimento o meno di un qualche criterio di selezione. Soluzioni più sofisticate possono invece prevedere l'impiego di un qualche algoritmo di ricerca in grado di stabilire, tra le ${}^oT_{a,n}$ matrici, quella che minimizza l'errore di registrazione tra tutte le n coppie $P_{0,n}P_{a,n}$ disponibili. Infine è possibile trattare il calcolo della matrice ${}^oT_a^*$ come un problema di ottimizzazione per la cui risoluzione possono essere impiegati, ad esempio, i ben noti algoritmi genetici [90].

Nella selezione della metodologia più efficiente da adottare per il calcolo della matrice ${}^oT_a^*$, affrontata nella sezione 2.5.5 mettendo a confronto diverse ipotesi implementative, è stata valutata la complessità della soluzione adottata e la bontà dei risultati ottenuti alla luce del fatto che esiste, comunque, un limite alla precisione e all'accuratezza con cui è possibile registrare la nuvola di punti 3D acustica con quella ottica, limite dato dall'errore che si ha nella ricostruzione della nuvola di punti 3D da parte del sistema stereo ottico e soprattutto dalla precisione con cui la camera acustica impiegata è in grado di ricostruire la scena osservata, considerando la sua bassa risoluzione spaziale. In effetti, risultati al di sotto di tale limite sarebbero del tutto casuali e legati allo specifico *dataset* utilizzato per determinare la matrice ${}^oT_a^*$.

2.5 Set-up sperimentale

La metodologia di calibrazione proposta è stata testata presso i laboratori della Whitehead Sistemi Subacquei S.p.A. di Pozzuoli (NA) in cui è presente una vasca (11 x 7 x 5 m) dotata di attrezzatura necessaria sia alla movimentazione del prototipo di camera ottico acustica realizzato sia dei target d'acquisire. Al fine di assicurare un supporto meccanico al sistema è stata progettata e realizzata una struttura costituita da profilati in alluminio, in cui sono stati alloggiati gli scafandri delle due camere ottiche e la testa della camera acustica *Echoscope*, quest'ultima connessa attraverso una piastra posteriore. L'intera struttura è stata poi fissata ad una staffa di supporto per la connessione ad un palo telescopico (Fig. 2.17).

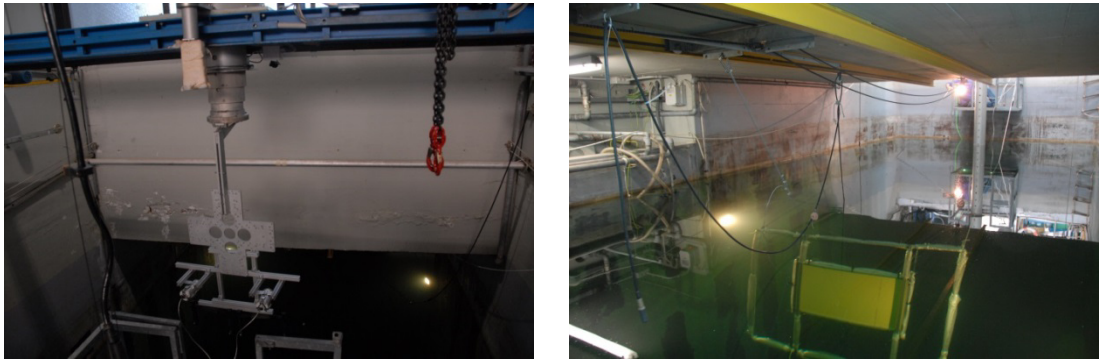


Fig. 2.17: vista palo telescopico per movimentazione prototipo camera ottico acustica (sinistra), sistema a cinghia per movimentazione pannello di calibrazione e *target* (destra) .

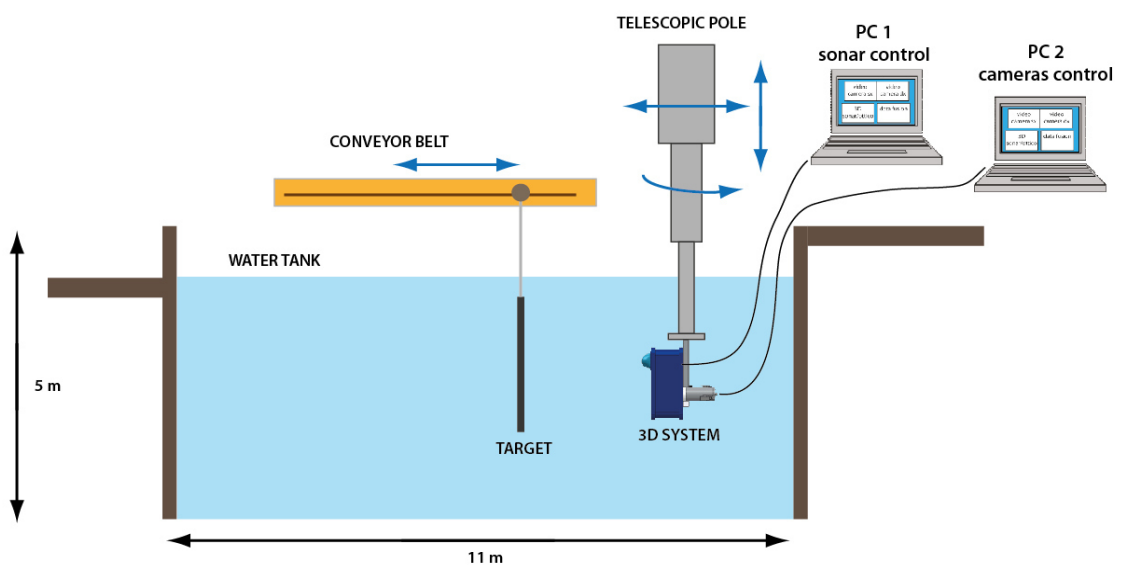


Fig. 2.18: Schematico del setup sperimentale realizzato presso i laboratori WASS.

Uno schematico del *layout* del *setup* sperimentale progettato è rappresentato in Fig. 2.18: in esso è possibile osservare sia le connessioni tra sistema d'acquisizione 3D (costituito dalla testa del sonar Echoscope e dagli scafandri in cui sono alloggiati le due camere ottiche) e PC, sia i vari sistemi di movimentazione del *target* e della camera ottico acustica stessa. Per quanto riguarda il sistema d'interconnessioni, la testa della camera acustica è collegata, tramite bus proprietario, al modulo di alimentazione Power Supply Unit (PSU), mentre la connessione alla *workstation* in cui vi sono installati i software di gestione dell'intero sottosistema acustico, cioè il Data Integration Unit (DIU) che gestisce sia il flusso dati proveniente dalla testa del sonar che l'invio e la ricezione dei comandi di controllo dello stesso e il software UIS Survey Explorer per la visualizzazione della nuvola acustica 3D, è realizzata tramite interfaccia seriale RS232 (per invio e ricezione comandi) ed Ethernet 10Mbps (per ricezione dati). Il sottosistema stereo ottico è invece connesso al PC in cui è installato il software *Flycapture* che ne consente la gestione ed il settaggio dei parametri d'acquisizione tramite interfaccia Firewire 800.

Il sistema ottico acustico è fissato ad un palo telescopico che consente di poter immergere il sistema stesso in vasca, offrendo inoltre la possibilità di poterlo movimentare secondo i tre gradi di libertà mostrati nello schematico, mentre il *target* d'acquisire è fissato ad un carrello che ne consente la sola traslazione longitudinale.

In Fig. 2. 19 sono invece mostrati i due pannelli utilizzati nelle varie fasi sperimentali.



Fig. 2.19 Pannello di calibrazione (sinistra), target realizzato con oggetti di materiale e forma diversa (destra).

Le loro dimensioni sono state determinate sulla base dell'analisi del campo di vista dei due sottosistemi e del range operativo previsto nei diversi test programmati (da 1.5 m

a 10 m circa). Oltre al *pattern* planare (una scacchiera di 9 x 7 quadrati ognuno di 100 x 100 mm) è stato realizzato un ulteriore pannello in cui sono stati inseriti oggetti di diversa forma e materiale. In particolare si possono osservare due vasi ceramici, piastrelle in marmo e tufo, mattoncini, tegola e maschera realizzati in terracotta.

2.6 Sperimentazione

In questa sezione andremo a descrivere le operazioni condotte nelle varie fasi che caratterizzano la campagna di test programmata al fine di valutare tre diversi aspetti: a) le *performance* del prototipo di camera ottico acustica proposto; b) la scelta del metodo di elaborazione statistica necessario per ottenere la migliore stima della matrice di calibrazione; c) validare il protocollo di calibrazione implementato.

A partire dalle operazioni preliminari di settaggio dei sottosistemi stereo ottico ed acustico la prima parte dei test è stata interamente incentrata sull'acquisizione sincrona dei *target* al fine di costruire i *dataset* d'immagini ottiche ed acustiche oggetto dell'elaborazioni delle successive fasi sperimentali attraverso le quali viene determinata la matrice di trasformazione rigida cercata. La metodologia proposta è stata infine validata applicando la matrice di trasformazione determinata a valle del processo di elaborazione alla nuvola acustica 3D rappresentante il pannello *target* costituito dagli oggetti di diversa forma e materiale.

2.6.1 Acquisizioni immagini

Prima dell'inizio della fase d'acquisizione vera e propria sono state effettuate operazioni preliminari di *setting* dei sensori della camera ottico acustica. In particolare, per quanto concerne il sottosistema stereo ottico, sono stati effettuati, in aria, i settaggi per la messa a fuoco delle camere e sono state effettuate delle prime prove di acquisizione in acqua per verificare la corretta sovrapposizione degli angoli di vista, come previsto in fase di progettazione del *layout* della struttura di supporto, mentre si è deciso di utilizzare la frequenza di lavoro di 610 KHZ per quanto riguarda il sonar Echoscope. In questa fase si è potuto verificare come, al fine di avere una corretta sovrapposizione dei campi di vista delle due camere nel range di operatività previsto

gli assi ottici non possono essere mantenuti paralleli, ma è stato necessario renderli leggermente convergenti. Per quanto concerne la metodologia di calibrazione, è stata acquisita una sequenza di 20 pose differenti del pannello di calibrazione posizionandolo a partire da una distanza di circa 1.5 m fino a circa 10 m, per diversi orientamenti, mantenendo fissa la posizione della camera ottica acustica agganciata al palo telescopico (Fig. 2.20).

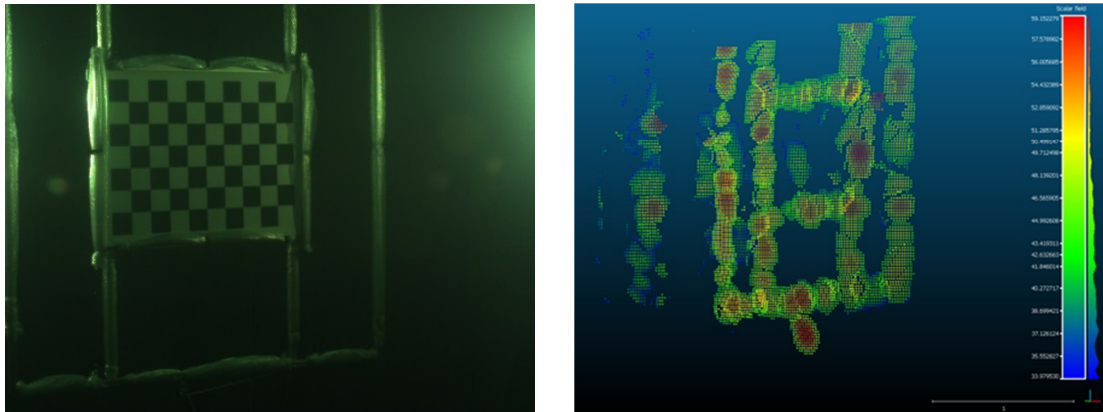


Fig. 2.20: acquisizione di una posa del pannello di calibrazione: ottica (sinistra) acustica (destra).

Per l'acquisizione del pannello contenente gli oggetti di varia forma e materiale, esso è stato invece posizionato sul bordo laterale della vasca (ad una distanza di circa 2.5 m dalla camera ottico acustica) movimentando in questo caso la camera ottico acustica stessa attraverso il palo telescopico. Il *dataset* creato è composto da 8 pose del pannello (Fig. 2.21).

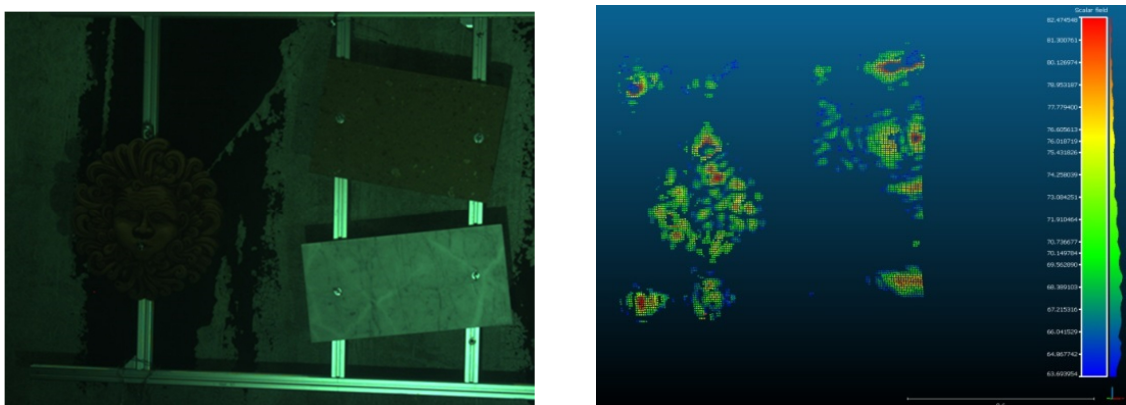


Fig. 2.21: Acquisizione target con oggetti: ottica (sinistra) acustica (destra).

Infine bisogna notare come le immagini acustiche forniscono, oltre alla nuvola di punti 3D, un campo scalare rappresentante l'intensità degli echo ricevuti per ogni punto 3D acquisito. In questa fase si sono evidenziati enormi limiti della camera acustica per la tipologia di applicazione prevista. In particolare, a differenza da quanto dichiarato dal costruttore, nel range di operatività minimo previsto (circa 1 m), le immagini acustiche presentano una rumorosità tale da non permettere una discriminazione del target acquisito rispetto al background, mentre a causa della scarsa risoluzione spaziale alla distanza di circa 2.5 m delle 8 pose del pannello con oggetti solo una presenta una qualità minima tale da poter essere utilizzata nella fase di validazione presentata nella sezione 2.6.

2.6.2 *Image Enhancement*

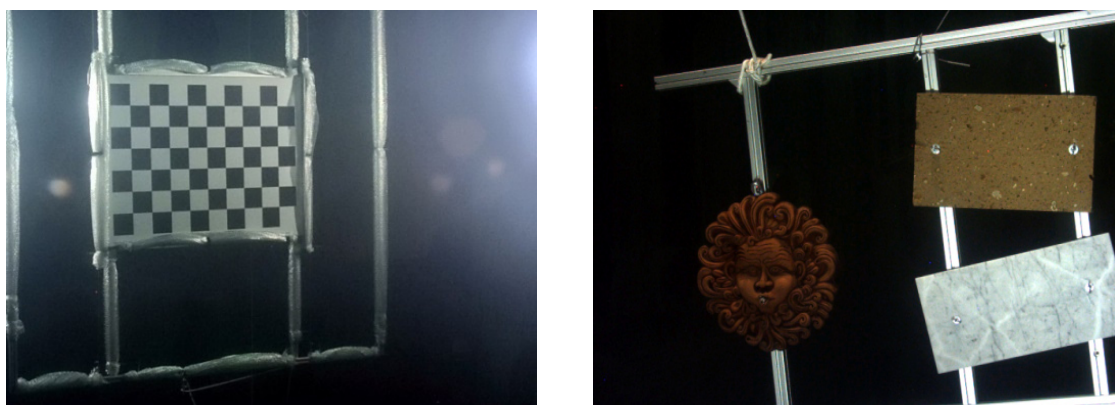


Fig. 2.22: immagini ottiche pannello di calibrazione e target con oggetti dopo applicazione algoritmi di miglioramento colore.

Al fine di garantire un'efficacia maggiore delle tecniche impiegate nelle successive fasi del protocollo di calibrazione progettato, ed in particolare la ricostruzione ottica 3D dei due pannelli nelle varie pose, è stata applicata una metodologia di *image enhancement* al fine di migliorare la qualità delle immagini che costituiscono il *data set* ottico creato nella precedente fase d'acquisizione. In Fig. 2.22 si possono osservare i risultati ottenuti dall'applicazione dell'algoritmo alle immagini ottiche di Figg. 2.20 e 2.21 rispettivamente.

2.6.3 Ricostruzione ottica 3D

Successivamente, le immagini ottiche acquisite sono state elaborate al fine di ottenere le nuvole di punti 3D del *pattern* planare e del pannello con oggetti nelle varie pose assunte durante la fase d'acquisizione. L'applicazione della metodologia implementata ha permesso di ricostruire 11 delle 20 pose acquisite del pannello di calibrazione, poiché nelle rimanenti il numero di *features* estratte non è stato sufficiente per ottenere una ricostruzione 3D dell'oggetto. Una fase di pulitura manuale delle nuvole *raw* effettuata con Cloudcompare è stata infine necessaria a valle del processo di ricostruzione per eliminare la presenza di diversi valori anomali. In Fig. 2.23 sono mostrati due esempi di nuvole di punti 3D ottiche del pannello di calibrazione e del target con oggetti rispettivamente.



Fig. 2.23: ricostruzione ottica 3D pannello di calibrazione e target con oggetti.

2.6.4 Registrazione ottico - acustica

Come descritto nella sezione 2.3.3, la matrice di trasformazione rigida che lega il sottosistema acustico a quello stereo ottico, assunto come riferimento, è ottenuta a valle di un'elaborazione statistica delle ${}^o\mathbf{T}_{a,n}$ stime determinate dalla registrazione delle n coppie $\mathbf{P}_{0,n}\mathbf{P}_{a,n}$ di nuvole di punti 3D ottiche ed acustiche. Tale operazione è stata eseguita per ognuna delle coppie $\mathbf{P}_{0,n}\mathbf{P}_{a,n}$ con $n \in [1,11]$ derivate dai precedenti step della metodologia di calibrazione implementata, ottenendo, dunque, come risultato ${}^o\mathbf{T}_{a,n}$ con $n \in [1,11]$ stime della matrice di trasformazione cercata. Poiché la procedura di registrazione, ed in particolare l'algoritmo ICP, è molto sensibile

alla presenza di valori anomali, è stato eseguito sulle $\mathbf{P}_{a,n}$ con $n \in [1, 11]$ nuvole acustiche un processo preliminare di filtraggio manuale.

In Fig.2.24 sono mostrati i risultati ottenuti dall'applicazione degli algoritmi per una delle 11 coppie acquisite.

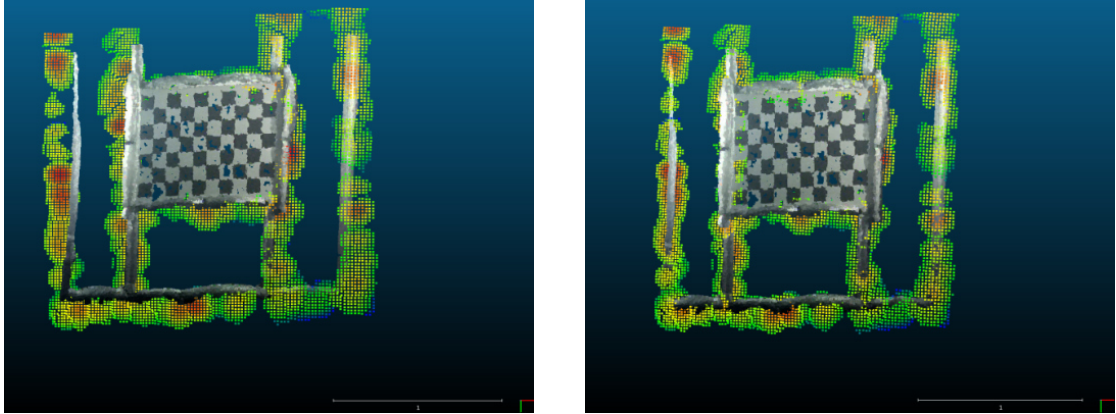


Fig. 2.24: nuvole di punti 3D ottiche ed acustiche dopo fase di *Coarse* (sinistra) and *Fine Registration* (destra).

2.6.5 Elaborazione statistica

Tre differenti metodologie di elaborazione statistica sono state implementate e successivamente messe a confronto al fine di determinare la matrice di trasformazione rigida ${}^o\mathbf{T}_a^*$ cercata a partire dalle matrici ${}^o\mathbf{T}_{a,n}$ con $n \in [1, 11]$ ottenute a valle dell'operazione di *Coarse* e *Fine Registration* applicata alle coppie $\mathbf{P}_{0,n}\mathbf{P}_{a,n}$ con $n \in [1, 11]$.

Una prima ipotesi è stata quella di ottenere gli elementi della matrice di trasformazione finale ${}^o\mathbf{T}_a^*$ attraverso un'operazione di media sulle ${}^o\mathbf{T}_{a,n}$ con $n \in [1, 11]$ matrici presenti nel dataset. Tuttavia questa operazione, se applicata direttamente alle matrici omogenee ${}^o\mathbf{T}_{a,n}$ con $n \in [1, 11]$, compromette l'ortonormalità della soluzione trovata, per cui si è pensato di applicare tale soluzione rappresentando le ${}^o\mathbf{T}_{a,n}$ con $n \in [1, 11]$ nelle corrispondenti \mathbf{R}_n con $n \in [1, 11]$ matrici di rotazione e t_n con $n \in [1, 11]$ vettori di traslazione. Successivamente, ottenuta la matrice di trasformazione media $\bar{\mathbf{R}}$ ed il vettore di traslazione medio \bar{t} ,

porre la matrice di trasformazione finale ${}^o\mathbf{T}_a^* = \begin{bmatrix} \bar{\mathbf{R}} & \bar{t} \\ 0 & 1 \end{bmatrix}_a^*$

Il vettore di traslazione medio \bar{t} è stato calcolato semplicemente come media aritmetica tra gli elementi dei vettori t_n con $n \in [1, 11]$. Per quanto concerne invece il calcolo della matrice di rotazione media $\bar{\mathbf{R}}$ le soluzioni possono essere molteplici. Infatti, come riportato in [91], esistono in letteratura diverse formulazioni per ottenere la rotazione media, alcune basate sulla metrica di distanza euclidea altre su quella di *Riemann*. Nello stesso articolo è inoltre dimostrato come, nel caso in cui i dati non abbiano una elevata variabilità (come è possibile ipotizzare per il caso in esame, essendo le ${}^o\mathbf{T}_{a,n}$ con $n \in [1, 11]$ stime della stessa matrice di trasformazione) la media aritmetica, applicata ai vettori di rotazione φ_n con $n \in [1, 11]$, ottenuti dalle corrispondenti matrici di rotazione \mathbf{R}_n con $n \in [1, 11]$, rappresenta una soluzione approssimata di entrambe le medie *riemmaniane* $\bar{\mathbf{R}}_{Riem}$ e *euclidea* $\bar{\mathbf{R}}_{Eucl}$, e come $\bar{\varphi}$ si discosti dalle precedenti per valori oltre la terza cifra decimale. Sulla base delle considerazioni precedenti, si è dunque deciso di determinare la matrice di rotazione media $\bar{\mathbf{R}}$ dal vettore di rotazione medio $\bar{\varphi}$, quest'ultimo calcolato come media aritmetica dei vettori di rotazione φ_n con $n \in [1, 11]$.

Da questa prima soluzione si è ottenuto la seguente matrice di trasformazione rigida:

$${}^o\mathbf{T}_{a,1}^* = {}^o \begin{bmatrix} \bar{\mathbf{R}} & \bar{t} \\ 0 & 1 \end{bmatrix}_{a,1}^* = \begin{bmatrix} 0.9918 & -0.1176 & 0.0506 & 0.3490 \\ 0.0497 & -0.0100 & -0.9987 & 0.2963 \\ 0.1179 & 0.9930 & -0.0041 & 0.1802 \\ 0 & 0 & 0 & 1.0000 \end{bmatrix} \quad (2.6)$$

Nella seconda ipotesi si è cercato di ottenere la matrice di trasformazione rigida ${}^o\mathbf{T}_a^*$ come fatto in precedenza, ma includendo nel calcolo solo un sottoinsieme delle ${}^o\mathbf{T}_{a,n}$ con $n \in [1, 11]$ presenti nel dataset iniziale. Questa selezione è stata effettuata al fine di eliminare dal calcolo dei vettori medi $\bar{\varphi}$ e \bar{t} i valori anomali che portano ad una polarizzazione dei risultati ottenuti dall'applicazione della media aritmetica ai vettori φ_n e t_n , con $n \in [1, 11]$. In particolare, è stato implementato un algoritmo in ambiente *Matlab* il quale determina automaticamente il set delle matrici ${}^o\mathbf{T}_{a,s}$ con $s < n$ da includere nel calcolo della ${}^o\mathbf{T}_a^*$. L'algoritmo inizialmente seleziona tra le ${}^o\mathbf{T}_{a,n}$ con $n \in [1, 11]$ matrici del dataset iniziale quelle a cui è associato un rms (errore quadratico medio ottenuto dall'applicazione dell'algoritmo ICP nella fase di *fine registration*) al di sotto di un valore di soglia rms* preimpostato. La scelta del valore

rms* è stata effettuata analizzando il grafico in Fig. 2.25 che riporta i valori rms associati per ognuna delle $P_{0,n}P_{a,n}$ con $n \in [1, 11]$ coppie di nuvole di punti 3D registrate. Esse sono state ordinate in maniera crescente lungo l'asse delle ascisse in funzione della distanza d'acquisizione da parte del sistema di *imaging* ottico - acustico. Come si evince dal grafico, le coppie $P_{0,n}P_{a,n}$ con $n = 1, 10, 11$ presentano un valore rms elevato rispetto alle rimanenti, per cui rms* è stato posto pari a 0.065 m in modo da eliminare le matrici ${}^oT_{a,n}$ associate dalle operazioni di media.

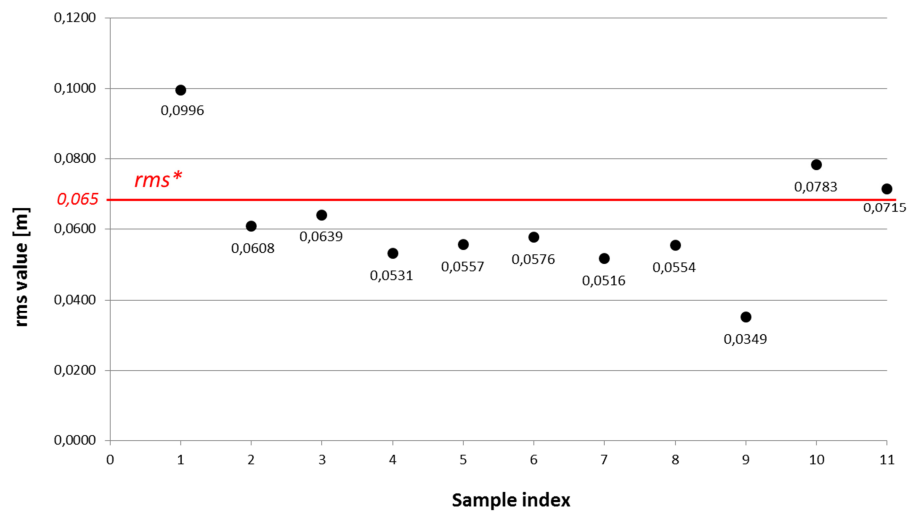


Fig. 2.25: grafico valori rms ottenuti a valle del processo di *fine registration* sulle $P_{0,n}P_{a,n}$ con $n \in [1, 11]$ coppie di nuvole di punti 3D.

