

Università degli Studi di Padova
Facoltà di Ingegneria



**Acquisizione di dati
con videocamere e
sensore ToF**

Candidato: Alvise Memo – matr. 1001986
anno accademico 2012/13

Relatore: Prof. Pietro Zanuttigh

Correlatore: Ing. Fabio Dominio

INDICE

1. Introduzione	pag.	1
2. I sensori ToF		
2.1. Generalità dei sensori ToF	pag.	3
2.2. Problematiche comuni	pag.	6
2.3. Caratteristiche del sensore SR4000 della MESA Imaging	pag.	9
3. Formato dei dati grezzi	pag.	11
3.1. Distanza	pag.	12
3.2. Ampiezza	pag.	13
3.3. Mappa di confidenza	pag.	14
4. Approntamento del software di base	pag.	15
5. Algoritmo di acquisizione	pag.	16
5.1. Classe FormMain	pag.	17
5.2. Classe SensorHelper	pag.	20
5.3. Classe DoubleCamera	pag.	21
6. Algoritmi di elaborazione	pag.	23
7. Risultati conseguiti e sviluppi futuri	pag.	25
Bibliografia/Sitografia	pag.	27

1. Introduzione

Descrizione del problema da affrontare, dell'ambito di impiego, degli obiettivi da raggiungere, e breve presentazione dell'hardware e del software a disposizione.

Questa tesi si propone di descrivere la creazione di un software specifico in grado di acquisire contemporaneamente da due telecamere ed un sensore ToF un insieme di dati successivamente utilizzabili per ricreare graficamente con tecnologia 3D l'ambiente e/o gli oggetti inquadrati. Inoltre il software deve garantire la massima possibilità di configurazione di tutti i parametri che caratterizzano le funzioni dei dispositivi in uso, in modo da avere il massimo controllo possibile della fase di acquisizione dei dati.



I dati grezzi acquisiti dalla sezione di software che si interfaccia direttamente con i dispositivi in uso potranno essere successivamente elaborati da una ulteriore sezione di software che, mediante il ricorso a specifiche librerie esterne quali **OpenCV**^[1] o simili, oppure tramite l'utilizzo di algoritmi e programmi dedicati all'elaborazione di dati 3D come **MeshLab**, permetterà di offrire un'analisi visiva sia dei grezzi che dei dati rielaborati al fine di dare loro un'interpretazione più vicina alla realtà o allo specifico obiettivo prefissato (quale ad esempio il riconoscimento e movimento di oggetti prestabiliti o parti del corpo umano inquadrato).

L'hardware utilizzato è composto da due telecamere modello scout **scA1000** della ditta BASLER montate su di un supporto meccanico stabile e regolabile, con al centro il sensore ToF **SR4000** della ditta MESA Imaging. Il software di partenza è composto dalle librerie e dai driver resi disponibili dalle due case costruttrici dei suddetti componenti, e dei relativi wrapper per accedervi mediante programmazione in **C#**^[2].

2. I sensori ToF

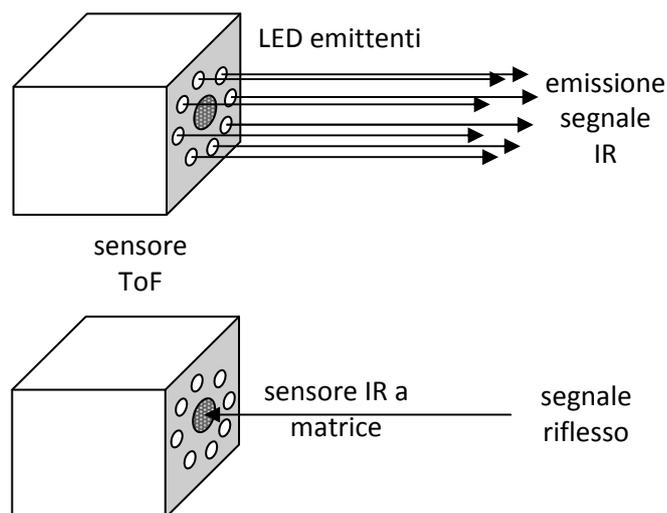
Descrizione, analisi del funzionamento, caratteristiche funzionali ed operative dei sensori Time of Flight.

2.1 Generalità dei sensori ToF

I sensori **ToF** (*Time of Flight* - tempo di volo)^[3] vengono utilizzati per catturare e ricostruire un'immagine 3D. I primi tentativi di acquisizione di immagini tridimensionali si sono basati sull'utilizzo di due sensori visivi (telecamere) solidali e posti ad una certa distanza tra loro, ricostruendo in qualche modo il funzionamento del sistema visivo umano, anche detto sistema stereoscopico.

Il sensore ToF invece è un dispositivo unico che stima in tempo reale la distanza tra il sensore e le superfici frontali degli oggetti inquadrati, fornendo una rappresentazione matriciale dell'immagine, dove ad ogni punto dell'oggetto inquadrato corrisponde la sua distanza dal sensore. Spesso i sensori ToF vengono chiamati anche telecamere ToF, anche se in realtà non forniscono un'immagine come siamo abituati a concepire, ma l'insieme delle distanze dei singoli punti degli oggetti inquadrati.

I sensori ToF possiedono una sorgente illuminante a frequenza elevata (generalmente IR) che modulano con una data frequenza portante (ad esempio 20 o 30 MHz).



L'impulso ottico IR viene generato da una serie di LED IR posti attorno al sensore. Questo segnale, opportunamente modulato da un segnale sinusoidale, raggiunge l'oggetto inquadrato e poi viene riflesso verso la telecamera. Il sensore è di struttura matriciale, e rileva per ogni punto due informazioni: l'intensità del segnale riflesso ed il tempo impiegato a ritornare sulla telecamera.

L'intensità del segnale riflesso è collegata alle caratteristiche fisiche (tipo di materiale, tipo di superficie, colore) del punto dell'oggetto colpito dal segnale IR. Di conseguenza per ogni punto varia la quantità di segnale riflesso, e questa informazione viene utilizzata per costruire una immagine in bianco e nero dell'oggetto. Lo sfasamento del segnale è invece proporzionale alla distanza di quel punto dell'oggetto dal sensore, e quindi serve per costruire una mappa 3D dell'oggetto.

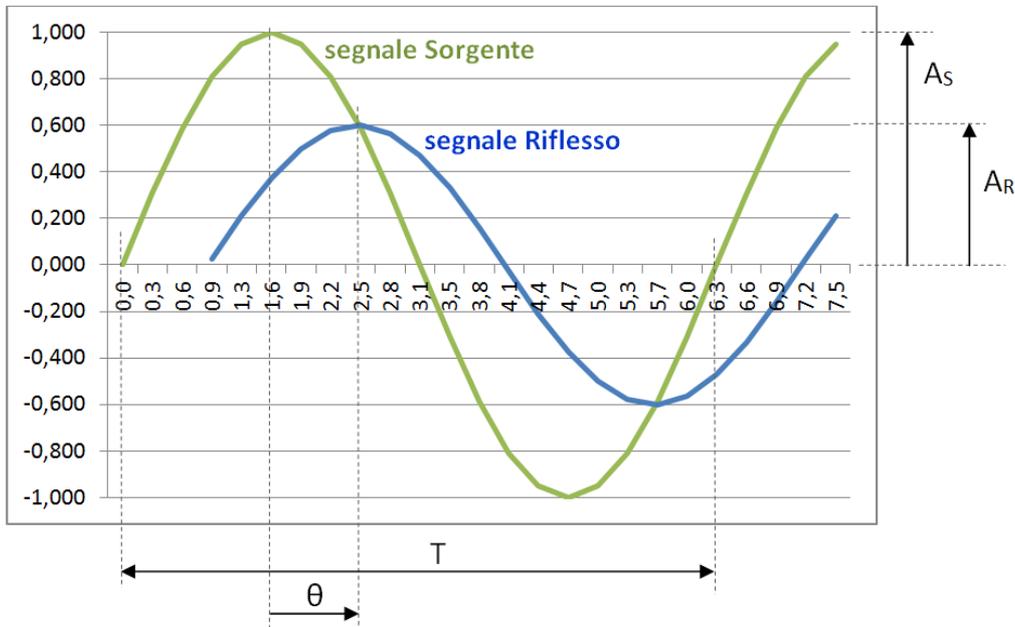
La parte trasmissiva del sensore si sarebbe potuta realizzare con un semplice trasmettitore ad impulsi IR, ma la presenza di altre superfici rimbalzanti (specie quelle al contorno) e delle diverse angolazioni del raggio riflesso (causate da inclinazioni quasi mai ortogonali al segnale incidente) avrebbe comportato numerose sovrapposizioni di segnali riflessi residui, con una conseguente instabilità delle misurazioni; problema che viene affrontato dai sistemi SPAD.

