



Agenzia Nazionale per le Nuove Tecnologie,
l'Energia e lo Sviluppo Economico Sostenibile



Ministero dello Sviluppo Economico

RICERCA DI SISTEMA ELETTRICO

Analisi immagini di scene urbane per applicazioni di smart lighting e smart city

M. Annunziato, C. Meloni, F. Moretti, S. Pizzuti, F. Bucci

ANALISI IMMAGINI DI SCENE URBANE PER APPLICAZIONI DI SMART LIGHTING E SMART CITY

M. Annunziato, C. Meloni, S. Pizzuti (ENEA)

F. Moretti (Università degli Studi "Roma Tre", Dipartimento di Informatica e Automazione)

F. Bucci (Università la Sapienza, Dipartimento di Ingegneria Astronautica, Elettrica ed Energetica)

Novembre 2011

Report Ricerca di Sistema Elettrico

Accordo di Programma Ministero dello Sviluppo Economico – ENEA

Area: Razionalizzazione e risparmio nell'uso dell'energia

Progetto: Tecnologie "smart" per l'integrazione della illuminazione pubblica con altre reti di servizi energetici e loro ottimizzazione

Responsabile Progetto: Mauro Annunziato, ENEA

Riferimenti

Questo rapporto tecnico si riferisce all'Accordo di Programma tra Ministero dello Sviluppo Economico ed ENEA, nel contesto della Ricerca di Sistema il cui obiettivo è il miglioramento del sistema elettrico nazionale sia in termini di nuove tecnologie che di efficienza energetica.

Il rapporto si inquadra nell'Area "Razionalizzazione e risparmio nell'uso dell'energia elettrica", nell'ambito del PAR 2010, in riferimento alla tematica di ricerca "Risparmio di energia elettrica nell'illuminazione pubblica" - Progetto 3.2.1, "Tecnologie smart per l'integrazione della illuminazione pubblica con altre reti di servizi energetici e loro ottimizzazione" - Task C, punto a) "Elaborazione delle immagini provenienti dai lampioni intelligenti con capacità di analisi della scena, conteggio dei veicoli nelle diverse corsie, analisi delle velocità medie, discriminazione dei mezzi pubblici, conteggio dei flussi delle persone passanti. Sperimentazione su registrazioni raccolte in aree urbane".

L'obiettivo prefissato è quello di realizzare delle metodologie che consentano di usufruire dell'infrastruttura tecnologica Smart riguardante l'illuminazione pubblica, i *lampioni intelligenti* e la relativa rete di connessione, per fornire servizi di infomobilità.

Il rapporto si focalizza in particolare su metodologie basate sull'elaborazione delle immagini delle scene urbane fornite dai lampioni intelligenti. Le informazioni rilevate dalle elaborazioni consentono quindi di erogare una serie di servizi orientati al risparmio energetico ed utili ai cittadini.

Sommario

1. Introduzione	6
1.1. I risultati dello studio di fattibilità sulla illuminazione adattiva effettuato nel PAR 2007	7
1.2 I risultati della analisi tecnico-economica del PAR 2008-2009	7
1.3 L'attività svolta nel contesto del PAR 2010 sullo sviluppo dell'analisi delle immagini	8
2. Sviluppo del sistema real time per l'analisi della scena	9
2.1 Struttura dell'applicativo per l'analisi delle immagini	9
2.2 Implementazione del sistema di analisi immagine	13
2.2.1 Estrazione del background	13
2.2.2 Interfaccia Grafica	14
2.3 Release dell'applicativo	16
2.3.1 Suddivisione in aree di interesse	16
2.3.2 Filtro sui bounding box	17
2.3.3 Diagramma di flusso del sistema di elaborazione	17
2.3.4 Scrittura degli output su file	19
2.3.5 Interfaccia grafica	20
3 Sviluppo degli indici di traffico	21
4 Sperimentazione	22
4.1 Aurelia	23
4.2 Stazione Centrale di Napoli	24
4.3 Roma Torre Vecchia	26
4.4 Riepilogo sperimentazioni	28
5 Conclusioni	28
6 Bibliografia	29

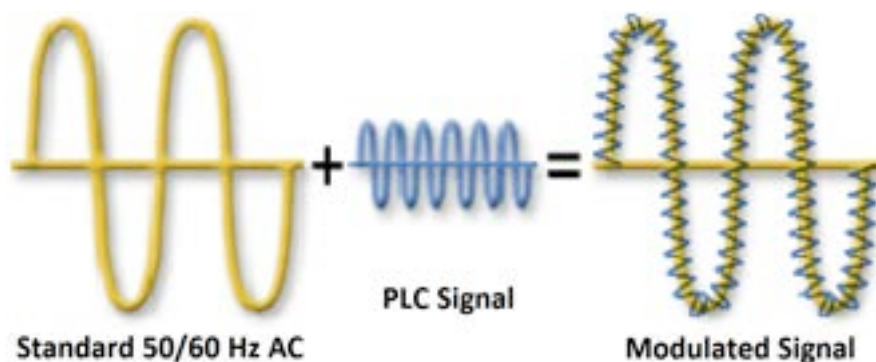
Indice delle figure

<u>Figura 1. Diagramma del processo del primo prototipo.....</u>	<u>10</u>
<u>Figura 2. Esempio di output in formato testo</u>	<u>12</u>
<u>Figura 3. Interfaccia Grafica.....</u>	<u>14</u>
<u>Figura 4. Analisi della correlazione tra due aree di interesse.....</u>	<u>17</u>
<u>Figura 5. Diagramma di flusso dell'applicativo.....</u>	<u>18</u>
<u>Figura 6. Frame Acquisito (a sinistra) e Background Corrente (destra).....</u>	<u>18</u>
<u>Figura 7. Frame filtrato (sinistra) e frame con contorni e bounding box estratti (destra).....</u>	<u>19</u>
<u>Figura 8. Esempio di output delle features su file di testo</u>	<u>19</u>
<u>Figura 9. Esempio di output degli indici su file di testo</u>	<u>19</u>
<u>Figura 10. Interfaccia utente</u>	<u>20</u>
<u>Figura 11. Conteggio tramite superamento della soglia.....</u>	<u>22</u>
<u>Figura 12. Aurelia.....</u>	<u>23</u>
<u>Figura 13. Stazione Centrale di Napoli.....</u>	<u>24</u>
<u>Figura 14. Soglia di attraversamento stazione centrale di Napoli.....</u>	<u>25</u>
<u>Figura 15. Roma, via Torre Vecchia</u>	<u>26</u>
<u>Figura 16. Soglia di attraversamento via Torre Vecchia.....</u>	<u>27</u>

1. Introduzione

La tecnologia principale su cui poggiano le metodologie presentate nel rapporto è il lampione intelligente con tele gestione e trasmissione dati in banda larga dotato di sensori di visione per svolgere attività di rilevamento del traffico della scena urbana ed altre funzionalità aggiuntive. Tale tipologia di lampione consente il controllo adattivo della potenza dell'illuminazione a distanza, sfruttando la tecnologia della *power line communication (PLC) a banda larga* la quale è in grado di portare al lampione una linea digitale in grado di supportare l'invio di immagini ad alta risoluzione.

PLC è una tecnologia che permette di trasferire dati in forma digitale, utilizzando una classica linea elettrica trasformandola in un supporto di comunicazione. Consiste nell'aggiungere al segnale elettrico in bassa frequenza (50/60 Hz) uno nuovo a frequenza più alta. Esistono due tipologie di tecnologie PLC, a bassa velocità, con frequenze intorno ai 112 kHz, e ad alta velocità, con frequenze tra i 2 e i 30 MHz.



Per quanto riguarda la sensoristica di rilevamento della attività stradale esistono diverse metodologie al fine di stabilire in maniera ragionevolmente precisa la quantità di traffico, sia veicolare che pedonale, rilevata nelle strade pubbliche.

- Sensori Piezoelettrici
- Sensori Infrarossi
- Sensori Ultrasuoni
- Radar
- Telecamere

I sensori piezoelettrici rilevano la pressione esercitata su di essi, quindi ponendo tali sensori in maniera opportuna lungo il manto stradale è possibile conteggiare con una buona precisione il numero di veicoli. I sensori a infrarossi si basano sulla radiazione infrarossa per la rilevazione di oggetti in movimento. I sensori a ultrasuoni si basano sul rilevamento delle onde meccaniche sonore. I sensori radar si basano sulle onde elettromagnetiche.

Le telecamere rispetto ai sensori precedentemente descritti offrono la possibilità di acquisire informazioni aggiuntive rispetto al solo conteggio dei veicoli. Infatti sebbene i sensori basati sui fenomeni fisici garantiscano una precisione maggiore nel rilevamento degli oggetti in movimento in determinate situazioni, spesso non consentono di acquisire informazioni di primaria importanza della scena urbana monitorata.