Successivamente l'algoritmo determina i vettori di traslazione t_n ed i vettori di rotazione φ_n dalle matrici ${}^oT_{a,n}$ che soddisfano il primo criterio di selezione ed applica a quest'ultimi un secondo criterio di selezione basato "sulla distanza dai quartili". Per ogni componente dei vettori t_n e φ_n l'algoritmo determina il range interquartile (i.e. la differenza tra terzo e primo quartile che rappresenta l'ampiezza dell'intervallo entro cui cade il 50% delle osservazioni che occupano le posizioni centrali nelle serie ordinate dei dati) e considera come anomalo un valore con scostamento (positivo) dal terzo quartile o (negativo) dal primo quartile superiore a 3 volte il range interquartile stesso, eliminando, di conseguenza, l'intero vettore a cui la specifica componente appartiene. Applicando questa metodologia sono state eliminate 5 delle 11 stime ${}^oT_{a,n}$ della matrice di trasformazione rigida ottenendo la seguente matrice media:

$${}^o\mathbf{T}_{a,2}^* = {}^o\begin{bmatrix} \bar{\mathbf{R}} & \bar{\mathbf{t}} \\ 0 & 1 \end{bmatrix}_{a,2}^* = \begin{bmatrix} 0.9940 & -0.1081 & 0.0174 & 0.2554 \\ 0.0220 & 0.0413 & -0.9989 & -0.0436 \\ 0.1072 & 0.9933 & 0.0434 & 0.1858 \\ 0 & 0 & 0 & 1.0000 \end{bmatrix} \quad (2.7)$$

Prima di andare a descrivere la terza metodologia di elaborazione implementata è interessante notare il legame esistente tra i valori rms dell'algorithm ICP e le pose del pannello di calibrazione. Il campione con indice $i = 1$ del grafico di Fig. 2.25 è associato all'unica coppia di nuvole $\mathbf{P}_{0,1}\mathbf{P}_{a,1}$ che si è stati in grado di ottenere dalle prime 9 pose del pannello di calibrazione sulle 20 totali, pose acquisite a distanza compresa tra 1.5 m e 4.5 m rispetto alla camera ottico - acustica. I campioni con indice $i \in [2, 9]$ si riferiscono invece alle 8 pose acquisite ad una distanza compresa tra 4.5 m e 7 m. Infine, i campioni con indice $i = 10, 11$ si riferiscono alle coppie di nuvole $\mathbf{P}_{0,10}\mathbf{P}_{a,10}$ e $\mathbf{P}_{0,11}\mathbf{P}_{a,11}$, ottenute da 2 delle 3 pose finali acquisite a distanza compresa tra 7 m e 10 m. Le coppie che presentano valori elevati di rms sono associate a pose acquisite a distanza minore di 4.5 m e superiore a 7 m da cui possiamo evincere che il comportamento ottimale della camera ottico acustica, in questo specifico ambiente di test, è per distanze comprese fra 4.5 m e 7 m circa. Come ci si può aspettare, a distanza minore di 4.5 m è la componente acustica del sistema a presentare i limiti maggiori nella capacità di ricostruire il target acquisito, mentre per distanze superiori ai 7 m la camera stereo ottica subisce un degrado delle sue performance non essendo più in grado di estrarre sufficienti features dalle immagini acquisite e dunque ricostruire il target stesso.

La terza soluzione proposta si basa su un algorithm implementato in ambiente Matlab, il quale determina, in maniera automatica, la matrice di trasformazione finale ${}^o\mathbf{T}_a^*$, selezionandola tra le ${}^o\mathbf{T}_{a,n}$ con $n \in [1, 11]$ matrici presenti nel dataset iniziale.

L'algorithm opera nel modo seguente:

1. per ogni ${}^o\mathbf{T}_{a,n}$ applica tale trasformazione alle $\mathbf{P}_{a,n}$ con $n \in [1, 11]$ nuvole di punti 3D acustiche per allinearle alle rispettive nuvole di punti 3D ottiche $\mathbf{P}_{o,n}$ con $n \in [1, 11]$;
2. Per ogni coppia $\mathbf{P}_{o,n}\mathbf{P}_{a,n}$ con $n \in [1, 11]$ viene calcolata la distanza media $d_{oa,n}$ fra i punti della nuvola 3D ottica e della corrispondente nuvola di punti 3D acustica;
3. calcola \bar{d}_n come media delle $d_{oa,n}$;

4. seleziona la $\overline{d_{n,min}}$ come valore minimo delle $\overline{d_n}$;
5. assume come ${}^o\mathbf{T}_a^*$ la trasformazione ${}^o\mathbf{T}_{a,n}$ corrispondente a $\overline{d_{n,min}}$.

Applicando quest'ultima metodologia si ottiene la seguente matrice di trasformazione:

$${}^o\mathbf{T}_{a,3}^* = {}^o\begin{bmatrix} \overline{\mathbf{R}} & \overline{\mathbf{t}} \\ 0 & 1 \end{bmatrix}_{a,3}^* = \begin{bmatrix} 0.9953 & -0.0968 & 0.0021 & 0.2131 \\ 0.0048 & 0.0276 & -0.9996 & 0.0038 \\ 0.0967 & 0.9949 & 0.0279 & 0.1686 \\ 0 & 0 & 0 & 1.0000 \end{bmatrix} \quad (2.8)$$

Nella sezione successiva saranno analizzati i risultati ottenuti, in termine di errore medio di registrazione, applicando le tre matrici di trasformazione ottenute con le metodologie sopraesposte alla coppia di nuvole di punti 3D ottico ed acustiche rappresentante il pannello con oggetti.

2.7 Risultati

Come si è detto all'inizio del paragrafo lo scopo dei test è duplice: si vuole da un lato valutare quale dei tre metodi ipotizzati per l'elaborazione statistica risulta essere più efficace; dall'altro si vuole ottenere una validazione del protocollo di calibrazione. La validazione può essere ottenuta stimando la distanza media μ (e la rispettiva deviazione standard σ) fra le nuvole ottiche ed acustiche allineate utilizzando la matrice di calibrazione estrinseca ottenuta con il metodo proposto. Il test come si è detto in 2.5 è stato condotto utilizzando il pannello target con oggetti. Il calcolo di μ e σ è stato effettuato con Cloudcompare.

In Fig. 2.26 è possibile osservare l'allineamento delle nuvole di punti 3D ottico ed acustiche del pannello target con oggetti ottenuto applicando le matrici di calibrazione estrinseca ottenuti dall'applicazione delle tre metodologie di elaborazione statistica proposte. Applicando la matrice ${}^o\mathbf{T}_{a,1}^*$ si ha un non perfetto allineamento tra le due nuvole (Fig. 2.26a), come confermato dalle successive analisi statistiche effettuate. La nuvola acustica 3D presenta infatti una "distanza media" da quella ottica di circa 4.6 cm con una deviazione *standard* σ di circa 3.6 cm.

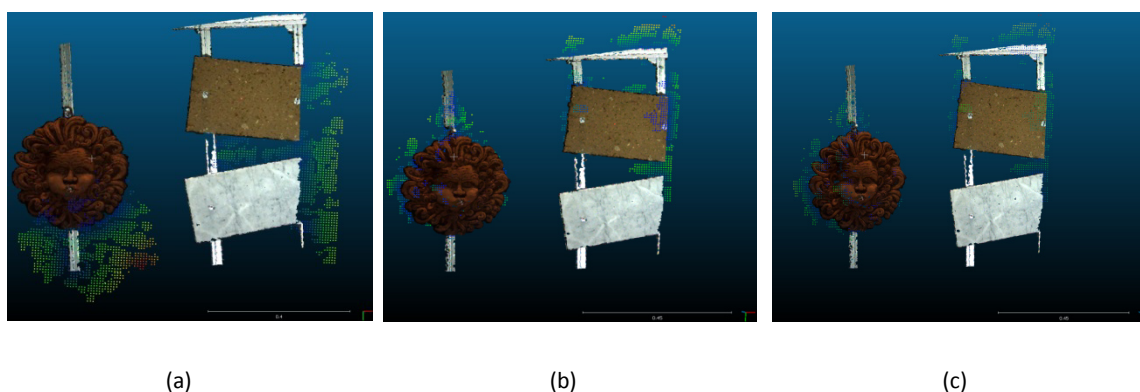


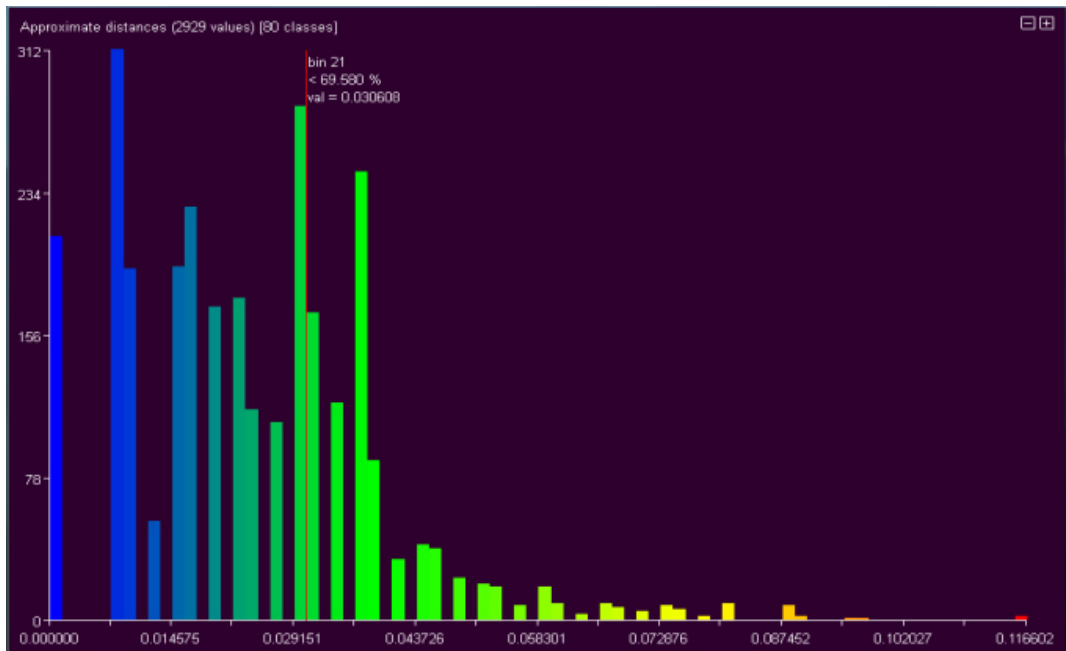
Fig.2.26: allineamento nuvole di punti 3D ottico acustiche pannello con oggetti applicando matrice ${}^oT_{a,1}^*$ (a), ${}^oT_{a,2}^*$ (b), ${}^oT_{a,3}^*$ (c)

La bontà degli algoritmi implementati è confermata dai risultati ottenuti applicando le matrici di trasformazione rigida ${}^oT_{a,2}^*$ e ${}^oT_{a,3}^*$ alla nuvola acustica 3D del pannello target con oggetti. In effetti, a differenza del caso precedente, applicando la matrice ${}^oT_{a,2}^*$ alla nuvola acustica 3D si osserva un netto miglioramento dell'allineamento delle due nuvole di punti 3D (Fig. 2.26b), quest'ultimo confermato dall'analisi statistica condotta. La "distanza media" della nuvola acustica 3D dalla controparte ottica è di circa 2.3 cm con una deviazione *standard* σ di circa 1,4 cm. Inoltre, andando ad analizzare l'istogramma in Fig. 2.27a si può notare come, circa il 70 % dei punti della nuvola acustica 3D presentano una distanza al di sotto dei 3 cm (*range resolution* del sonar), e come i valori con distanza maggiore siano dovuti al rumore presente sulla nuvola acustica, in particolare nell'intorno della maschera in terracotta e del profilo di alluminio con cui è stato realizzato il pannello stesso.

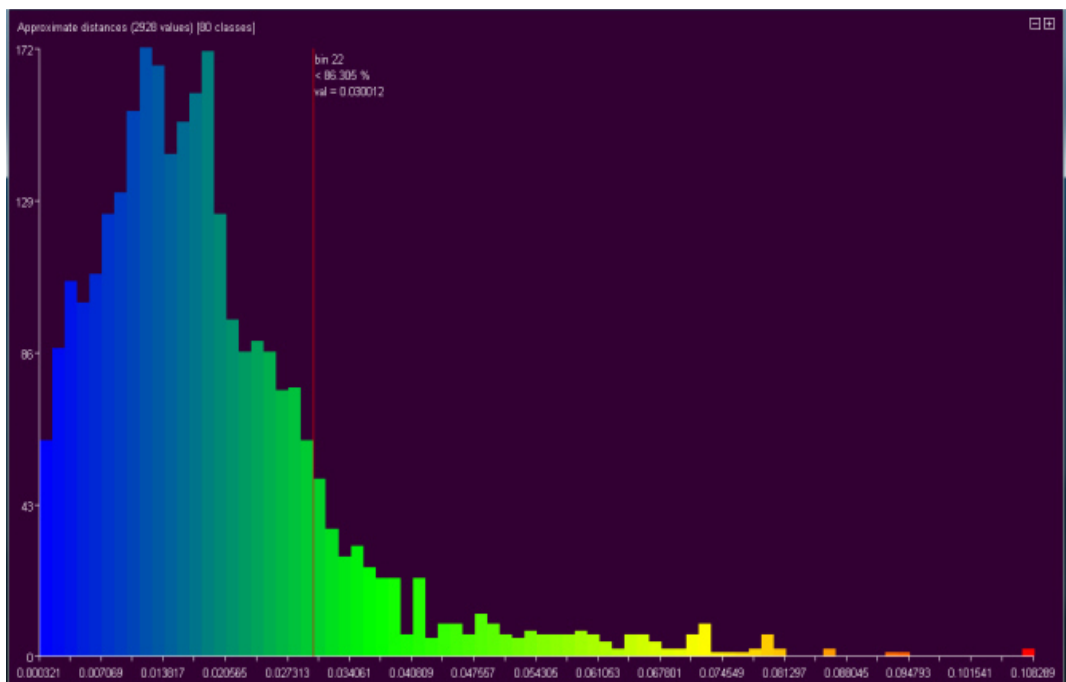
Infine, risultati ancora migliori sono stati ottenuti applicando la matrice di trasformazione ${}^oT_{a,3}^*$ ottenuta dall'algoritmo di selezione (Fig. 2.26c). La "distanza media" della nuvola acustica 3D dalla controparte ottica è, in questo caso di circa 1,8 cm mentre la deviazione *standard* σ di circa 1,3 cm. Inoltre, dall'istogramma in figura 2.27b è possibile notare come circa 86,3% dei punti presenta una distanza al di sotto della *range resolution* del sonar sotto il quale i risultati ottenuti sarebbero del tutto casuali.

Appare quindi evidente che il terzo metodo di elaborazione statistica proposto offre i risultati migliori e questi risultati non sarebbero migliorabili adottando un diverso

metodo poiché l'errore medio ottenuto è comparabile con la *range resolution* del sonar.



(a)



(b)

Fig. 2.27: istogramma distanza nuvole 3D ottico acustiche pannello con oggetti ottenuto applicando matrice ${}^oT_{a,2}^*$ (a), ${}^oT_{a,3}^*$ (b).

Prototipo camera 3D ottico - acustica

L'oggetto del presente capitolo riguarda la descrizione di un prototipo di sistema composto da una camera ottico-acustica per l'acquisizione in *real-time* di immagini 3D in ambiente subacqueo in condizioni di torbidità variabile e da un *software* per la visualizzazione in realtà aumentata (AR) dei dati acquisiti. L'architettura del sistema prevede un'unità d'interfacciamento con l'operatore costituita da un computer equipaggiato con il software per la visualizzazione e il controllo della camera e da un'unità di acquisizione installata a bordo di un ROV subacqueo. Le due unità sono collegate attraverso un cavo ombelicale impiegato per la trasmissione/ricezione dati ed eventualmente anche per alimentare l'unità d'acquisizione (qualora il ROV non sia dotato di accumulatori di energia adeguati). L'unità di acquisizione è costituita, a sua volta, da una camera acustica tridimensionale e da una camera ottica stereoscopica vincolati ad una struttura rigida di sostegno progettata al fine di garantire la massima sovrapposizione dei loro campi di vista (FOVs). Entrambe le camere sono in grado di generare una nuvola di punti 3D rappresentativa della scena osservata. I dati 3D ottici ed acustici possono essere visualizzati separatamente o fusi in un unico data-set. In particolare, l'utilizzo di differenti metodologie di *data fusion* e *data merging* consente di integrare in maniera efficiente i dati provenienti dai sensori ottici ed acustici e di ottenere una visualizzazione "aumentata" in *real-time* della scena acquisita utilizzando algoritmi e metodologie tipiche della *computer vision*. La scena può quindi essere rappresentata come una nuvola di punti in uno spazio tridimensionale o come una mappa di profondità (a trasparenza variabile) in cui la profondità della scena è codificata attraverso un gradiente colore RGB. La mappa di profondità può essere sovrapposta all'immagine 2D fornita da uno dei dispositivi ottici. Il *software* di visualizzazione e controllo presenta un'interfaccia utente configurabile che consente di variare la trasparenza della mappa e gli altri parametri che influenzano la modalità di fusione e rappresentazione dei dati 3D per facilitarne la comprensione e

interpretazione da parte dell'utente. Il sistema è stato concepito come strumento di supporto ad un operatore ROV durante le fasi di navigazione in condizioni di acque torbide o in operazioni di esplorazione/ manutenzione condotte attraverso l'uso, ad esempio, di uno o più manipolatori. In questo contesto applicativo, dopo la fase di detezione della scena, eseguita dal sottosistema acustico, l'operatore è in grado di posizionare il ROV alla corretta distanza di lavoro (facendo uso di diverse modalità di visualizzazione). A questo punto, durante tutte le varie fasi d'intervento, esso può scegliere la migliore modalità di visualizzazione (ottica od acustica) o fonderle in un'unica immagine in modo tale che le parti mancanti della ricostruzione ottica sia compensate da quelle acustiche.

Dopo una descrizione generale dell'architettura dell'intero sistema vengono presentate le principali funzionalità della camera ed un primo prototipo del suo *software* di visualizzazione e controllo. Successivamente si riporta una breve descrizione e le specifiche tecniche dei dispositivi che compongono il sistema in esame, cioè camera acustica e camera stereo ottica rispettivamente.

3.1 Descrizione del sistema

In questo paragrafo vengono dapprima presentate l'architettura e le unità che compongono la camera ottico - acustica tridimensionale ed a seguire, una descrizione delle principali funzionalità del sistema.

3.1.1 Architettura *hardware*

L'architettura completa del sistema prevede un'unità d'interfacciamento con l'operatore (posta a bordo di una nave appoggio) costituita da un computer equipaggiato con il *software* di visualizzazione e controllo della camera ed un'unità di acquisizione installata a bordo di un *Remotely Operated Vehicle* (ROV) subacqueo. Le due unità sono collegate attraverso un cavo ombelicale impiegato sia per alimentare l'unità d'acquisizione che per la trasmissione \ ricezione dati (Fig. 3.1).

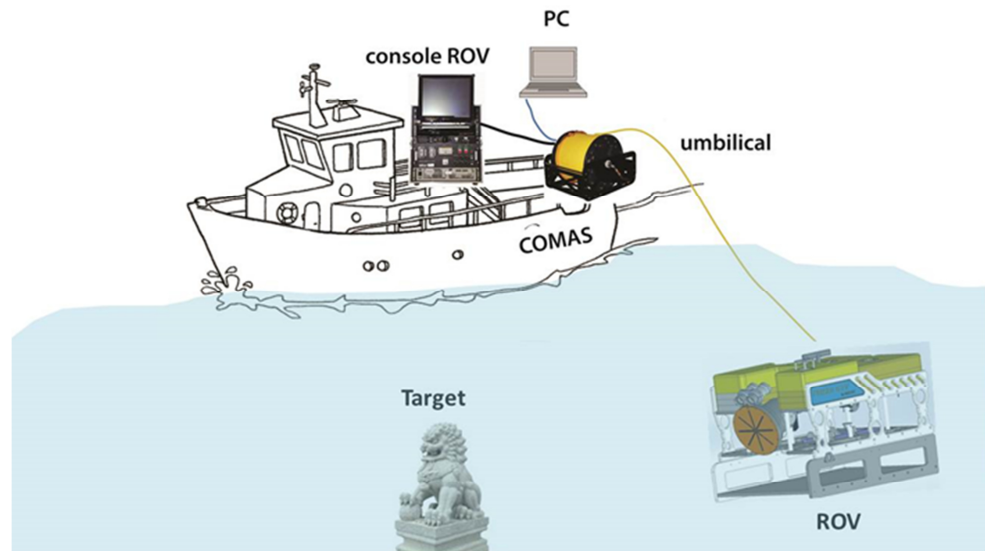


Fig. 3.1: sistema completo camera ottico - acustica tridimensionale.

L'unità di acquisizione è costituita, a sua volta, da una camera acustica tridimensionale e da due camere ottiche in configurazione stereo. Il *layout* dell'unità di acquisizione ed il suo alloggiamento a bordo del ROV è stato progettato al fine di garantire la massima sovrapposizione dei campi di vista dei due sottosistemi (Fig. 3.2).

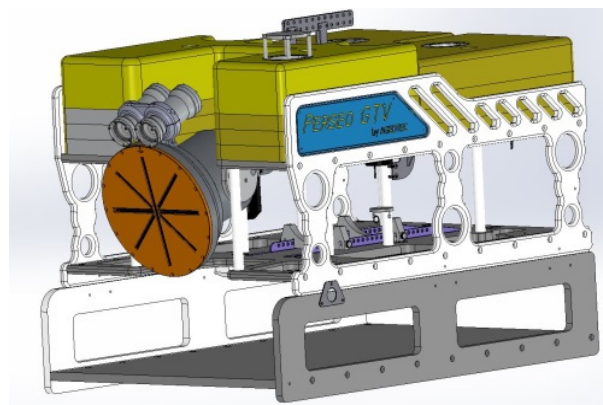


Fig. 3.2: modello CAD alloggiamento camera ottico - acustica tridimensionale su ROV Perseo GTV.

Come riportato in Fig. 3.3, l'unità di acquisizione ottico-acustica comprende una camera stereo ottica passiva (4) costituita da 2 videocamere digitali ultracompatte equipaggiate con un pacchetto ottico grandangolare alloggiate all'interno di due case subacquee. Questi sono realizzati con un corpo in alluminio tale da garantire una efficiente dissipazione di calore e finestre ottiche (*flat port*) in policarbonato.

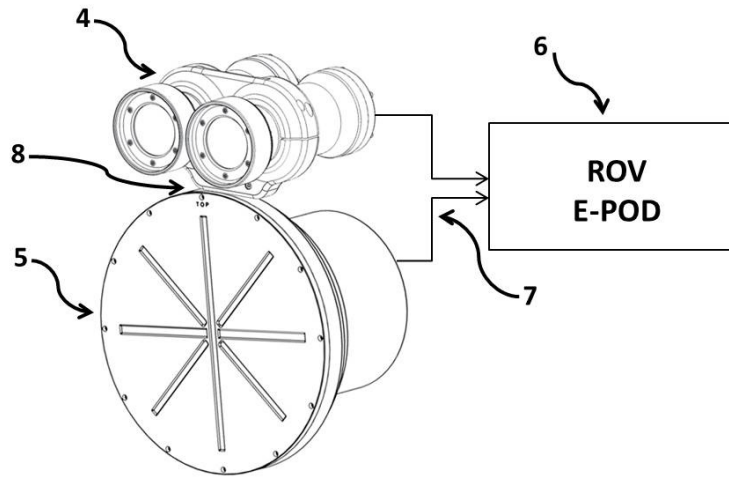


Fig. 3.3: unità di acquisizione ottico-acustica tridimensionale.

La camera acustica tridimensionale (5) è un sistema *custom* che presenta canali di trasmissione e ricezione disposti su *array* lineari a spaziatura logaritmica ed elaborazione digitale dei segnali tramite algoritmi di *beamforming* e *beampower evaluation*. I due sottosistemi sono vincolati ad una struttura rigida di sostegno (8) regolabile tramite un sistema meccanico od eventualmente elettronico e tale da garantire la massima sovrapposizione dei loro FOVs. I dati acquisiti dai due sensori sono trasmessi, tramite un bus di comunicazione dati ad alta velocità (7) (Gigabit Ethernet) all'EPOD del ROV (6), il quale provvede a multiplexarli, tramite un dispositivo multiplexer (non mostrato in figura), cioè uno *switch* ad alta velocità, sulla fibra ottica dell'ombelicale e dunque ad inviarli all'interfaccia di comunicazione ubicata nella console dell'operatore. L'EPOD del ROV contiene inoltre, al suo interno, l'elettronica (non illustrata) per l'alimentazione dell'unità d'acquisizione.

3.1.1.1 Camera acustica

La camera acustica tridimensionale è un sistema *custom* sviluppato dalla Whitehead Sistemi Subacquei S.p.A., Pozzuoli (NA), suddiviso in un sottosistema contenente la sensoristica di trasmissione/ricezione e un sottosistema contenente l'elettronica di generazione/elaborazione dei segnali (Fig. 3.4).

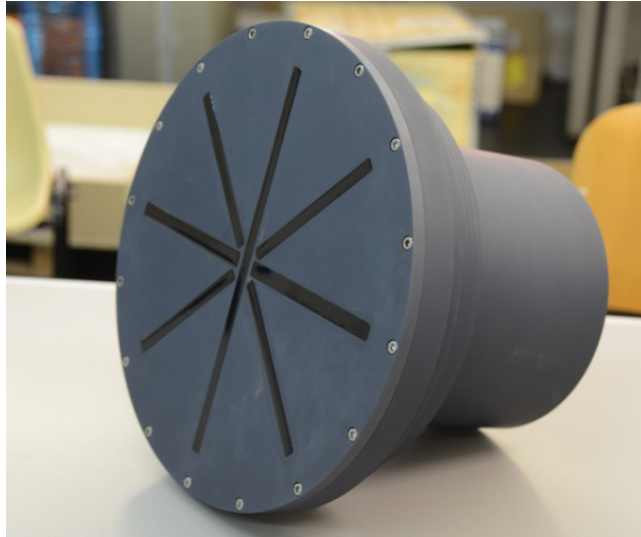


Fig.3.4: camera acustica tridimensionale.

Il sottosistema contenente la sensoristica è composto da una sezione di trasmissione costituita da 40 canali indipendenti disposti su un array lineare a spaziatura logaritmica, mentre la sezione ricevente è costituita da 256 canali disposti su 4 array lineari anch'essi a spaziatura logaritmica. La sezione acustica genera segnali che focalizzano l'energia in opportune direzioni di puntamento e riceve il segnale scatterato da un *target* posizionato nella regione di focalizzazione. L'elettronica di elaborazione dei segnali elabora i segnali ricevuti a livello digitale tramite algoritmi di *beamforming* e *beam power evaluation*, al fine di rilevare la provenienza dell'eco in termini di coordinate polari (distanze ed angoli).