Per ridurre sensibilmente questi fenomeni, il trasmettitore del sensore emette con continuità un segnale IR detto portante modulato da un segnale sinusoidale di ampiezza A e, invece del ritardo, il sensore ToF misura lo sfasamento tra gli involuppi del segnale trasmesso e quello ricevuto.

Ovviamente ci sono delle condizioni a cui questo processo deve soddisfare:

- la frequenza del segnale modulante deve essere molto più bassa del segnale della portante, per garantire una facile ricostruzione dell'involuppo; nel caso dei sensori ToF il rapporto è ampiamente garantito dall'uso di una portante IR (circa 850 nm) ed una modulante a qualche decina di MHz
- Il periodo del segnale modulante deve essere abbastanza lungo per garantire che tutti i segnali riflessi ritornino al sensore prima che sia passato un intervallo di tempo pari al valore di questo periodo; in caso

contrario si avrebbero punti a distanze multiple rappresentati dallo stesso valore di ritardo (vedere successivamente i punti di ambiguità).



Questo ultimo vincolo comporta la determinazione di un range ben preciso all'interno del quale il sensore riesce ad acquisire correttamente le distanze. Calcoliamo questo intervallo: sapendo che un segnale ottico si propaga nello spazio alla velocità della luce c , se il segnale viene modulato da una sinusoide alla frequenza F , in un intervallo di tempo T (pari ad un periodo del segnale modulante) percorrerà un percorso P pari a:

$$P = cT = \frac{c}{F} = \frac{299.792.458 \text{ [m/s]}}{30 \cdot 10^6 \text{ [Hz]}} = 9,99 \text{ [m]}$$

Dove:

c = velocità della luce nel vuoto, pari a 299.752.458 [m/s]

F = frequenza del segnale modulante, nell'esempio pari a 30 [MHz]

$T = 1/F$ = periodo del segnale modulante, nell'esempio pari a 33,3 nS

Se si desidera che il sensore ToF sia in grado di rilevare un oggetto senza ambiguità, questo dovrà essere posto ad una distanza massima pari a $P/2$, dato che il segnale dovrà percorrere il tragitto diretto (da sensore ad oggetto) e quello riflesso (da oggetto a sensore) all'interno di un periodo T .

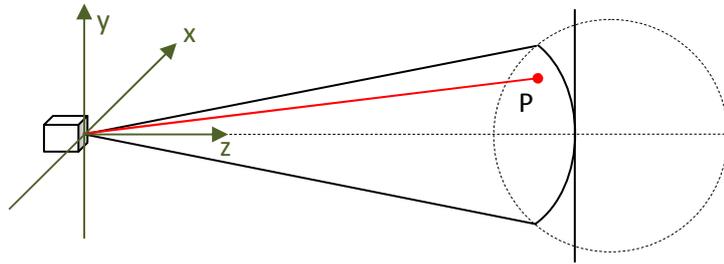
In definitiva, con una frequenza modulante di 30 MHz, un sensore TOF sarà in grado di rilevare oggetti posti ad una distanza massima di $9,99/2 \approx 5$ m.

Ovviamente con una modulazione a 20 MHz si otterrebbe un aumento della distanza raggiungibile, pari a 7,5 m.

2.2 Problematiche comuni

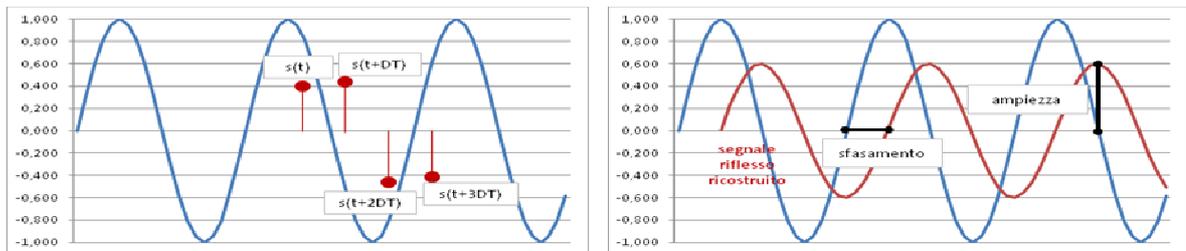
I principali problemi che contraddistinguono l'adozione della tecnologia ToF sono:

- Diversa qualità di **riflessione** dei vari tipi di materiali
L'impulso luminoso emesso dai LED della telecamera urta l'oggetto inquadrato e viene riflesso con caratteristiche specifiche dell'oggetto stesso. Dipende dal colore del materiale, dalla superficie e dall'inclinazione dell'oggetto. È evidente che un foglio bianco riflette molto di più di un foglio grigio scuro. Altra caratteristica essenziale è la distribuzione del raggio riflesso: quanto si distribuisce (determinando una distribuzione dell'energia riflessa) e in quale direzione (verso il sensore della telecamera o verso pareti laterali). Le misure rilevate ovviamente dipendono fortemente da questi elementi.
- Necessità di passare da **coordinate sferiche** a coordinate cartesiane
Le distanze acquisite dal sensore TOF sono espresse in coordinate sferiche di natura radiale aventi come origine il sensore. Per ovvie comodità di impiego occorre effettuare una conversione in coordinate cartesiane aventi l'origine nello stesso punto. Di norma questa trasformazione avviene via software direttamente nel driver del dispositivo.



- Impostazione del **tempo di integrazione**

In maniera analoga all'apertura del diaframma di una macchina fotografica, ogni singolo sensore della telecamera deve rimanere esposto al segnale riflesso per un certo intervallo di tempo (Intervallo di Integrazione, o *Integration Time*), per dar tempo alle cariche di accumularsi. Tenendo presente che ogni singola valutazione dello sfasamento (e quindi della distanza) richiede la lettura di 4 valori ad intervalli di tempo prefissati (ad esempio sfasati ciascuno di 90°), l'acquisizione totale di una immagine richiede almeno quattro volte l'intervallo di integrazione. Questo parametro è impostabile da software e la sua scelta è critica in quanto direttamente collegato al frame rate desiderato, alla precisione delle misure, alle caratteristiche ambientali e degli oggetti da riprendere.

