



UNIVERSITÀ  
DEGLI STUDI  
FIRENZE

SCUOLA DI INGEGNERIA

CORSO DI LAUREA IN INGEGNERIA INFORMATICA

---

# MIGLIORAMENTO DEL SISTEMA DI ANNOTAZIONE WATSS CON TECNICHE DI COMPUTER VISION

**Basi di Dati Multimediali**

Lorenzo Cioni

---

ANNO ACCADEMICO 2015/2016

# Indice

<b>Indice</b>	<b>i</b>
<b>Introduzione</b>	<b>1</b>
<b>1 Stato dell'arte</b>	<b>2</b>
1.1 Analisi comparativa di alcuni sistemi di annotazione . . . . .	2
1.1.1 LabelMe . . . . .	2
1.1.2 Viper-GT . . . . .	3
1.1.3 VATIC . . . . .	3
1.1.4 WATSS . . . . .	3
<b>2 Obiettivi</b>	<b>4</b>
<b>3 Pedestrian tracking</b>	<b>5</b>
3.1 Motion detection . . . . .	5
3.2 Pedestrian detection . . . . .	5
3.3 Filtro di Kalman . . . . .	5
<b>4 Implementazione</b>	<b>6</b>
<b>Conclusioni</b>	<b>7</b>
<b>Bibliografia</b>	<b>8</b>

# Introduzione

L'obiettivo di questo elaborato è di migliorare, anche tramite tecniche della *Computer Vision*, il sistema di annotazione web *WATSS* [1].

WATSS, abbreviazione per *Web Annotation Tool for Surveillance Scenarios*, è un sistema di annotazione web per la creazione di un groundtruth di scenari di sorveglianza. Il sistema consente infatti di annotare persone all'interno dei singoli frame di un video, assegnandogli una posizione (determinata tramite una *bounding box*), una identità (tramite *avatar*), la parte visibile e le orientazioni del corpo e dello sguardo. E' inoltre possibile associare più persone ad un medesimo gruppo e il punto di interesse presso il quale la persona si trova.

Uno dei principali obiettivi è quello di introdurre nel sistema un meccanismo di predizione delle annotazioni, andando a generare, a partire da una o più annotazioni consecutive di una stessa persona, una serie di *proposals* per i frames successivi. L'elaborazione dell'immagine a questi scopi viene effettuata tramite l'utilizzo di OpenCV, una libreria open source, nativa per C++, per la Computer Vision e l'Image Analysis.

Nelle sezioni successive viene presentata inizialmente un'analisi comparativa tra i vari sistemi di annotazione che vanno a costituire l'attuale stato dell'arte. Viene poi presentata la parte relativa alle migliorie apportate al sistema e le tecniche di Computer Vision utilizzate. Infine viene presentata un'analisi di usabilità a posteriori, mettendo in evidenza anche eventuali sviluppi futuri per l'applicazione.

# Capitolo 1

## Stato dell'arte

In questa sezione viene presentata un'analisi comparativa di alcuni dei più famosi sistemi di annotazione esistenti. Lo scopo di questa ricerca è quello di individuare le caratteristiche comuni ai vari strumenti e le loro limitazioni. Da questa analisi poi è stata stilata una lista di requisiti che portano WATSS ad essere in accordo con gli altri sistemi introducendo allo stesso tempo nuove caratteristiche.

### 1.1 Analisi comparativa di alcuni sistemi di annotazione

L'analisi dei sistemi si è incentrata principalmente su 3 strumenti open source esistenti: *LabelMe*, *ViPER-GT* e *VATIC*. Per ciascuno dei sistemi è stata stilata una lista di caratteristiche offerte ed evidenziate le eventuali limitazioni. Infine è presentata un'analisi anche con il sistema WATSS.

#### 1.1.1 *LabelMe*

*LabelMe* è un sistema Web che consente l'annotazione di oggetti all'interno di immagini. Le singole annotazioni sono effettuate mediante la definizione di aree poligonali nell'immagine e l'assegnazione di una label. Il tool offre la possibilità di indicare se un oggetto annotato è occluso o meno da altri oggetti presenti nella scena (non consente però di individuare la parte occlusa o visibile).

Le annotazioni possono essere annidate, è possibile dunque etichettare oggetti che sono inclusi gli uni negli altri.

In aggiunta alle annotazioni di oggetti il sistema consente di annotare intere aree dell'immagine: questo è reso possibile andando inizialmente a delimitare una porzione di immagine ed associando ad essa una label. In questo caso l'area così definita viene colorata interamente ed è necessario stabilire se si tratta di un'area interna o esterna.



Figura 1.1: Annotazione di un'immagine tramite LabelMe

In fase di esportazione delle annotazioni viene generata una struttura in formato XML che è possibile importare nuovamente in un'altra immagine.

Il sistema non prevede la possibilità di generare *proposals* per le annotazioni, tutto il lavoro è a carico dell'utente.

### 1.1.2 ViPER-GT

ViPER-GT è un sistema di annotazione per video e la generazione di un *ground-truth*.

ViPER, the Video Performance Evaluation resource, is a set of tools that make video and document algorithm evaluation possible. The two main components, ViPER-Ground Truth and ViPER-Performance Evaluation, provide the ability to mark up videos and documents with ground truth and the ability

to compare result data with ground truth, respectively. The extended toolkit, which currently only works on Solaris, provides a more complete solution, with scripting and the ability to generate charts.

is a video annotation tool, a viewer of algorithmically generated markup, a tool for assisting performance evaluation of such markup, and more. This document presents the programs, and details both what it does and does not do. It first contains a tutorial, then a reference, and finally a set of useful information for developers who wish to extend ViPER-GT or use its data format.

### ***1.1.3 VATIC***

### ***1.1.4 WATSS***

## Capitolo 2

### Obiettivi

# Capitolo 3

## Pedestrian tracking

### 3.1 Motion detection

### 3.2 Pedestrian detection

### 3.3 Filtro di Kalman



# Capitolo 4

## Implementazione

# Conclusioni

Conclusioni dell'elaborato

# Bibliografia

- [1] F. Bartoli, L. Seidenari, G. Lisanti, S. Karaman, and A. Del Bimbo. Watts: A web annotation tool for surveillance scenarios. In *Proceedings of the 23rd ACM International Conference on Multimedia*, MM '15, pages 701–704, New York, NY, USA, 2015. ACM.