



UNIVERSITÀ DEGLI STUDI DI SALERNO

Dottorato in Informatica e Ingegneria dell'Informazione  
Curriculum Informatica

Coordinatore:  
Prof. Alfredo De Santis  
XV Nuovo Ciclo

About the Development of Visual Search Algorithms and  
their Hardware Implementations

Relatori:  
Ch.mo Prof. Giancarlo Raiconi  
Ph.D. Mario Vigliar

Candidato:  
Luca Puglia  
Matr. 8888100007

A handwritten signature in black ink, appearing to read 'Luca Puglia', is written in a cursive style.

Anno Accademico 2016/2017

## Abstract

L'obiettivo principale di questo lavoro e' quello di esplorare i benefici di una implementazione hardware per una pipeline di visual search 3D. Il termine visual search si riferisce al problema di ricerca di oggetti nell'ambiente. L'object recognition ai giorni nostri e' principalmente basato sull'uso di descrittori della scena, una descrizione univoca per i punti salienti. Questo compito e' stato implementato per anni utilizzando immagini: il descrittore di un punto dell'immagine e' un semplice vettore di caratteristiche. Accoppiando i descrittori presenti in differenti viste della stessa scena permette di trovare punti nello spazio visibili da entrambe le viste. Chiaramente, utilizzando immagini 2D non e' possibile avere descrittori che sono robusti a cambiamenti della prospettiva, per questo motivo, molte tecniche sono state proposte per risolvere questo problema utilizzando descrittori 3D. Questa categoria di descrittori si avvale di 3D point cloud e mappe di disparita'. Ovviamente tecniche piu' sofisticate sono necessarie per ottenere la rappresentazione 3D della scena. Il campo dell'acquisizione 3D e' molto vasto ed e' possibile distinguere tra due categorie di sensori: sensori attivi e passivi. Tra i sensori attivi possiamo annoverare dispositivi in grado di proiettare un pattern di luce infrarossa sulla scena, questo pattern noto presenta delle variazioni dovute agli oggetti presenti nella scena. Una camera infrarossa riceve l'immagine distorta del pattern e deduce la geometria della scena. Questo tipo di dispositivi non sono molto efficienti dal punto di vista energetico dato che un sacco di corrente viene consumata per proiettare il pattern. Un altro modo per ottenere un modello 3D e' quello di usare sensori passivi, una coppia di telecamere puo' essere utilizzata per ottenere informazioni utilizzando metodi di triangolazione. Questi metodi pero' richiedono un sacco di potenza computazionale nel caso di applicazioni real time, per questo motivo e' necessario utilizzare dispositivi ad-hoc quali architetture hardware dedicate implementate mediante l'uso di FPGA e ASIC.

In questo lavoro ho esplorato gli step principali di una pipeline per la visual search composta da un sistema di visione 3D e uno per la descrizione di punti. Entrambi i sistemi si avvalgono di architetture hardware dedicate

prototipate in RTL e implementate su FPGA. Questo e' un grosso campo di lavoro e provo ad esplorare i benefici di una implementazione hardware per l'accelerazione degli algoritmi stessi e il risparmio di energia elettrica.