La camera acustica genera un'immagine tridimensionale ad alta risoluzione, operando in due distinte modalità:

- *Modalità di Detezione*, in cui il sistema esegue una scansione isotropa della scena, al fine di determinare la posizione "grossolana" di un oggetto da rivelare;
- *Modalità di Classifica*, in cui il sistema passa ad una scansione più fine del suo "campo di vista" ("panorama acustico"). Questa modalità si attiva manualmente (tramite impostazione della distanza di messa a fuoco da parte dell'operatore) o automaticamente, quando nella *modalità di Detezione* è stato individuato un oggetto a distanza inferiore a 2m.

Nella seguente tabella sono riportate le sue principali specifiche tecniche:

	<i>Detection mode</i>	<i>Classification mode</i>
Frequency	700 KHz	700 KHz
Number of beams	1	41x84x61 total
Maximum range	20 m ca.	Focusing distance + 30 cm
Minimum range	1 m	Focusing distance - 30 cm
Update Rate	//	Up to 1 Hz
Beam spacing	//	1°
Angular coverage	60° x 40°	60° x 40°
Dimensions (max diam., min diam., l)	410 mm, 260 mm, 287 mm	410 mm, 260 mm, 287 mm
Weight	27 Kg (air), 5 Kg (water)	27 Kg (air), 5Kg (water)
Supply voltage	12 VDC	12 VDC
Power consumption	200 W ca.	200 W ca.
Interface	Gigabit Ethernet	Gigabit Ethernet

Tab.2 : principali parametri della camera acustica 3D.

3.1.1.2 Camera stereo ottica

Per quanto concerne la componente ottica del sistema è stato sviluppato, presso i laboratori del Dipartimento di Ingegneria meccanica energetica e gestionale (DIMEG) dell'Università della Calabria, un prototipo di camera ottica stereoscopica (Fig. 3.5). Essa è composta essenzialmente da due corpi cilindrici in alluminio (camera destra e sinistra), fissati mediante un opportuno supporto. Ognuno dei cilindri, lavorati in modo da ridurre al massimo il peso, contiene al suo interno una camera GigE ImagingSource DFK 23G445 munita di ottica CS ultra wide 1.8-3 mm a bassissima distorsione.

Tale configurazione consente di ottenere immagini che possono essere efficacemente processate dal *software* di ricostruzione 3D. La camera è fissata all'interno del *case* mediante un opportuno supporto. Sostituendo il supporto è possibile alloggiare all'interno del *case* camere ed obiettivi differenti. Il supporto, oltre a garantire il

montaggio ottimale della camera, funge anche da dissipatore per la camera stessa, esso infatti trasferisce il calore generato dalla camera verso il case in alluminio e quindi verso l'esterno. Al fine di garantire un perfetto allineamento delle camere, sul case in alluminio è stato ricavato un foro per una opportuna spina di centraggio.



Fig.3.5: camera ottica stereoscopica

Nelle seguenti tabelle sono riportati i principali parametri del sistema:

Working distance (min.)	1 m (baseline 100 mm)
FOV (water) (diag.)	73.6° (min. working distance)
Dim. Point cloud	200.000 ca
Dimensions (w x h x l)	210 x 100 x 207 mm
Weight	2.6 Kg (air), 1.2 Kg (water)
trigger	Yes
Supply voltage	11 – 13 VDC
Power consumption	4.8 W ca.
Interface	Gigabit Ethernet

Tab.3: specifiche tecniche camera ottica stereoscopica.

3.1.2 Architettura software

Con riferimento alla Fig. 3.6, è stato proposto un metodo per l'acquisizione, la generazione, la memorizzazione e la visualizzazione in realtà aumentata di immagini 3D ottico - acustiche.

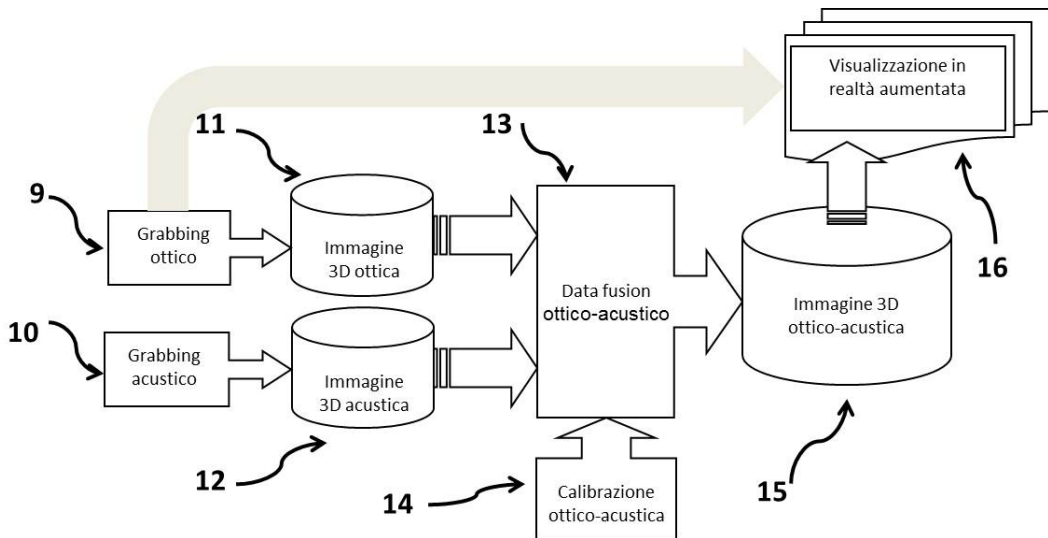


Fig. 3.6.: metodo per l'acquisizione, la generazione, la memorizzazione e la visualizzazione in realtà aumentata di immagini 3D ottico – acustiche.

Il metodo si compone dei seguenti moduli: un modulo di grabbing ottico (9) ed uno di grabbing acustico (10) per l'acquisizione sincronizzata, rispettivamente, dei dati ottici ed acustici della scena osservata. Un modulo (11) per l'elaborazione, attraverso algoritmi di *stereo matching*, dei dati ottici grezzi (immagini ottiche 2D) in uscita al modulo (9) e la generazione e memorizzazione dell'immagine ottica 3D risultante. Un modulo (12) per l'elaborazione, attraverso algoritmi di *beamforming* e/o *beam power evaluation* dei dati acustici grezzi (segnali di eco) in uscita al modulo (10) e la generazione e memorizzazione dell'immagine acustica 3D risultante. Un modulo (13) di *data fusion* tale da generare una immagine 3D ottico – acustica ricevendo in *input* la coppia di immagini 3D generata dai moduli (11) e (12) e i dati di calibrazione della camera ottico – acustica dal modulo (14). Un modulo (15) per la memorizzazione e la gestione dei dati 3D ottici ed acustici e i parametri che settano la modalità di fusione e rappresentazione dei dati 3D. Un modulo (16) per la visualizzazione e la

rappresentazione dei dati 3D ottici ed acustici secondo un determinato modello rappresentativo dell'informazione selezionabile da parte dell'utente.

3.1.2.1 Generazione immagine ottica 3D

I moduli (9) e (11) operano secondo l'architettura *software* mostrata in Fig. 3.7:

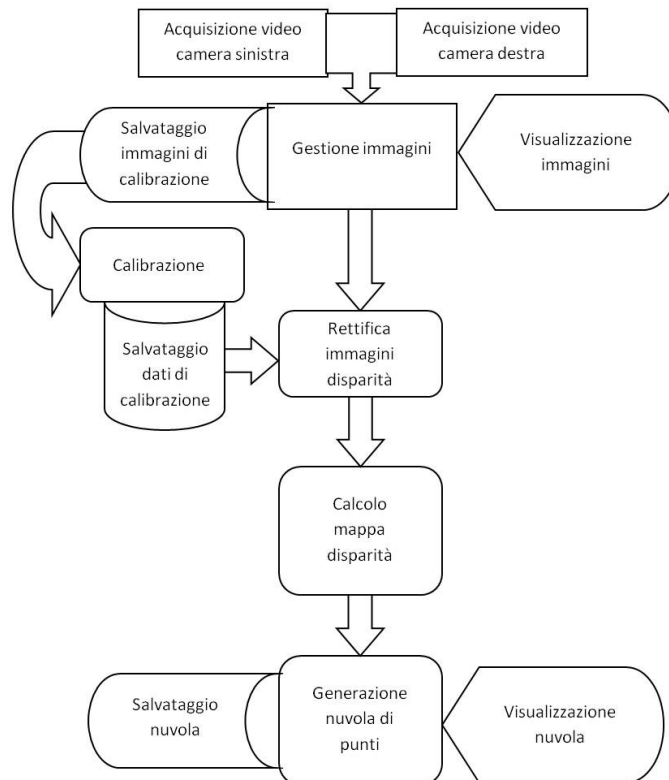


Fig. 3.7: Diagramma a blocchi dell'architettura software per la generazione dell'immagine ottica 3D

- Acquisizioni frame video.** La routine che si occupa dell'acquisizione video riceve come parametro di input la risoluzione di acquisizione e, collegandosi ai *driver* delle camere, ne gestisce la modalità impostandone il *frame rate* e la tipologia (Mono oppure RGB). La coppia di immagini acquisita può essere quindi visualizzata in *real-time* su schermo o eventualmente salvata per costruire il set di calibrazione. I parametri di controllo delle camere possono essere modificati a *runtime* permettendo così di variare facilmente tempo di esposizione e *frame rate* al fine eseguire un *tuning* di fine per ottimizzare la ricostruzione.

- **Calibrazione stereo.** La calibrazione del sottosistema stereo ottico permette di stimare sia la posizione relativa tra i sistemi di riferimento delle due camere sia le distorsioni ottiche dovute allo scostamento del modello di camera ideale (modello *pinhole*) da quello reale, in cui sono considerati gli effetti delle distorsioni prodotte dal pacchetto ottico. Tale metodologia restituisce in output la matrice Q necessaria per la riproiezione della mappa di disparità e la generazione della nuvola di punti 3D ed inoltre i parametri di distorsione (sia radiale che tangenziale) per la correzione e quindi la rettifica delle immagini 2D utilizzate per generare la mappa di disparità stessa.
- **Generazione mappa di disparità.** Il calcolo della mappa di disparità è la porzione di codice che richiede il maggior numero di risorse e tempo di esecuzione. Per applicazioni in *real-time* occorre utilizzare algoritmi veloci che introducono necessariamente un certo grado di approssimazione. Tra questi i più veloci sono quelli eseguiti su GPU sfruttando la potenza di calcolo di migliaia di processori. I risultati dei test effettuati hanno comunque dimostrato come in termini qualitativi gli algoritmi in CPU sono più performanti. Il calcolo della mappa di disparità richiede l'utilizzo di un algoritmo di *stereo matching* ad approccio globale o, eventualmente locale per velocizzare i tempi di calcolo. Questo ha lo scopo di individuare le *features* comuni nelle due immagini della stereo coppia per associarvi un valore di profondità. Per fare ciò si esegue una ricerca nell'intorno di ciascun pixel. Ecco perchè per velocizzare il calcolo occorre eseguire prima la rettifica delle immagini cosicché ciascun punto dell'immagine sinistra abbia il suo equivalente sulla stessa riga dell'immagine destra. In questo modo la ricerca delle *features* comuni viene eseguita unicamente sulla riga.
- **Generazione nuvola di punti 3D.** Calcolata la mappa di disparità è eseguita la riproiezione dei punti in uno spazio 3D attraverso la matrice Q di calibrazione. La nuvola di punti 3D così ottenuta può essere memorizzata e/o visualizzata a schermo.

3.1.2.2 Generazione immagine acustica 3D

Per quanto riguarda invece i moduli (10) e (12) operano secondo l'architettura software mostrata in Fig. 3.8:

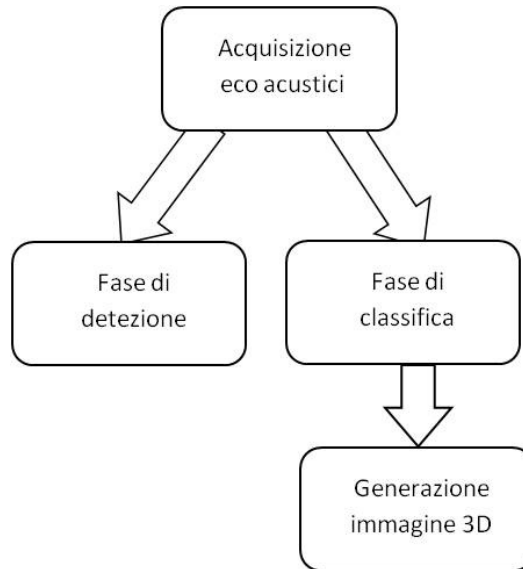


Fig. 3.8: Diagramma a blocchi dell'architettura software per la generazione dell'immagine acustica 3D.

- **Acquisizione eco acustici.** Nella fase di deteazione saranno impiegati solo pochi trasduttori centrali dell'array di trasmissione ed il PRT (*Pulse Rate Time*) sarà maggiore rispetto alla fase di classifica per coprire una maggiore portata. Anche l'impulso avrà una durata maggiore in questa modalità operativa. Nella fase di classifica sono utilizzati tutti i trasduttori attivi del supporto ceramico. L'antenna trasmittente invierà un segnale acustico di breve durata (5 cicli di portante) rispetto a quello utilizzato in deteazione attraverso un *fan beam* puntato in modo tale da spazzare tutto il FOV della camera acustica. Per ciascun *fan beam* il tempo di ascolto sarà legato ad una prefissata finestra temporale intorno al tempo corrispondente alla distanza di focalizzazione determinata in fase di deteazione.
- **Fase di deteazione.** L'eco (somma dei segnali) ricevuto dal gruppo dei sensori centrali attivi dell'array di ricezione viene analizzato con un filtro stimatore della potenza media di rumore che sarà utilizzata per stimare in maniera adattiva la soglia di deteazione. Il segnale passa attraverso un integratore dei campioni

quadrati a finestra mobile ottenendo così la seguente successioni di campioni di potenze medie:

$$P(t_k) = e^{\beta(t_k - t_G)} \sum_{i=1}^{N_c} x^2(t_{k-i}) \quad t_k \geq t_G \quad (3.1)$$

La finestra comprenderà un numero N_c di campioni legato alla durata dell'impulso trasmesso e si muoverà con una sovrapposizione del 75%. I campioni di potenza media sono pesati attraverso una funzione che tiene conto dell'attenuazione del livello di riverbero con la distanza e quindi con il tempo. Per la detezione dell'eventuale target si procede secondo il seguente algoritmo: si memorizzano le ampiezze ed i tempi relativi ai primi due campioni di potenza più elevata, distanti nel tempo più di quattro volte la durata dell'impulso. Si trascurano tali campioni e quelli adiacenti per un'estensione di ± 4 e sui restanti si ricerca il valore mediano. I due valori massimi di potenza sono attribuiti a echi provenienti da target se superano di 2 dB il valore mediano. Se uno solo dei massimi supera il test il suo tempo permetterà di terminare la distanza di focalizzazione; se entrambi superano il test sarà considerato il massimo cui corrisponde il tempo minore; se nessuno dei due supera il test si assume che non vi siano target per l'impulso trasmesso. La distanza di focalizzazione può essere determinata in base a due possibili interventi. Elaborando in maniera completamente automatica le potenze dei segnali secondo l'algoritmo sopra indicato o eventualmente lasciando all'operatore alla console la decisione di individuare il tempo dell'eventuale eco di un target sulla base delle potenze ricevute. In alternativa il dato relativo alla distanza di focalizzazione può essere fornito direttamente dalla console nell'ambito dei comandi trasmessi alla telecamera senza alcuna relazione con il segnale ricevuto. L'acquisizione delle potenze è realizzata dal sistema di acquisizione dati integrato nella telecamera mentre l'elaborazione delle potenze avviene nell'unità di superficie in maniera automatica o eventualmente manuale.

- **Fase di classifica.** I trasduttori del proiettore sono eccitati con segnali ad onda quadra, opportunamente ritardati ed opportunamente sfasati rispetto ad una griglia prefissata per poter generare in acqua le armoniche fondamentali di

opportuna fase che possono produrre per interferenza in corrispondenza del punto di focalizzazione un'onda di ampiezza pari N_s volte l'ampiezza della fondamentale. I ritardi temporali fra gli impulsi e le fasi delle portanti armoniche sono determinate in base al dato distanza di focalizzazione. La conoscenza di questo dato permetterà anche di calcolare le matrici di indici di ritardo da applicare ai segnali RX campionati, per effettuare i fan beam puntati agli angoli di elevazione.

La simmetria geometrica delle antenne consente un notevole risparmio del numero di operazioni necessario per l'elaborazione del *beamforming*.

- **Generazione immagine 3D.** Ogni impulso trasmesso determina in ricezione una matrice di potenze relativa ad una griglia di punti che si estendono in elevazione ed in *range* intorno al punto di focalizzazione. Dopo un tempo T pari alla durata di ricezione dei *ping* emessi sulla doppia distanza (andata e ritorno) di focalizzazione, l'algorithmo di acquisizione ha consentito di memorizzare le relative matrici di potenze. Si applica alle matrici il test a doppia soglia ottenendo una matrice completa di Nx4 elementi; nella prima colonna sono riportate le ascisse dei punti oggetto rivelati; nella seconda colonna le ordinate; nella terza le quote; nella quarta la potenza dell'eco. Il numero N di righe dipende dal particolare oggetto perché rappresentano i punti dell'oggetto stesso di massima riflettività che pertanto superano entrambi i test sulle soglie. Dalle coordinate x,y,z si può ricavare il dato di distanza del generico punto target e gli angoli di azimuth e di elevazione corrispondenti. In tal modo, nelle elaborazioni in cui è vantaggioso, il dato di uscita dell'algorithmo di *imaging* può fornire una matrice di dati contenente solo distanza e potenza nelle direzioni di puntamento lungo le quali si è avuta la detezione di un punto oggetto.

3.1.2.3 Funzionalità software del sistema

Successivamente sono stati definiti e implementati, come mostrato in Fig. 3.9, i servizi (16) accessibili da parte dell'utente durante la fase di esplorazione e di ricostruzione 3D della scena subacquea:

- **Stereo Camera Calibration.** E' il processo mediante il quale vengono calcolati i parametri intrinseci (lunghezza focale, centro ottico) ed estrinseci (posizione e orientazione tra le 2 camere).
- **Stereo Camera Control.** E' lo strumento mediante il quale l'utente può avere il controllo in tempo reale dei vari parametri di configurazione della stereo camera (modifica del tempo di esposizione, del guadagno, etc.)
- **Views.** Rappresentano il modo in cui i dati vengono mostrati, e gestiti da parte dell'utente. Dipendentemente dalle informazioni che l'utente ha la necessità di analizzare è possibile, tramite le *view*, selezionare o eventualmente combinare le informazioni utili da analizzare in un determinato momento.
- **Data Storage.** E' il processo che si occupa del salvataggio su *file-system* di tutti i dati acquisiti ed elaborati dal software.
- **Windows Layout Management.** E' lo strumento mediante il quale l'utente può gestire le informazioni da visualizzare. La selezione delle informazioni ai fini dell'esplorazione o dell'analisi di ricostruzione in *real-time* avverrà in base alla selezione di alcuni dei vari *layout* messi a disposizione dal software.
- **3D Registration Control (opzionale).** E' l'insieme delle funzionalità riferite alla fase di registrazione/tracking dell'ambiente 3D da ricostruire; l'applicazione dovrà fornire la possibilità di abilitare/disabilitare i vari sensori di tracking e fornire degli opportuni controlli in grado di poter gestire il processo di registrazione delle nuvole di punti.
- **Application Setting Configuration.** Rappresenta il punto di accesso per tutti i parametri di configurazione globali dell'applicazione.
- **Stereo Matching Algorithm Configuration.** E' lo strumento mediante il quale è possibile modificare in qualsiasi momento i parametri ed i vincoli degli algoritmi di *stereo matching* di cui il software fa uso ai fini della generazione della nuvole di punti 3D.

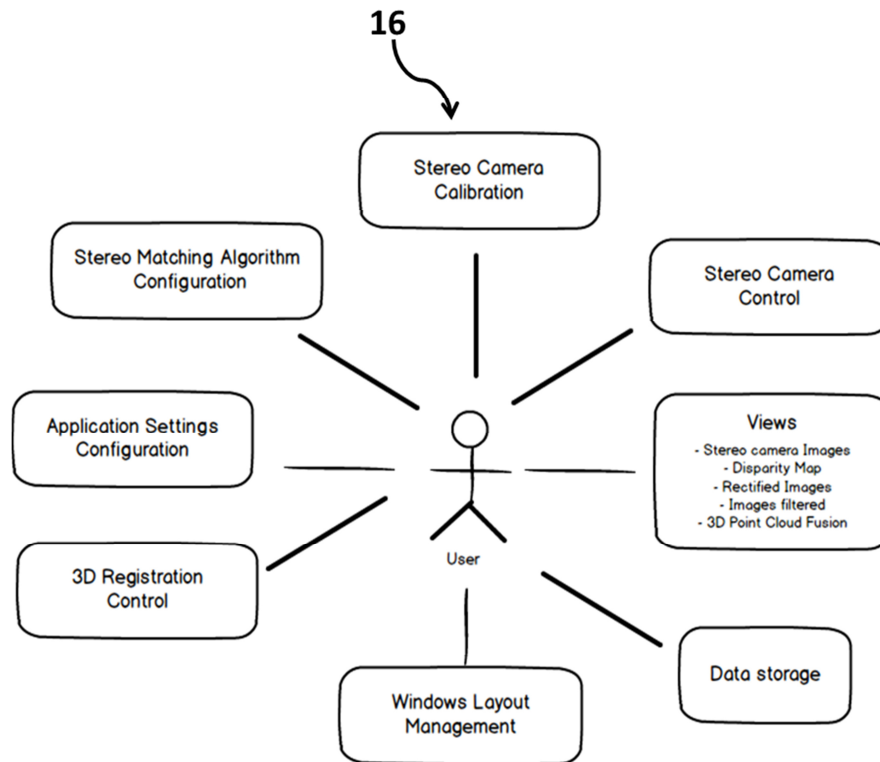


Fig. 3.9: schematico delle principali funzionalità software del sistema.

3.1.2.4 Interfaccia utente

Come mostrato in Fig. 3.10, l'interfaccia di visualizzazione (16) prevede una toolbar rappresentante il punto di accesso a tutte le funzionalità del software descritte in precedenza, compresa la gestione dei vari layout delle finestre di *rendering* posizionate nell'area di lavoro sottostante. Nella Fig. 3.10 si osserva la presenza di 3 *view* collocate all'interno dell'area di lavoro. Ciascuna di esse mostra all'utente l'avanzamento del sistema secondo diversi punti di vista. All'interno della *view* numero 1 vengono mostrate le immagini acquisite dalla stereo camera del ROV; tramite questa finestra l'utente potrà avere un riscontro sulla posizione e l'orientamento attuale del ROV e potrà analizzare la qualità delle immagini correntemente acquisite. La *view* numero 2 mostra all'utente un livello di dettaglio ben diverso dalla precedente; all'interno di essa potranno essere visualizzate diverse informazioni in maniera singola o combinata; il primo *layer* mostrerà l'immagine rettificata, il secondo invece mostrerà la mappa di disparità riferita all'immagine rettificata contenente le informazioni di profondità di tutte le corrispondenze trovate all'interno della stereo coppia di immagini, ed infine il

terzo *layer* conterrà i punti 3D rilevati dalla camera acustica in corrispondenza dell'immagine osservata.

La sovrapposizione delle informazioni riguardanti la disparità assieme all'immagine acustica prodotta dal sonar permette all'utente di stimare visivamente le informazioni sulla distanza tra il ROV e gli oggetti della scena in prossimità, utile per avere un percezione maggiore sulla profondità durante la guida del ROV.

La view numero 3 infine mostra una vista globale della ricostruzione 3D che si sta effettuando passo dopo passo, offrendo anche la possibilità di poter esplorare e navigare il modello 3D in ricostruzione.

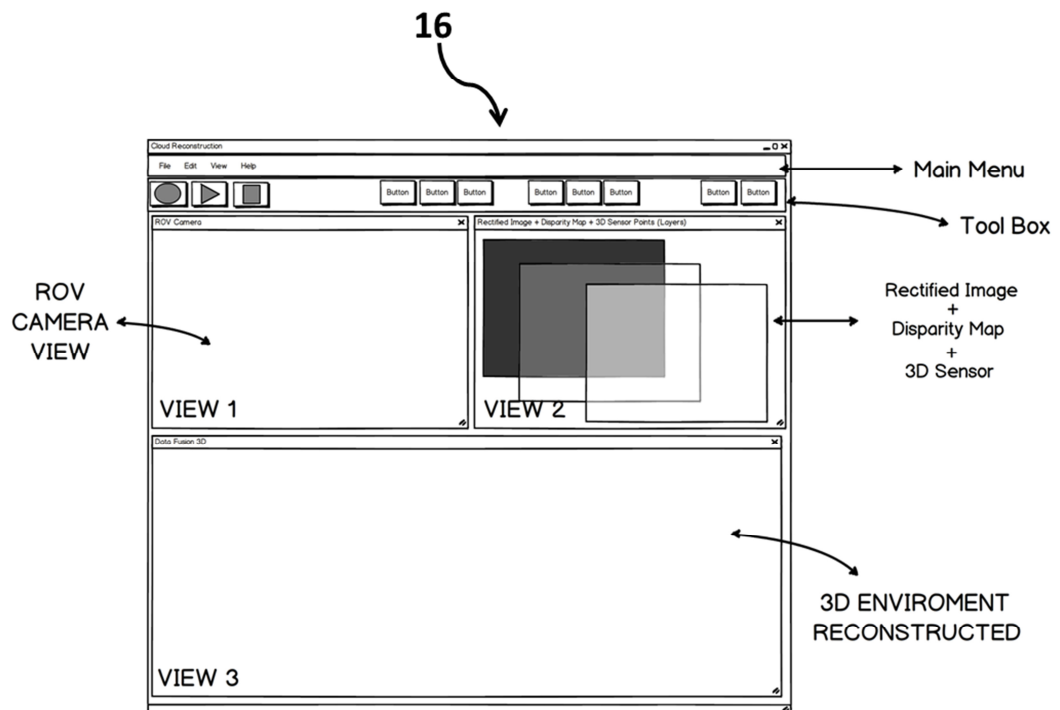


Fig. 3.10: interfaccia utente e gestione layout di visualizzazione in realtà aumentata

Il *software* di visualizzazione e controllo, esso è stato concepito e progettato con l'obiettivo di soddisfare i seguenti requisiti:

- interfaccia utente tale da implementare un modello rappresentativo delle informazioni visive che risulti comprensibile e di facile interpretazione;
- elevata flessibilità tale per cui l'utente sia in grado di intervenire in tempo reale sui principali parametri di settaggio della camera ottico - acustica in modo da far

fronte ai repentini cambiamenti delle condizioni ambientali tipici di un ambiente sottomarino.