Il connubio tra la tecnologia PLC e la sensoristica del lampione intelligente ha come obiettivo la realizzazione di una metodologia di illuminazione pubblica adattiva: erogare una potenza sulle lampade dei lampioni proporzionale alla domanda, ovvero alla quantità di traffico rilevata. Tale metodologia è stata realizzata da ENEA creando un sistema intelligente che elabora i dati provenienti dai lampioni, trasmessi attraverso la PLC ad elaboratori e fornisce al variare del traffico la relativa potenza da erogare a ciascun lampione.

Tale metodologia quindi si basa su tre moduli fondamentali di seguito descritti.

1. Sistema di analisi immagini: effettua delle elaborazioni real time delle immagini per produrre degli indici di traffico nelle postazioni dove sono presenti sistemi di visione.
2. Sistema predittivo della smart street: gli indici prodotti dal sistema di analisi immagini creando dei modelli previsionali adattivi per predire la situazione di traffico nell'ora successiva in ogni punto della smart street dove è localizzato un lampione tele controllato; essendo modelli adattivi si adattano in maniera ragionevolmente repentina alle variazioni della quantità di traffico, situazione abbastanza usuale soprattutto in situazioni di traffico urbano.
3. Sistema di controllo della smart street: elabora, per ogni lampione, la quantità di potenza necessaria al lampione nel breve e nel lungo termine compatibilmente con le normative, i flussi di veicoli e di persone, la categoria della strada, gli agenti meteorologici, la autodiagnostica del sistema stesso.

Nel presente rapporto viene discusso in particolare il sistema di analisi immagini (punto 1), partendo dai primi sviluppi (par 2007 e par 2008-2009), l'implementazione, le metodologie per la produzione degli indici, le problematiche riscontrate e le sperimentazioni effettuate.

1.1. I risultati dello studio di fattibilità sulla illuminazione adattiva effettuato nel PAR 2007

L'idea di utilizzare l'analisi delle immagini di sensori di visione installati sui pali della illuminazione pubblica al fine di regolare attivamente il flusso luminoso fu introdotta da ENEA ed Università di Roma Tre per la prima volta nel rapporto "SVILUPPO DI UN SISTEMA DI CONTROLLO INTEGRATO ED ADATTIVO PER L'ILLUMINAZIONE PUBBLICA", Autori: Fabio Moretti, Mauro Annunziato, Stefano Panzieri, pubblicato a settembre 2010. Report RdS/2010/248.

http://www.enea.it/it/Ricerca_sviluppo/documenti/ricerca-di-sistema-elettrico/illuminazione-pubblica/16-rapporto-tecnico-enea-un-rm3.pdf

In questo rapporto furono rilevati i limiti principali che si evidenziano nell'uso dei sistemi di regolazione del flusso che in sintesi potremmo riferire al fatto che nella maggioranza dei casi l'informazione dei flussi reali di veicoli e persone non sono conosciuti se non a livello statistico e molto spesso con grande aleatorietà. Questa situazione rischia di non sfruttare a pieno il potenziale di risparmio energetico e nei casi limite porre il sistema illuminotecnico temporaneamente fuori norma quando si verificano picchi di attività sulla strada al di sopra delle condizioni normali nelle ore centrali della notte.

Nel rapporto fu sviluppato un approccio complessivo e gettate le basi per una prima forma prototipale di analisi delle immagini che in tempo reale potesse produrre dei dati di attività della strada utili per il controllo (smart sensor). Venne ipotizzata una struttura algoritmica, descritta in grandi linee nel successivo paragrafo 2.1.

1.2 I risultati della analisi tecnico-economica del PAR 2008-2009

Il successivo sviluppo (par 2008-2009) si concentrò sugli algoritmi per il sistema di predizione della illuminazione (in coerenza con le normative esistenti) di una "Smart Street" in cui i lampioni sono "intelligenti e multifunzionali", ossia equipaggiati con sensori di varia natura ed attraverso sistemi di comunicazione avanzati interagiscono con un sistema intelligente in grado di ritracciare continuamente il profilo di attività della strada in base al quale attivare una regolazione adattiva ed automatica per l'intero anno, del flusso luminoso punto-punto. Il documento "Sviluppo sistemi intelligenti per la gestione della Smart Street". Autori M. Annunziato, F. Bucci, C. Meloni, F. Moretti, S. Pizzuti. Report RdS/2011/198

http://www.enea.it/it/Ricerca_sviluppo/documenti/ricerca-di-sistema-elettrico/illuminazione-pubblica/rds-179.pdf illustra i risultati ottenuti articolati su quattro punti base:

- Sviluppo di una metodologia di modellazione della domanda di illuminazione in termini di predizione dei flussi di traffico a breve termine (1 ora) per arterie stradali. Tale sistema predittivo va associato ai

risultati delle analisi delle immagini e costituisce l'informazione di base su cui il sistema di controllo deve elaborare le strategie (passo 2 della metodologia).

- Sviluppo di una metodologia di controllo adattivo del flusso luminoso in relazione alla richiesta dell'utenza coerente con le normative in vigore (primo prototipo passo 3 della metodologia).
- Validazione della metodologia sviluppata su dati reali di flussi di traffico. La validazione è stata condotta su dati provenienti da Terni ed hanno evidenziato dei risparmi energetici molto elevati (da 30 al 45 %) rispetto al controllo "statico" (non adattivo attualmente diffuso che registra un risparmio energetico intorno al 15 % per effetto della regolazione di flusso) e delle aumenti di sicurezza molto elevati rispetto all'approccio "tutta notte – mezzanotte" attualmente molto diffuso benché piuttosto rischioso nel caso che l'impianto non sia sovradimensionato.
- Comparazione tecnico-economica dell'approccio proposto rispetto all'approccio di mercato più diffuso su una realtà urbana di medie dimensioni (Castelnuovo Magra). Tale analisi ha mostrato tempi di ritorno dell'investimento e VAN di qualità superiore rispetto all'approccio del controllo a tele gestione punto-punto non adattivo e di qualità del servizio molto più interessante rispetto all'approccio del controllo di linea (grazie ai servizi smart offerti sul monitoraggio del traffico e sulla sicurezza).

1.3 L'attività svolta nel contesto del PAR 2010 sullo sviluppo dell'analisi delle immagini

L'attività svolta nella attuale fase si è articolata su due linee di attività. La prima si è concentrata sulla realizzazione del sistema di controllo della illuminazione sulla base di significativi miglioramenti del prototipo sviluppato nel PAR 2008-2009. Uno specifico rapporto del PAR 2010 è dedicato a questo tema.

La seconda linea di attività si riferisce al completamento della analisi di immagine di un singolo sistema di visione di cui fu sviluppato un prototipo nel par 2007 con la successiva conversione dei dati di elaborazione di base (relativi ad oggetti in movimento) in "indici di attività" utilizzabili per il sistema predittivo e per il sistema di controllo. In questa fase si è quindi passati da un prototipo off-line ad un sistema robusto per analisi di flussi continui di immagini con capacità di autoregolazione, predisposizione ad operare su tempi lunghi e senza supervisione, capacità di generazione real time degli indici di attività della strada.

Tale sviluppo ha richiesto un significativo sforzo sperimentale compiuto sia in condizioni notturne che in condizioni diurne per valorizzare il servizio di monitoraggio del traffico anche aldilà degli aspetti di controllo della illuminazione pubblica.

La sperimentazione evidenzierà dei limiti ancora presenti in situazioni di traffico molto intenso, per problematiche dovute a situazioni tipiche come la sovrapposizione di due veicoli e la proiezione delle ombre. Di conseguenza tali problematiche sono rimaste ancora irrisolte ed il sistema risulta ulteriormente migliorabile. Si può dire che benché il sistema racchiude tutte le funzionalità necessarie alla analisi del singolo palo intelligente ancora non ha raggiunto la maturità di rispondere a tutta la ricchezza di avvenimenti che possono verificarsi su strade di diversa natura (discriminazione dell'ombra, rimozione disturbi dovuti a condizioni meteo quali pioggia o neve o vibrazioni dovute al vento).

Tali elaborazioni sono rimandate ad un successivo sviluppo che prenda in considerazione anche la possibilità di includere altri servizi smart tra cui:

- Supporto efficiente per la gestione del traffico urbano con classificazione veicoli e persone
- Videosorveglianza
- Rilevamento di inquinamento acustico e ambientale
- Rilevamento condizioni meteorologiche
- Rilevamento e modellazione parcheggi
- Rilevamento criticità di traffico, di incidenti, di flussi di persone

Tali tipologie di funzionalità possono essere fornite solamente a seguito di un accurata elaborazione delle immagini fornite da telecamere posizionate in punti strategici dell'area urbana, distribuite in maniera capillare ed uniforme.

2. Sviluppo del sistema real time per l'analisi della scena

Il sistema di elaborazione delle immagini è un applicativo software realizzato in linguaggio C/C++ e si basa sulle librerie OpenCV per quanto riguarda l'elaborazione delle immagini sulle librerie Vlc per la connessione con le telecamere IP e sulle librerie QT per quanto riguarda l'interfacce grafiche.

