



UNIVERSITÀ DEGLI STUDI DI PADOVA

---

Dipartimento di Ingegneria dell'Informazione

Corso di Laurea Triennale in Ingegneria Informatica

Tesi di Laurea

FUSIONE DI DATI  
STEREO E TIME-OF-FLIGHT MEDIANTE  
TECNICHE DI DEEP LEARNING

Relatore:

Prof. PIETRO ZANUTTIGH

Laureando:

FRANCESCO PHAM

---

ANNO ACCADEMICO 2018-2019

Ai miei genitori,  
che mi hanno sempre sostenuto.

# Indice

<b>1</b>	<b>Introduzione</b>	<b>1</b>
1.1	Descrizione del progetto . . . . .	2
1.2	Principi di funzionamento del sistema di acquisizione . . . . .	2
1.2.1	Sistema di visione stereo . . . . .	2
1.2.2	Sensore Time-Of-Flight . . . . .	2
1.3	Il Machine learning . . . . .	2
1.3.1	Reti neurali artificiali . . . . .	2
1.3.2	Reti neurali convoluzionali . . . . .	2
<b>2</b>	<b>Acquisizione e pre-elaborazione dei dati</b>	<b>3</b>
2.1	Preparazione dei dati ToF . . . . .	3
2.1.1	Riproiezione e interpolazione . . . . .	3
2.1.2	Calcolo della disparità . . . . .	3
2.2	Elaborazione dei dati stereo . . . . .	3
2.2.1	Disparità stereo . . . . .	3
<b>3</b>	<b>Fusione tramite deep learning</b>	<b>4</b>
3.1	Le mappe in input . . . . .	4
3.2	Convoluzione dilatata . . . . .	4
3.3	Reti neurali residuali . . . . .	4

<b>INDICE</b>	<b>ii</b>
3.4 Confronti tra le varie architetture . . . . .	4
3.5 La rete neurale selezionata . . . . .	4
<b>4 Analisi dei risultati</b>	<b>5</b>
4.1 Tensorflow . . . . .	5
4.2 Il processo di training . . . . .	5
4.3 Valutazione dei risultati . . . . .	5
4.3.1 Valutazione sul dataset sintetico . . . . .	5
4.3.2 Valutazione sul dataset reale . . . . .	5
<b>5 Conclusioni</b>	<b>6</b>
<b>Bibliografia</b>	<b>7</b>

# Capitolo 1

## Introduzione

La stima della profondità ha da sempre rappresentato un problema di massimo interesse. L'informazione sulla profondità è importante, ed in alcuni casi essenziale, per molteplici applicazioni pratiche della visione artificiale: guida autonoma, robotica, ricostruzione 3D e realtà aumentata sono soltanto alcuni. L'acquisizione delle geometrie tridimensionali di scene del mondo reale rappresenta da sempre un problema complesso, e gli strumenti sono stati accessibili soltanto in grosse compagnie e centri di ricerca. Nel corso degli anni, molte tecniche sono state sviluppate e nuovi dispositivi, dai costi più ridotti, sono stati introdotti nel mercato.

Le principali tecnologie per la percezione dell'ambiente sfruttano sensori ottici che catturano la luce visibile o infrarossa che viene riflessa dalla scena. Questo lavoro di tesi tratta due categorie di sensori in particolare, entrambi molto diffusi e accessibili: i sensori a visione stereoscopica e i sensori basati sul tempo di volo (Time-Of-Flight, ToF).

Nello stesso periodo, nel mondo della computer vision, il machine learning (e deep learning in particolare) si è dimostrato uno strumento che ha ampiamente incrementato le prestazioni rispetto ad altri metodi deterministici, in molti

problemi ritenuti complessi, fino a superare le prestazioni umane. E proprio grazie al machine learning è stato possibile sviluppare sistemi efficaci per la ricostruzione delle informazioni catturate dai sensori.

## **1.1 Descrizione del progetto**

L'obiettivo del lavoro di questa tesi è lo sviluppo di un sistema di machine learning per la fusione dei dati tridimensionali forniti dai sensori, cercando di fornire una ricostruzione più accurata. La scelta dei sensori gioca perciò un ruolo fondamentale: le telecamere stereo e il sensore ToF tendono ad avere caratteristiche complementari, e la loro combinazione si è rivelato efficace.

## **1.2 Principi di funzionamento del sistema di acquisizione**

### **1.2.1 Sistema di visione stereo**

### **1.2.2 Sensore Time-Of-Flight**

## **1.3 Il Machine learning**

### **1.3.1 Reti neurali artificiali**

### **1.3.2 Reti neurali convoluzionali**

## Capitolo 2

# Acquisizione e pre-elaborazione dei dati

### 2.1 Preparazione dei dati ToF

#### 2.1.1 Riproiezione e interpolazione

#### 2.1.2 Calcolo della disparità

### 2.2 Elaborazione dei dati stereo

#### 2.2.1 Disparità stereo

## Capitolo 3

# Fusione tramite deep learning

3.1 Le mappe in input

3.2 Convoluzione dilatata

3.3 Reti neurali residuali

3.4 Confronti tra le varie architetture

3.5 La rete neurale selezionata



# Capitolo 4

## Analisi dei risultati

### 4.1 Tensorflow

### 4.2 Il processo di training

### 4.3 Valutazione dei risultati

#### 4.3.1 Valutazione sul dataset sintetico

#### 4.3.2 Valutazione sul dataset reale

Capitolo 5

Conclusioni

# Bibliografia

- [1] Geitner F.K. Block H.P. *Machinery Failure Analysis and Troubleshooting*. 1999.
- [2] R. Zaltieri D. De Pasquale, L. Troiano. A bayesian approach supporting troubleshooting.
- [3] K. Rommelse D. Heckerman, J. S. Breese. Troubleshooting under uncertainty. Technical report, 1994.
- [4] J. Gertler. *Fault detection and diagnosis in engineering systems*. CRC Press, 1998.
- [5] T. Gustavsson. Troubleshooting using cost effective algorithms and bayesian networks. Master's thesis.
- [6] Z. Kowalczyk J. Korbicz, J. M. Koscielny. *Fault diagnosis: models, artificial intelligence, applications*. Springer, 2004.
- [7] R. M. Jagt. *Support for Multiple Cause Diagnosis with Bayesian Networks*. PhD thesis.
- [8] F. V. Jensen. *Bayesian Networks and Decision Graphs*. 2000.
- [9] R. D. Braatz L. H. Chiang, E. Russell. *Fault detection and diagnosis in industrial systems*. Springer, 2001.

- [10] D. Lesage N. Burrus. Theory of evidence (draft). Technical report.
- [11] Hobbs A. Reason J. *Managing Maintenance Error*. 2003.
- [12] K. Sentz. *Combination of Evidence in Dempster-Shafer Theory*. PhD thesis.