I requisiti sopra esposti sono stati coniugati realizzando un'interfaccia utente altamente configurabile e flessibile, con la possibilità di poter selezionare diverse modalità di visualizzazione. In tal modo, le diverse informazioni 2D e 3D disponibili possono essere visualizzate in un'unica schermata (Fig. 3.11) o in apposite schermate in cui i dati visivi sono corredati da informazioni aggiuntive di cui l'utente potrebbe aver bisogno durante lo svolgimento di una specifica attività. Nelle Figg. 3. 12 e 3.13 è possibile osservare una particolare modalità di visualizzazione, in cui la profondità della scena (espressa in m) è codificata sulla mappa di profondità generata dai dati ottici, acustici o dalla fusione di entrambi, utilizzando un gradiente colore RGB a trasparenza variabile.

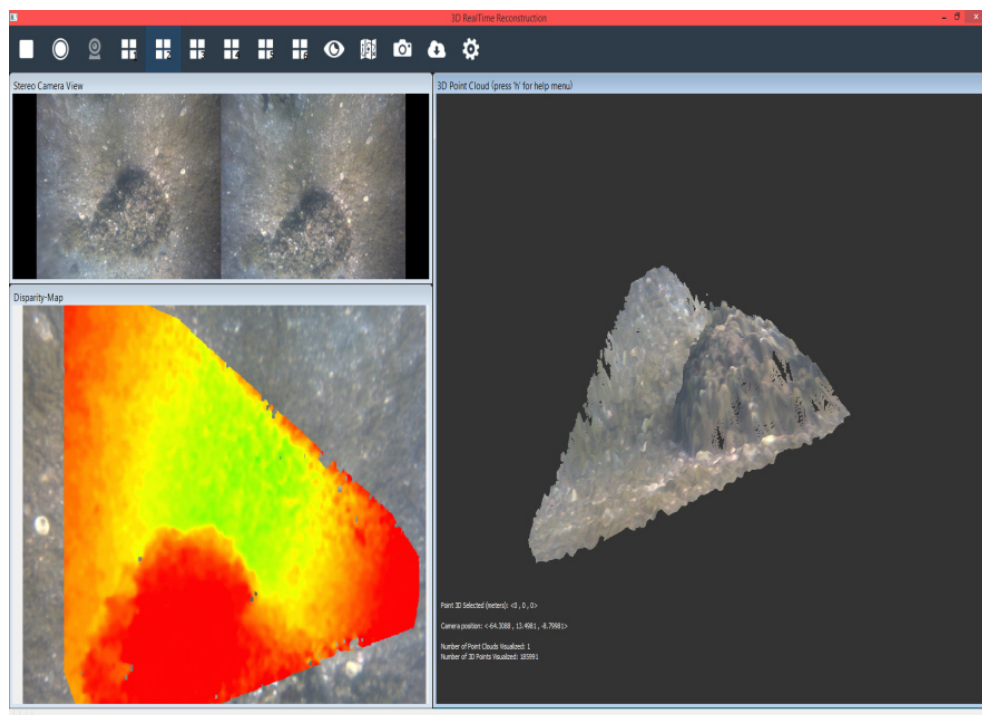


Fig.3.11: prototipo interfaccia utente del software di visualizzazione e controllo camera ottico - acustica tridimensionale.

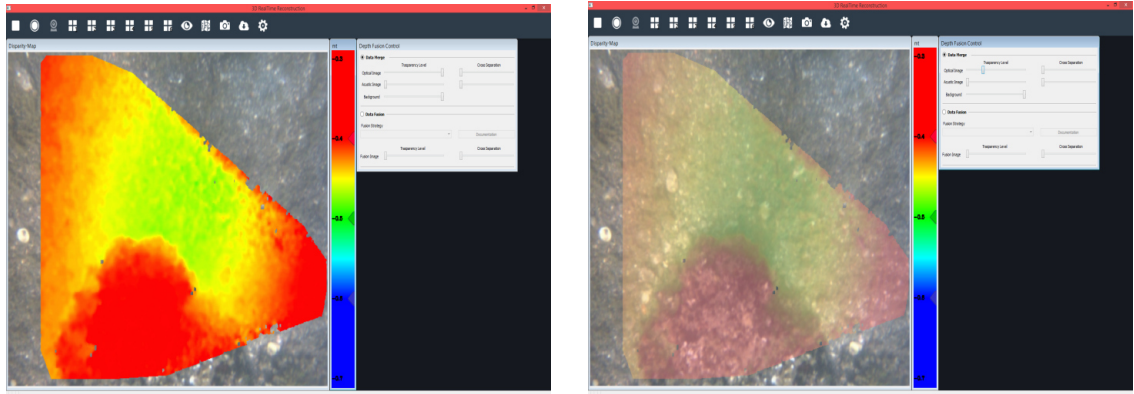


Fig. 3.12: 3sempio modalità di visualizzazione. Mappa di profondità codificata attraverso un gradiente colore RGB a trasparenza variabile.

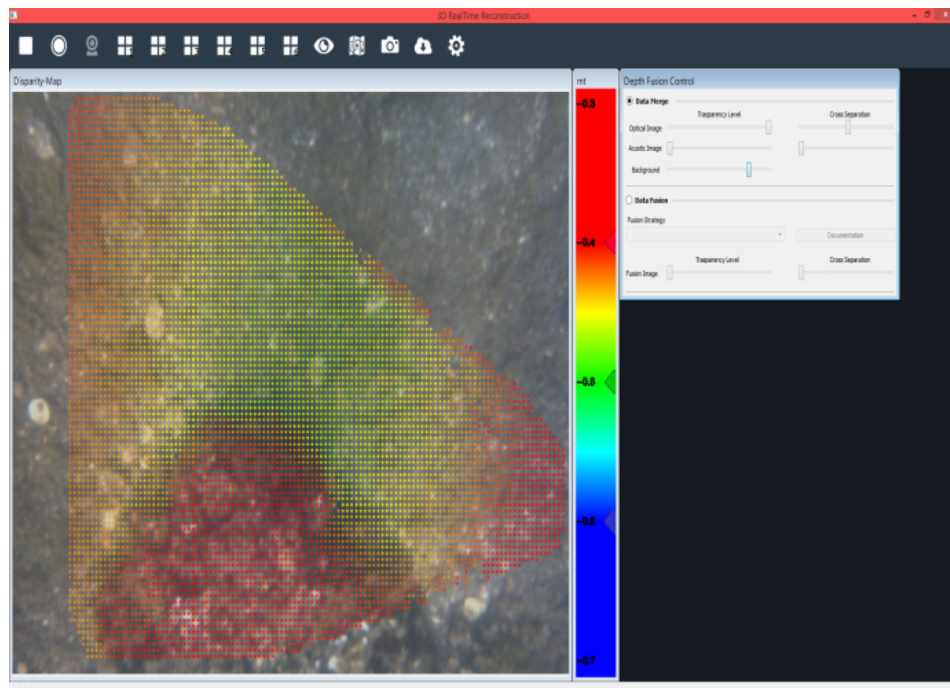


Fig.3.12: modalità di visualizzazione con menu per la selezione della metodologia di data fusion o data merging.

Metodologia per la modellazione 3D di siti archeologici sommersi

Lo stato dell'arte sulle tecniche e sulle tecnologie per la raccolta dati ed il rilievo 3D effettuato attraverso sistemi ottici ed acustici, presentato nel capitolo 1, ha messo in evidenza come molteplici siano le soluzioni che possono essere impiegate allo scopo. Dunque, al fine di poter sviluppare una piattaforma *hardware/software* in grado di integrare in maniera efficiente i due approcci, la prima fase del lavoro condotto ha riguardato la definizione di un protocollo operativo per il rilievo, la ricostruzione 3D e la geo – localizzazione sia di reperti che di interi siti archeologici sommersi. In effetti, poiché le diverse attività che caratterizzano l'indagine di un sito archeologico subacqueo sono fortemente correlate a diversi fattori strettamente dipendenti dalla singola missione, non esiste uno standard che definisca analiticamente le fasi di una missione di rilievo e documentazione subacquea. Per tale motivo, lo scopo è stato quello di definire una metodologia in grado di potersi adattare ai diversi contesti operativi definendo in maniera chiara le modalità di impiego delle tecniche e degli strumenti innovativi sviluppati.

Definita l'architettura della metodologia e le variabili da considerare nell'adottare una determinata soluzione tecnologica, sono state successivamente effettuate diverse sperimentazioni in campo al fine di poter confrontare in maniera analitica le tecniche e le tecnologie candidate ad essere impiegate nelle diverse fasi di cui si compone il protocollo proposto.

4.1 Metodologia di rilievo, ricostruzione 3D e geo-localizzazione di siti archeologici sommersi

L'indagine e la documentazione di un sito archeologico subacqueo è da sempre stata strettamente dipendente, molto più rispetto a quella "di terraferma", dai metodi, dalle tecniche, dagli strumenti e dalla programmazione, a causa dei limiti fisici ed ambientali in cui si opera. Il primo obiettivo è stato, dunque, quello di definire una serie di linee guida per acquisire ed elaborare i dati in modo efficiente, economico e sicuro:

- l'efficienza è correlata alla raccolta di una esaustiva mole di dati di qualità tale da consentire una ricostruzione 3D ottico – acustica soddisfacente considerando un predeterminato livello di accuratezza e dettaglio;
- l'economicità si riferisce alla possibilità di utilizzare, con opportuni adattamenti, apparecchiature *off-the-shelf* che gli operatori possono facilmente ottenere e impiegare;
- la sicurezza ha a che fare con la riduzione di rischi e disagi per gli operatori durante le varie fasi del protocollo soprattutto ad elevate profondità e/o in un ambiente ostile.

Il primo passo nella definizione della metodologia è stato quello di individuare le variabili che influenzano l'adozione di una determinata soluzione tecnologica ed operativa. In effetti, i siti archeologici subacquei possono essere di differenti tipologie e dimensioni, alcuni si presentano meglio conservati di altri, possono essere ubicati in ambienti più o meno ostili o a profondità tali per cui il tempo a disposizione per le operazioni in loco è molto limitato. Una delle prime scelte da effettuare in fase di pianificazione delle attività è quello di determinare il livello di accuratezza voluto. Questo parametro rappresenta una misura di quanto bene si vuole conoscere la posizione e le dimensioni di un oggetto rispetto ad un determinato sistema di riferimento (globale o locale) e di conseguenza definisce quanto fedelmente il modello 3D rappresenterà il sito reale. Strettamente associato a questo parametro è il livello di dettaglio voluto, cioè quante e quali caratteristiche del sito archeologico in esame è necessario acquisire e ricostruire. Un ulteriore fattore che incide fortemente sull'adozione di una determinata soluzione tecnica e tecnologica è senza dubbio quello

economico. In linea di principio, qualsiasi sito archeologico e il suo ambiente circostante, può essere riprodotto con il massimo livello di dettaglio e accuratezza consentita dalle tecnologie allo state dell'arte presenti attualmente in letteratura scientifica. La domanda da porsi è se le risorse economiche impiegate al fine di attuare queste soluzioni rientrano nel budget a disposizione e se sono giustificate dalla tipologia di modello 3D e dalle informazioni che si intendono ottenere.

Accanto a queste variabili di tipo decisionale ve ne sono altre che pongono vincoli fisici nell'adozione di una data soluzione tecnologica. Ad esempio, una piattaforma hardware con un sistema GPS o DGPS può essere impiegata per la geo-localizzazione di un sito archeologico o dei reperti in esso contenuti solo in condizioni di acque poco profonde (max 10 – 15 m circa), così come alcune tecnologie di acquisizione immagini possono essere usate solo in un ambiente subacqueo dove le condizioni di visibilità lo consentono. Un ulteriore aspetto da considerare riguarda, inoltre, la gestione della mole di dati acquisiti e delle informazioni da essi ottenute. Infatti, se da un lato il modello 3D ottico – acustico generato deve rappresentare in maniera quanto più fedele possibile il sito archeologico investigato, dall'altro, esso deve avere una dimensione (in Megabyte) tale per cui la sua visualizzazione e navigazione, così come la possibilità offerta agli utenti di poter interagire con esso avvenga in maniera fluida ed accattivante. Questo vincolo sulla dimensione massima del modello 3D ottico – acustico generato, così come tutti gli altri che derivano dai limiti delle risorse *hardware* e dai vincoli progettuali pongono un limite al livello di dettaglio che è possibile ottenere dell'intero sito archeologico e impongono di introdurre nella metodologia una politica di gestione delle risorse *hardware* e *software* disponibili basata sulla decisione di quali aree acquisire, con quale livello di dettaglio e con quale tecnica, tra quelle possibili, ottenere la fusione dei dati ottico – acustici.

In Tab. 1 sono riportate le principali variabili individuate insieme ad una loro breve descrizione:

Parametri	Descrizione
Profondità	La profondità del sito archeologico pone vincoli stringenti sull'adozione di una determinata soluzione tecnologica sia per quanto riguarda la piattaforma hardware impiegata in fase di acquisizione dati che per la georeferenziazione degli stessi. Ad esempio, profondità superiori ai 40 m pongono un limite severo sul tempo di fondo consentito agli operatori sub limitando dunque la possibilità di poter acquisire e ricostruire tali siti solo attraverso l'impiego di veicoli ROV o AUV equipaggiati con opportuni sistemi di acquisizione immagini (camere digitali ad alta risoluzione) e georeferenziazione (tipicamente sistemi LBL o USBL).
Estensione	Le dimensioni dell'area d'interesse archeologico incidono sulla scelta del sistema di acquisizione da adottare in quanto determinano sia i tempi di acquisizione che la mole di dati da archiviare e successivamente elaborare. Siti archeologici di elevata estensione (ad es. superiore ai 250 m ² circa) richiedono l'uso di sistemi dotati di propulsione o a traino e di conseguenza l'impiego di camere con una elevata velocità d'acquisizione e di archiviazione.
Accuratezza	Questo parametro rappresenta una misura di quanto bene vogliamo conoscere la posizione e la dimensione di un oggetto rispetto ad un determinato sistema di riferimento (globale o locale) e di conseguenza definisce quanto fedelmente il modello 3D rappresenta il sito reale. Livelli di accuratezza elevati (25 mm circa) richiedono l'uso di strumenti e metodologie molto più costose rispetto a modelli con accuratezza minore (300 mm).
Livello di dettaglio	Questo parametro definisce quante e quali caratteristiche del sito archeologico e dell'ambiente circostante è necessario acquisire e ricostruire, e dunque, le tecniche, gli strumenti e il budget da impiegare per ottenere tali risultati.
Tipologia di fondale	La tipologia di fondale (sabbia, fango, rocce affioranti, sabbia e roccia irregolari, sedimenti con increspature ecc.) determina la metodologia e gli strumenti da utilizzare in fase di acquisizione immagini. Ad esempio, un fondale piatto può essere acquisito attraverso tipiche tecniche di acquisizione aerea, mentre la presenza di rocce e scogli richiede necessariamente l'acquisizione della scena da posizione obliqua in modo da acquisire le zone non visibili dall'alto.
Tipologia di sito archeologico	Relitti di antiche navi, anfore e altri manufatti più o meno sparsi sul fondale o intere città sommerse sono tipici esempi di siti archeologici subacquei. Le loro peculiarità ed il loro interesse archeologico definiscono in maniera univoca le tecniche e le metodologie impiegate per il loro rilievo e la loro documentazione.

Flora e fauna	La fedeltà nella riproduzione dell'ambiente marino in cui il sito archeologico è ubicato implica la necessità di acquisire, modellare e ricostruire anche la flora e la fauna ivi presente. Un sito archeologico in un mare tropicale avrà delle peculiarità (ad es. presenza di coralli a basse profondità, ecc.) completamente differenti da siti ubicati nel mar mediterraneo (caratterizzati ad es. da ampie praterie di posidonia oceanica).
Budget e tempo	Queste due variabili incidono in maniera determinante sulle tecniche, gli strumenti ed i sistemi impiegati per il rilievo e la documentazione del sito archeologico. Le risorse temporali ed economiche impiegate nelle varie attività dovranno essere sempre tali da garantire il massimo grado di efficacia ed efficienza possibile.

Tab. 1: principali parametri di valutazione per la definizione del protocollo operativo.

Definiti i parametri decisionali attraverso i quali determinare l'impiego di una determinata soluzione tecnica e tecnologica per l'esecuzione di una specifica attività tra quelle previste per il rilievo e la documentazione dei siti archeologici subacquei, sono state successivamente definite le varie fasi del protocollo operativo. In Fig. 4.1 è mostrata l'architettura della metodologia implementata.

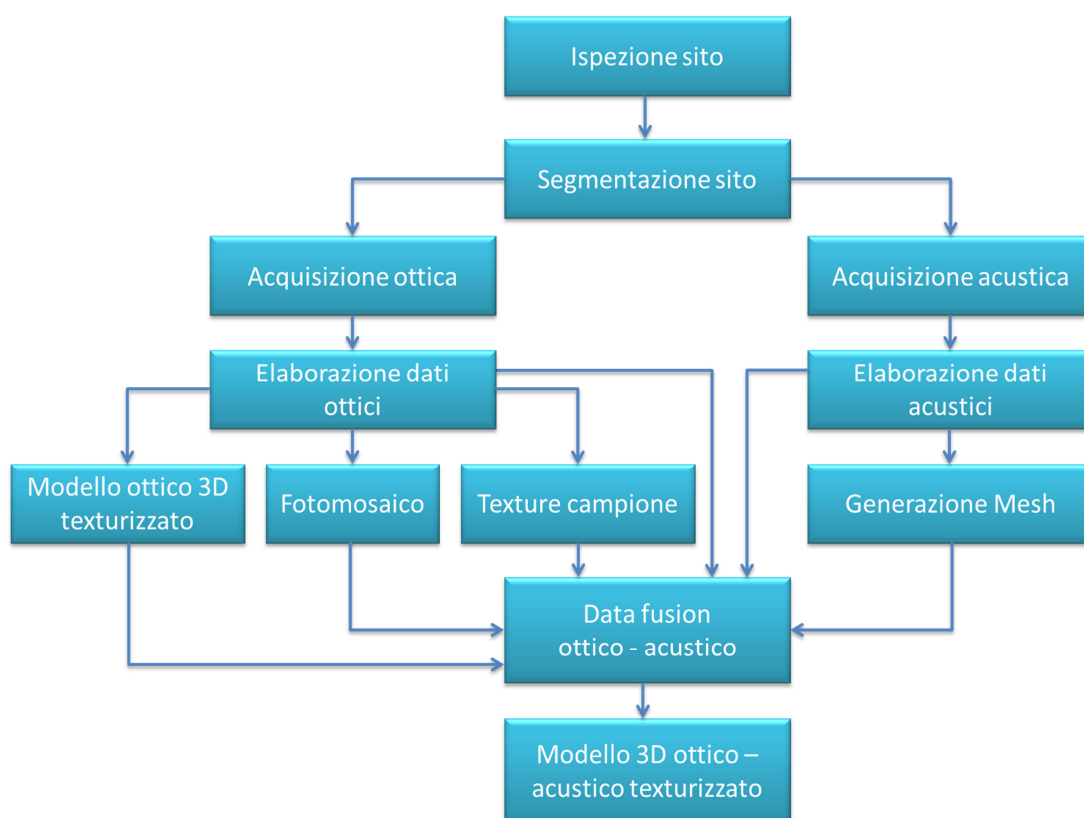


Fig. 4.1: Architettura della metodologia proposta per la generazione dei modelli 3D ottico – acustici.

Il primo passo del protocollo operativo consiste in una osservazione preliminare al fine di avere una visione generale del sito archeologico d'interesse ed assegnare in tal modo un valore (sia qualitativo o, ove possibile, quantitativo) a molte delle variabili decisionali sopra elencate: tipologia e profondità del sito archeologico, posizione dei manufatti, orientazione relativa tra gli oggetti, tipologia di fondale, flora e fauna presenti, livello di luminosità e torbidità dell'acqua. Questi dati verranno successivamente analizzati al fine di effettuare una "segmentazione" del sito archeologico in "zone d'interesse." In sostanza si tratterà di individuare all'interno del sito le aree di maggiore importanza dal punto di vista storico-archeologico, quelle che presentano le stesse caratteristiche biologiche e geomorfologiche, quelle alle quali saranno applicate le stesse tecniche di acquisizione, ricostruzione 3D e fusione dei dati ottico – acustici. In queste fasi preliminari sarà inoltre definita la(e) metodologia(e) utilizzata per la geo – localizzazione del sito archeologico e dei reperti in esso contenuti. Questa scelta dovrà essere necessariamente effettuata in questa fase poiché, come vedremo in dettaglio nelle successive sezioni, alcune delle tecniche proposte prevedono il dislocamento all'interno del sito archeologico di particolari *marker* ottico – acustici attraverso i quali sarà possibile assegnare in maniera univoca, a valle del processo di acquisizione, precise coordinate spaziali a determinati punti d'interesse (posizione di reperti, perimetro del sito archeologico ecc.) secondo un sistema di riferimento locale o globale.

Le fasi operative "in campo" si concludono con l' acquisizione sia ottica che acustica. Queste attività rappresentano il punto più critico del protocollo operativo proposto in quanto determinano la qualità dei dati acquisiti e di conseguenza la qualità del modello 3D finale e delle informazioni in esso contenute. A valle del processo di acquisizione, i data set ottico - acustici disponibili verranno elaborati al fine di generare il modello 3D ottico – acustico del sito archeologico investigato. Indipendentemente dalla particolare soluzione tecnica e tecnologica adottata, l'idea alla base per la generazione di modelli 3D multi-risoluzione consiste nell'ottenere una *mesh* poligonale a bassa risoluzione (dell'ordine delle decine di centimetri) a partire dalla batimetria acustica generale del sito sommerso e delle aree circostanti a cui saranno sovrapposti modelli ad alta definizione (a risoluzione subcentimetrica) delle aree di maggiore interesse del sito subacqueo attraverso una fusione che renda impercettibile all'utente

la giunzione fra i due modelli. Il modello 3D così generato sarà infine texturizzato attraverso una *texture* ad alta definizione ottenuta elaborando le immagini acquisite durante il rilievo ottico e opportunamente mappata sul modello stesso.

Le varie soluzioni tecniche e tecnologiche candidate ad essere applicate nelle varie fasi operative previste dalla metodologia proposta sono state testate ed analizzate in diverse sperimentazioni sia in campo che in laboratorio. I risultati ottenuti e le valutazioni che ne derivano sono presentate nelle successive sezioni.

4.1.1 Ispezione e segmentazione sito archeologico

Le diverse attività condotte durante le fasi iniziali della metodologia per il rilievo, la ricostruzione 3D e la geo-localizzazione di siti archeologici sommersi consistono in una osservazione preliminare al fine di avere una visione generale del sito archeologico d'interesse e definire in tal modo i parametri decisionali in funzione dei quali attuare una determinata soluzione tecnica e tecnologica nelle successive fasi del protocollo. Successivamente il sito verrà delimitato e segmentato in zone d'interesse attraverso l'uso di boe di segnalazione e marker adagiati sul fondale (Fig. 4.2).

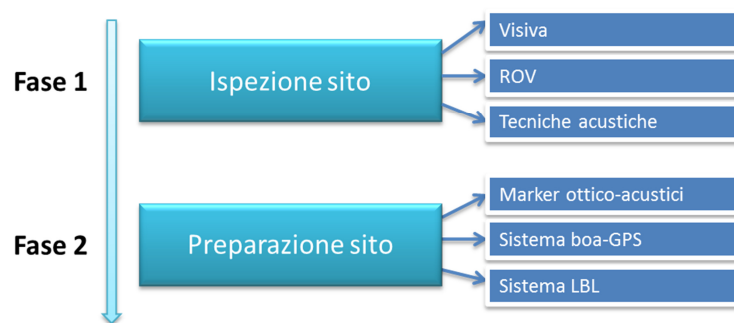


Fig. 4.2: attività condotte nell'ambito dell'ispezione e della segmentazione del sito archeologico e relative tecniche

In caso di basse profondità (< 20 m) e per estensioni limitate (< 400 m² c.a.) l'attività di indagine può essere una semplice ispezione visiva condotta attraverso l'ausilio di operatori subacquei. Questa tecnica è stata adottata in una prima sperimentazione in campo condotta sul "relitto punta Scifo D" presso il sito archeologico di Capo Colonna (KR) (Fig. 4.3). Si tratta dei resti di una grande nave romana di età imperiale, ancora in assetto di navigazione, con il suo carico di marmi grezzi o semilavorati, di notevoli

dimensioni. Il carico si estende per un'area larga 22 m x 16 m circa, ad una profondità media di 6,5 m.

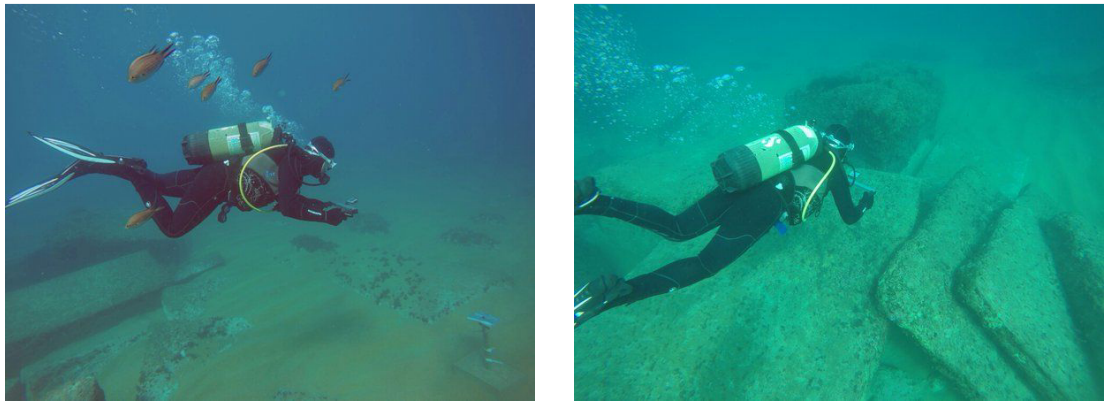


Fig. 4.3: ispezione con tecnica visiva da parte di operatori sub del "relitto punta Scifo D" presso il sito archeologico di Capo Colonna.

Per siti archeologici ubicati a grande profondità (> 40 m) e/o di elevata estensione (> 400 m² c.a.) dovranno essere necessariamente impiegate soluzioni che prevedono l'utilizzo di piattaforme *hardware* più o meno complesse. In caso di buone condizioni di visibilità, l'ispezione può essere realizzata attraverso l'ausilio di un *Remotely Operated underwater Vehicle* (ROV) equipaggiato con un sistema di camere ad elevato *frame rate*. Se il *budget* a disposizione per la missione di rilievo è elevato, possono essere impiegati ROV di tipo *Light Workclass* opportunamente equipaggiati con una o più camere ottiche, sonar, sensori di navigazione per il rilevamento e la scansione del fondale marino. La dotazione di sensori a bordo di questa tipologia di veicoli è tale da consentire non solo l'ispezione del sito ma anche di acquisire dati ottici ed acustici geo – referenziati con conseguente miglioramento della fusione dei *dataset*. Per missioni a budget limitato, l'ispezione visiva può essere condotta utilizzando ROV di tipo *Micro Class*. Recentemente l'Università della Calabria (Unical) si è dotata di un veicolo subacqueo appartenente a questa classe, l'OpenROV. Questo è un robot subacqueo *open source* alimentato a batterie, dotato di una *webcam* a risoluzione *fullHD* (1920 x 1080) in grado di raggiungere i 70 m di profondità. Primi test sono stati condotti in vasca per verificarne l'affidabilità, la stabilità ed investigare, inoltre, la possibilità di poter installare a bordo del veicolo camere ottiche ad elevata risoluzione e *frame rate* (in particolare 2 GoPro hero 4 silver edition con risoluzione massima 4K e *frame rate* di 120 fps). In tal modo esso potrà essere impiegato non solo in fase di ispezione ma

anche di acquisizione dei dati ottici (Fig. 4.4). Nei prossimi mesi saranno effettuati dei test in mare per verificare il comportamento del veicolo (in termini di assetto e stabilità) in condizioni operative reali e definire il giusto settaggio delle camere in modo tale da poter acquisire immagini subacquee con la giusta messa a fuoco e nitidezza di colore.