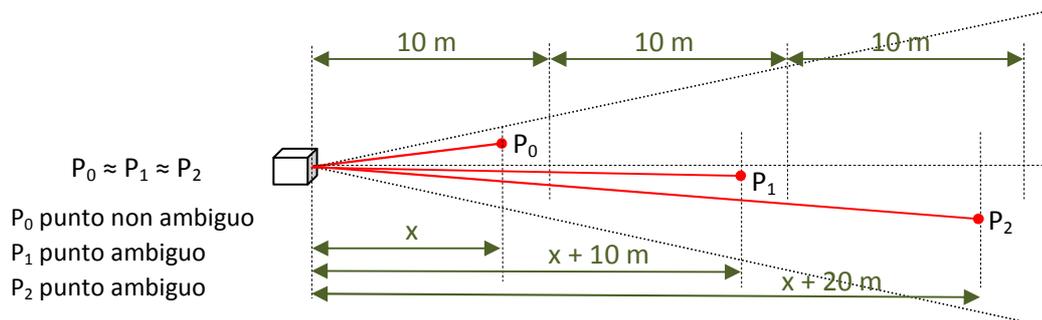


- Gestione di **oggetti in movimento**

Per migliorare la rilevazione degli oggetti in movimento riducendo il tempo di integrazione, ovviamente a discapito della qualità generale.

- Misurazioni **ambigue**

La scelta di valutare la distanza in base allo sfasamento piuttosto che al fronte d'onda di ritorno migliora notevolmente la stabilità e la precisione della misura, ma introduce possibili ambiguità nel rilievo di punti al di fuori dell'intervallo di attendibilità di un periodo della frequenza modulante. Il fenomeno si può presentare per punti di contorno o sullo sfondo.



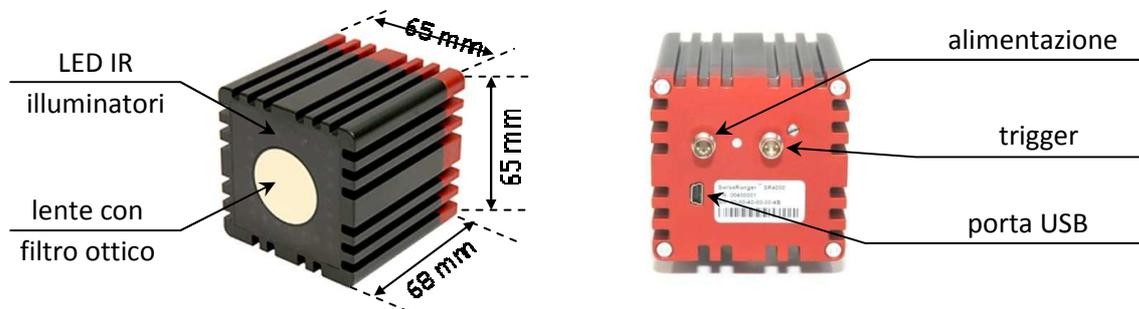
- **Mappa di confidenza** (*confidence map*)

La mappa di confidenza (o mappa di fiducia o di affidabilità) è una matrice che riporta, per ogni pixel della telecamera, un valore correlato alla probabilità che la misura di quel punto sia più o meno attendibile. Viene calcolata via software in tempo reale, in base alle variazioni medie e di scarto effettuate in precedenza sul singolo punto, e di norma avviene direttamente nel driver del dispositivo.

- Difficoltà **ambientali**

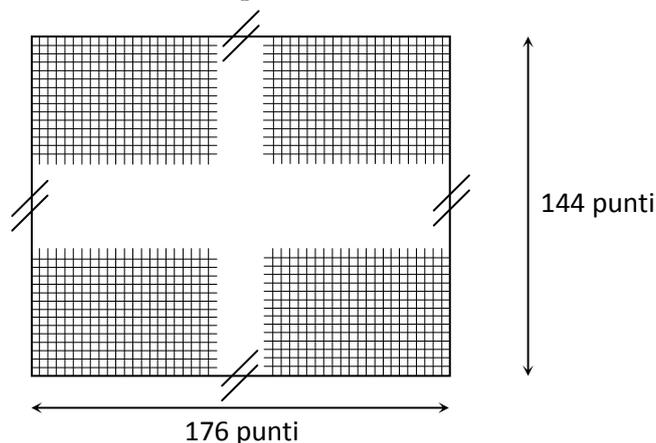
Date le frequenze in gioco, il sistema è abbastanza critico e sensibile alle condizioni ambientali. Ad esempio occorre limitare gli effetti termici indotti sul sensore, e quindi tener presente la temperatura ambiente e cercare di mantenere a temperatura costante la telecamera. Ma soprattutto occorre evitare totalmente eventuali riflessioni ambientali, che possono introdurre oggetti fantasma o deformare quelli reali. Si pensi al fenomeno del multipath, cioè riflessioni su oggetti concavi o sui muri che delimitano la scena, che associano ad un punto percorsi diversi, e quindi distanze misurate diverse, o a riflessioni dirette dei muri periferici che inducono la presenza di oggetti inesistenti.

2.3 Caratteristiche fisiche del sensore sr4000 della MESA Imaging



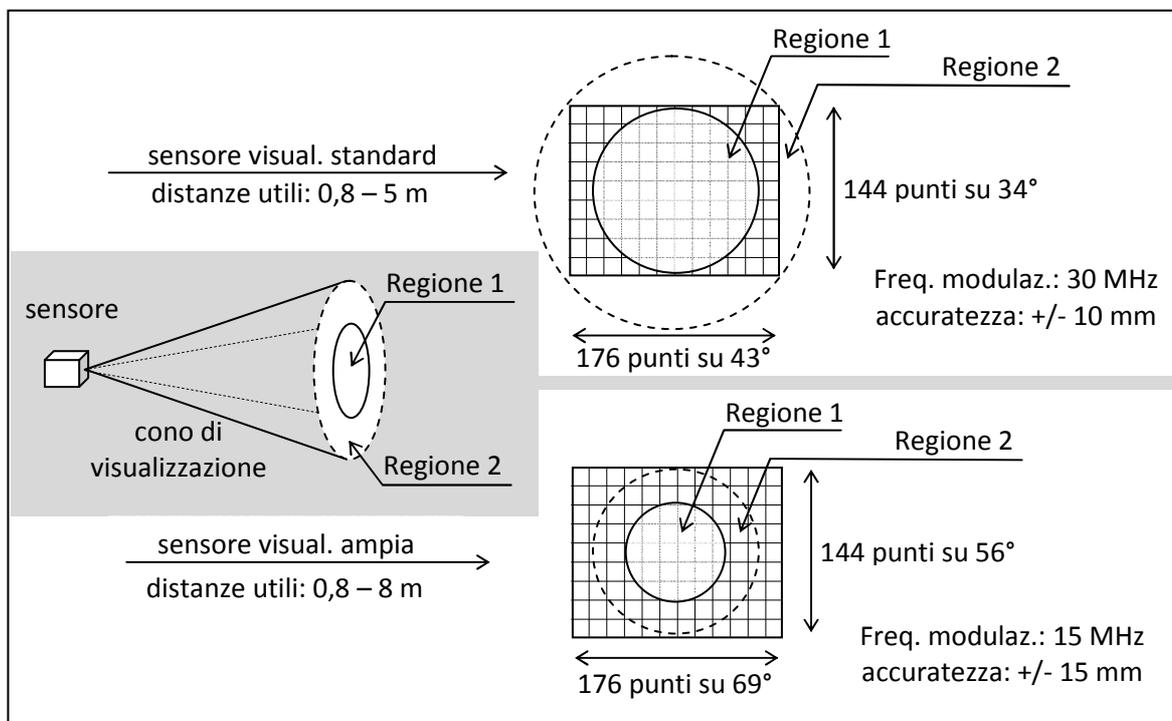
Le caratteristiche funzionali essenziali del sensore/telecamera SwissRanger SR4000 della MESA Imaging sono:

- Risoluzione di 176 x 144 punti di rilevazione



- Frequenza di modulazione selezionabile in tre valori distinti all'interno del range scelto: per l'intervallo di base (30 MHz) i valori selezionabili sono 29, 30 e 31 MHz, mentre per l'intervallo ridotto (15 MHz) i valori possibili sono 14.5, 15.0 e 15.5 MHz. In tal modo è possibile impiegare nella stessa struttura di acquisizione tre sensori distinti operanti contemporaneamente con segnali di modulazione diversi (impianto multicamera)
- Cono di visualizzazione di 43°(oriz.) x 34°(vert.) per il modello a visualizzazione standard, e di 69°(oriz.) x 56°(vert.) per il modello a visualizzazione ampia; in entrambi i modelli si identificano due zone di

acquisizione, quella centrale, denominata Regione 1, e quella periferica denominata Regione 2



- Distanza radiale minima di 0,8 m
- Distanza radiale massima di circa 5 o 10 m (in base alla frequenza di modulazione); l'approssimazione è dovuta agli effetti di bordo ed alle diverse frequenze di modulazione selezionabili all'interno del range.
- Accuratezza tipica di +/- 10 mm nel range 0-5 m (corrispondente ad una frequenza di modulazione di 30 MHz), e di +/- 15 mm nel range 0-10 m (corrispondente ad una frequenza di modulazione di 15 MHz); i valori sono espressi con una codifica floating-point ad una risoluzione di 16 bit
- Deriva termica inferiore a 0,5 mm per temperature inferiori ai 30°, inferiori a 1,5 mm per temperature fino a 50° (si tenga presente che il problema si riferisce sia alla temperatura ambiente, ma anche all'auto-surriscaldamento del dispositivo se acceso per lunghi periodi)
- Lunghezza focale della lente pari a 10 mm per il modello a visualizzazione standard, e di 5,8 mm per quello a visualizzazione ampia; la regolazione è manuale all'interno delle distanze ammesse

- La lente posta davanti al sensore è realizzata anche per fare funzioni di filtro passabanda centrato sulla gamma delle frequenze IR per i segnali ottici in ingresso: in tal modo si limitano eventuali interferenze determinate da illuminazioni ambientali particolari, anche se è consigliabile non esporre la zona di ripresa ai raggi solari
- Frame rate massimo di 50 FPS (selezionabile via software)
- Lunghezza d'onda segnale IR di circa 850 nm
- Valore di soglia della qualità del segnale utilizzato per la determinazione della mappa di confidenza, selezionabile via software
- Acquisizione in modalità continua e/o a comando esterno (trigger)
- Tempo di integrazione selezionabile tra 0,3 e 5,8 ms
- Valori d'uscita digitali prodotti da convertitori a 16 bit per
 - Distanza sferica e distanza in coordinate cartesiane
 - Ampiezza del segnale riflesso
 - Immagine risultante in scale di grigio
 - Mappa di confidenza
- Interfacciamento dipendente dal modello, o tramite porta USB o tramite interfaccia di rete RJ45 Fast Ethernet

3. Formato dei dati grezzi

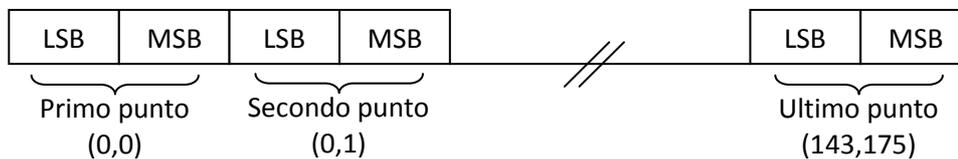
Descrizione delle strutture dati prodotte dal sensore, distanze, ampiezze, mappa di confidenza.

I dati grezzi forniti dal sensore/telecamera SR4000 sono relativi alle distanze radiali ed alle ampiezze dei segnali di ritorno relativi a tutti i punti del sensore. Inoltre il driver elabora questi dati e mediante opportune trasformazioni e correzioni fornisce le distanze in coordinate cartesiane spaziali, ed inoltre predispone la mappa di confidenza per poter interpretare correttamente i dati forniti.

3.1 Distanze

Le distanze radiali rilevate dai singoli punti del sensore vengono fornite in uscita in modalità seriale, a partire dal primo punto in alto a sinistra fino all'ultimo in basso a destra. Dato che abbiamo a che fare con un sensore matriciale con risoluzione 144 x 176, in uscita otteniamo un vettore di 25.344 dati a 16 bit ciascuno. Per ogni punto i dati vengono organizzati in modalità little-endian, ed i due bit meno significativi vengono dedicati alla misurazione degli sfasamenti nei quattro punti di lettura, quindi vanno trascurati.

MMMMMMMM LLLLLLxx : formato binario (16 bit, i 2 lsb non significativi)
0x0000 → sfasamento nullo, cioè oggetto posto immediatamente davanti al sensore
0xFFFF → sfasamento massimo, cioè oggetto posto al limite estremo del range utile



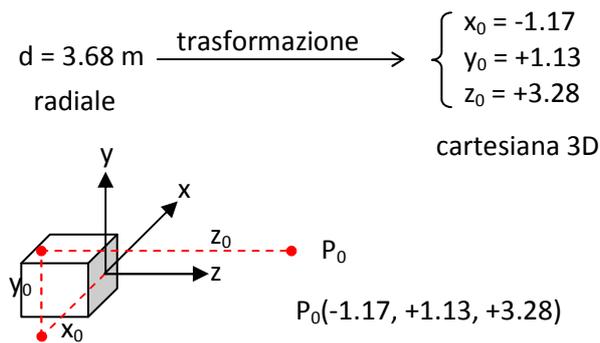
Le misure grezze delle distanze prodotte dal sensore vengono utilizzate direttamente dal driver per generare delle misure più significative per l'utente mediante trasformazione da sferico-radiale a cartesiana tridimensionale.

Contemporaneamente il software del driver effettua anche operazioni correttive dovute alle distorsioni dell'ottica del dispositivo.

Il sistema cartesiano è ovviamente centrato sul sensore (il punto di coordinate 0,0,0 è posto nel suo centro), mentre la direzione dell'asse Z va verso l'oggetto inquadrato, quella dell'asse X verso la sinistra del sensore e quella dell'asse Y verso l'alto.

Tre sono le possibili trasformazioni applicabili dal software del driver, selezionabili via software:

- mediante interi senza segno proporzionali al range ammesso
- mediante valori in virgola mobile a precisione singola
- mediante valori in virgola mobile a doppia precisione



Ad esempio, nel caso del punto riportato nella figura precedente, la quintupla di valori generata dal driver con rappresentazione intera delle coordinate, è:

$$x=-1.170185 \quad y=1.138125 \quad z=3.288972 \quad a=48256 \quad d=367$$

corrispondente ad una distanza radiale effettiva di 3,68 m.

3.2 Ampiezza

Anche i dati relativi all'ampiezza dei vari punti, cioè un valore legato all'intensità del segnale riflesso in quel punto, è rappresentata da una quantità numerica senza segno a 16 bit. In realtà il bit più significativo indica saturazione (segnale troppo elevata e quindi non attendibile), mentre i due bit meno significativi vengono utilizzati per l'acquisizione dei campioni e quindi non vengono considerati. Di conseguenza i bit utili sono solo 13, con una risoluzione dovuta al formato di $1/2^{13}$, cioè valori utili da 0 a 4.095.

Anche in questo caso il driver apporta delle correzioni significative ai dati grezzi rilevati, in funzione della distanza rilevata e dell'illuminazione.