Al fine di ottenere un servizio di monitoraggio del traffico è stato realizzato un applicativo dedicato all'elaborazione delle immagini ed all'estrazione di indici di flusso di traffico veicolare e di conteggio.

Tale applicativo per quanto riguarda l'elaborazione delle immagini, sulle librerie QT per quanto riguarda l'interfaccia grafica e sulle librerie VLC per la connessione con le telecamere IP. In particolare le telecamere IP svolgono un ruolo fondamentale ai fini del monitoraggio del traffico in quanto in condizioni di banda sufficientemente larga risulta possibile monitorare diversi scenari urbani in luoghi geograficamente distanziati tramite lo stesso server di elaborazione.

2.1 Struttura dell'applicativo per l'analisi delle immagini

Occorre tenere in considerazione le esigenze operative stringenti cui l'applicativo per le analisi delle immagini deve rispondere, ovvero bassa latenza ed elevata velocità di calcolo, caratteristiche indispensabili per poter effettuare elaborazioni real time di elevate quantità di dati quali i flussi video. Allo stesso tempo l'applicativo deve essere realizzato con un linguaggio portatile, in quanto destinato alla sperimentazione e divulgazione, e solo successivamente ad una ingegnerizzazione.

Si è scelto di realizzare un codice ad un ragionevole livello di astrazione, quindi la scelta si è rivolta verso i linguaggi OO (Object Oriented), in particolare Java e C++. L'implementazione in Java, sebbene abbia una elevata portabilità, è stata scartata per esigenze di tempo di calcolo, in quanto è un linguaggio interpretato, con ovvie conseguenze sulla velocità di elaborazione. L'applicativo è stato quindi realizzato in C++.

Successivamente si è passato alla identificazione delle classi kernel dell'applicativo, e le relative funzioni che esse dovessero fornire. Ad una prima analisi si è palesata l'esigenza di almeno tre classi: una dedicata alla gestione del flusso video, una dedicata al processo di analisi delle immagini, una dedicata all'interazione tra l'utente e l'applicativo. Inoltre sono state progettate altre classi con funzionalità a contorno come il parsing con i file di testo ed operazioni su strutture dati.

L'applicativo è strutturato in una serie di procedure sequenziali, a partire da uno step iniziale di configurazione e inizializzazione dei parametri vengono successivamente eseguiti dei task di acquisizione del frame, elaborazione e produzione degli indici di attività, come mostrato nella figura 1.

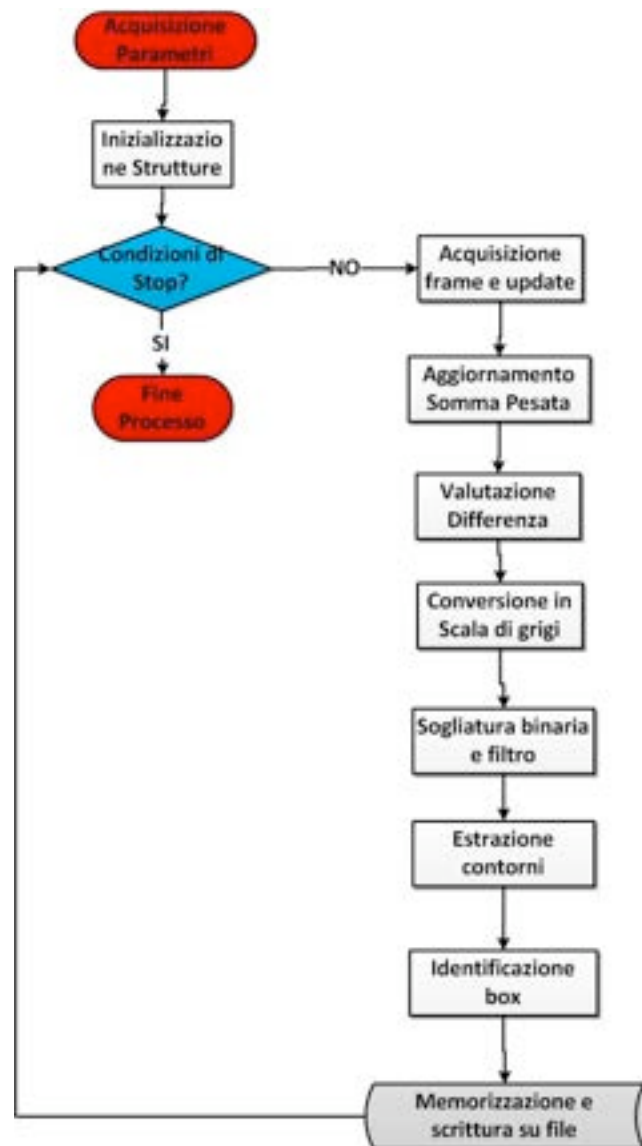


Figura 1. Diagramma del processo del primo prototipo

La prima fase consiste nell'acquisizione dei parametri da un file di configurazione in formato testo. Tale file di configurazione specifica i valori iniziali dei parametri del processo. Tali parametri possono essere suddivisi in tre categorie:

- *Video Properties*. Caratteristiche peculiari riguardanti la sorgente video
- *Analysis Properties*. Caratteristiche riguardanti i parametri di analisi delle immagini.
- *Writing Properties*. Caratteristiche riguardanti i parametri di scrittura su file degli output.

Una volta acquisiti i parametri di configurazione il processo esegue una fase di inizializzazione delle strutture dati necessarie. Le strutture dati dell'applicativo sono di diversa natura, ma possono essere raggruppate in tre macrocategorie:

- *Immagini*. Struttura dati in forma matriciale. Contiene le informazioni di un dato frame, pixel per pixel.
- *Liste*. Struttura dati in forma sequenziale. Utilizzata per contenere i dati di output.
- *Mappe*. Struttura dati in forma chiave-valore. Utilizzata per contenere ed aggiornare i parametri del processo.

Tali strutture non sempre prevedono la possibilità di essere allocate dinamicamente, per cui è necessario che determinati valori, come ad esempio la risoluzione dell'immagine, siano noti prima dell'esecuzione dell'applicativo.

La fase successiva esegue il processo di elaborazione del frame. L'idea è quella di identificare gli oggetti in movimento facendo una differenza tra il frame corrente (l'ultimo frame acquisito) ed una somma pesata dei frame precedenti. La somma pesata è stata pensata tenendo conto del fatto che un frame immediatamente precedente a quello corrente deve avere un peso maggiore rispetto ad un frame meno recente. Per cui l'immagine ottenuta dalla somma pesata tiene conto della seguente formula:

$$sum = (\alpha)frame_t + (1 - \alpha)frame_{t-1} \text{ con } 0 \leq \alpha \leq 1$$

Per cui al variare del parametro α è possibile regolare il peso che dei frame recenti rispetto a quelli passati: un valore di α tendente a uno implica un peso pressoché unilaterale ai frame recenti, la conseguenza è che vengono messi in risalto anche movimenti repentini dell'immagine e con tempi di reazione quasi istantanei, il che potrebbe favorire l'identificazione degli oggetti in movimento ma il processo sarebbe soggetto ad errori anche grossolani. Al contrario un valore troppo basso ha come conseguenza dei tempi di reazione all'identificazione degli oggetti ritardata nel tempo. Il tuning di tale parametro risulta quindi molto importante ai fini del processo, così come altri parametri che saranno presentati successivamente.

Una volta aggiornata l'immagine differenza viene effettuata una differenza assoluta con il frame corrente. L'immagine risultante contiene quindi quei pixel che sono diversi tra le due immagini e che quindi si sono spostati.

L'elaborazione successiva consiste nella conversione dell'immagine differenza in scala di grigi. Tale trasformazione è propedeutica alla sogliatura binaria che viene effettuata successivamente. La conversione da immagine RGB a scala di grigi è stata implementata secondo tre metodologie differenti:

- Distanza euclidea
- Massimo
- Media.

Il massimo prende il valore maggiore tra i valori di verde rosso e blu, quindi al momento della conversione non viene più distinta la differenza tra un blu ed un verde, ad esempio. La media valuta il valor medio dei tre valori, è una tipologia di conversione più fedele all'originale ma rimane una valutazione abbastanza approssimata. La distanza euclidea valuta il valore ottenuto facendo la distanza euclidea tra i tre valori RGB, rispetto agli altri due metodi fornisce una conversione più fedele dell'immagine originaria, ma richiede un dispendio computazionale esponenzialmente maggiore, per cui nella maggior parte dei casi si preferiscono le altre due metodologie di conversione.

La sogliatura binaria consente di risaltare solo le componenti in movimento della fonte video. Tale sogliatura viene effettuata sull'immagine in scala di grigi, se la differenza non supera un certo valore il pixel viene considerato spento, quindi non vi è stato movimento, ed il suo valore viene portato a 0, altrimenti viene considerato acceso, attivo ed il suo valore viene portato al valore massimo (255 nella scala di grigi).