In condizioni di scarsa visibilità, o comunque in tutti i casi in cui risulta impossibile impiegare un veicolo subacqueo per l'ispezione del sito archeologico, una prima classificazione del fondale può essere effettuata a partire dai dati di *backscatter* ottenuti utilizzando strumentazione e tecniche acustiche.

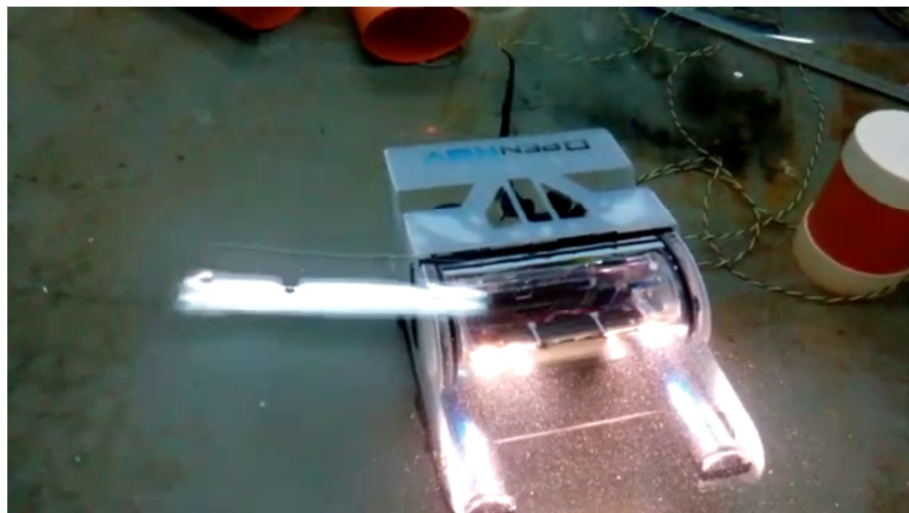


Fig. 4.4: primi test in vasca veicolo subacqueo OpenROV (sinistra). Console di comando su PC (destra).

Nell'ambito delle attività di test condotte presso il sito archeologico di Capo Colonna, le basse profondità operative, hanno consentito di effettuare una sperimentazione in campo della metodologia di geo – localizzazione dei siti archeologici subacquei al fine di delimitare il sito archeologico e l'area circostante ed attuare la fase di segmentazione in zone d'interesse (Fig. 4.5).



Fig. 4.5: geo – referenziazione “relitto punta Scifo D” presso il sito archeologico di Capo Colonna con l'ausilio di una boa di superficie equipaggiata con un sistema GPS.

A partire dal punto centrale del relitto (marker 1) è stata delimitata un'area di circa 70 m x 65 m e geo – referenziati i suoi quattro angoli (marker A, B, C, D). Il fondale in questo tratto di mare si presenta prevalentemente sabbioso. Il sito è stato dunque segmentato in due zone: la zona 1 in cui sono presenti i blocchi di marmo e la zona 2 che circonda il relitto fino ad arrivare ai confini dell'area delimitata.

La realizzazione di un “percorso archeologico subacqueo” da parte dell'Area Marina Protetta (AMP) di Isola Capo Rizzuto sul relitto “Punta Scifo D” è stato utilizzato nell'attività di sperimentazione per suddividere la zona 1 dalla zona 2. In effetti, tale percorso, delimita l'area in cui è adagiato il carico lapideo attraverso l'uso di cime ancorate a pali di sostegno. Quest'ultimi sono, inoltre, classificati attraverso l'uso di marker numerati (percorso subacqueo) e marker con lettere (percorso snorkeling) in modo da orientare l'utente durante la visita del sito (Fig. 4.6). Essendo questi marker

geo – referenziati, i dati relativi sono stati utilizzati nelle nostra attività di sperimentazione.

In Fig. 4.6 (sinistra) si può osservare, oltre ai pali di sostegno delle cime che delimitano il percorso archeologico, la presenza di diversi marker adagiati sul fondale oltre ad una *color chart*. Come vedremo in dettaglio nelle successive sezioni, i primi sono stati utilizzati al fine di ottenere un preciso *scaling* del modello 3D ricostruito attraverso una tecnica *Structure From Motion* (SFM), mentre la *color chart* è uno strumento impiegato per effettuare la correzione colore attraverso tecniche di *image enhancement* sia del data set ottico acquisito che della *texture* finale del modello 3D.

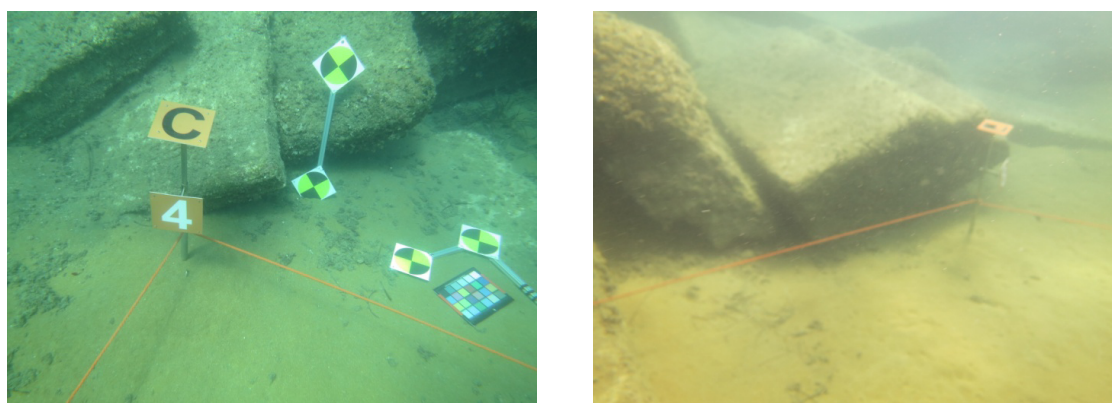


Fig. 4.6: delimitazione zona 1 del “relitto punta Scifo D” presso il sito archeologico di Capo Colonna.

Una seconda tecnica per la geo – localizzazione dei siti archeologici subacquei e dei reperti in esso contenuti prevede l’utilizzo ed il posizionamento all’interno del sito archeologico di particolari marker ottico – acustici. Questi sono formati da piccole boe riempite d’aria agganciate a corpi morti che saranno collocate in prossimità del reperto d’interesse (Fig. 4.7, alto). L’acquisizione ad alta risoluzione del fondale marino attraverso una scansione *multibeam* con opzione di *backscatter* restituirà una batimetria acustica in cui in corrispondenza dei marker si avranno valori di riflettività acustica molto superiori alla media del fondale, essendo l’aria il miglior riflettente acustico esistente. Le coordinate spaziali del punto sulla batimetria rappresentante un determinato marker saranno dunque associate al reperto relativo. Un primo aspetto analizzato per l’implementazione della metodologia ha riguardato il dimensionamento dei marker. Essendo la risoluzione spaziale della mappa batimetrica generata utilizzando il sistema multibeam Seabat 7125, (individuato come strumento idoneo ad

essere impiegato per l'acquisizione dei dati acustici) dell'ordine di 10 cm, al fine di poter discriminare il marker attraverso un solo punto sulla mappa è necessario che la boa abbia un diametro della stessa dimensione. Successivamente è stata realizzata una simulazione al PC a partire dalla batimetria acustica di un tratto di fondale presso il sito archeologico di Baia (NA) (Fig. 4.7, destra) al fine di determinare in che modo discriminare i marker sulla batimetria stessa.

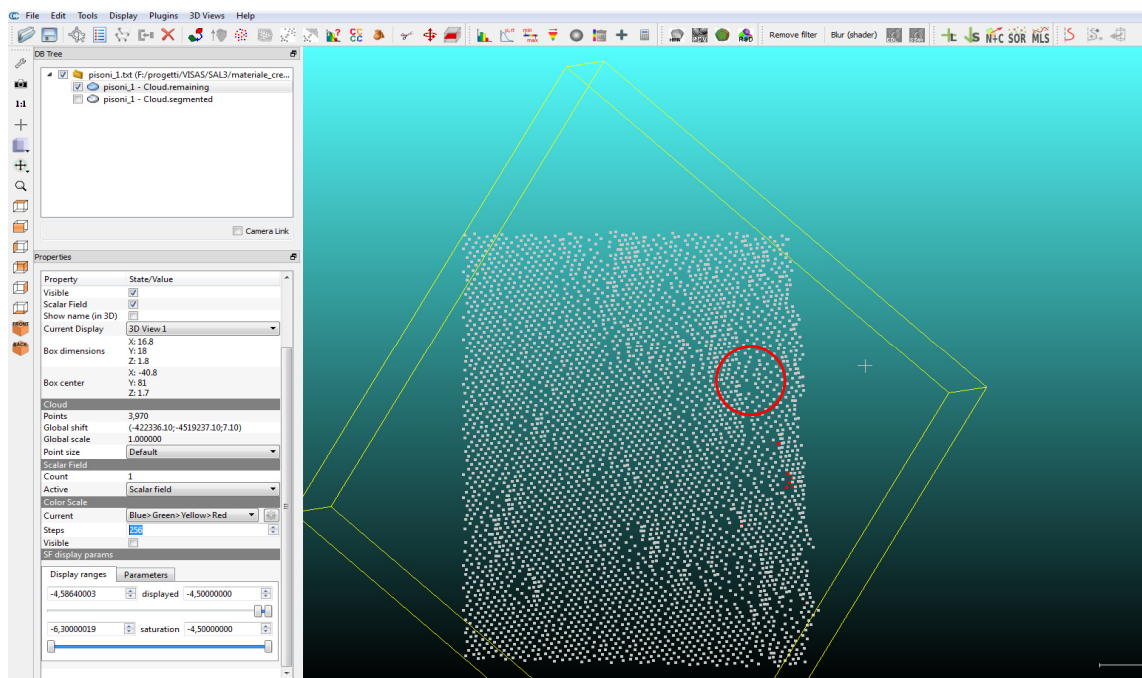


Fig. 4.7: marker ottico – acustico per la geo – localizzazione di reperti subacquei (alto). Simulazione al computer visualizzazione marker su porzione mappa batimetrica del sito archeologico subacqueo di Baia (NA) (destra).

I valori di *backscatter* sono stati codificati attraverso l'utilizzo di un campo scalare RGB ed associati ai punti spaziali relativi. Stimando un aumento di circa 2-3 volte del valore

di riflettanza acustica dei punti in corrispondenza dei quali sono ubicati i marker la simulazione dimostra come, agendo sui cursori che delimitano il *range* in cui i valori del campo scalare sono visualizzati, sia possibile discriminare in maniera netta tali punti (di colore rosso) dal resto della batimetria (punti bianchi).

In caso di siti archeologici di elevata estensione o con un numero elevato di reperti in esso contenuti l'impiego delle tecniche sopra esposte può risultare di difficile attuazione. In queste circostanze risulterà necessario impiegare sistemi hardware più complessi. Una delle soluzioni candidate è rappresentata dall'impiego di transponder LBL o USBL. A tal fine potrà essere impiegato, dove possibile, un sistema di posizionamento acustico direttamente installato a bordo di un ROV di tipo *Light Workclass*. Un'ulteriore soluzione tecnologica, applicabile su fondali a bassa profondità (< 6-7 m), è rappresentata dal sistema mostrato in Fig. 4.8 (destra). Si tratta di una slitta di superficie a traino progettata e realizzata presso l'Istituto per la Ricerca e la Protezione Ambientale (ISPRA). Questa è dotata di due camere GoPro hero 3 + black edition (risoluzione di 12MP e campo di vista (FOV) di 108°) e un GPS con *data logger* Blumax 4404. Le camere sono posizionate in modo tale da acquisire immagini zenitali del fondale sottostante. Il tracciato GPS relativo all'acquisizione delle immagini dei transetti acquisiti viene associato ai dati EXIF delle relative immagini attraverso l'utilizzo del *software* RoboGeo[®]. Il risultato di tale operazione è un data set costituito da immagini geo – referenziate (Fig. 4.8, sinistra). Tale sistema, come vedremo in dettaglio nella prossima sezione, è stato utilizzato per acquisire la zona 2 del sito archeologico di Capo Colonna.

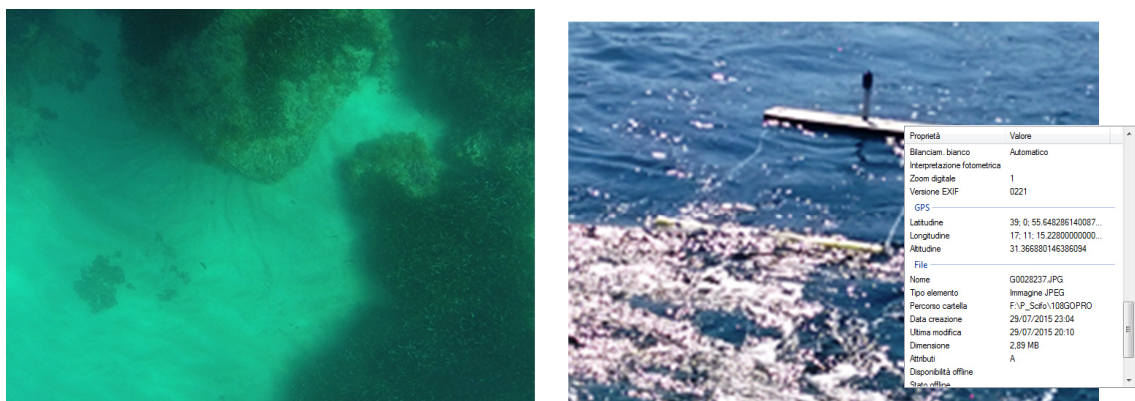


Fig.4.8: slitta di superficie a traino per l'acquisizione di immagini geo – referenziate del fondale sottostante (sinistra). Esempio di immagine ottenuta con relativi dati EXIF.

4.1.2 Acquisizione ed elaborazione dati ottici

In questa sezione del documento vengono presentate le tecniche impiegate ed i risultati ottenuti in diverse sperimentazioni in campo relative alle fasi di acquisizione ed elaborazione dei dati ottici (Fig. 4.9):

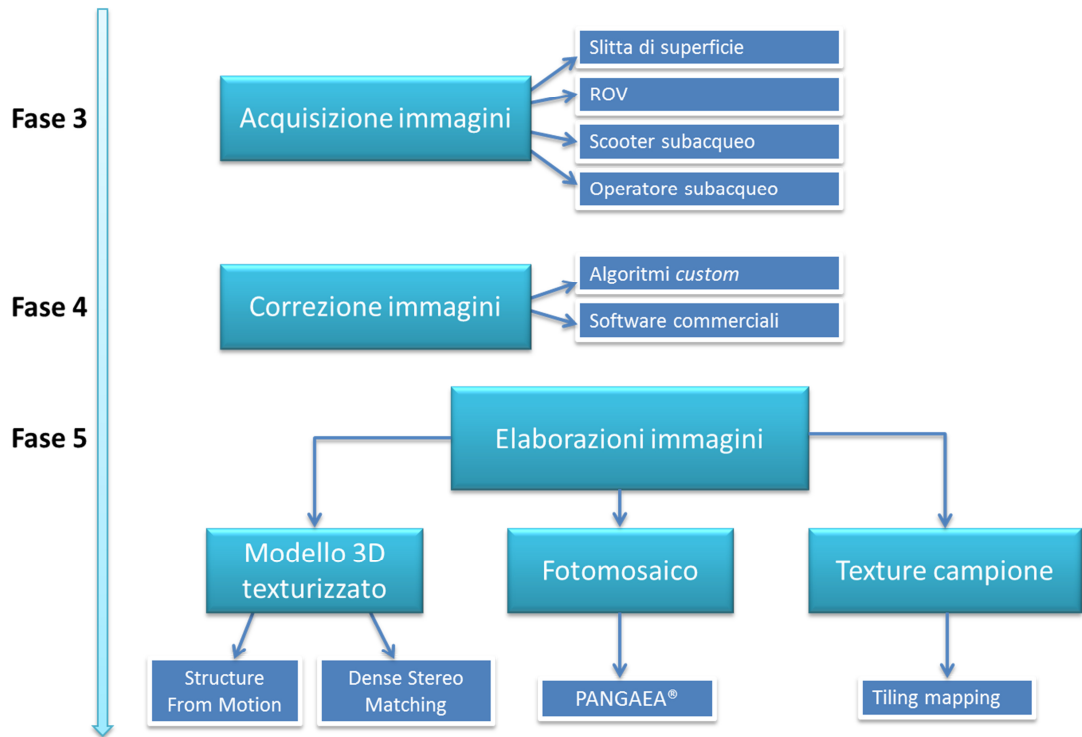


Fig. 4.9: attività condotte nell'ambito delle fasi di acquisizione ed elaborazione dati ottici e relative tecniche e tecnologie impiegate.

La fase di acquisizione immagini rappresenta uno dei punti più critici del protocollo operativo proposto in quanto determina la qualità dei dati ottici che successivamente verranno elaborati ed allineati con quelli acustici per la generazione del modello 3D ottico – acustico. Al fine di ottenere un adeguato *mapping* del sito archeologico è necessario utilizzare un'adeguata configurazione di camere tale da ottenere un elevato rapporto segnale rumore, in cui gli effetti di elementi non statici, *motion blur*, attenuazione della luce lungo la colonna d'acqua e rumore digitale sono ridotti al minimo e la profondità di campo massimizzata. Un'attenta pianificazione delle attività di rilievo è inoltre necessaria al fine di garantire un'adeguata copertura e sovrapposizione tra le immagini acquisite, operazione non sempre di semplice attuazione in ambiente subacqueo. Di conseguenza l'attività di sperimentazione è stata

focalizzata più che sugli algoritmi di ricostruzione vera e propria (oramai ben definiti e consolidati in letteratura scientifica) alla ricerca di soluzioni tecniche e tecnologiche in grado di garantire un'efficace ed efficiente acquisizione dei dati ottici nelle diverse condizioni operative.

4.2 Ricostruzione ottico- acustica sito archeologico Capo Colonna

4.2.1 Acquisizione ed elaborazione dati zona 2: fondale.

Il tratto di fondale che costituisce la zona 2 del sito archeologico di capo Colonna è prevalentemente sabbioso con presenza in alcune aree di tratti rocciosi ricoperti da Posidonia Oceanica. Poiché l'Unical, in collaborazione con ISPRA, ha realizzato nell'Area Marina Protetta di Isola Capo Rizzuto alcune sperimentazioni in campo sull'applicazione di tecniche di *imaging* 3D per la mappatura di praterie di Posidonia Oceanica, si è deciso di utilizzare tale metodologia e le soluzioni tecnologiche ivi proposte per il rilievo e l'acquisizione dei dati ottici della suddetta zona e realizzare in tal modo sia un fotomosaico che una ricostruzione 3D dei transetti acquisiti. La piattaforma *hardware* impiegata per l'acquisizione è stata la slitta di superficie a traino presentata nella sezione 4.1.1. Al fine di evitare eccessive distorsioni ottiche sulle immagini acquisite ed ottenere una buona risoluzione delle stesse, le GoPro hero 3 + black edition installate a bordo della slitta sono state settate in modalità *time lapse* (intervallo 0.5 s), risoluzione 12 MP (4000 x 3000 pix) e FOV medium (94.4° x 72.2°). Il primo passo della metodologia consiste nel dimensionamento della griglia attraverso la quale definire la rotta che dovrà percorrere l'imbarcazione da traino durante l'acquisizione dei vari transetti. L'obiettivo è quello di garantire una sovrapposizione tra le immagini del singolo transetto di circa il 60% e una sovrapposizione di circa il 20% - 30% tra i vari transetti acquisiti (in maniera parallela). I parametri necessari ad un corretto dimensionamento della griglia sono:

- stima FOV delle camere in acqua;
- profondità del fondale da acquisire;
- velocità massima dell'imbarcazione.

Per quanto concerne il primo parametro, precedenti sperimentazioni hanno dimostrato che il FOV in acqua della camera GoPro nel settaggio proposto è di $63^{\circ} \times 48^{\circ}$ (riduzione di circa il 30%). Essendo la profondità media del fondale di 6.5 m il footprint C sarà di circa 7.2 m x 5.7 m (Fig. 4.10, sinistra). Una velocità massima dell'imbarcazione di circa 1 Kn/h, con un *time lapse* di 0.5 s, garantisce una sovrapposizione tra le immagini del singolo transetto superiore al 75%. Si è dunque deciso di creare una griglia di 5 m x 5 m in modo tale da avere una sovrapposizione tra i transetti di circa il 50%. Essendo l'area delimitata di circa 70 m x 65 m e eseguendo un'acquisizione parallela alla costa (direzione sud-ovest nord-est) in modo da avere una profondità di fondale circa uguale per ogni transetto, sono stati acquisiti un totale di 13 transetti (Fig. 4.10, destra). La griglia così generata è stata successivamente geo-referenziata utilizzando i dati ottenuti nella fase 2 del protocollo operativo e caricati su un GPS cartografico della serie Trimble.

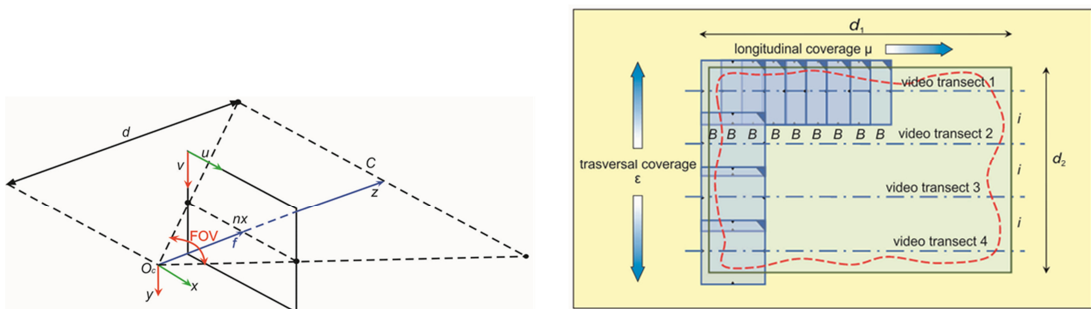


Fig.4.10: schematico per il calcolo del footprint sul fondale della camera GoPro (sinistra). Geometria per la determinazione della dimensione e del numero di transetti d'acquisizione (destra).

Il data set ottico è stato successivamente elaborato utilizzando una metodologia *custom* di *image enhancement* sviluppata presso il DIMEG dell'Unical, impiegata al fine di eliminare gli effetti di sfocatura dovuti allo *scattering* e correggere le dominanti di colore (componente colore verde - blu) [92] (Fig. 4.11).

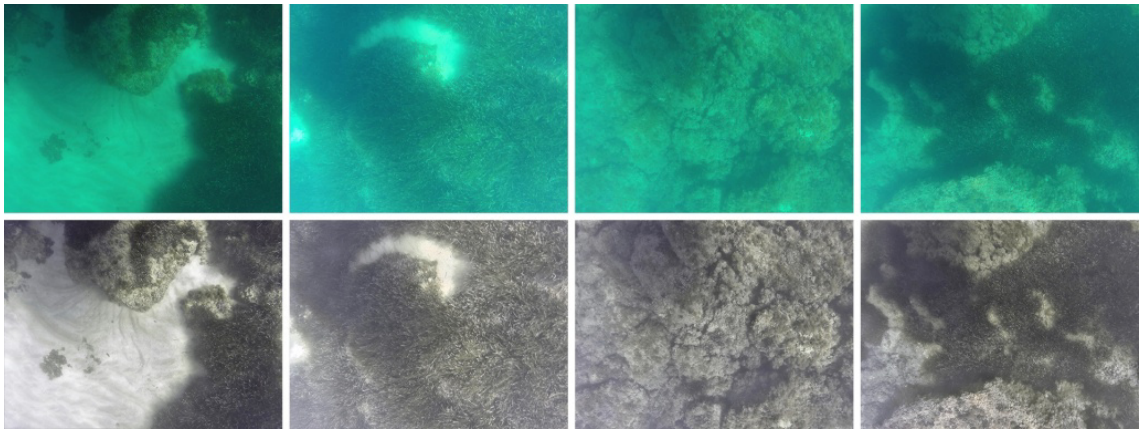


Fig. 4.11: risultati applicazione algoritmo *custom* di correzione immagini [4] su immagini di un tratto di fondale della zona 2 con una notevole estensione di Posidonia oceanica.

Il primo passo della metodologia consiste nell'applicare uno *sharp filter* al *data set* d'immagini al fine di attenuare l'effetto "nebbia" causato dallo *scattering*, che porta ad una diminuzione del contrasto ed ad un aumento della sfocatura delle immagini stesse. Successivamente, il colore delle immagini viene migliorato attraverso un bilanciamento del bianco effettuato nello spazio colore $\alpha\beta$. Questa tecnica consente di eliminare la dominante colore delle immagini subacquee (soggette tipicamente a fenomeni di alterazione dei colori) bilanciando le componenti cromatiche α e β , mentre la componente di luminanza I è usata per migliorare il contrasto delle immagini attraverso una tecnica di *cutoff* e *histogram stretching*. Per quanto concerne la ricostruzione 3D, la metodologia verte su un algoritmo Multi View Stereo (MVS) utilizzato per risolvere la corrispondenza di punti omologhi su una sequenza di immagini non ordinate.

La prima fase della ricostruzione 3D è stata effettuata mediante il *software open source* Bundler (pacchetto di strumenti per lo *structure from motion*) che riceve in *input* un *set* di immagini (scattate in sequenza oppure anche disordinate) e produce in *output* una ricostruzione geometrica preliminare della scena, nonché la posizione ed i parametri di calibrazione della fotocamera. La ricostruzione 3D finale dei transetti è stata poi portata a termine mediante PMVS2, un *software open source* che riceve in *input* i parametri calcolati da Bundler e fornisce in *output* una nuvola di punti 3D densa. Le nuvole di punti 3D dense dei 13 transetti, dopo una fase preliminare di pulitura per la riduzione del rumore dovuto alla ricostruzione, eseguita applicando filtri automatici presenti all'interno del *software open source* MeshLab, sono state

processate mediante l'algoritmo *Poisson Surface Reconstruction* incluso in Meshlab stesso, ottenendo in tal modo una superficie 3D continua (mesh triangolarizzata). Successivamente è stata effettuata la mappatura delle texture sui modelli 3D risultanti. Tale operazione, definita di *UV Mapping*, può essere realizzata utilizzando diverse tecniche. Nel caso in esame, vista la conoscenza dei punti di presa della camera, la tecnica utilizzata consiste nella proiezione delle immagini originali ad alta risoluzione direttamente sulla superficie. In questo modo si riesce ad ottenere un modello 3D la cui *texture* ha una risoluzione comparabile con le immagini originali ed inoltre si avrà l'informazione sul colore anche per le aree aggiunte in *post-processing*. Applicando tale metodologia sono stati generati i modelli 3D testurizzati per i 13 transetti acquisiti (Fig. 4.12, sinistra). Essendo le immagini del data set geo – referenziate, i modelli 3D generati risultano anch'essi geo – referenziati e già scalati. Grazie ad una sovrapposizione del 50% tra le immagini appartenenti ai 13 transetti, i modelli 3D generati presentano un'elevata percentuale di aree in comune. Il modello 3D finale della zona 2 del sito archeologico di capo Colonna è stato dunque ottenuto registrando tra loro i modelli 3D dei singoli transetti attraverso un'applicazione sistematica dell'algoritmo *Iterative Closest Point* (ICP).