Ancora una volta è utile ricordare che i valori così ottenuti non sono una rappresentazione ottica della realtà inquadrata (così come potrebbe essere ottenuta da una videocamera), ma sono una mappa monocromatica delle riflessioni ai raggi infrarossi degli oggetti illuminati.

Inoltre i suoi valori dipendono direttamente dal tempo di integrazione adottato: un tempo elevato permetterà ai singoli punti di accumulare cariche (proporzionali al raggio riflesso) e quindi di raggiungere valori più elevati e determinando un contrasto maggiore tra i valori minimi e massimi acquisiti. Ma in tal modo si rischia anche di saturare l'immagine.

3.3 Mappa di confidenza

Questa mappa riporta, in ogni suo punto, la probabilità che la distanza misurata per quel punto sia effettivamente attendibile. I suoi valori vengono calcolati utilizzando i dati grezzi della distanza e dell'ampiezza, sia in riferimento ai valori dei punti adiacenti che alla loro evoluzione temporale (ad esempio se l'andamento nel tempo è costante, varia in maniera contenuta o se invece assume valori molto diversi tra loro in intervalli brevi).

Ancora una volta i dati sono rappresentati da un vettore di 144×176 elementi da 16 bit ciascuno, ed ovviamente un valore molto basso rappresenta un punto

poco attendibile e le misure relative sono poco affidabili, mentre un valore elevato presuppone misurazioni corrette. Il valore per decidere quando considerare attendibile una misura si chiama soglia di qualità, ed è configurabile dall'utente via software.



4. Approntamento del software di base

Descrizione del software da installare, driver dei sensori, versioni testate, compatibilità, ambiente di sviluppo adottato, librerie richieste, configurazione dei percorsi di ricerca, compilazione ed esecuzione del programma demo e di uno semplice di test.

La predisposizione dell'ambiente software per la produzione del codice ha richiesto:

- Installazione dell'ambiente di sviluppo della Microsoft per linguaggi visuali, limitatamente al linguaggio c#, cioè Microsoft Visual Studio C# Express versione 10
- Installazione dei driver del sensore ToF della MESA Imaging SwissRanger™ Driver versione 1.0.14.706, e del software di visualizzazione 3D associato, medesima versione^[4]
- Installazione della Basler Pylon Camera Software Suite ^[5], versione 3.2.3, necessaria per un trasferimento dei dati forniti dalle telecamere che non sovraccarichi eccessivamente il PC di acquisizione

Nell'installazione delle librerie Pylon bisogna fare attenzione ad installare anche il supporto delle librerie .NET (per programmare in c#) a 64 bit (se ovviamente si sviluppa su sistemi a 64 bit). Per semplicità e standardizzazione, si è scelto di adottare la configurazione automatica dei percorsi.

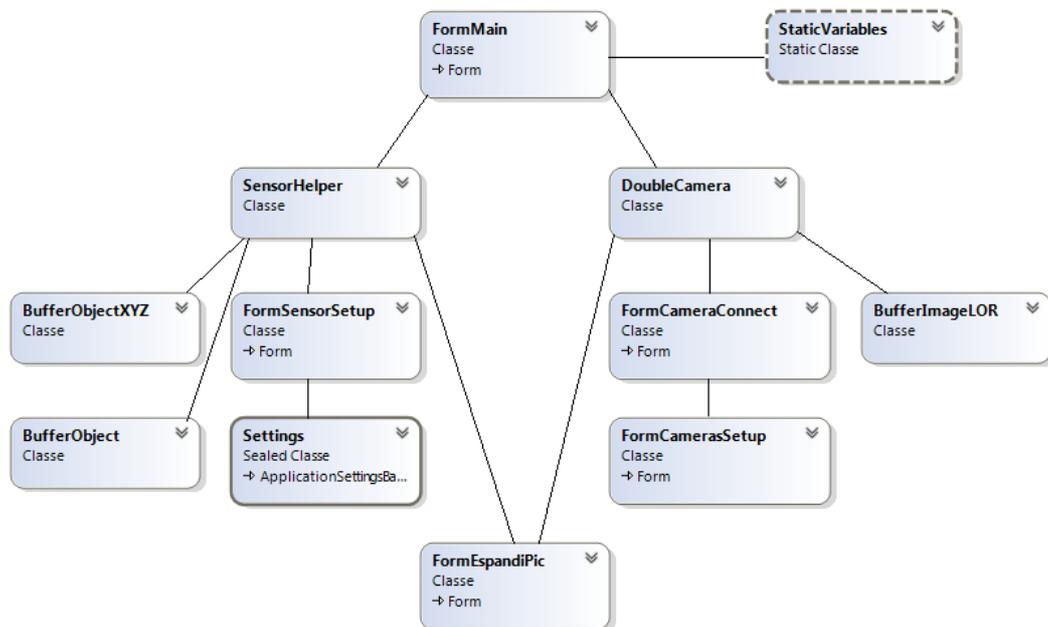
5. Algoritmo di acquisizione

Struttura del software minimo per acquisire i dati dal sensore (core dell'applicazione finale), con un semplice modulo per rappresentarli in forma grezza

Il software di acquisizione è essenzialmente suddiviso in due parti distinte:

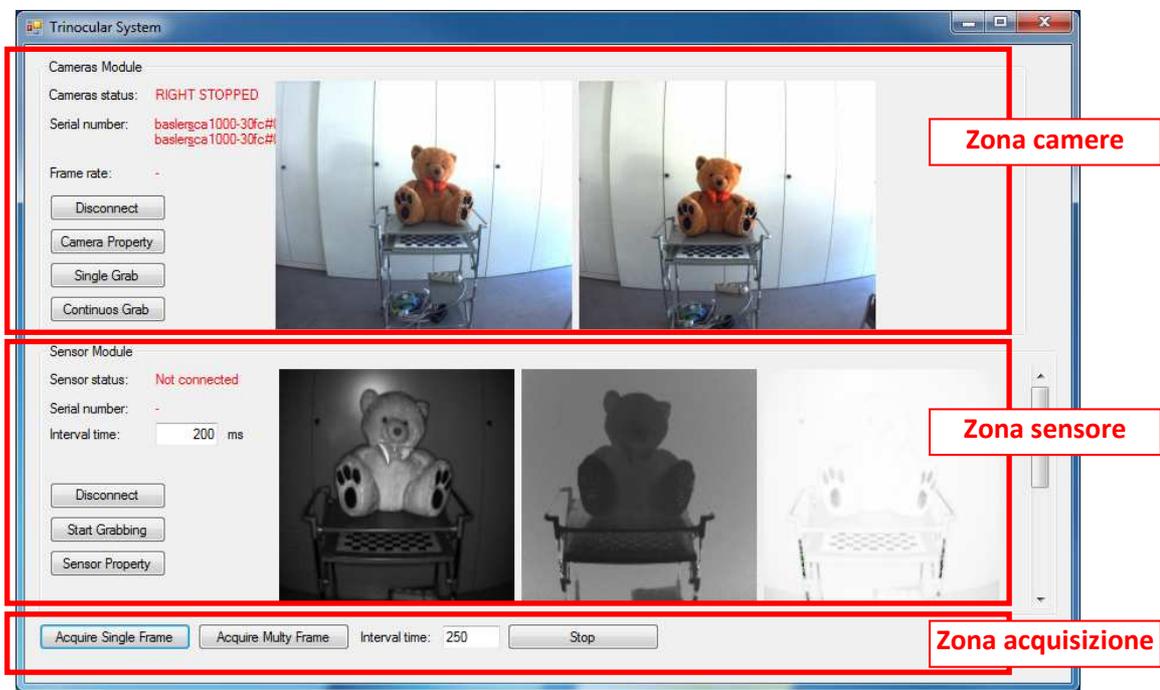
- Gestione delle camere,
- Gestione del sensore.