Chiaramente l'immagine risultante è soggetta ad errori dovuti ad imprecisioni originate dall'approssimazione dell'immagine ottenuta con la somma pesata e da rumori originati dalla sorgente video (es. movimento, jitter..) o da fenomeni atmosferici (es. pioggia, vento..). Il passo successivo è quindi quello di rendere l'analisi indipendente dal disturbo.

Per limitare quindi il rumore viene applicato un filtro. Sono state implementate differenti tipologie di filtro.

- Mediano
- Gaussiano
- Bilaterale.

Il filtro mediano per ogni pixel esegue semplicemente la media tra il valore del pixel ed i suoi NxN vicini. Il filtro Gaussiano invece della media esegue una convoluzione del pixel con una matrice NxN di vicini. Il filtro

gaussiano è più dispendioso computazionalmente, ma consente di preservare una certa continuità di valori tra pixel vicini, quindi tale filtro eviterà degli sbalzi improvvisi di valori per pixel vicini. Il problema di tale metodologia è che a volte tale discontinuità di valori se presente nell'immagine originale deve essere preservata in quanto potrebbe indicare la presenza di due entità fisiche differenti (si pensi ad un veicolo bianco su una strada scura). Il filtro bilaterale si basa sullo stesso concetto di kernel gaussiano ma prevede una funzionalità aggiuntiva che consente di preservare i bordi delle immagini. Come parametro aggiuntivo infatti prevede un valore che pesa in maniera maggiore i pixel simili, rispetto a quelli fortemente differenti, in questa maniera al momento della valutazione del kernel una forte discontinuità di valori dei pixel viene mantenuta.

La fase successiva prevede l'estrazione dei contorni degli oggetti in movimento dall'immagine binaria filtrata risultante. Tale processo identifica un perimetro andando a trovare le sequenze non interrotte di pixel accesi. Il contenuto di tale perimetro si presuppone quindi sia un'entità fisica in movimento (veicolo o pedone).

Una volta estratto il contorno viene identificato il bounding box relativo ad ogni oggetto. Il bounding box è una struttura dati che identifica che in quel determinato frame è stato rilevato un determinato oggetto in movimento. Le caratteristiche di tale bounding box sono:

- Coordinate del box (x,y) rispetto al frame
- Lunghezza e larghezza
- Valori RGB medi all'interno.

Tali valori costituiscono le features caratterizzanti di ogni oggetto in movimento.

La fase successiva prevede la scrittura periodica di tali box su file di testo, in cui ad ogni frame vengono elencati i box identificati e le relative features ed il timestamp di rilevamento, espresso in millisecondi trascorsi rispetto all'istante iniziale (l'avvio del processo). Viene mostrato in figura 2 un esempio:

```
66 20 BOX0 21 436 72 462 153 155 160 BOX1 718 354 778 429 88 92 90 BOX2 734 336 761 360 76 80 78 BOX3
100 29 BOX0 382 429 420 449 93 94 91 BOX1 20 426 64 466 153 156 160 BOX2 1 349 19 385 132 132 133 BOX3
133 26 BOX0 447 449 513 470 97 99 95 BOX1 111 445 151 478 101 125 136 BOX2 173 425 209 437 74 88 88 BOX3
146 23 BOX0 447 451 505 466 99 100 96 BOX1 368 430 412 447 92 93 89 BOX2 257 430 298 448 105 106 103 BOX3
200 24 BOX0 401 461 631 479 108 112 110 BOX1 487 447 714 470 108 111 109 BOX2 348 428 417 461 97 98 94 BOX3
233 30 BOX0 479 451 715 474 108 112 110 BOX1 591 441 631 479 108 113 111 BOX2 1 424 68 479 141 165 169 BOX3
246 33 BOX0 481 459 702 479 105 110 108 BOX1 446 448 556 470 98 101 97 BOX2 185 443 251 479 108 109 107 BOX3
300 25 BOX0 720 436 782 479 125 130 128 BOX1 550 424 632 479 109 113 111 BOX2 536 416 556 460 102 105 103 BOX3
333 19 BOX0 720 438 778 479 118 124 122 BOX1 536 425 631 479 108 112 109 BOX2 659 410 719 479 111 116 114 BOX3
367 2 BOX0 1 91 32 124 61 62 50 BOX1 113 1 847 479 109 112 111 BOX2 1 90 94 125 61 61 50 BOX3
400 4 BOX0 1 408 101 479 144 170 175 BOX1 2 139 142 239 71 68 60 BOX2 562 421 647 479 110 115 113 BOX3
433 22 BOX0 470 450 714 479 109 113 112 BOX1 453 443 520 479 97 98 95 BOX2 576 446 632 479 111 115 113 BOX3
467 24 BOX0 445 454 512 470 99 101 98 BOX1 680 451 700 478 107 111 110 BOX2 271 440 301 479 102 104 101 BOX3
500 27 BOX0 304 461 333 479 100 101 97 BOX1 266 460 300 479 102 104 101 BOX2 583 438 632 479 109 113 111 BOX3
533 40 BOX0 301 460 320 479 103 104 101 BOX1 474 381 736 440 116 121 119 BOX2 827 222 847 249 139 136 125 BOX3
567 24 BOX0 386 428 422 451 95 96 92 BOX1 734 326 786 422 97 101 99 BOX2 1 238 237 437 104 107 105 BOX3
600 16 BOX0 494 393 724 427 119 124 121 BOX1 734 327 797 432 101 105 103 BOX2 162 121 186 158 71 69 62 BOX3
633 11 BOX0 24 445 52 466 144 147 151 BOX1 1 244 235 462 106 109 108 BOX2 783 131 807 200 100 123 137 BOX3
667 9 BOX0 734 322 796 422 98 102 100 BOX1 735 320 776 373 72 75 74 BOX2 1 248 228 466 106 109 108 BOX3
700 9 BOX0 742 322 796 421 94 95 96 BOX1 735 320 776 370 71 74 72 BOX2 1 257 220 467 106 109 108 BOX3
734 10 BOX0 747 368 786 420 101 105 104 BOX1 738 319 777 374 70 73 71 BOX2 1 259 219 473 106 110 109 BOX3
767 11 BOX0 745 368 797 418 102 106 104 BOX1 738 320 777 375 68 71 69 BOX2 1 262 214 479 107 110 110 BOX3
800 13 BOX0 745 370 797 412 99 103 101 BOX1 738 320 777 375 68 71 69 BOX2 1 262 214 479 107 110 110 BOX3
834 11 BOX0 746 370 797 411 99 103 101 BOX1 738 320 777 375 68 71 69 BOX2 1 262 214 479 107 110 110 BOX3
```

Figura 2. Esempio di output in formato testo

Il file di output è stato realizzato tenendo in considerazione la grossa quantità di dati che viene memorizzata, quindi ogni valore memorizzato è di tipo *unsigned byte*, per salvaguardare lo spazio in memoria. Inoltre si è cercato di ottenere il miglior compromesso tra quantità di informazioni conservate e estensione del file prodotto. Questo tipo di formato produce circa 1Mbyte al minuto di file. Inoltre è prevista la funzionalità di creazione di un file giornaliero, in modo tale da non produrre file eccessivamente estesi nel caso di riproduzioni per periodi di tempo molto estesi (nell'ordine delle settimane).

2.2 Implementazione del sistema di analisi immagine

Nel seguito vengono descritti i metodi con cui sono state implementate le funzionalità più importanti su cui si basa il sistema di analisi al fine di operare con auto calibrazione per tempi lunghi producendo in uscita gli indici di attività della strada.

2.2.1 Estrazione del background

L'elaborazione video dell'applicativo è risultata in molti casi insoddisfacente in quanto l'identificazione degli oggetti in movimento tramite la differenza tra il frame corrente e la somma pesata dei frame precedenti comporta degli errori in varie situazioni ricorrenti durante l'analisi del traffico urbano. Inoltre la staticità della posizione delle telecamere consente di presupporre che lo scenario urbano monitorato non subisca cambiamenti così repentini a livello scenario di sfondo.

Si è pensato quindi di introdurre come funzionalità aggiuntiva l'estrazione del background. L'idea è stata quella di realizzare un istogramma che memorizzasse per ogni pixel il numero di occorrenze rilevate dei valori RGB. Di conseguenza se per una serie di frame consecutivi in valore RGB di una porzione di area dell'immagine è sempre lo stesso, il numero di occorrenze di quella terna di valori RGB aumenta progressivamente, quindi al momento della periodica valutazione del background, viene estratta la terna di valori con il maggior numero di occorrenze. Si può quindi presupporre che prendendo per ogni pixel la terna RGB con il maggior numero di occorrenze si ottenga un'immagine di sfondo indipendentemente dal movimento rilevato.