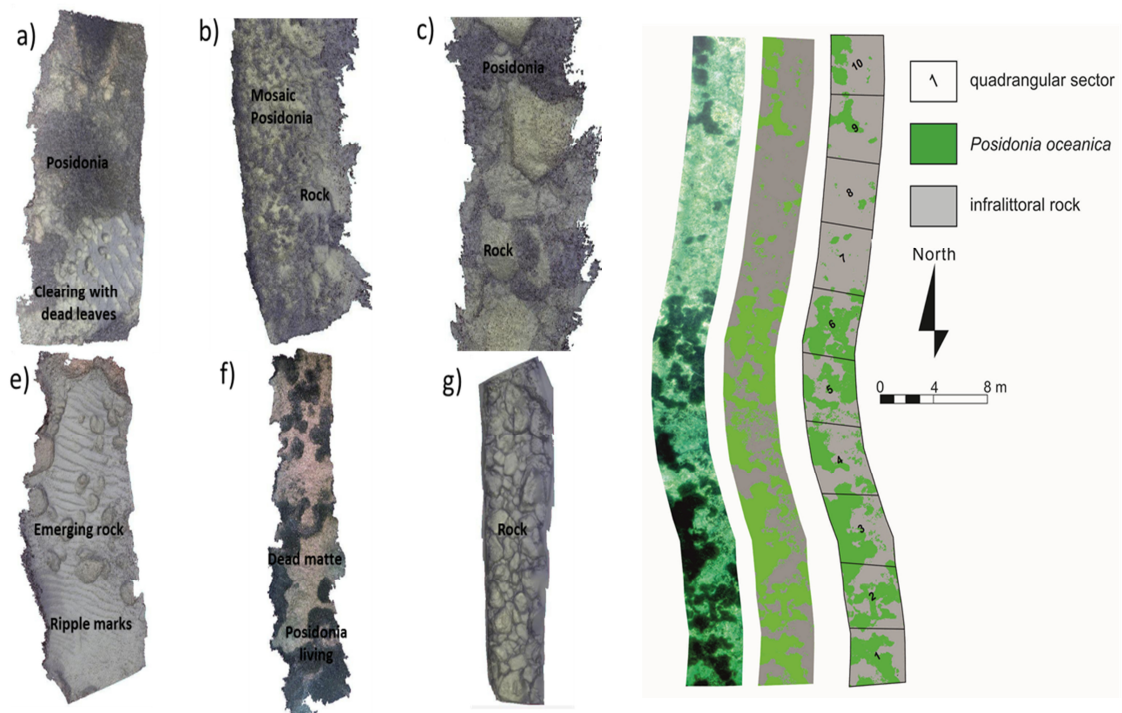


Fig. 4.12: immagini ortografiche modelli 3D di alcuni transetti della zona 2 del sito archeologico di capo Colonna (sinistra). Fotomosaico transetto 1 (destra).

Il fotomosaico del transetto 1 (Fig. 12, destra) è stato realizzato utilizzando il *software open source* PANGAEA, strumento per la creazione di fotomosaici specificatamente progettato per applicazioni subacquee. Dato un set d'immagini della scena acquisita con una sovrapposizione del 60% circa, l'applicazione crea un fotomosaico ad alta risoluzione combinando le varie immagini in maniera tale da non far percepire all'utente la giunzione tra le stesse. La fase di caricamento e di allineamento delle foto è del tutto automatica. Le proiezioni applicabili alle immagini sono presentate in termini di "movimenti della camera" (*rotating camera* e *planar camera*) mentre tutta una serie di strumenti sono presenti al fine di correggere le foto caricate (esposizione, *vignetting*, bilanciamento del bianco).

4.2.2 Acquisizione ed elaborazione dati zona 1: "relitto Punta Scifo D".

Il relitto Punta Scifo D è costituito da un carico lapideo di enormi dimensioni (stimato in recenti studi di circa 350 tonnellate) che si estende per un'area di 22 m x 16 m, su di un modesto spessore sabbioso. Il carico è composto da 54 manufatti, tra blocchi e lastre di marmo di dimensioni variabili, molti dei quali di grande pezzatura, anche fino a 4 metri di lunghezza. Molti di essi si trovano comunque al di sotto del fondale sabbioso e non visibili in superficie.

La tipologia di sito archeologico in esame, costituito da manufatti di grande dimensioni con pochi dettagli costruttivi, ci ha consentito di effettuare una sperimentazione in campo sulla possibilità di poter impiegare nella fase di rilievo ottico camere GoPro in modalità video. Tale soluzione tecnologica risulterà di grande impatto nell'applicazione del protocollo operativo implementato, in quanto ci consentirà di abbattere notevolmente i tempi di acquisizione e dunque i tempi di fondo degli operatori subacquei impiegati. Al fine di poter effettuare una prova comparativa sulla qualità del data set ottico ottenuto, il sito è stato acquisito contemporaneamente utilizzando una tipica camera DSLR.

L'acquisizione è stata condotta con l'ausilio di due operatori subacquei secondo un layout standard di acquisizione aerea nuotando ad una altezza media di 2.5 m rispetto al fondale, acquisendo transetti consecutivi con una sovrapposizione di circa 50% -

60% (Fig. 4.13). Ulteriori rilievi sono stati successivamente realizzati in posizione obliqua, in modo da acquisire le aree non visibili dall'alto ed ulteriori dettagli della scena.

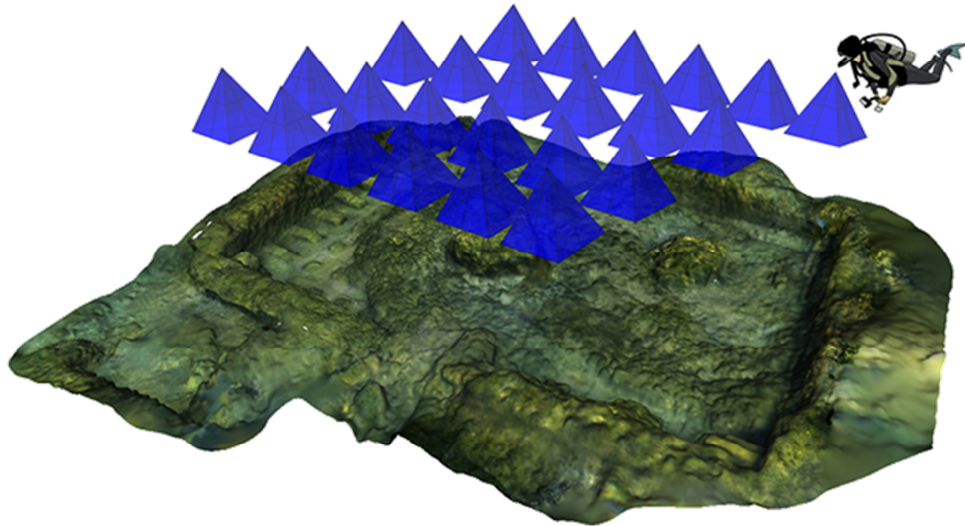


Fig. 4.13: layout acquisizione immagini e video.

Il primo sub ha eseguito l'acquisizione utilizzando una Canon PowerShot S100, lunghezza focale 5mm f/3.2, ISO 8, shutter 1/125 s, a risoluzione di 12MP (4000 x 3000 pix), alloggiata all'interno di un case subacqueo con una porta di tipo sferico. Il data set così acquisito è stato di 850 immagini.

Il secondo sub ha eseguito il rilievo utilizzando una GoPro hero 4 silver edition in modalità video, disabilitando tutti gli algoritmi di *image enhancement* presenti al suo interno attraverso l'attivazione della funzione ProTune. Sebbene la camera sia in grado di acquisire video a risoluzione 4K (4000 x 3000 pix) a 7.5 fps, è stato verificato, in sperimentazioni precedenti, che a causa del basso *frame rate* le immagini estratte dalla sequenza video presentano una non perfetta messa a fuoco, problema che incide notevolmente sulla qualità della ricostruzione 3D finale. Per ovviare a tale inconveniente si è deciso allora di acquisire in modalità fullHD (1920 x 1080 pix) aumentando al contempo il *frame rate* a 47 fps. Il video acquisito è stato successivamente processato al fine di poter estrarre immagini a tale risoluzione (1 frame su 10). Il data set costruito è composto da 900 immagini.

Al fine di valutare in maniera comparativa la qualità delle immagini ottenute dalla camera GoPro rispetto a quelle ottenute con la Canon S100 due immagini campione

della stessa porzione di sito (Fig. 4.14) sono state elaborate applicando due diverse metodologie di *image enhancement* sviluppate presso il DIMEG dell'Unical: la prima è una tecnica basata sull'*Histogram Stretching* e su una procedura di ritocco manuale del colore eseguita su un campione di immagini, mentre l'altra è la tecnica descritta in [92].



Fig. 4.14: immagini di alcuni blocchi di marmo del relitto Punta Scifo D: GoPro hero 4 (sinistra), Canon PowerShot S100 (destra).

In Fig. 4.14 sono mostrate le immagini ottenute a valle dell'applicazione della metodologia 1 basata sulla tecnica dell'*Histogram Stretching*. In Fig. 15 quelle ottenute applicando la metodologia 2.

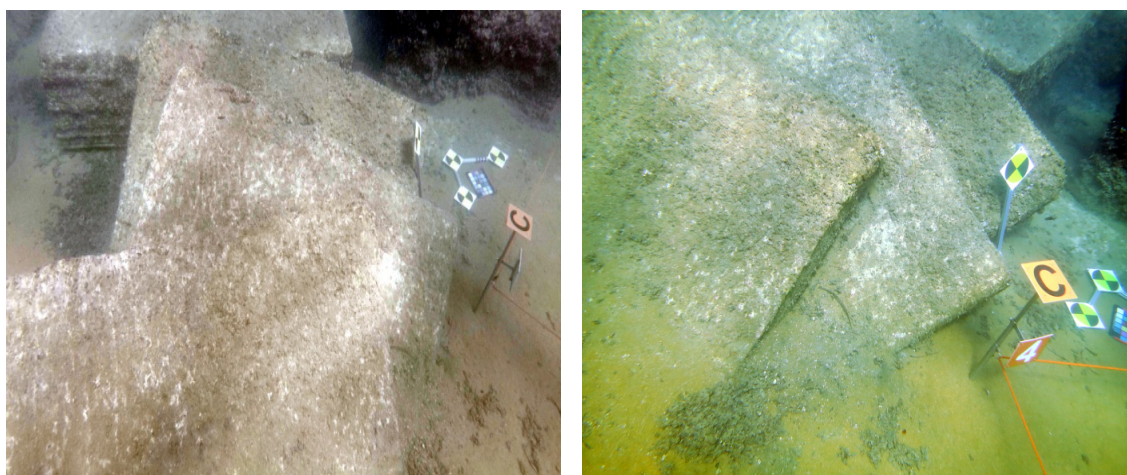


Fig. 4.15: immagini dopo applicazione metodologia 1. GoPro hero 4 (sinistra), Canon PowerShot S100 (destra).

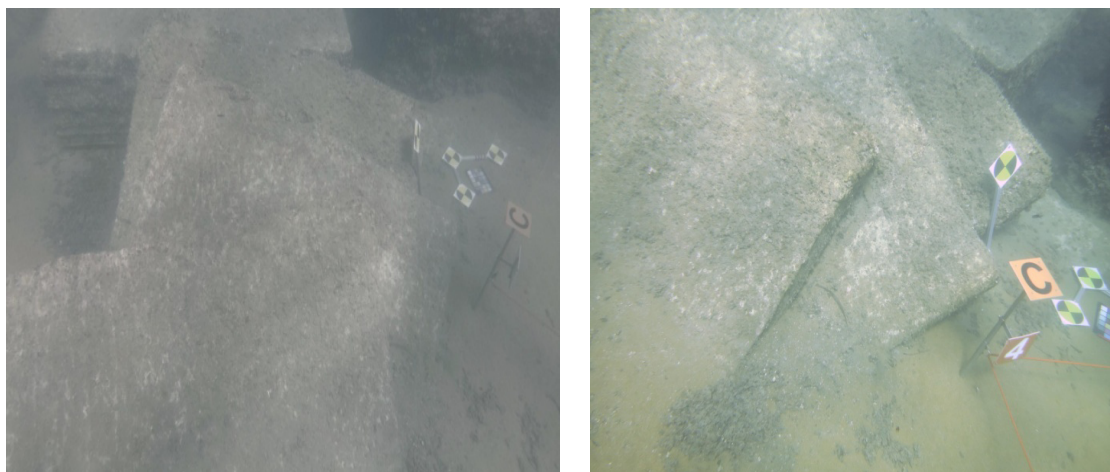


Fig.4.16: immagini dopo applicazione metodologia 2. GoPro hero 4 (sinistra), Canon PowerShot S100 (destra).

Da un'analisi visiva delle immagini ottenute è possibile verificare come, almeno nelle condizioni operative in cui i rilievi sono stati condotti, le immagini estratte dalla sequenza video della camera GoPro non presentano una dominanza di colore verde – blu tipica delle acquisizioni subacquee. Questa è invece ben marcata nell'immagine acquisita tramite la Canon S100. Inoltre, entrambi gli algoritmi di *image enhancement* proposti consentono di ottenere una buona correzione colore che comunque risultano più naturali applicando la metodologia 2.

Il data set d'immagini acquisito tramite la camera GoPro è stato successivamente elaborato al fine di ottenere il modello 3D *texturizzato* ad alta definizione del relitto Punta Scifo D.

Allo scopo è stata utilizzata la medesima tecnica MVS presentata in precedenza utilizzando, però in questo caso il *software* commerciale *Photoscan*[®]. I risultati di tale elaborazione sono mostrati nelle Fig. 4.17, 4.18. Infine, in Fig. 4.19 sono mostrate 4 ortofoto ottenute dal modello 3D applicando differenti effetti cromatici alla texture ad alta definizione (4096 x 4096) ivi mappata.

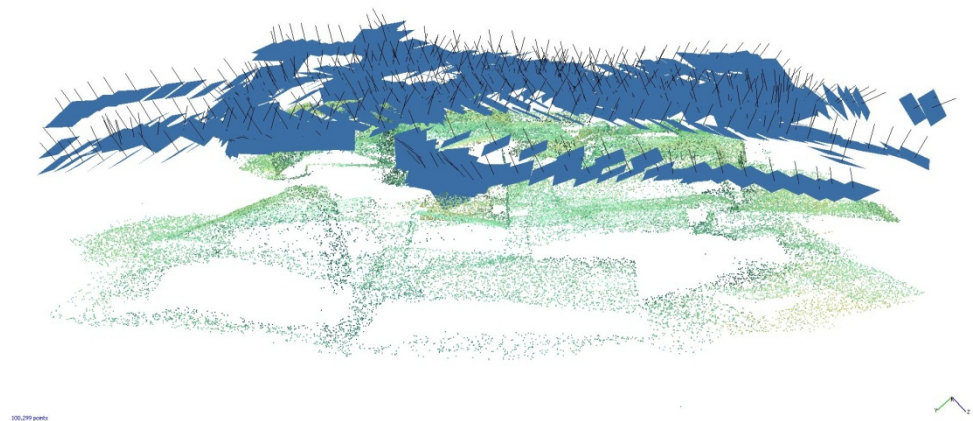


Fig. 4. 17: nuvola di punti 3D sparsa relitto Punta Scifo D (100K punti).

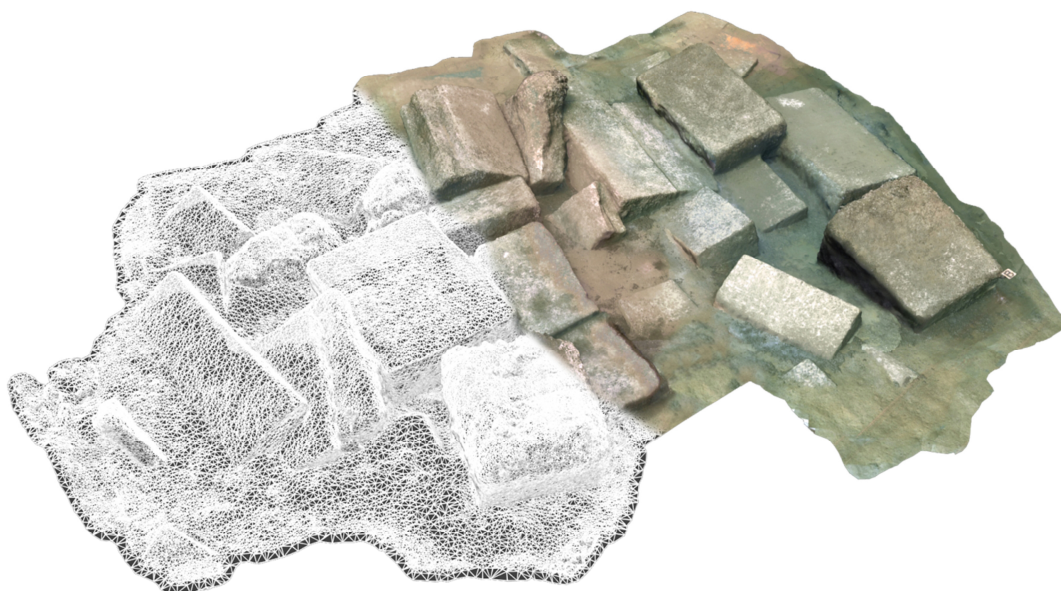


Fig. 4.18: modello 3D texturizzato relitto Punto Scifo D (32 milioni di geometrie).





Fig. 4.19: ortofoto modello 3D relitto Punta Scifo D con diverse equalizzazioni degli effetti cromatici sulla texture finale.

4.2.3 Acquisizione ed elaborazione dati acustici

La ricerca di oggetti di interesse archeologico adagiati al fondo mare o parzialmente sepolti viene attuata utilizzando strumentazioni acustiche (ecoscandagli multifascio e sidescan sonar). È ben noto l'utilizzo di tecniche di indagine morfobatimetriche (side-scan sonar e multibeam) per il ritrovamento di target emersi o parzialmente emersi rispetto al fondo, in quanto le rugosità intrinseche dei materiali, e di conseguenza il contrasto di backscatter (side-scan sonar), oppure un'accurata ricostruzione del DTM (Digital Terrain Model) del fondale che esalti le anomalie morfologiche dovute all'emersione degli oggetti stessi, sono segni di distinzione ben marcati e caratterizzanti. Di seguito vengono riportate le specifiche della strumentazione che verrà utilizzata per l'esecuzione delle attività in mare (in merito alla parte acustica) e le fasi di elaborazione del dato finalizzate ad ottenere la migliore rappresentazione acustica del fondale e, di conseguenza, la migliore sovrapposizione e rappresentazione ottico - acustica possibile sul sito di interesse scelto come caso - studio. Tali sito di interesse archeologico, Capo Colonna (Crotone) è situato in acque basse (*onshore*, ovvero nei primi 30 m di profondità).

In generale, le indagini geofisiche ad alta risoluzione in ambienti di acqua bassa sono un obiettivo di difficile realizzazione dal punto di vista tecnologico, sia per la limitata operatività dei natanti, sia per le difficoltà dovute alle configurazioni strumentali. Per ottimizzare la scelta dello strumento più opportuno bisogna tenere presenti le caratteristiche del DTM da generare, che richiede un passo di cella dell'ordine dei 5-10 cm allo scopo di ottimizzare il *match* con i dati ottici. Ne deriva la necessità di utilizzare

un sistema con sorgente con una frequenza di emissione di almeno 100 KHz, frequenza che riesce a minimizzare i possibili errori sulla verticale, e con almeno 200 *beam* su un'apertura complessiva minima del fascio di 150°. Per l'acquisizione su entrambi i siti è stato scelto, quindi, il sistema Seabat 7125 (Reson[®]Inc, Fig. 4.20), che risponde perfettamente ai requisiti individuati per le caratteristiche di sistema necessarie. Tale sistema, infatti, è:

- Piccolo e leggero, e quindi facilmente installabile su imbarcazioni a basso pescaggio e di facile manovrabilità;
- ha una sorgente (modulabile) ad elevata frequenza (200-400 KHz);
- elevatissimo *ping-rate* (ovvero, numero di insonificazioni al secondo);
- elevatissimo numero di *beam* (Fino a 512 nella configurazione a 400 KHz);
- elevatissima risoluzione potenziale (fino a 6 mm);
- consente l'acquisizione del *backscatter*, ovvero della riflettività del fondo mare dovuta alle caratteristiche di rugosità intrinseche dei materiali, analogamente a quanto avviene con le strumentazioni sidescan sonar.

SEABAT 7125 SYSTEM SPECIFICATIONS

	7125 SV2	7125 ROV2	7125 AUV
Power requirement	Typical: 110-220 VAC, 50/60 Hz, 250 W. Max: 110-220 VAC, 50/60 Hz, 700 W.	Processor Typical: 110-220 VAC, 50/60 Hz, 110 W. Processor Max: 110-220 VAC, 50/60 Hz, 400 W. Wet end Typical: 48 VDC (+/-10%), 115 W. Wet end Max: 48 VDC (+/-10%) 250 W. Power requirements when Wet-ends are powered from sonar processor: 110-220 VAC, 50/60 Hz, 700 W.	48V DC (± 10%)
Transducer cable length	25m standard	3m standard 10m optional	3m standard 10m optional
LCU to processor cable length	N/A	25m (st), 3 m	N/A
System depth rating	25m	6000m	6000m optional
Frequency	200kHz or 400kHz (dual frequency available)		
Along-track transmit beamwidth	2° at 200kHz & 1° at 400kHz		
Across-track receive beamwidth	1° at 200kHz & 0.5° at 400kHz		
Max ping rate	50Hz (±1Hz)		
Pulse length	30µs – 300µs Continuous Wave; 300µs – 20ms Frequency Modulated (X-Range)		
Number of beams	512EA/ED at 400kHz, 256EA/ED at 200kHz		
Max swath angle	140° in Equi-Distant Mode; 165° in Equi-Angle Mode		
Typical depth ¹⁾	0.5m to 150m at 400kHz, 0.5m to 400m at 200kHz		
Max depth ²⁾	>175m at 400kHz; 450m at 200kHz		
Depth resolution	6mm		
Data output	Bathmetry, sidescan and snippets 7K data format		
Temperature:	-2° to +35°C		
Flexmode:	Optional		
Full Rate Dual Head	400 KHz for ROV/ AUV		

Fig. 4.20: specifiche tecniche del Reson Seabat 7125.

A proposito dell'opzione di *backscatter*, tale parametro (opportunamente processato e cartografato) viene utilizzato per la determinazione delle caratteristiche morfoacustiche dei fondali: esso rappresenta un'applicazione differenziata del tradizionale principio dell'ecoscandaglio utilizzato per i sondaggi acustici batimetrici verticali, che viene adattato all'impiego per i sondaggi acustici con scansione areale. Ogni impulso acustico emesso dal trasduttore genera una spazzata che parte e ritorna come eco di riflessione per poi essere tradotta in immagine, nella quale ogni pixel assume una gradazione in toni di grigio proporzionale all'ampiezza del segnale di *backscatter* di ritorno. L'immagine morfoacustica del fondale (sonogramma Fig. 4.21) esprime le variazioni di intensità del *backscatter*, in termini di toni di grigio, associate a variazioni morfologiche, della tessitura/composizione dei sedimenti e alla presenza di affioramenti rocciosi.

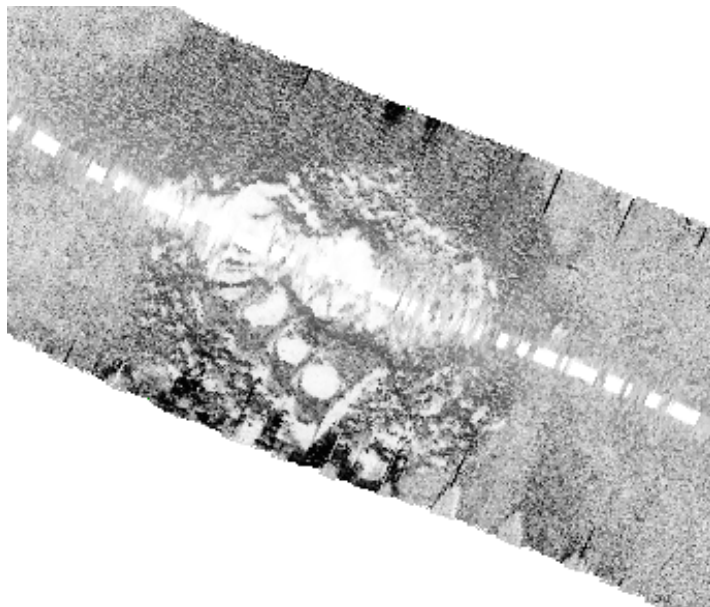


Fig.4. 21: esempio di acquisizione di un sonogramma con misure di backscatter.

Il sistema emette un ventaglio di onde acustiche orientato trasversalmente alla direzione di navigazione e registra le onde provenienti dal fondale (*backscatter*); in questo caso le onde ricevute vengono generate sul fondale da fenomeni di diffrazione che si manifestano quando la lunghezza d'onda del segnale incidente è confrontabile con le discontinuità del fondale (rugosità). Le elevate frequenze che caratterizzano il sistema assicurano l'assenza di penetrazione e l'elevata risoluzione dei dati. Nelle fasi

di acquisizione si proverà, quindi, ad acquisire e cartografare il *backscatter* fondamentalmente per due motivi:

- Costituisce un'immagine "sonica" del fondale e, in tal senso, faciliterà la sovrapposizione con le immagini acquisite per la parte ottica;
- fornirà ulteriori elementi di conoscenza sui fondali da investigare, da integrare potenzialmente nella ricostruzione virtuale dei manufatti presenti nei siti da acquisire (composizione del fondale, eventuali coperture degli stessi, etc.).

Tali caratteristiche (Fig. 4.20) rendono il Seabat 7125 una strumentazione assolutamente adatta agli scopi in termini di risoluzione in acque basse. I dati sono elaborati in base alle normative indicate dell'Organizzazione Idrografica Internazionale, utilizzando strumenti per la rimozione del rumore relativi a misurazioni errate, correzione di assetto sui tre assi nello spazio del vettore, salti di navigazione e profilo di velocità del suono. Questi consistono in una "nuvola di punti" di misure di profondità, che devono essere trasformate in una matrice regolare, di più facile gestione, chiamata Digital Elevation Model (DEM) o Digital Terrain Model (DTM).

4.2.3.1 Elaborazione dei dati multibeam

La naturale conseguenza della scelta del sistema multibeam Seabat 7125 è l'utilizzo del pacchetto software PDS2000 per l'elaborazione dei dati. Il PDS2000 (sviluppato dalla RESON BV) è un *software* per operazioni di acquisizione, elaborazione e restituzione cartografica (2D e 2.5 D) ampiamente utilizzato sia in ambito industriale che di ricerca. Il pacchetto completo permette di:

- gestire, editare e correggere la navigazione e l'acquisizione dei dati, Multibeam;
- editare ed elaborare le spazzate singolarmente (swath editor);
- editare ed elaborare il DTM;
- produrre la cartografia o interfacciarsi con altre piattaforme software.