Queste due funzioni distinte sono collegate tra loro dalla classe **FormMain** che racchiude sia l'interfaccia utente che il software necessario per la gestione diretta dei suddetti device.



5.1 Classe FormMain

La classe **FormMain** è la classe essenziale del progetto, ed ha come scopo principale quello di fornire un'interfaccia semplice e intuitiva all'utente, in modo da permettere acquisizioni ed impostazioni dei relativi parametri, operazioni queste generalmente complesse, con pochi click. L'interfaccia si suddivide in due blocchi essenziali, quello delle telecamere e quello del sensore.



L'area relativa alle telecamere, posta nella parte alta, contiene:

- Il pulsante Connect/Disconnect, che permette di avviare la procedura software di collegamento alle due telecamere tramite i relativi driver, con possibilità di distinguere tra telecamera destra e sinistra, indipendentemente dal collegamento fisico,

- Il pulsante Camera Property che permette di entrare nella procedura di impostazione delle proprietà e dei parametri dei dispositivi,
- I due pulsanti Single e Continuous Grab. Questi permettono di avviare i metodi fondamentali della classe **DoubleCamera** che gestiranno il prelevamento delle immagini dalle camere in modalità a frame singolo o in modalità continua.

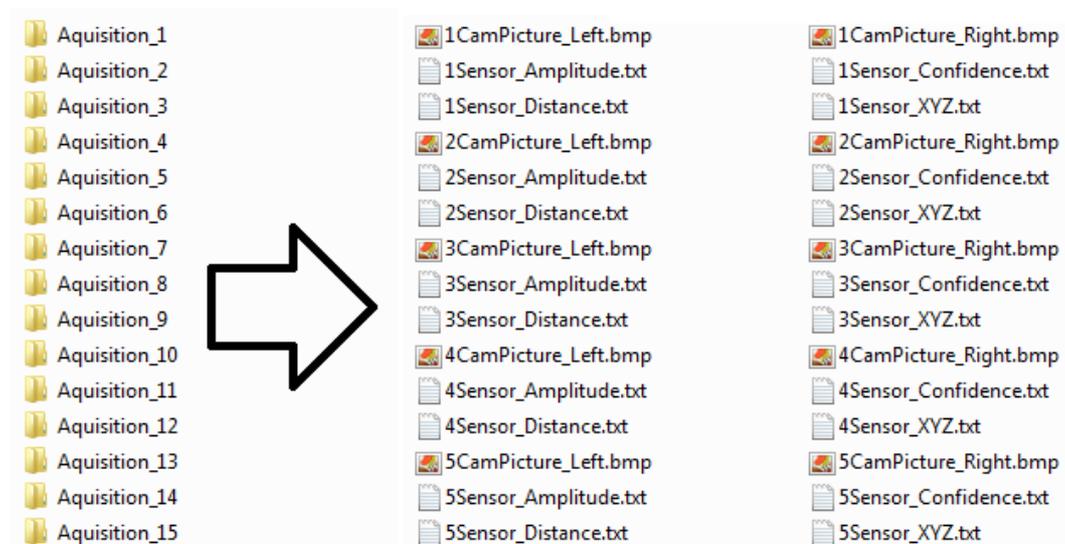
Nella figura precedente notiamo le immagini acquisite da due camere, già bilanciate per quanto riguarda il colore.

Il secondo blocco, posto nella parte medio/bassa dell'interfaccia, è associato al sensore ToF, e contiene:

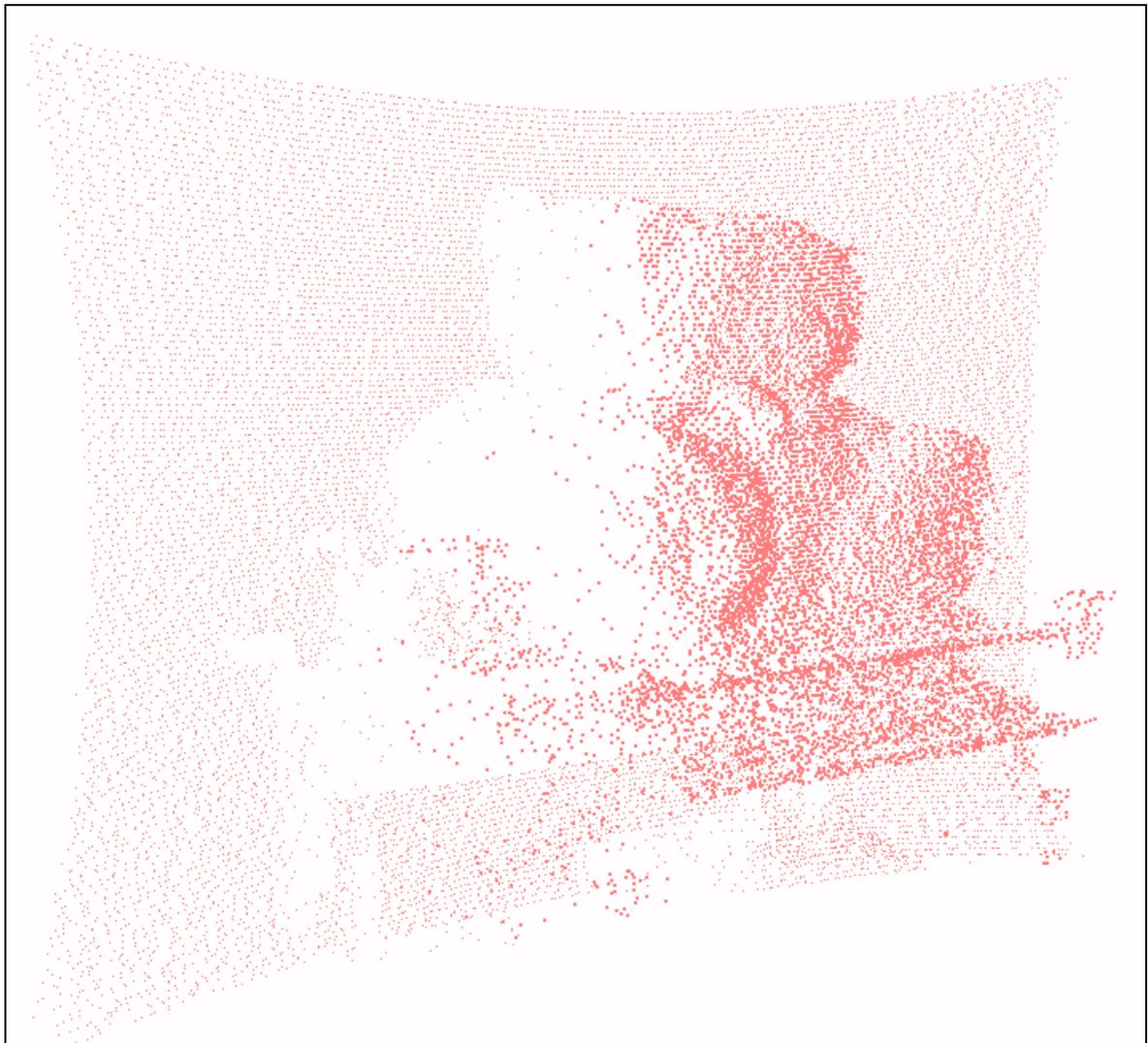
- Il pulsante Connect/Disconnect, che permette di instaurare la connessione con il sensore,
- Il pulsante Start Grabbing che permette di avviare il rilevamento e l'acquisizione dei dati da parte del sensore; con questo pulsante si richiama un form associato che permette di impostare i parametri del sensore prima dell'acquisizione,
- La visualizzazione delle tre mappe (o meglio le strutture dati matriciali) che rappresentano rispettivamente l'ampiezza, la distanza e la confidenza dei vari punti.