L'estrazione del background comporta però grossi problemi di allocazione dinamica della memoria. Lo spazio occupato infatti è fortemente dipendente dalla risoluzione dell'immagine, in quanto il numero dei pixel cresce proporzionalmente, e dal numero di possibili valori che ogni pixel può assumere: per un'immagine RGB ogni pixel può assumere circa 16 milioni di valori, quindi è necessario allocare un'elevata porzione di memoria per creare un istogramma sufficientemente ampio. Per ovviare al problema dell'allocazione eccessiva della memoria è stata realizzata l'estrazione del background direttamente in scala di grigi, riducendo il numero di possibili valori dei pixel a sole 256 possibilità. La tabella 1 mostra gli ordini di grandezza di differenza fra un background estratto in RGB ed uno in scala di grigi.

Risoluzione	Tipo di immagine	Memoria Allocata (GB)
640x480	RGB	9600
480x360	RGB	5400
320x240	RGB	2400
640x480	Scala di grigi	0,15
480x360	Scala di grigi	0,07
320x240	Scala di grigi	0,03

Tabella 1. Allocazione della memoria dinamica per il background

Il passo successivo per la riduzione dell'allocazione della memoria è stato quello di suddividere i possibili valori della scala di grigi in sottolivelli. Questa suddivisione ha consentito non solo un ulteriore risparmio sulla memoria allocata ma anche di evitare fenomeni di jitter a livello di background.

Per evitare problemi di overflow dell'istogramma dopo un periodo sufficientemente lungo di tempo quando l'istogramma sta per arrivare all'overflow viene effettuato un processo di svuotamento della memoria dell'istogramma.

Il processo di valutazione della differenza dell'immagine descritto precedentemente non avviene quindi più con la sola somma pesata ma con una fusione logica tra la somma pesata dei frame precedenti e il background. La fusione logica può essere fatta eseguendo un AND od un OR logico pixel per pixel tra le due immagini. La scelta ricade il più delle volte sull'AND logico, per cui viene rilevato un movimento solamente quando viene rilevata una differenza tra il frame corrente ed il background e tra il frame corrente e la somma pesata dei frame passati.

2.2.2 Interfaccia Grafica

Durante lo sviluppo dell'applicativo il solo uso delle librerie grafiche di OpenCV non si è rivelato sufficiente alla caratterizzazione della grossa quantità di parametri configurabili prima e durante il processo di elaborazione delle immagini. E' stato quindi necessario implementare un'interfaccia più completa. L'interfaccia grafica è stata realizzata usando le librerie QT. La figura 3 mostra l'interfaccia realizzata:

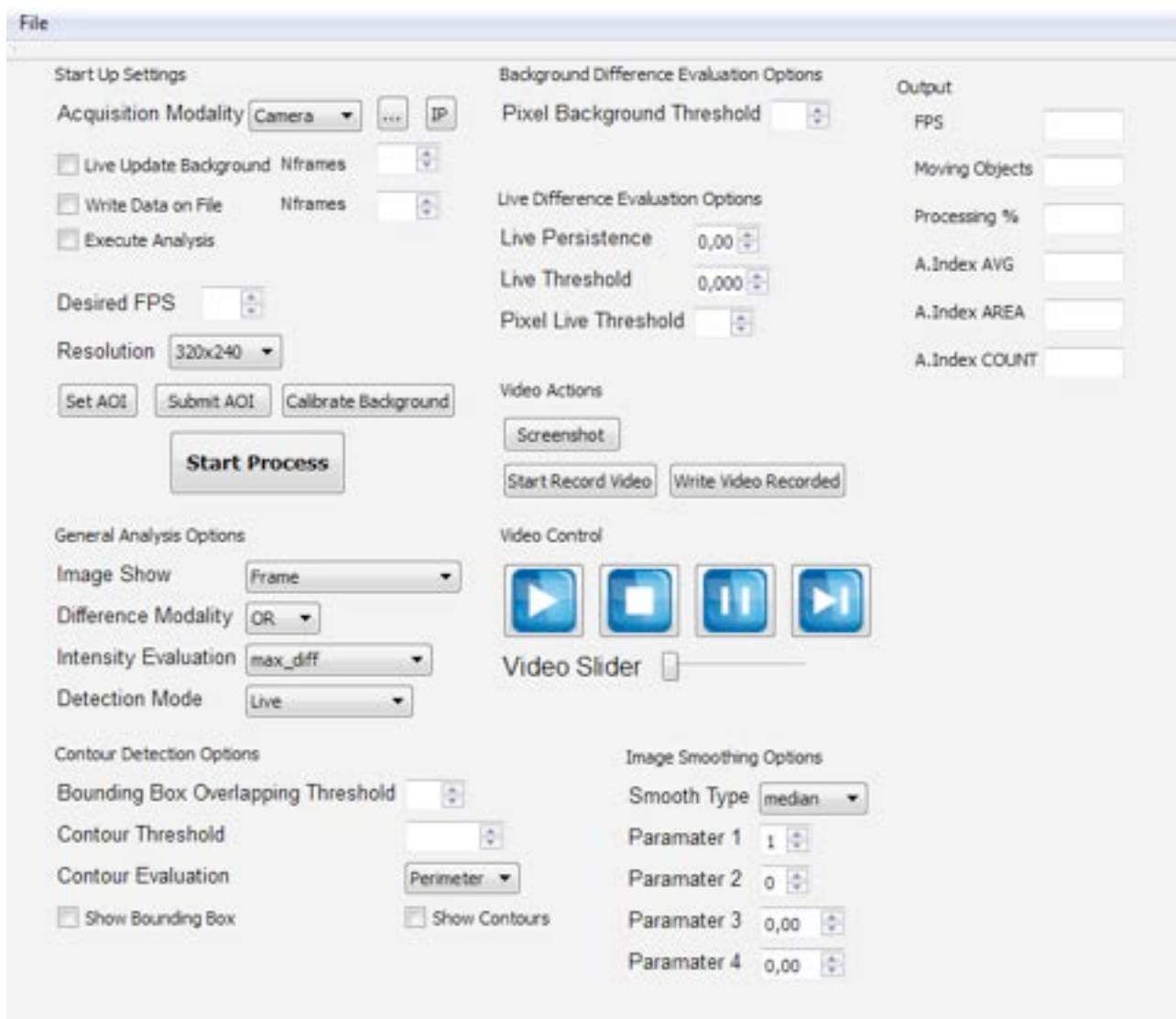


Figura 3. Interfaccia Grafica

Come mostrato tale interfaccia consente di manipolare una serie di parametri, la maggior parte dei quali in tempo reale. I parametri infatti sono divisi in sottocategorie:

- **Startup Settings.** In questa categoria rientrano i parametri che non possono successivamente essere modificati durante il processo.
 - **Acquisition Modality.** Descrive la tipologia di sorgente del flusso video, che può provenire da una telecamera locale (es. USB), un file preregistrato o da una telecamera IP.
 - **Live update background.** Checkbox tramite cui è possibile stabilire se si desidera un aggiornamento continuo del background durante il processo o meno. Nframes specifica ogni quanti frame deve essere eseguito l'update del background.
 - **Write on file.** Checkbox tramite cui è possibile stabilire se si desidera o meno la scrittura su file.
 - **Execute Analysis.** Checkbox tramite cui è possibile stabilire se si desidera o meno eseguire l'analisi o solo l'esecuzione del flusso video senza elaborazione.
 - **Resolution.** Permette di selezionare il tipo di risoluzione di riproduzione del video desiderata, tale opzione ovviamente non va ad agire sulla sorgente, la cui risoluzione è un' proprietà intrinseca, ma solo a livello di riproduzione ed elaborazione.
 - **Desired FPS.** Permette di regolare il throughput di calcolo del processo, espresso in numero di frame elaborati al secondo. Anche in questo caso non si va a modificare l'FPS della sorgente video, il quale è una proprietà intrinseca.
 - **Save/Load CFG.** Permette di salvare/caricare i file di configurazione in formato testo.
 - **Set/Submit AOI.** Permette di impostare aree di interesse sull'immagine, funzionalità introdotta successivamente e che sarà discussa in seguito.
 - **Calibrate Background.** Permette di calibrare un'immagine di background in fase di preprocessing. Tale possibilità risulta utile per svolgere un'analisi accurata anche quando il processo è appena avviato.
- **General Analysis Options.** In questa categoria rientrano i parametri generici dell'elaborazione:
 - **Image Show.** Permette di selezionare quale immagine mostrare durante il processo (es. il frame acquisito, il background, l'immagine differenza, i soli contorni ecc..)
 - **Difference Modality.** Permette di selezionare la modalità di fusione logica tra l'immagine di background e l'immagine somma pesata.
 - **Intensity Evaluation.** Permette di selezionare la modalità di conversione in scala di grigi di un'immagine RGB (distanza euclidea, media, massimo).
 - **Detection mode.** Permette di selezionare la modalità di rilevamento del movimento (solo live, ovvero con la somma pesata, solo background o entrambe)
- **Contour Detection Options.** In questa categoria rientrano i parametri relativi all'estrazione dei contorni.
 - **Bounding overlapping threshold.** Permette di selezionare una sogliatura sulla sovrapposizione dei bounding box estratti. Se una % sopra la soglia impostata di un selezionato box è sovrapposta ad un altro box, il box selezionato viene scartato. Questa è una funzionalità aggiuntiva che sarà descritta in seguito.
 - **Contour Evaluation.** Permette di selezionare la modalità di valutazione della misura di un contorno, se in base al perimetro o l'area.
 - **Contour Threshold.** Permette di filtrare i contorni inferiori ad una determinata soglia, in termini di perimetro o area.
 - **Show Bounding Box.** Attiva o disattiva la visualizzazione a video dei bounding box
 - **Show Contour.** Attiva o disattiva la visualizzazione a video dei contorni
- **Background difference evaluation options.** In questa categoria rientrano i parametri inerenti la valutazione della differenza del frame rispetto al background.
 - **Pixel Background Threshold.** Determina la soglia al di sotto della quale un pixel del background viene considerato spento e quando invece acceso in fase di sogliatura binaria del background.