E' inoltre suddiviso in moduli che permettono di effettuare tutte le fasi del rilievo ed elaborazione, ovvero:

- progettazione ed esecuzione dei percorsi specifici di acquisizione, con rilevazione quantitativa dell'errore rispetto a quanto previsto in fase di progetto;
- controllo qualitativo e quantitativo dei *raw data* strumentali sia in fase di acquisizione che nelle successive elaborazioni;
- correzione di marea;
- editing della posizione;
- modulo 3D e cartografico.

Durante la navigazione, sulle linee teoriche impostate, il *software* è in grado di controllare tutti i dati provenienti dai sensori connessi al sistema, di gestire le connessioni in *input/output* con tutti gli strumenti interfacciati; tutte le informazioni acquisite possono essere visualizzate in formato numerico o grafico, tramite la generazione in tempo reale di rappresentazioni planimetriche. Le fasi finali per la restituzione cartografica dei siti da acquisire saranno:

- creazione e pulizia DTM;
- estrazione dei file ascii (xyz) processati, da utilizzare come formato di interscambio con le piattaforme software dove si eseguirà la fusione ottico - acustica dei dati.

4.2.3.2 “Footprint” e Risoluzione orizzontale e verticale dei record multibeam.

Come conseguenza diretta della risoluzione estrema richiesta, la fase di calibrazione del sistema *multibeam* può essere considerata uno degli aspetti critici. La calibrazione del sistema viene eseguita dal *software* di acquisizione mediante un modulo per il calcolo dei parametri di compensazione necessari al corretto funzionamento del *multibeam*. Dal punto di vista operativo, viene eseguita effettuando delle linee di acquisizione specifiche con un confronto ricorsivo di tipo prova - errore sui profili acquisiti che consentiranno di valutare i corretti parametri iniziali di “yaw”, “pitch”, “roll” ed “heave” dell'imbarcazione, e di correggere il posizionamento georeferenziato di ogni singolo beam. Sono previste:

- 1) La misura degli offset della strumentazione in relazione allo scafo (offset orizzontali e verticali dell'antenna GPS rispetto ad un sistema di coordinate scelto, etc.);
- 2) Una calibrazione della strumentazione e del sistema di acquisizione per roll, pitch, yaw, tempo di ritardo e velocità del suono, operata svolgendo specifici percorsi di navigazione.

L'area dell'elemento di fondale insonificata da ogni singolo beam dipende dalla profondità del fondo e dall'ampiezza acustica dei fasci e, quindi, dalla geometria del trasduttore. Nel caso di un trasduttore piatto, l'ampiezza acustica trasversale di ogni beam non è costante, ma aumenta con l'aumentare dell'angolo di inclinazione rispetto alla verticale. Tale valore, che viene detto "*footprint*", può essere calcolato con le seguenti formule:

- footprint angolare (Nadir) = $\text{tg}(\text{numero di beam}/\text{ampiezza del fascio})$;
- footprint areale (Nadir) = Profondità* footprint angolare.

La misura di profondità acquisita da un beam, quindi, rispecchia la media dell'area insonificata, per cui non può essere ritenuta strettamente puntuale. Il footprint, sostanzialmente, rappresenta per questo motivo una misura della potenziale risoluzione strumentale orizzontale ad una determinata profondità. La risoluzione verticale viene migliorata applicando un Profilo di velocità del suono (SVP) lungo la colonna d'acqua, registrato in tempo reale ed applicato ogni 6-8 ore durante l'intera durata dell'acquisizione. Tali misure saranno effettuate utilizzando una sonda MiniSVP. Una volta corretto il profilo tramite misure con MiniSVP, la risoluzione verticale dipenderà principalmente dalla frequenza strumentale (nel nostro caso, nella configurazione a 400 KHz, l'errore potenziale si riduce a 6 mm), mentre la risoluzione orizzontale dipenderà principalmente dal *footprint*. Sia la risoluzione orizzontale che verticale sono funzione della precisione del sistema di posizionamento, che può essere valutata mediante l'uso dei parametri di Diluizione Di Precisione Verticale (VDOP) o Diluizione di Precisione Orizzontale (HDOP). Questi parametri permettono la stima quantitativa dell'errore commesso dal ricevitore GPS durante l'indagine.

Il posizionamento dei natanti verrà garantito da un sistema DGPS Skyfix con margine di errore inferiore ai 50 cm. Naturalmente, anche le acquisizioni multibeam che

potenzialmente possono produrre un DTM ad altissima risoluzione sono vincolate alla scelta del sistema di posizionamento, in particolare se la HDOP è superiore al footprint in termini di dimensione areale. Nel caso delle acquisizioni da effettuare, in particolare, la presenza di forme significative di alterazione della morfologia del fondale (gavitelli di perimetrazione, spigoli dei manufatti da acquisire, etc.) consentono, grazie agli strumenti del PDS2000, la ri-collocazione delle spazzate finalizzata al *best fit* delle misure. Tale tecnica che ci consentirà di ridurre ulteriormente l'errore sulla precisione orizzontale introdotto dal rilevamento satellitare, ottenendo un DTM a risoluzione maggiore.

4.2.3.3 Elaborazione (processing) dei dati.

Lo scopo finale dell'elaborazione dei dati multibeam e quello di raggiungere la più onesta e quanto più accurata possibile rappresentazione del fondale marino sotto forma di modello digitale del terreno (DTM). La dimensione della cella elementare che caratterizza il DTM finale deve necessariamente essere coerente con i limiti dell'accuratezza, spaziale e verticale, imposti dalle caratteristiche strumentali, nonché dai parametri di acquisizione [8]. Come già definito in precedenza, la risoluzione verticale dipende solo dalla frequenza della strumentazione e dalla velocità del suono nella colonna d'acqua, mentre quella orizzontale dipende dalle caratteristiche strumentali e dalla profondità. Come già definito in precedenza, la risoluzione verticale dipende solo dalla frequenza della strumentazione e dalla velocità del suono nella colonna d'acqua, mentre quella orizzontale dipende dalle caratteristiche strumentali e dalla profondità. Tipiche fonti di errore possono essere [9]:

- 1) Un'errata calibrazione strumentale;
- 2) Un errore nel posizionamento dell'imbarcazione (salti di GPS o salti della correzione differenziale);
- 3) La mancata (oppure incompleta) correzione dell'assetto dell'imbarcazione (Pitch, roll, yaw, heave);
- 4) Problemi legati alla marea ed al profilo di velocità del suono;
- 5) Problemi connessi all'insonificazione in senso stretto.

Di seguito (Tab. 2) viene riportato uno schema riassuntivo del tipo di errore, delle possibili cause e delle soluzioni che possono essere applicate per rimuovere o ridurre gli effetti:

Tipo di errore	Effetti sull'acquisizione e sul DTM	Possibili accorgimenti
Errata calibrazione dei sensori (di moto e di offset)	Spazzate "tiltate" rispetto alle reali inclinazioni	Ripetizione della calibrazione ed applicazione dei nuovi parametri a tutti i records
Errato posizionamento	"Salti" nella ricezione, bassa qualità o salti nella correzione differenziale del segnale GPS	Editing della navigazione con software specifici
mancata o incompleta correzione dell'assetto dell'imbarcazione (Pitch, roll, yaw, heave)	Distorsione della contiguità delle spazzate. Oscillazioni perpendicolari al "vessel track" (percorso seguito dalla nave)	Aumento delle dimensioni di cella del Grid oltre quelli suggeriti dal footprint
Insonificazione acustica (bassa energia di ritorno)	Dislivelli fra spazzate contigue	Rimozione delle misure errate ("spikes") con l'utilizzo di softwares specifici (Ad es., PDS2000 Swath-model editor)
Errori connessi al profilo di velocità del suono nella colonna d'acqua (CTD)	Problemi di "spikes" in particolare nella parte distale della spazzata Possibili forme curve nelle spazzate (concave o convesse)	Plyback dei dati ed applicazione di nuove CTD
Marea	Dislivelli verticali, in particolare su spazzate vicine in senso spaziale ma acquisite in momenti differenti	Correzione del disturbo mareale effettuato con l'utilizzo di softwares specifici (Ad es., PDS2000 tide editor)
Spikes random occasionali (Ostacoli sul "beam pattern", cattiva qualità)	Falsi picchi sul DTM	Eliminazione effettuata con l'utilizzo di softwares specifici (Ad es., PDS2000 Swath-model editor)

generale del beam in acquisizione)		
------------------------------------	--	--

Tab. 2: tipici errori (e possibili cause) delle acquisizioni con sistemi multibeam.

Classi	“Special”	1°Ordine	2°Ordine	3°Ordine
Esempi di aree tipiche	Porti, zone di ormeggio e canali tipici associati	Porti, percorsi raccomandati ed alcune aree costiere con profondità maggiori di 100	Aree non implementate nelle classi “Special” e “1° Ordine”, oppure aree con profondità maggiore di 200 m	Aree non descritte in “Special”, “1° Ordine” e “2° Ordine”
Precisione in orizzontale	2m	5m + 5% della profondità	20m + 5% della profondità	150m + della profondità
Precisione in verticale	a=0.25 m b=0.0075	a=0.5m b=0.013	a=1.0m b=0.023	Come nel “2°Ordine”
100% di copertura del fondale	Obbligatoria	Richiesto in determinate aree	Potrebbe essere richiesta in determinate aree	Non richiesta
Capacità di detezione del sistema	Oggetti cubici > 1 m	Oggetti cubici > 1 m a profondità minori di 40 m; 10% della profondità oltre i 40m	Come nel “1°Ordine”	Non richiesta
Massima spaziatura fra le linee	Non definita	3 x la profondità media (massima 25m)	3-4 x profondità media (massima 200m)	4 x profondità media

Tab. 3: Classificazione dei rilievi secondo gli standard IHO

La classificazione della qualità del rilievo può essere valutata ancora rispetto agli standard IHO [6], come riportato in Tab. 3.

Contrariamente al rilievo terrestre, la compilazione del DTM marino normalmente richiede tempi più lunghi sia per quanto concerne l’acquisizione a bordo sia per la fase di processing successivo. Tale elaborazione è realizzata sostanzialmente in tre fasi principali (tutte in ambiente software PDS2000):

- 1) I dati acquisiti verranno corretti, per quanto attiene il Datum verticale, rispetto al livello medio delle basse maree sigiziali (Fig. 4. 24);

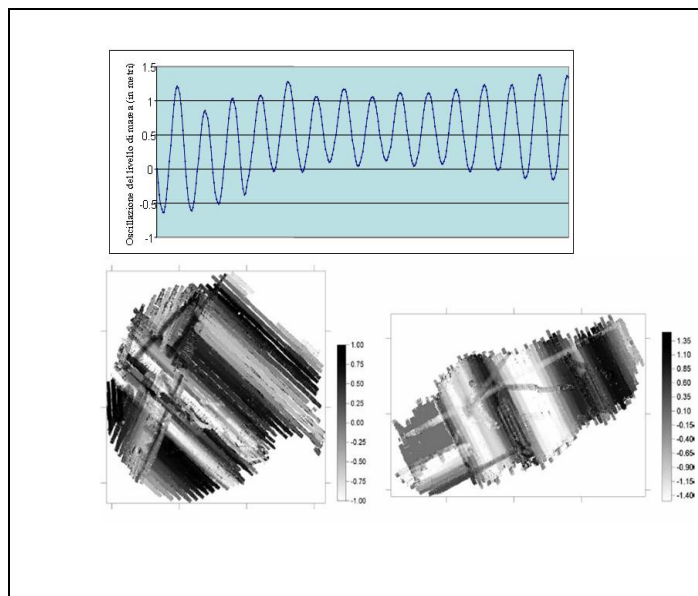


Fig. 4. 24: esempio di correzione del disturbo mareale su misure batimetriche.

- 2) I dati di tutte le campagne previste dalle attività del progetto verranno acquisiti con sistema di posizionamento satellitare differenziale DGPS. La correzione delle linee di navigazione, qualora necessaria, sarà effettuata tramite *"Position editor view"*, che consente di riposizionare automaticamente le spazzate una volta ricollocati i singoli punti di navigazione;
- 3) infine, la rimozione degli spikes locali e laterali (Fig.4.25), che consiste nella rimozione delle misure errate grazie ad una fase di editing specifica, evitando quindi l'utilizzo di filtri di tipo statistico in accordo con gli standard internazionali IHO.

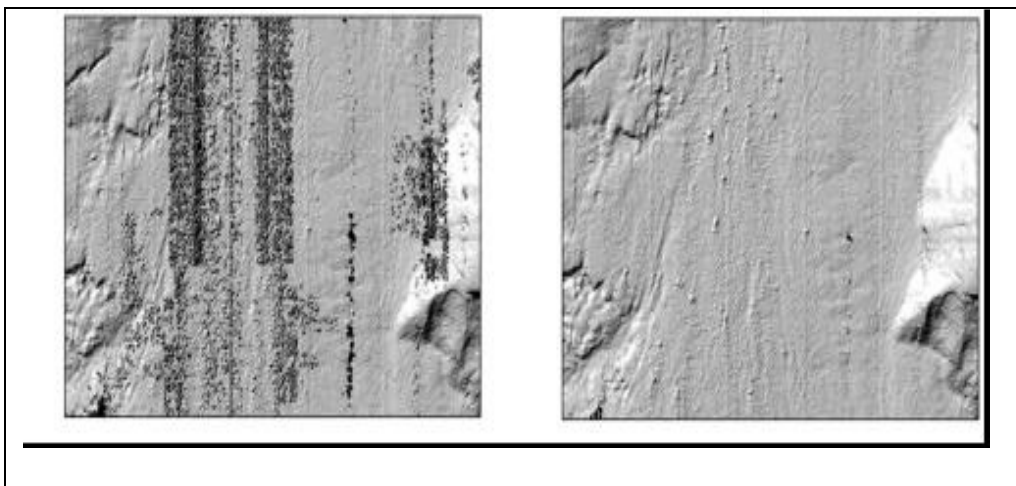


Fig. 4. 25: esempio di elaborazione di un DTM derivato da dati batimetrici: a sinistra il DTM acquisito, a destra dopo l'elaborazione in ambiente software PDS2000.

4.2.4 Data fusion ottico – acustico

Il problema principale da affrontare nella fusione dei dati acquisiti dai sensori ottici ed acustici è che essi operano a risoluzioni altamente differenti. La fusione di modelli con una risoluzione spaziale così differente pone serie problematiche nell'applicazione delle tecniche e degli algoritmi di registrazione automatici, come l'*Iterative Closest Point* (ICP). Un ulteriore problema riguarda la gestione della mole di dati acquisiti. Come già ampiamente discusso, in linea di principio, qualsiasi sito archeologico e il suo ambiente circostante può essere riprodotto con il massimo livello di dettaglio e accuratezza consentita dalle tecnologie allo state dell'arte presenti attualmente in letteratura scientifica, sia nel campo ottico che in quello acustico. Il modello 3D ottico – acustico generato deve, comunque, avere una dimensione tale per cui la sua visualizzazione e navigazione, avvenga in maniera fluida ed accattivante. Considerando le elevate risorse hardware richieste per la resa fotorealistica dell'ambiente subacqueo attraverso gli shader e gli effetti grafici una prima sperimentazione ha messo in evidenza come tali modelli non dovrebbero eccedere la dimensione di 200 MB se fatti girare su pc di fascia medio – alta (PC con processore I7 e scheda grafica Nvidia della serie GTX). Questo vincolo sulla dimensione massima del modello 3D ottico – acustico generato pone un limite al livello di dettaglio che è possibile ottenere dell'intero sito archeologico e impone di implementare nella metodologia proposta un'efficiente politica di gestione delle risorse disponibili basata sulla decisione di quale area

acquisire, con quale livello di dettaglio e con quale tecnica tra quelle possibili ottenere la fusione dei dati ottico - acustici.

Il primo problema affrontato nell'ambito dello studio delle tecniche di fusione ottico - acustiche è stato quello di riuscire ad ottenere una *mesh* poligonale a partire dal DTM generato dall'acquisizione acustica. In effetti, anche se la batimetria acustica può essere rappresentata come punti 3D nello spazio, la sua struttura dati è notevolmente differente da quella di una classica nuvola di punti 3D. Mentre per quest'ultime, ad ogni coppia (x,y) di coordinate, è possibile associare più coordinate z, nel caso di una batimetria acustica ottenuta con un sonar multibeam questo non è più vero. Il piano di equazione $z = 0$ è infatti suddiviso in una griglia di dimensione comparabile alla risoluzione dello strumento (10 x 10 cm nel nostro caso) a cui è possibile associare un solo valore z.

Gli algoritmi presenti in letteratura scientifica normalmente impiegati per la generazione di una mesh poligonale a partire da punti nello spazio 3D sono sostanzialmente due: l'algoritmo di Poisson e l'algoritmo di Delaunay.

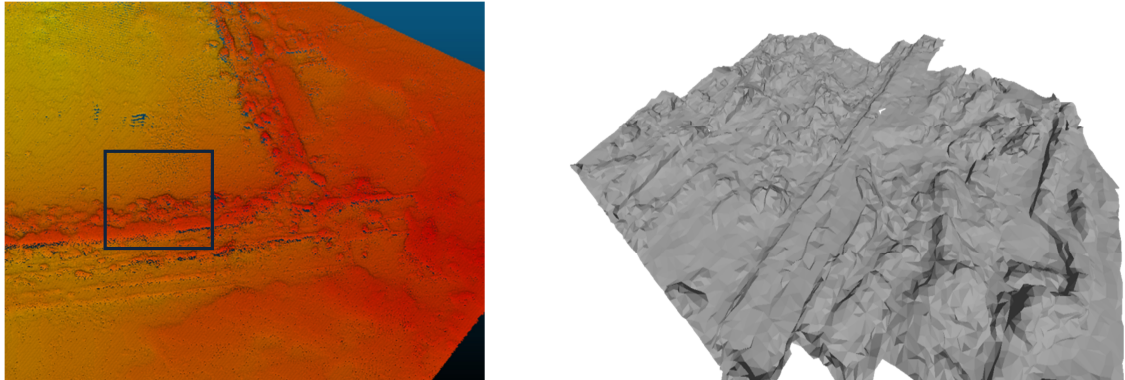


Fig. 4.26: Villa dei Pisoni (Baia): nuvola acustica 3D (sinistra), mesh generata con algoritmo Delaunay (destra).

Da una prima sperimentazione effettuata a partire dalla nuvola acustica 3D di una porzione del sito archeologico subacqueo di Baia (Villa dei Pisoni) è emerso che, a causa della bassa risoluzione spaziale dei dati forniti (10 cm) se la si compara a quella di una tipica nuvola 3D acquisita con tecniche ottiche (ordine del mm) e della tipologia di struttura che essi presentano, l'algoritmo di Poisson, nella sua implementazione sul software Meshlab, fallisce nella generazione della mesh poligonale. L'algoritmo di

Delaunay, implementato sul software JRC Reconstructor 3.0 fornisce invece, per questa tipologia di dati, dei risultati soddisfacenti, come è possibile osservare in Fig. 4.26.

Definita la tecnica attraverso la quale generare una mesh poligonale a partire da una batimetria acustica, i dati per la fusione ottico – acustica saranno entrambi rappresentati sotto la medesima forma. Il problema da affrontare diviene allora quello di poter registrare/allineare mesh poligonali di risoluzione, accuratezza e livello di dettaglio notevolmente differenti. Una prima soluzione è rappresentata da una classica tecnica di registrazione effettuata con l'ausilio dell'algoritmo ICP. Comunque, l'applicazione di tale algoritmo è limitata al caso in cui le due mesh presentano una geometria strutturata con una grande area in comune. Questa tecnica è stata applicata al modello 3D texturizzato di Villa dei Pisoni ottenendo una fusione dello stesso sulla mappa batimetrica dell'intero sito archeologico di Baia (NA) (Fig. 4.27).



Fig. 4.27: Villa dei Pisoni (Baia): registrazione del fotomosaico su mesh generata a partire dalla batimetria acustica.

Un obiettivo primario delle metodologie di data fusion che saranno utilizzate nell'ambito del servizio di Ricostruzione 3D è quello di poter efficientare le risorse hardware e software disponibili al fine di poter modellare con un elevato livello di dettaglio e accuratezza solo le aree di maggiore interesse, secondo un approccio di

tipo Level Of Detail (LOD). A tale scopo si è implementata e testata sui dati acquisiti durante la campagna di sperimentazione presso il sito di capo Colonna una metodologia per la modellazione 3D di fondali non strutturati (tipicamente sabbiosi) come lo è l'area circostante al relitto di Punta Scifo D.

Il primo passo della metodologia consiste nell'acquisizione e nell'elaborazione tramite algoritmi di *image enhancement* di una texture campione del fondale da modellare. Da questa, attraverso uno strumento presente all'interno del pacchetto software Photoshop[®], è stata ricavata una displacement map (Fig. 4.28).

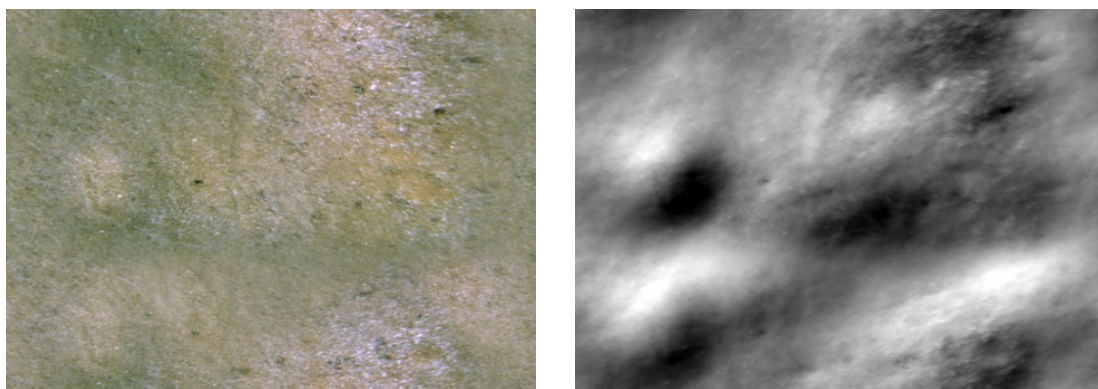


Fig. 4.28: texture campione del fondale che circonda il relitto Punta Scifo D (sinistra). Displacement map generata (destra).

Questa è un'immagine in scala di grigi utilizzata nel campo della modellazione 3D al fine di generare una variazione della geometria della mesh poligonale in funzione di una determinata codifica associata al valore del pixel considerato (0 – 255). Tale tecnica permette di ottenere, in maniera semplice ed automatica, mesh poligonali che modellano bene terreni con ondulazioni come lo è un fondale sabbioso.

Nella nostra sperimentazione questa tecnica è stata implementata utilizzando il software open source Blender. Successivamente, la texture campione è stata mappata sulla mesh poligonale rappresentante il fondale sabbioso attraverso la tecnica dell'UV mapping e ripetuta su tutta la superficie utilizzando una tecnica di tiling mapping (Fig .4.30)

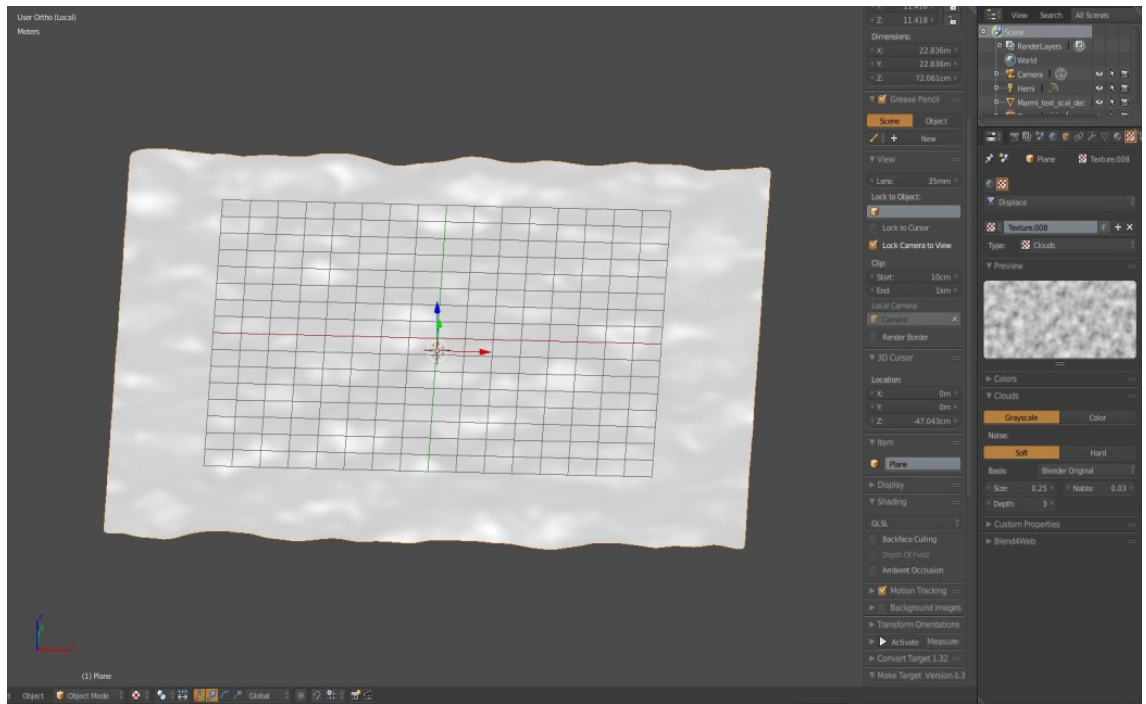


Fig. 4.29: interfaccia software open source Blender per generazione piano fondale.

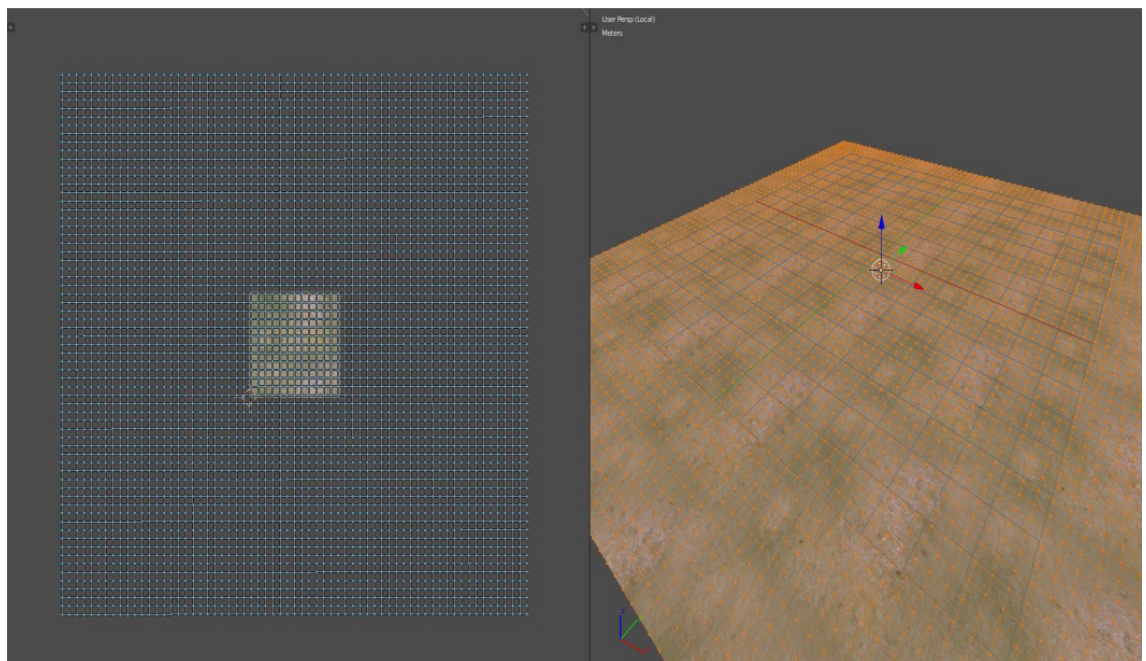


Fig. 4.30 : interfaccia software open source Blender per operazione di UV mapping e tiling mapping.