L'ultima parte dell'interfaccia utente, la più importante del programma, posizionata in fondo al form, è designata invece alla gestione dei flussi di dati acquisiti da parte del software, ovvero specifica se l'acquisizione deve essere a singoli frame o continua (composta da frame multipli). Essa garantisce la possibilità di acquisire contemporaneamente da tutti e tre i device.

Il software è organizzato in modo da creare una singola cartella per ogni acquisizione, distinguendo al suo interno i singoli frame catturati (se multipli), numerati sequenzialmente e distinti tra immagine della camera destra e immagine della camera sinistra, mappa delle ampiezze, mappa delle distanze grezze e standardizzate XYZ, ed infine la mappa di confidenza.



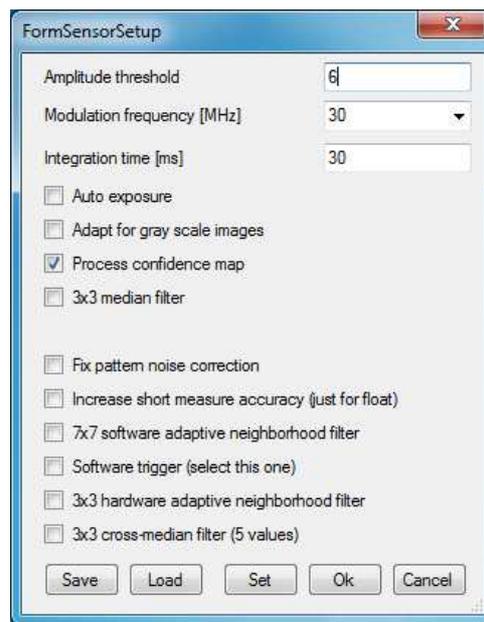
A titolo di esempio viene riportata un'immagine frutto dell'elaborazione di un dato file Sensor_XYZ.txt relativo alla mappa in dimensioni reali dell'ambiente circostante catturata dal sensore ToF.



5.2 Classe **SensorHelper**

La classe **SensorHelper**, che racchiude tutte le funzionalità relative al sensore Mesa SR4000, incorpora il wrapper per il linguaggio di programmazione C# delle librerie proprietarie della Mesa Imaging relative al suddetto device. Dispone inoltre di funzionalità quali una GUI per l'impostazione dei parametri

- Amplitude threshold, per evitare che valori troppo bassi di ampiezza possano confondere pesantemente l'attendibilità dei risultati (è una specie di eliminazione di rumore di fondo sulle distanze rilevate),
- La frequenza di modulazione, per la determinazione dell'intervallo di non-ambiguità (5 o 10 metri) e di aggiustamento nel caso di contemporaneità di più sensori (non è questo il caso),
- Il tempo di integrazione, per adeguare la durata di esposizione dei singoli punti del sensore al segnale riflesso,
- L'attivazione di tutti i possibili filtri selezionabili.



Il programma prevede inoltre un sistema di salvataggio su thread separati, in modo tale da consentire una memorizzazione dei dati sicura, asincrona e indipendente dal thread di acquisizione. In tal modo il salvataggio dei dati non influenzerà mai la velocità di acquisizione.

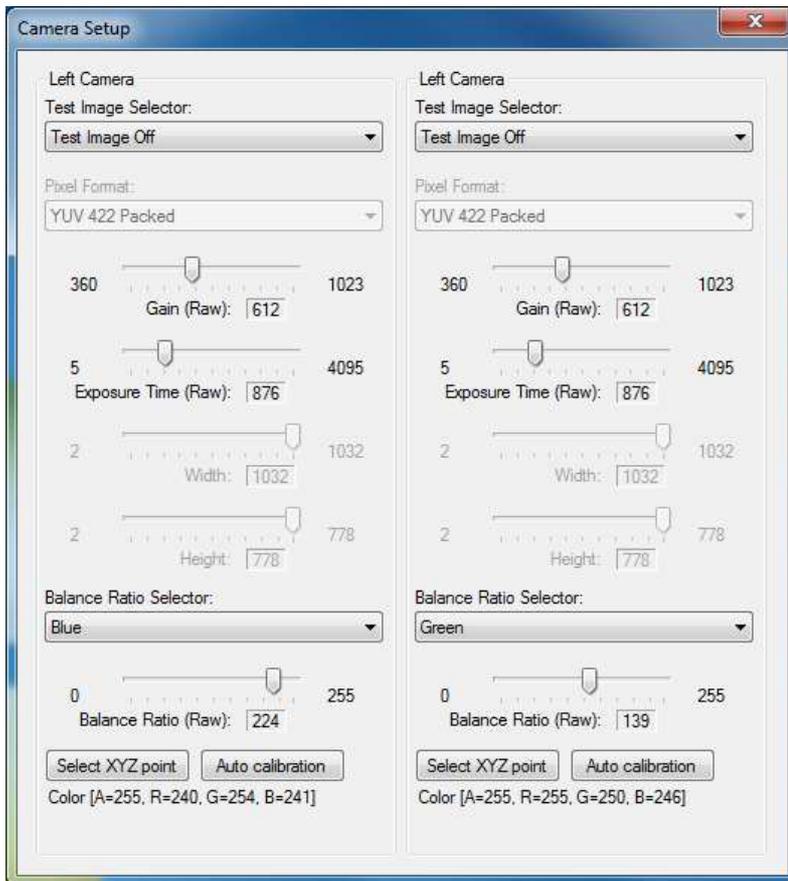
5.3 Classe **DoubleCamera**

Particolare attenzione va posta sulla classe **DoubleCamera** che gestisce le camere e ne implementa l'acquisizione e salvataggio dei frame. Si è deciso di non creare una generica classe 'camera' e di istanziarne due distinti oggetti, ma bensì di incorporare in un'unica classe entrambe le camere: questa scelta è stata dettata dalla necessità di garantire un'acquisizione contemporanea (invece di dover sincronizzare due eventuali flussi distinti) e di mantenere la classe compatta e solida.

La classe inizializza la libreria PylonC, messa a disposizione dalla Basler, casa produttrice delle due camere, e successivamente si appoggia al wrapper dei driver (sempre della Basler) che è implementata nel file PylonC.NETSupportLibrary.dll.

Inoltre permette sia la selezione delle camere in modo da impostare quale delle due dovrà essere la sinistra o la destra (indipendentemente dai collegamenti fisici), come l'impostazione dei parametri fondamentali per la calibrazione:

- Selezione del tipo di output utilizzato (rappresentato dal valore PixelFormat),
- Dimensione ed esposizione del frame acquisito,
- Regolazioni necessarie alla calibrazione del colore.



Anche questa classe dispone di un proprio sistema di salvataggio su thread separato, ed inoltre cura particolarmente la gestione della memoria liberandola il più possibile appena i dati sono stati salvati, per far sì che il programma possa funzionare adeguatamente anche su macchine non performanti.