- **Video Actions.** In questa categoria rientrano le operazioni di scrittura immagini/video.
 - **Screenshot.** Esegue uno screenshot dell'immagine corrente visualizzata
 - **Start Record Video.** Inizializza la scrittura del video su file avi.
 - **Write video recorded.** Esegue la scrittura del file video.
- **Video Controls.** Opzioni che consentono la manipolazione del flusso video (start/stop/pausa/forward)
- **Image Smoothing Options.** In questa categoria rientrano le opzioni inerenti il filtro dell'immagine.
 - **Smooth Type.** Permette di selezionare la tipologia di filtro sull'immagine (mediana, gaussiana, bilaterale)
 - **Parameter.** Consente di impostare i valori dei parametri delle funzioni di smooth.
- **Output.** Vengono mostrati i valori in tempo reale degli output rilevati.
 - **FPS.** Numero di frame al secondo che l'applicativo riesce a processare.
 - **Moving Objects.** Numero di oggetti in movimento rilevati nel frame corrente.
 - **Processing %.** Percentuale di processore dedicata al processo di elaborazione.
 - **A.Index AVG/AREA/Count.** Indici di attività rilevati correnti (il significato verrà esplicitato nel capitolo 3).

2.3 Release dell'applicativo

2.3.1 Suddivisione in aree di interesse

La funzionalità più interessante introdotta nella revisione dell'applicativo è l'introduzione della suddivisione dell'area di analisi in sottoaree di interesse. Tale funzionalità è stata introdotta con l'intento in primo luogo di scartare alcune zone dell'area monitorata: è infatti plausibile ipotizzare che di un'area urbana monitorata ci siano delle zone in cui si esclude a priori che possa verificarsi qualche evento da monitorare (es. palazzi). In secondo luogo focalizzare l'elaborazione solo su determinate zone del flusso video consente uno sgravio computazionale elevato. Inoltre impostando in maniera appropriata le aree di interesse è possibile creare dei "traguardi" utili ai fini del conteggio dei veicoli.

La suddivisione in aree di interesse consente inoltre di analizzare separatamente le carreggiate stradali, facilitando ad esempio il compito dell'individuazione della direzione di marcia dei veicoli, o ancora consente di focalizzare l'attenzione su aree pedonali, sgravando così l'applicativo dal compito di individuare la differenza tra veicolo o pedone.

Infine focalizzandosi su due esigue porzioni di area di interesse è possibile facilitare l'individuazione della correlazione tra gli oggetti riconosciuti nelle due differenti aree e valutare conseguentemente informazioni quali velocità media e tempo di percorrenza dei veicoli.

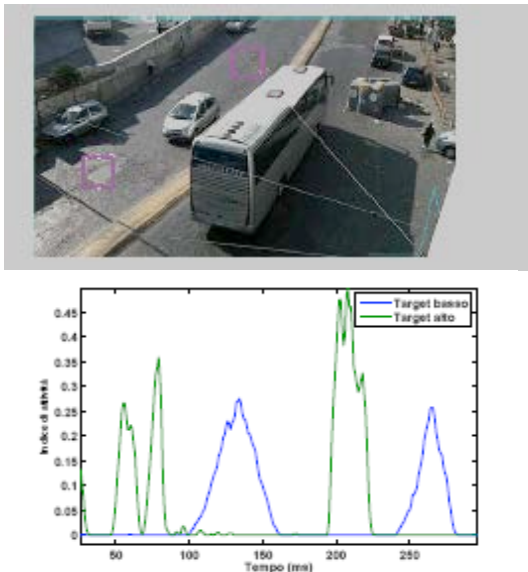


Figura 4. Analisi della correlazione tra due aree di interesse

La figura mostra come il segnale rilevato dovuto al passaggio del veicolo sul primo target (quello in alto) seppur con una quantità non trascurabile di rumore, viene rilevato a distanza di 50ms nel secondo target (quello in basso).

2.3.2 Filtro sui bounding box

Il sistema di filtro sui bounding box consiste nel confrontare ad ogni frame ogni box rilevato con tutti gli altri, se l'area occupata dal bounding box selezionato è sovrapposta di una certa % a quella dei box confrontati allora il box selezionato viene scartato.

Questo tipo di filtro è stato introdotto in quanto accade frequentemente, soprattutto in situazioni di traffico elevato, che per ogni veicolo si sovrappongano molteplici box spesso tale situazione di verifica con veicoli di grosse dimensioni, i quali vengano identificati come due o più bounding box (solitamente la parte anteriore e la parte posteriore del veicolo).

Inoltre tale situazione può presentarsi in presenza di ombre proiettate dai veicoli o i pedoni stessi: viene infatti identificata una divergenza di colorazione tra i veicoli o le persone e le rispettive ombre e di conseguenza viene rilevata una discontinuità durante l'estrazione dei contorni. Tale filtro in parte risolve questa problematica.

2.3.3 Diagramma di flusso del sistema di elaborazione

La figura 5 mostra il diagramma di flusso dell'applicativo. Rispetto al diagramma precedente è stata introdotta l'estrazione del background e la fusione logica tra l'immagine background e l'immagine somma pesata. Inoltre la conversione in scala di grigi e la soglia binaria viene applicata singolarmente alle due immagini, una volta effettuata la fusione logica viene valutata la differenza con il frame corrente ed applicato il filtro.

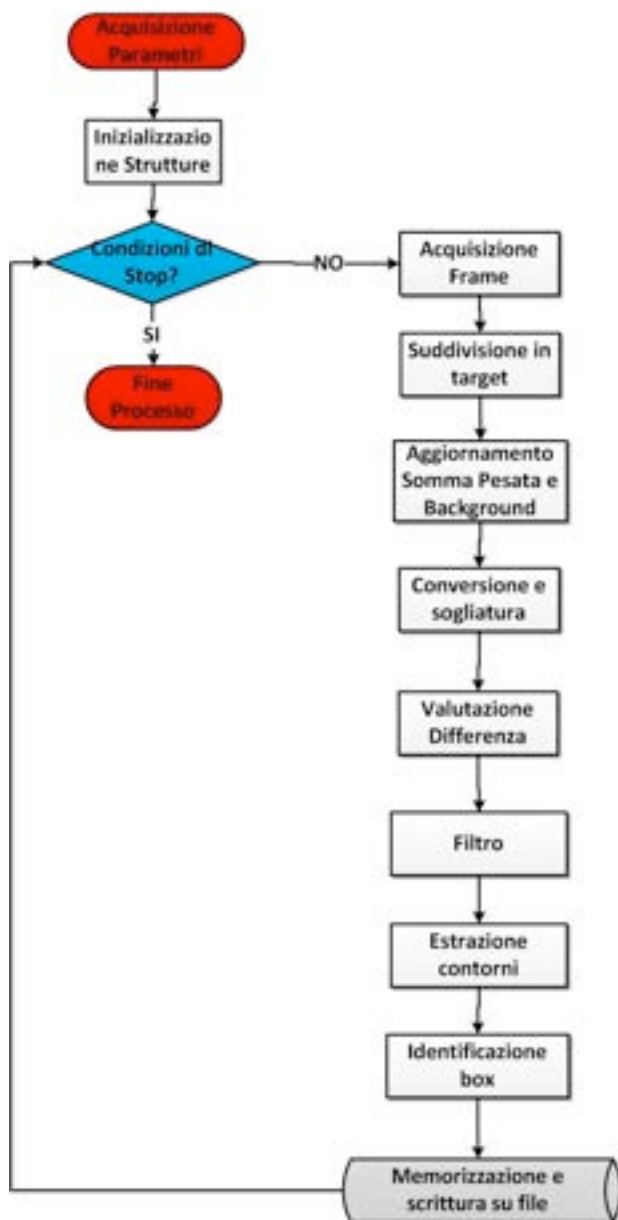


Figura 5. Diagramma di flusso dell'applicativo

Le figure 6 e 7 mostrano visivamente il processo di elaborazione del frame: l'acquisizione, l'estrazione del background, l'immagine filtrata con i soli oggetti in movimento e l'estrazione dei contorni e dei bounding box.