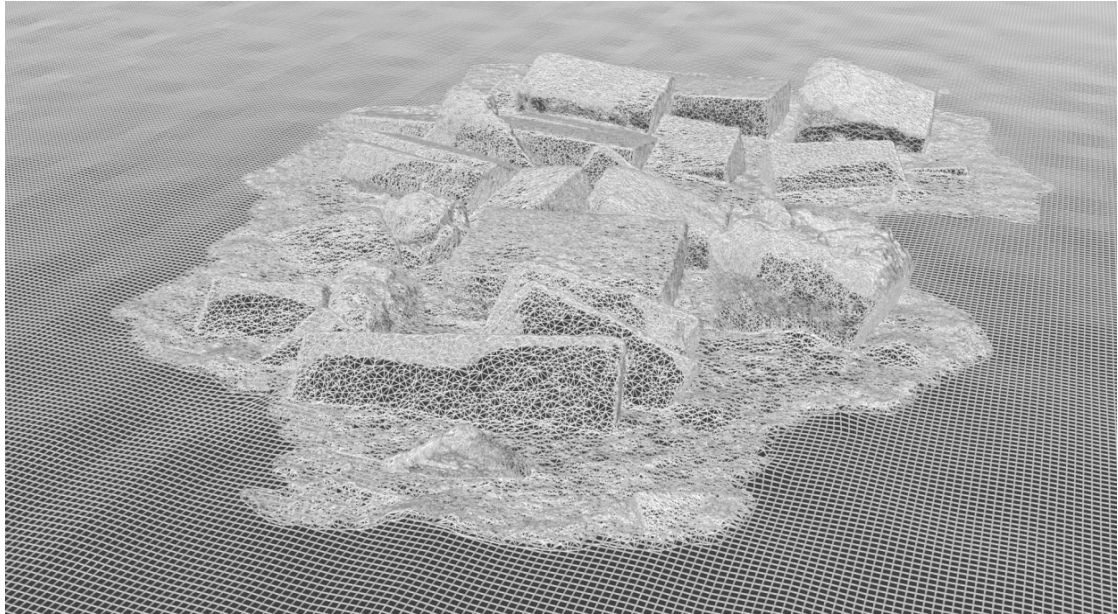


Fig. 4.31: modello 3D visualizzato in modalità wireframe del relitto punta scifo D con fondale circostante.

L'ultimo passo della tecnica proposta consiste in un allineamento manuale tra le 2 mesh poligonali. Questo è ottenuto selezionando i contorni dei due modelli ed effettuando una giunzione della mesh attraverso lo strumento *joint vertices* del software Blender. Il risultato finale è una mesh poligonale a differente risoluzione (Fig.4.31). Un'operazione finale di ritocco della texture alla giunzione è successivamente effettuata in modo tale che essa sia impercettibile per l'utente (Fig. 4.32).



Fig. 4.32 modello 3D texturizzato relitto Punta Scifo D con fondale circostante.

Conclusioni finali

In questa tesi sono state investigate diverse soluzioni per l'integrazione di tecniche ottiche e acustiche di *imaging* 3D in ambiente subacqueo. Sebbene un enorme quantità di lavoro è stato compiuto nel rendere più robusti gli algoritmi di ricostruzione 3D puramente visivi, la comunità robotica subacquea è alla continua ricerca di sistemi che integrano diverse tipologie di sensori al fine di ottenere sistemi più affidabili. È stato dimostrato che l'idea di combinare sensori ottici ed acustici è ragionevole, grazie alle proprietà complementari mostrate da entrambe le modalità d'acquisizione. In questo senso è stata eseguita un'analisi dello stato dell'arte attuale per quanto concerne le metodologie che combinano dati ottici ed acustici per la ricostruzione 3D di scene subacquee. È stato mostrato come gli approcci esistenti al *sensor fusion* per la ricostruzione 3D sono di tipo *mapped-oriented*, intendendo con tale dicitura che la fusione è eseguita a livello di *features*, generalmente attraverso corrispondenze geometriche o registrazione di dati. Pertanto, la calibrazione è un prerequisito fondamentale per questa tipologia di fusione poiché consente ai dati provenienti da un sensore di essere associati ai dati corrispondenti provenienti dagli altri sensori. Tuttavia, la procedura di calibrazione è altamente dipendente dai vincoli imposti dalla tipologia di sensori utilizzati (sensori coinvolti nella fusione, tipologia di dati forniti dagli stessi ecc.). Dunque, il primo passo nello sviluppo del prototipo di camera ottico acustica proposto è stato un'analisi dettagliata dei sensori che lo realizzano. La soluzione proposta consiste in un sottosistema ottico costituito da due camere in configurazione stereo, alloggiato all'interno di due scafandri opportunamente progettati e in un sottosistema acustico formato da una camera acustica 3D. Successivamente, è stata condotta un'analisi su differenti approcci alla calibrazione estrinseca del sistema progettato. A valle di tale processo è stato proposto un primo protocollo di calibrazione. A partire da un'acquisizione sincrona delle varie pose assunte da un pannello di calibrazione a scacchiera durante la fase iniziale della metodologia proposta, l'intero processo ha avuto come obiettivo quello di ottenere due nuvole di punti 3D, rispettivamente dal sottosistema stereo ottico e da quello acustico, opportunamente elaborate, dalle quali è stato possibile determinare successivamente i parametri estrinseci di calibrazione Φ e Δ attraverso un'operazione

di registrazione, realizzata applicando opportuni algoritmi d'allineamento alle nuvole di punti 3D stesse. L'intera *pipeline* di elaborazione si compone di alcuni tipici stadi di *image processing*, sebbene essi siano stati adattati e resi più robusti al fine di gestire in maniera più efficiente la particolare natura dei dati su cui vengono applicati, in particolar modo i dati rumorosi e a bassa risoluzione provenienti dal sottosistema acustico. Una preliminare sperimentazione in vasca presso i laboratori di Pozzuoli (NA) del partner di progetto WASS (Whitehead Sistemi Subacquei S.p.A.) è stata condotta al fine di validare il protocollo di calibrazione progettato. A partire dalle operazioni preliminari di settaggio dei sottosistemi stereo ottico ed acustico e la loro integrazione nella struttura di supporto connessa ad un palo telescopico di movimentazione, la prima parte dei test è stata interamente incentrata sull'acquisizione sincrona dei *target* progettati (sia del pannello di calibrazione che uno composto dai diversi oggetti di forma e materiale), al fine di costruire i *data set* d'immagini ottiche ed acustiche oggetto dell'elaborazioni delle successive fasi sperimentali. Ottenute le nuvole di punti 3D ottiche ed acustiche, si è proceduto infine all'esecuzione delle fasi finali previste nel protocollo di calibrazione progettato, cioè i moduli di *Coarse registration* e *Fine registration*, ottenendo in tal modo, dopo un'elaborazione statistica sui dati ottenuti, la matrice di trasformazione rigida cercata. I risultati ottenuti dai test effettuati risultano perfettamente in linea con i dati presenti in letteratura scientifica confermando in tal modo la bontà della soluzione proposta.

E' stato successivamente progettata e testata una metodologia per il rilievo, la modellazione 3D e la geo-localizzazione di reperti e siti archeologici sommersi impiegando tecniche ottiche e acustiche secondo un approccio di tipo *data level fusion*. Sono anche state effettuate alcune sperimentazioni presso il sito archeologico di Capo Colonna, "relitto Punta Scifo D", al fine di acquisire dati ed informazioni attraverso le quali validare la metodologia implementata. Il modello 3D ottico – acustico generato ha dimostrato la possibilità di poter usare queste tecniche al fine di poter realizzare modelli 3D multi –risoluzione dotati di texture ad alta definizione.

Riferimenti

- [1] Bythell, J., Pan, P., & Lee, J. (2001). Three-dimensional morphometric measurements of reef corals using underwater photogrammetry techniques. *Coral Reefs*, 20(3), 193-199.
- [2] Rende, F. S., Irving, A. D., Lagudi, A., Bruno, F., Scalise, S., Cappa, P., ... & Cicero, A. M. (2015). Pilot Application of 3d Underwater Imaging Techniques for Mapping Posidonia Oceanica (L.) Delile Meadows. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1, 177-181.
- [3] Bräger, S.; Chong, A.K. An application of close range photogrammetry in dolphin studies. *Photogramm. Record* 1999, 16, 503–517.
- [4] Gallo, A.; Angilica, A.; Bianco, G.; de Filippo, F.; Muzzupappa, M.; Davidde, B.; Bruno, F. 3D Reconstruction and Virtual Exploration of Submerged Structures: A Case Study in the Underwater Archaeological Site of Baia (Italy). In Proceedings of 13th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST 2012), Brighton, UK, 19–21 November 2012; pp. 121–128.
- [5] S. Duntley, "Light in the sea.", *Journal of the Optical Society of America*, vol. 53, no. 2, pp. 214-233, 1963.
- [6] M. Jenkin, B. Verzijlberg and A. Hogue, "Progress towards underwater 3D scene recovery.", in *Canadian Conference on Computer Science & Software Engineering*, 2010.
- [7] Gilbert, G. D., & Pernicka, J. C. (1966, June). Improvement of underwater visibility by reduction of backscatter with a circular polarization technique. In *Underwater Photo Optics I* (pp. 15-25). International Society for Optics and Photonics.
- [8] Detection of Interest Points in Turbid Underwater Images.", in *IEEE OCEANS Conference 2011*, Santander, Spain, 2011.
- [9] Singh, H., Roman, C., Whitcomb, L., & Yoerger, D. (2000). Advances in fusion of high resolution underwater optical and acoustic data. In *Underwater Technology, 2000. UT 00. Proceedings of the 2000 International Symposium on* (pp. 206-211). IEEE.
- [10] Fusiello, A., & Murino, V. (2004). Augmented scene modeling and visualization by optical and acoustic sensor integration. *Visualization and Computer Graphics, IEEE Transactions on*, 10(6), 625-636.
- [11] Singh, H., Roman, C., Pizarro, O., Eustice, R., & Can, A. (2007). Towards high-resolution imaging from underwater vehicles. *The International journal of robotics research*, 26(1), 55-74.
- [12] Singh, H., Salgian, G., Eustice, R., Mandelbaum, R., (2002). Sensor fusion of structure-from-motion, bathymetric 3D, and beacon-based navigation modalities. In *Robotics and Automation*, . Proceedings. ICRA'02. IEEE International Conference on, Vol. 4, pp. 4024-4031.
- [13] Murino, V., & Trucco, A. (2000). Three-dimensional image generation and processing in underwater acoustic vision. *Proceedings of the IEEE*, 88(12), 1903-1948.
- [14] Karaman, M., Kutay, M. A., & Bozdagi, G. (1995). An adaptive speckle suppression filter for medical ultrasonic imaging. *Medical Imaging, IEEE Transactions on*, 14(2), 283-292.

- [15] Hokland, J. H., & Tact, T. (1994). Ultrasound speckle reduction using harmonic oscillator models. *Ultrasonics, Ferroelectrics, and Frequency Control, IEEE Transactions on*, 41(2), 215-224.
- [16] Sound Metrics Corp, DIDSON Product Options, 2008.
- [17] Gross, M. H., Staadt, O. G., & Gatti, R. (1996). Efficient triangular surface approximations using wavelets and quadtree data structures. *Visualization and Computer Graphics, IEEE Transactions on*, 2(2), 130-143.
- [18] Canepa, G., Bergem, O., & Pace, N. G. (2003). A new algorithm for automatic processing of bathymetric data. *Oceanic Engineering, IEEE Journal of*, 28(1), 62-77.
- [19] Canepa, G., & Bergem, O. (1998, September). Testing a new algorithm for automatic processing of bathymetric data. In *OCEANS'98 Conference Proceedings* (Vol. 1, pp. 313-317). IEEE.
- [20] Lurton, X. (1998, September). Precision analysis of bathymetry measurements using phase difference. In *OCEANS'98 Conference Proceedings* (Vol. 2, pp. 1131-1134). IEEE.
- [21] Kocak, D. M., & Caimi, F. M. (2005). The current art of underwater imaging—with a glimpse of the past and vision of the future. *Marine Technology Society Journal*, 39(3), 5-26.
- [22] Jaffe, J. S., Moore, K. D., McLean, J., & Strand, M. P. (2001). Underwater optical imaging: status and prospects. *Oceanography*, 14(3), 66-76.
- [23] Narayanan, R. M., & Kalshoven, J. E. (1998, July). Advances in Laser Remote Sensing for Terrestrial and Hydrographic Applications. In *Advances in Laser Remote Sensing for Terrestrial and Hydrographic Applications* (Vol. 3382).
- [24] Churnside, J. H., Wilson, J. J., & Tatarskii, V. V. (1997). Lidar profiles of fish schools. *Applied Optics*, 36(24), 6011-6020.
- [25] Narasimhan, S. G., & Nayar, S. K. (2005). Structured light methods for underwater imaging: light stripe scanning and photometric stereo. In *OCEANS, 2005. Proceedings of MTS/IEEE* (pp. 2610-2617). IEEE.
- [26] Grossberg, M. D., & Nayar, S. K. (2005). The raxel imaging model and ray-based calibration. *International Journal of Computer Vision*, 61(2), 119-137.
- [27] Will, P. M., & Pennington, K. S. (1972). Grid coding: A preprocessing technique for robot and machine vision. *Artificial Intelligence*, 2(3), 319-329.
- [28] Fournier, G. R., Bonnier, D., Forand, J. L., & Pace, P. W. (1993). Range-gated underwater laser imaging system. *Optical Engineering*, 32(9), 2185-2190.
- [29] Gilbert, G. D., & Pernicka, J. C. (1966, June). Improvement of underwater visibility by reduction of backscatter with a circular polarization technique. In *Underwater Photo Optics I* (pp. 15-25). International Society for Optics and Photonics.
- [30] Walker, J. G., Chang, P. C., & Hopcraft, K. I. (2000). Visibility depth improvement in active polarization imaging in scattering media. *Applied Optics*, 39(27), 4933-4941.
- [31] Barone, S., Curcio, A., & Rationale, A. V. (2004). Epipolar Geometry and Structured Light for Reverse Engineering of Free Form Surfaces. In *Visualization, Imaging, And Image Processing: Fourth IASTED International Conference Proceedings*.

- [32] Bythell, J., Pan, P., & Lee, J. (2001). Three-dimensional morphometric measurements of reef corals using underwater photogrammetry techniques. *Coral Reefs*, 20(3), 193-199.
- [33] Kocak, D. M., Jagielo, T. H., Wallace, F., & Kloske, J. (2004, September). Remote sensing using laser projection photogrammetry for underwater surveys. In *Geoscience and Remote Sensing Symposium, 2004. IGARSS'04. Proceedings. 2004 IEEE International* (Vol. 2, pp. 1451-1454). IEEE.
- [34] Chotiros, N. P., Piper, J. N., Moore, J., & Weigl, D. F. (2001). Small scale seafloor roughness measurement using a ROV. In *OCEANS, 2001. MTS/IEEE Conference and Exhibition* (Vol. 4, pp. 2525-2529). IEEE.
- [35] Wang, C. C., Shyue, S. W., Hsu, H. C., Sue, J. S., & Huang, T. C. (2001). CCD camera calibration for underwater laser scanning system. In *OCEANS, 2001. MTS/IEEE Conference and Exhibition* (Vol. 4, pp. 2511-2517). IEEE.
- [36] Pilgrim, D. A., D. M. Parry e S. Rimmer, (2001), *The underwater optics of Abiss (Autonomous Benthic Image Scaling System)*, da "Ocean Optics VI", Institute of Physics, London.
- [37] Pinkard, D. R., Kocak, D. M., & Butler, J. L. (2005). Use of a video and laser system to quantify transect area for remotely operated vehicle (ROV) rockfish and abalone surveys. In *OCEANS, 2005. Proceedings of MTS/IEEE* (pp. 2824-2829). IEEE.
- [38] Woodham, R. J. (1980). Photometric method for determining surface orientation from multiple images. *Optical engineering*, 19(1), 191139-191139.
- [39] Hertzmann, A., & Seitz, S. M. (2003, June). Shape and materials by example: A photometric stereo approach. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on* (Vol. 1, pp. 1-533). IEEE.
- [40] Basri, R., Jacobs, D., & Kemelmacher, I. (2007). Photometric stereo with general, unknown lighting. *International Journal of Computer Vision*, 72(3), 239-257.
- [41] Khaleghi, B., Khamis, A., Karray, F. O., & Razavi, S. N. (2013). Multisensor data fusion: A review of the state-of-the-art. *Information Fusion*, 14(1), 28-44.
- [42] White, F. E. (1991). Data fusion lexicon. JOINT DIRECTORS OF LABS WASHINGTON DC.
- [43] Dasarathy, B. V. (1994). Decision fusion (Vol. 1994). Los Alamitos, CA: IEEE Computer Society Press.
- [44] Goodman, I. R., Mahler, R. P., & Nguyen, H. T. (1997). Mathematics of Data Fusion, volume 39 of Series B: Mathematical and Statistical Methods.
- [45] Kokar, M. M., Tomasik, J. A., & Weyman, J. (2004). Formalizing classes of information fusion systems. *Information Fusion*, 5(3), 189-202.
- [46] Kumar, M., Garg, D. P., & Zachery, R. A. (2006, June). A generalized approach for inconsistency detection in data fusion from multiple sensors. In *American Control Conference, 2006* (pp. 6-pp). IEEE.
- [47] Smets, P. (2007). Analyzing the combination of conflicting belief functions. *Information Fusion*, 8(4), 387-412. (Smith & Singh, 2006) Smith, D., & Singh, S. (2006). Approaches to multisensor data fusion in target tracking: A survey. *Knowledge and Data Engineering, IEEE Transactions on*, 18(12), 1696-1710.

- [48] Smith, D., & Singh, S. (2006). Approaches to multisensor data fusion in target tracking: A survey. *Knowledge and Data Engineering, IEEE Transactions on*, 18(12), 1696-1710.
- [49] Joshi, R., & Sanderson, A. C. (1999). Minimal representation multisensor fusion using differential evolution. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 29(1), 63-76.
- [50] Dong, X. L., Berti-Equille, L., & Srivastava, D. (2009). Truth discovery and copying detection in a dynamic world. *Proceedings of the VLDB Endowment*, 2(1), 562-573.
- [51] Zhu, Y., Song, E., Zhou, J., & You, Z. (2005). Optimal dimensionality reduction of sensor data in multisensor estimation fusion. *Signal Processing, IEEE Transactions on*, 53(5), 1631-1639.
- [52] Milenova, B. L., & Campos, M. M. (2005, July). Mining high-dimensional data for information fusion: a database-centric approach. In *Information Fusion, 2005 8th International Conference on* (Vol. 1, pp. 7-pp). IEEE.
- [53] Williams, S., & Mahon, I. (2004, April). Simultaneous localisation and mapping on the great barrier reef.
- [54] Fusiello, A., Giannitrapani, R., Isaia, V., & Murino, V. (1999). Virtual environment modeling by integrated optical and acoustic sensing. In *3-D Digital Imaging and Modeling, 1999. Proceedings. Second International Conference on* (pp. 437-446). IEEE.
- [55] Fusiello, A., & Murino, V. (2000). Calibration of an Optical-Acoustic Sensor for Underwater Applications.
- [56] Besl, P. J., & McKay, N. D. (1992, April). Method for registration of 3-D shapes. In *Robotics-DL tentative* (pp. 586-606). International Society for Optics and Photonics.
- [57] Lowe, D. G. (1991). Fitting parameterized three-dimensional models to images. *IEEE transactions on pattern analysis and machine intelligence*, 13(5), 441-450.
- [58] Negahdaripour, S., Pirsiavash, H., Sekkati, H. (2007, June). Integration of Motion Cues in Optical and Sonar Videos for 3-D Positioning. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on* (pp. 1-8). IEEE.
- [59] Sekkati, H., & Negahdaripour, S. (2006, June). Direct and indirect 3-D reconstruction from opti-acoustic stereo imaging. In *3D Data Processing, Visualization, and Transmission, Third International Symposium on* (pp. 615-622). IEEE.
- [60] Negahdaripour, S. (2007). Epipolar geometry of opti-acoustic stereo imaging. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(10), 1776-1788.
- [61] Negahdaripour, S. (2005, September). Calibration of DIDSON forward-scan acoustic video camera. In *OCEANS, 2005. Proceedings of MTS/IEEE* (pp. 1287-1294). IEEE.
- [62] Johnson-Roberson, M., Pizarro, O., & Williams, S. (2007, June). Towards Three-Dimensional Heterogeneous Imaging Sensor Correspondence and Registration for Visualization. In *OCEANS 2007-Europe* (pp. 1-6). IEEE.
- [63] Huber, D. F., & Hebert, M. (1999). A new approach to 3D terrain mapping. In *Intelligent Robots and Systems, 1999. IROS'99. Proceedings. 1999 IEEE/RSJ International Conference on* (Vol. 2, pp. 1121-1127). IEEE.

- [64] Horn, B. K., & Harris, J. G. (1991). Rigid body motion from range image sequences. *CVGIP: Image Understanding*, 53(1), 1-13.
- [65] Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6), 381-395.
- [66] Drap, P., Merad, D., Boï, J. M., Boubguira, W., Mahiddine, A., Chemisky, B., ... & Bianchimani, O. (2011, October). ROV-3D: 3D underwater survey combining optical and acoustic sensor. In *Proceedings of the 12th International conference on Virtual Reality, Archaeology and Cultural Heritage* (pp. 177-184). Eurographics Association.
- [67] Hurtós, N., Cufi, X., & Salvi, J. (2010, May). Calibration of optical camera coupled to acoustic multibeam for underwater 3D scene reconstruction. In *OCEANS 2010 IEEE-Sydney* (pp. 1-7). IEEE.
- [68] Zhang, Q., & Pless, R. (2004, September). Extrinsic calibration of a camera and laser range finder (improves camera calibration). In *Intelligent Robots and Systems, 2004.(IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on* (Vol. 3, pp. 2301-2306). IEEE.
- [69] Nicosevici, T., & Garcia, R. (2008, April). Online robust 3D mapping using structure from motion cues. In *OCEANS 2008-MTS/IEEE Kobe Techno-Ocean*(pp. 1-7). IEEE.
- [70] Murino, V., & Trucco, A. (2000). Three-dimensional image generation and processing in underwater acoustic vision. *Proceedings of the IEEE*, 88(12), 1903-1948.
- [71] Murino, V. (2001). Reconstruction and segmentation of underwater acoustic images combining confidence information in MRF models. *Pattern Recognition*,34(5), 981-997.
- [72] CloudCompare. Available online: <http://www.danielgm.net/cc/> (accessed on 10 June 2015).
- [73] Hartley, R.; Zisserman, A. *Multiple View Geometry in Computer Vision*; Cambridge University Press: Cambridge, UK, 2000; p. 655.
- [74] Leone, A., Diraco, G., & Distanto, C. (2007, September). Stereoscopic system for 3D seabed mosaic reconstruction. In *Image Processing, 2007. ICIP 2007. IEEE International Conference on* (Vol. 2, pp. II-541). IEEE.
- [75] Remondino, F., & El-Hakim, S. (2006). Image-based 3D Modelling: A Review. *The Photogrammetric Record*, 21(115), 269-291.
- [76] Bianco, G., Gallo, A., Bruno, F., Muzzupappa, M., A comparative analysis between active and passive techniques for underwater 3D reconstruction of close-range objects. *Sensors (Switzerland)*, Vol. 13, N.8, pp. 11007-11031.
- [77] Schettini R., & Corchs S. (2000). Imaging for underwater archaeology. *American Journal of Field Archaeology* 27(3), 319–328.
- [78] Mahiddine A., Seinturier J., Peloso D., Boï J.-M., Drap P., Merad D. (2012). Underwater image pre-processing for automated photogrammetry in high turbidity water. In *VSMM2012*.
- [79] Rizzi, A., Gatta, C., & Marini, D. (2003). A new algorithm for unsupervised global and local color correction. *Pattern Recognition Letters*, 24(11), 1663-1677.
- [80] Chambah, M. (2004). Automatic color restoration of faded pictures and motion pictures. In *The 8th World Multi-Conference on Systemics, Cybernetics and Informatics SCI*.

- [81] Bazeille, S., Quidu, I., Jaulin, L., & Malkasse, J. P. (2006). Automatic underwater image pre-processing. *Proceedings of CMM'06*.
- [82] Iqbal, K., Salam, R. A., Osman, A., & Talib, A. Z. (2007). Underwater Image Enhancement Using an Integrated Colour Model. *IAENG International Journal of Computer Science*, 34(2).
- [83] Robin Kalia, Keun-Dong Lee, Samir B.V.R., Sung-Kwan Je, & Weon-Geun Oh. (2011, 9-11 Feb.). An analysis of the effect of different image preprocessing techniques on the performance of SURF: Speeded Up Robust Feature. *Paper presented at the 17th Korea-Japan Joint Workshop on Frontiers of Computer Vision (FCV)*. pp.1-6.
- [84] Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2008). Speeded-up robust features (SURF). *Computer vision and image understanding*, 110(3), 346-359.
- [85] Corke, P., Detweiler, C., Dunbabin, M., Hamilton, M., Rus, D., & Vasilescu, I. (2007, April). Experiments with underwater robot localization and tracking. In *Robotics and Automation, 2007 IEEE International Conference on* (pp. 4556-4561). IEEE.
- [86] Bouguet, J. Y. (2004). Camera calibration toolbox for matlab.
- [87] <http://opencv.org/>
- [88] Geiger, A., Roser, M., & Urtasun, R. (2011). Efficient large-scale stereo matching. In *Computer Vision—ACCV 2010* (pp. 25-38). Springer Berlin Heidelberg
- [89] Matabosch, C., Salvi, J., Fofi, D., & Meriaudeau, F. (2005, February). Range image registration for industrial inspection. In *Electronic Imaging 2005* (pp. 216-227). International Society for Optics and Photonics.
- [90] Vose, M. D. (1999). *The simple genetic algorithm: foundations and theory* (Vol. 12). MIT press.
- [91] Sharf, I., Wolf, A., & Rubin, M. B. (2010). Arithmetic and geometric solutions for average rigid-body rotation. *Mechanism and Machine Theory*, 45(9), 1239-1251.
- [92] Biancoa, G., Muzzupappaa, M., Brunoa, F., Garciaab, R., & Neumann, L. (2015). a New Color Correction Method for Underwater Imaging. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1, 25-32.