6. Algoritmi di elaborazione

<Software di elaborazione dei dati grezzi, per loro interpretazione grafica, algoritmi di correzione dei dati in base alla mappa di confidenza, algoritmo di colorazione delle immagini, altri >

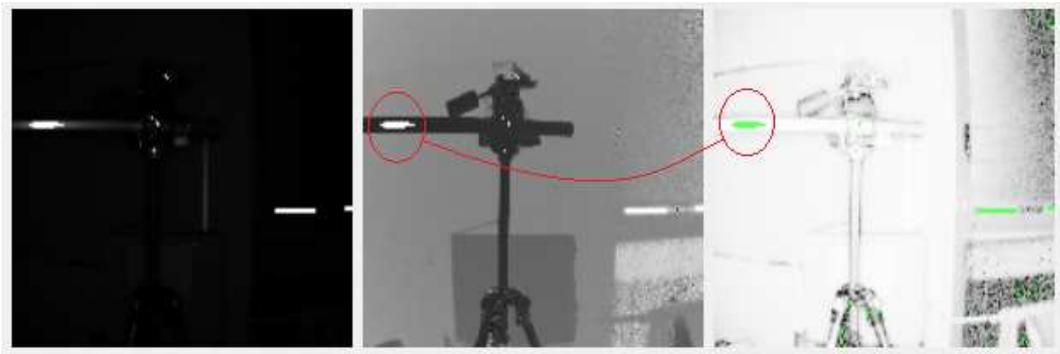
I dati ricevuti in input, sono di forma grezza sia per le camere a colori che per il sensore ToF, ma mentre per le telecamere non sono necessarie ulteriori operazioni per ottenere una prima visualizzazione abbastanza chiara tale da essere facilmente interpretabile anche solo visivamente, per i dati relativi al sensore a infrarossi sono necessarie alcune operazioni di elaborazione prima di poterli visualizzare.

Come evidenziato in precedenza, i dati ricevuti da tale sensore si ricevono sotto forma di vettori di valori 'float', di dimensione data da 'altezza x larghezza' della matrice di ricezione. Questi valori si riferiscono principalmente a quattro diverse gruppi di informazioni:

- Punti XYZ: ovvero tre vettori rappresentanti le coordinate spaziali in 3D, espressa in metri,
- Ampiezza: vettore rappresentante l'ampiezza del segnale ricevuto,
- Distanza: vettore rappresentante la distanza di ogni punto, espresso con un valore tra 0 e 65535,
- Confidenza: vettore rappresentante il valore che quantitativamente indica quanto punto sia da ritenere attendibile.

Le ultime tre classi di punti vengono automaticamente visualizzate dall'utente dopo una trasformazione che riproduce ogni vettore su di un'immagine in due dimensioni in scala di grigio, dove al valore più basso possibile viene assegnato

il colore nero, mentre il valore più alto possibile è contrassegnato con il bianco. Inoltre quando si incontra un valore nullo (ovvero 0), che rappresenterebbe un problema sia che si riferisca ad una ampiezza, ad una distanza o alla confidenza di un punto, il colore assegnato è il verde acceso, in modo da comunicare all'utente in maniera evidente che c'è un problema con quel punto.



Nell'acquisizione in figura si nota come l'elevato riflesso dell'oggetto inquadrato (e del bordo della finestra), produca una confidenza nulla (terza immagine da sinistra) e quindi una colorazione verde. Questo fenomeno segnala l'errore nell'interpretazione del punto, ed infatti l'immagine rappresentante la distanza, segnala tale zona di colore bianco, e quindi di massima distanza (evidentemente falso in quando c'è un naturale prolungamento dell'oggetto).

7. Risultati conseguiti e sviluppi futuri

|| *Analisi dei risultati ottenuti e proposte di ulteriori e successivi migliorie ed estensioni*

I risultati ottenuti con il software prodotto sono da considerarsi buoni.

Buoni in quanto il software risponde ai requisiti iniziali, ovvero acquisisce il set di frame desiderato, cattura egregiamente l'output del sensore Mesa, e permette una totale personalizzazione dei valori di acquisizione per poter effettuare in tempo reale tutte le prove desiderate.

L'aspetto negativo del programma è invece rappresentato da un fastidioso fenomeno di interruzione dell'acquisizione da parte delle camere a colori: si nota che con un basso tempo di intervallo tra l'acquisizione dei singoli frame (ovvero un alto valore di frame-rate) esse tendono a saturare il buffer interno e quindi tendono a rallentare il programma. Per ovviare a questo inconveniente bisognerà scegliere un intervallo tra frame abbastanza elevato (nell'ordine dei 100 millisecondi) che comporta quindi un utilizzo ancora rudimentale del sistema complessivo. In pratica, mentre l'acquisizione e l'elaborazione di singoli frame (immagini statiche e quindi ambiente circostante immobile) o con una frequenza di frame ridotta (ambiente circostante con mobilità limitata) funziona correttamente, nel caso di riprese di ambienti con una mobilità elevata (quali potrebbero essere un uomo in movimento) i risultati sono insoddisfacenti.

Tra gli sviluppi futuri quindi il primo passo fondamentale sarà quello di contattare direttamente il costruttore del sensore e verificare se, tramite comandi software, sia possibile svuotare il buffer interno (triggerabile dall'esterno oppure in modalità automatica) e risolvere così il problema della saturazione della memoria interna. Successivamente si può migliorare

l'interfaccia grafica, con l'aggiunta di un piccolo modulo di visualizzazione 3D, che potrà rappresentare i dati dei punti XYZ e garantire una migliore interpretazione dell'output (nonché del settaggio dei filtri).

Come ultima miglioria, si potrebbe predisporre un ulteriore modulo software di fusione tra i valori registrati dal sensore ToF e le foto a colori delle telecamere, individuandone le possibili correlazioni con gli oggetti inquadrati, e l'implementazione di filtri avanzati per l'eliminazione degli eventuali punti indesiderati, magari elaborando il tutto in real-time.

Bibliografia/Sitografia/Riferimenti

- [1] OpenCV, Open source Computer Vision, libreria orientata alla computer vision.
<http://opencv.org>

- [2] Visual C# (C sharp), linguaggio di programmazione orientato agli oggetti sviluppato da Microsoft, sia in versione CLI che visuale.
<http://msdn.microsoft.com/en-us/library/vstudio/ms228593.aspx>
<http://msdn.microsoft.com/it-it/library/kx37x362.aspx>

- [3] T.o.F., Time of Flight camera, sensore atto a rilevare le distanze dei punti che compongono l'inquadratura
http://en.wikipedia.org/wiki/Time-of-flight_camera

- [4] pacchetto completo contenente i driver ed il Software di Visualizzazione 3D
http://www.mesa-imaging.ch/dlm.php?fname=.%2Fcustomer%2Farchive%2F%2FSwissRanger_Customer_CD10.zip

- [5] pylon Camera Software Suite, raccoglie più framework per applicazioni web
<http://baslerweb.com/pylon>

- [9] libro che descrive con accuratezza i sensori ToF:
C. Dal Mutto, P. Zanuttigh, G. M. Cortelazzo (2013), Time-of-Flight Cameras and Microsoft Kinect™ – A user perspective on technology and applications, Springer
Miles Hansard, Seungkyu Lee, Ouk Choi,; Radu Horaud (2012)
Time-of-flight cameras: Principles, Methods and Applications
<http://hal.inria.fr/docs/00/72/56/54/PDF/TOF.pdf>

- [10] document web tratto da una tesi di C. Schaller, (2011), con breve trattazione sulla tecnologia ToF
<http://www.metrilus.de/range-imaging/time-of-flight-cameras/>

- [11] tesi di Yingzhu LI presso la University of Central Lancashire, Development of Immersive and Interactive Virtual Reality Environment for Two-Player Table Tennis, (2012)
[http://clou.uclan.ac.uk/5316/2/Li_Yingzhu_Final_e-Thesis_\(Master_Copy\).pdf](http://clou.uclan.ac.uk/5316/2/Li_Yingzhu_Final_e-Thesis_(Master_Copy).pdf)