Figura 6. Frame Acquisito (a sinistra) e Background Corrente (destra)



Figura 7.Frame filtrato (sinistra) e frame con contorni e bounding box estratti (destra)

2.3.4 Scrittura degli output su file

La scrittura sul file è suddivisa in due file differenti: uno contiene per ogni frame le informazioni delle features di ogni box, suddivisi per target (figura 8), l'altro contiene informazioni sugli indici di attività implementati (figura 9), che verranno presentati nella sezione successiva. Entrambi i file presentano un'intestazione indicante l'orario di creazione ed il significato dei valori elencati.

FILE HEADER														
Wed May 04 07:46:11 2011														
Targets Number	Target Label	Points Number	Points Coordinates(x,y)											
0														
Timestamp(ms)	Target Number	Box Found on Target	X-coordinate			Y-coordinate			Width	Height	RED avg			GREEN avg
66 TUTO	17 BOX0	18 442 52 468 155 151 148 BOX1	741	406	778	429	91	94	90	BOX2	747	364	762	400 69 71 68 BOX3
100 TUTO	26 BOX0	382 429 420 448 90 94 93 BOX1	20	426	64	464	160	156	153	BOX2	17	351	64	389 153 152 150 BOX3
133 TUTO	24 BOX0	447 449 513 470 95 99 97 BOX1	110	447	151	479	132	122	98	BOX2	1	425	71	479 170 165 161 BOX3
166 TUTO	24 BOX0	368 430 412 447 89 93 92 BOX1	257	430	296	448	103	106	105	BOX2	13	425	68	476 166 162 158 BOX3
200 TUTO	24 BOX0	401 441 431 479 110 112 108 BOX1	368	428	417	461	94	98	97	BOX2	237	426	311	447 104 104 105 BOX3
233 TUTO	36 BOX0	678 451 710 474 110 112 108 BOX1	591	441	631	479	111	113	108	BOX2	188	441	241	476 104 108 107 BOX3
266 TUTO	41 BOX0	681 460 702 479 108 110 108 BOX1	185	454	234	479	111	113	111	BOX2	746	450	766	477 114 116 110 BOX3
300 TUTO	24 BOX0	101 453 143 479 151 138 110 BOX1	720	440	745	460	114	117	112	BOX3	746	439	778	479 128 130 124 BOX3
333 TUTO	29 BOX0	647 456 669 479 114 114 109 BOX1	749	449	778	479	131	132	126	BOX2	726	439	756	463 118 120 114 BOX3
367 TUTO	17 BOX0	654 411 714 479 114 116 111 BOX1	1	405	101	479	175	170	164	BOX2	450	376	486	403 112 113 108 BOX3
400 TUTO	2 BOX0	2 139 134 239 62 70 72 BOX1	143	1	847	479	112	113	109					
433 TUTO	18 BOX0	670 451 714 479 111 113 109 BOX1	453	444	519	473	95	99	97	BOX2	751	442	782	479 138 139 133 BOX3
467 TUTO	21 BOX0	578 459 492 479 111 113 109 BOX1	465	454	505	467	99	103	101	BOX2	362	416	421	469 95 99 97 BOX3
500 TUTO	29 BOX0	595 462 635 479 112 114 110 BOX1	304	461	333	479	97	101	100	BOX2	271	460	301	479 101 104 102 BOX3
533 TUTO	30 BOX0	268 460 300 479 101 104 102 BOX1	583	458	632	479	111	113	109	BOX2	381	455	401	479 96 101 99 BOX3
567 TUTO	24 BOX0	386 428 422 451 92 96 95 BOX1	675	381	735	440	119	121	116	BOX2	734	341	797	434 102 104 100 BOX3
600 TUTO	14 BOX0	734 327 797 428 102 103 99 BOX1	341	269	384	305	140	122	98	BOX2	1	221	239	479 110 111 108 BOX3
633 TUTO	12 BOX0	24 445 49 466 151 147 144 BOX1	734	326	796	422	99	101	97	BOX2	1	236	236	436 104 106 104 BOX3

Figura 8. Esempio di output delle features su file di testo

FILE HEADER				
Wed May 04 07:46:11 2011				
Timestamp(s)	Target Number	Index Activity AVG	Index Activity AREA	Index Activity COUNT
2 TUTO	16.316668	4.759512	20	
4 TUTO	21.466667	6.355064	18	
6 TUTO	12.288134	4.416070	34	
8 TUTO	10.766668	3.916130	20	
10 TUTO	14.933333	5.432522	20	
12 TUTO	16.890000	7.319200	30	
14 TUTO	12.150000	5.284239	21	
16 TUTO	19.500000	4.872039	20	
18 TUTO	17.828301	4.187665	38	
20 TUTO	16.016668	4.480005	18	
22 TUTO	17.983334	4.662287	14	
24 TUTO	12.882383	3.944374	40	
26 TUTO	13.986668	4.332368	19	
28 TUTO	13.433333	4.019211	18	
30 TUTO	17.240000	4.767465	29	
32 TUTO	19.550000	4.366929	19	
34 TUTO	14.833333	5.211607	19	
36 TUTO	15.847824	4.816234	29	
38 TUTO	17.966667	4.921588	17	

Figura 9. Esempio di output degli indici su file di testo

2.3.5 Interfaccia grafica

Come mostra la figura è stata un'interfaccia utente semplificata. Solo in caso di autenticazione è consentito l'accesso all'interfaccia avanzata che permette la modifica di tutti i parametri dell'elaborazione.

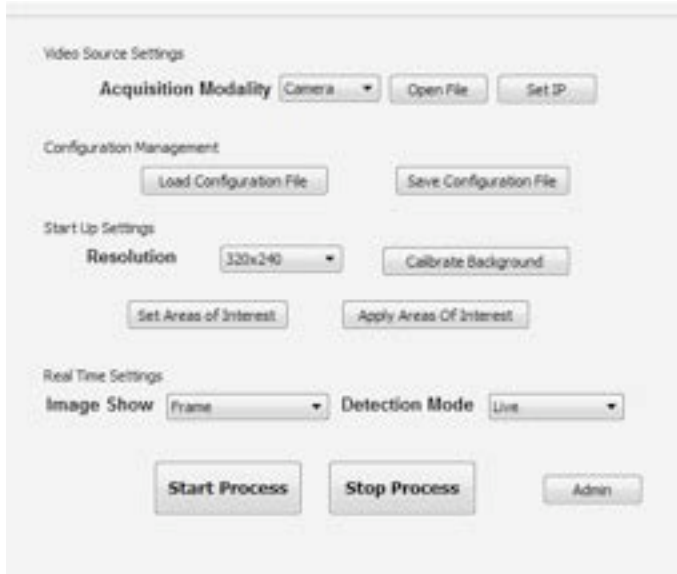


Figura 10. Interfaccia utente

3 Sviluppo degli indici di traffico

Al fine di fornire degli indici qualitativi sulla misura della quantità di traffico rilevata sono stati realizzati degli algoritmi di estrazione di indici di attività. Sono stati dapprima realizzati degli indici che non esprimono grandezze quantificabili in concreto come il conteggio o la velocità, ma delle grandezze adimensionali che esprimono in maniera indicativa la rilevazione del flusso di traffico.

In particolare sono stati implementati:

- *Index AVG*
- *Index Area*.

Index AVG fornisce il numero medio di bounding box per frame rilevato in un intervallo di integrazione.

Index Area fornisce la percentuale media di area dell'immagine occupata dai bounding box rispetto all'area dell'intero frame in un intervallo di integrazione.

Successivamente è stato sviluppato un algoritmo che fornisce il numero di veicoli conteggiati durante il processo di elaborazione (*Index Count*). Tale algoritmo tiene in memoria il numero di bounding box rilevati per ogni frame, se tale numero risulta incrementato per un certo numero di frame consecutivi l'algoritmo assume che un nuovo veicolo è stato rilevato e quindi incrementa il conteggio. Per evitare che vengano conteggiati più volte gli stessi veicoli è importante impostare opportunamente filtro, ovvero il numero minimo di frame consecutivi in cui si deve rilevare un incremento del numero di bounding box affinché si possa conteggiare un nuovo oggetto in movimento.

Le features contenenti i valori RGB medi di un bounding box sono dedicate alla discriminazione dei mezzi pubblici rispetto a quelli privati. Infatti sapendo a priori i colori dei mezzi pubblici è possibile fare un matching tra i valori RGB medi del bounding box e quelli relativi al mezzo pubblico.

Le features contenenti la dimensione del bounding box sono dedicate invece alla discriminazione tra veicolo e pedone.

L'utilizzo di tutte le features mostrate nel file di output in figura 8 (timestamp, suddivisione in target, posizione dei box e valori RGB medi) permettono la valutazione delle velocità medie dei veicoli identificati. Infatti creando delle aree di interesse in posizioni opportune del flusso video (ad esempio all'inizio ed alla fine di una carreggiata) è plausibile ipotizzare che un veicolo identificato nella prima area di interesse, che abbia determinati valori RGB medi, sia identificato, a meno di errori, anche nella seconda area di interesse, ad un determinato shift temporale, ma con valori RGB medi uguali o molto simili. Valutando quindi la differenza temporale tramite il timestamp e la differenza spaziale tramite le coordinate dei bounding box è possibile valutare il tempo di percorrenza e la velocità media del veicolo nella tratta compresa tra le due aree di interesse.

4 Sperimentazione

La sperimentazione dell'applicativo è stata effettuata su una situazione di traffico urbano periferico (Aurelia), di traffico urbano intenso diurno (stazione centrale di Napoli) e di traffico urbano notturno (Roma, zona Torre Vecchia).

Le sperimentazioni effettuate su questi casi studi consistono nella valutazione dell'indice di conteggio degli oggetti in movimento attraverso il superamento di una soglia. Gli indici di attività precedentemente descritto non hanno un riferimento quantitativo significativo con cui essere confrontati per cui non sono mostrati tra i risultati ottenuti, mentre l'indice di conteggio è stato confrontato con il numero reale di veicoli conteggiati nel relativo intervallo di tempo considerato.

Per effettuare il conteggio è stato introdotto un target in modo tale da focalizzare l'attenzione su una determinata area di interesse. Tale area di interesse costituisce la soglia di attraversamento per il conteggio dei veicoli. La Figura 11 mostra un esempio di attività di rilevamento di oggetti in movimento con relativo conteggio. Si può notare come l'applicativo rilevi gli oggetti in movimento nell'immagine, circondandoli con un bounding box di colore verde. Se tali bounding box entrano all'interno dell'area di interesse viene rilevato un nuovo oggetto in movimento da conteggiare con l'algoritmo descritto nella precedente sezione.



Figura 11. Conteggio tramite superamento della soglia

Sono stati effettuati una serie di test per trovare delle configurazioni dei parametri descritti nel paragrafo 2.2.2 efficaci per un rilevamento soddisfacente degli oggetti in movimento. In particolare sono risultati di importanza critica i valori relativi a *bounding box threshold*, *pixel background threshold*, *pixel live threshold* ed i parametri di *smoothing*. Effettuando in maniera opportuna il tuning dei parametri è stato possibile ottenere risultati molto soddisfacenti.

In particolare la qualità del conteggio è risultata sensibile alla variazione del pixel background threshold, soprattutto in relazione all'intensità dell'illuminazione naturale. Infatti in condizioni diurne il contributo fornito dal background ai fini del rilevamento degli oggetti in movimento è soddisfacente, ma in condizioni notturne la differenza fra gli oggetti in movimento ed il background in un'ampia gamma di situazioni è bassa, per cui impostare una soglia di background sui pixel troppo alta comporta una notevole difficoltà nel distinguere gli oggetti statici da quelli dinamici.

Di conseguenza sono state realizzate due tipologie di configurazione dei parametri differenti, una impostata per le situazioni di traffico diurno, l'altra per situazioni di traffico notturno.

4.1 Aurelia

Nel caso di studio preso in esame in questa sezione la zona monitorata è l'Aurelia. Il traffico è di tipo periferico, quindi non molto intenso, in situazione diurna. E' un caso di studio relativamente semplice in quanto il traffico è poco intenso e regolare.



Figura 12. Aurelia

E' stato calcolato l'index activity COUNT con cadenza di 30 secondi su un filmato di circa 180 secondi. I risultati sono mostrati nelle seguenti tabelle.

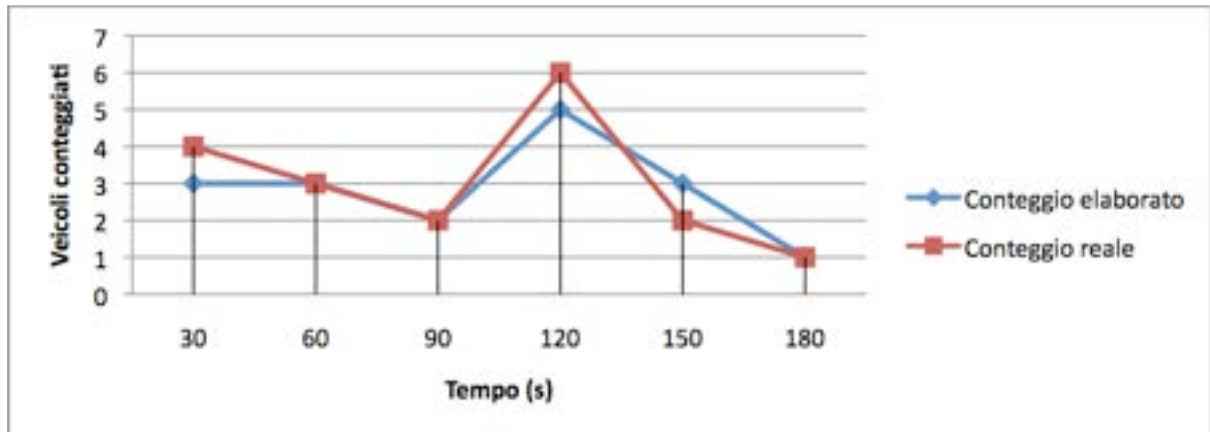
Timestamp(s)	Conteggio elaborato	Conteggio reale
30	3	4
60	3	3
90	2	2
120	5	6
150	3	2
180	1	1

Tabella 2. Test conteggio aurelia

Conteggio totale elaborato	Conteggio totale reale	% Errore
17	18	5,56%

Tabella 3. Conteggio totale aurelia

Come possiamo vedere dalle tabelle e dal grafico successivo la percentuale di errore complessiva è relativamente bassa.



4.2 Stazione Centrale di Napoli

Nel caso di studio preso in esame in questa sezione la zona monitorata è una via della stazione centrale di Napoli. Il traffico è di tipo urbano intenso, quindi caotico e irregolare, in situazione diurna. E' un caso di studio complesso in quanto presenta elementi di disturbo come la proiezione delle ombre e la sovrapposizione di automobili durante i sorpassi o gli affiancamenti.

Tale caso di studio presenta tipologie differenti di oggetti in movimento: automobili, pedoni, motocicli e autocarri.

Nello studio effettuato il conteggio è relativo agli oggetti in movimento relativi alle carreggiate, quindi i pedoni sono comunque rilevati dall'applicativo, ma sono esclusi dal conteggio.



Figura 13. Stazione Centrale di Napoli

Anche in questo caso l'area di interesse è stata impostata in modo tale da costituire una soglia di attraversamento per il conteggio dei veicoli, come mostrato dalla Figura 14. Tale figura mostra come l'applicativo rilevi tutti i veicoli presenti nell'area e li circoscriva in bounding box, comprendendo anche le ombre. Al momento del passaggio all'interno dell'area di interesse il bounding box viene evidenziato e successivamente conteggiato. La figura mostra come anche i pedoni vengano rilevati, sebbene in quelli in lontananza siano rilevati in raggruppamenti.

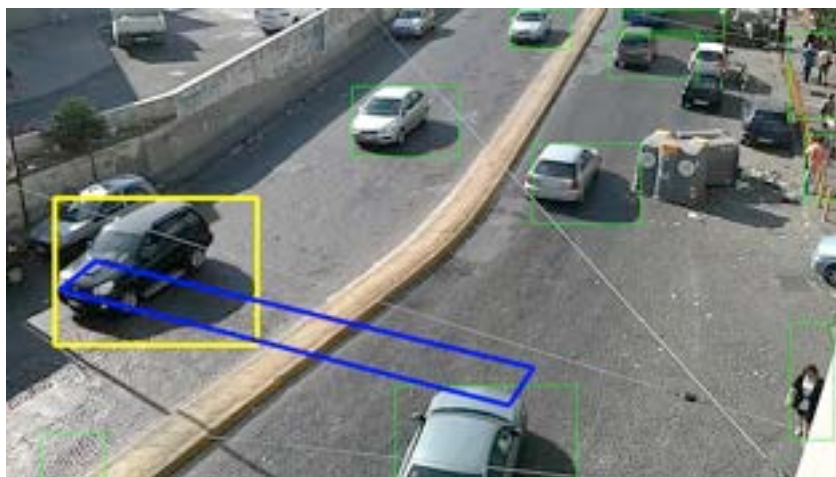


Figura 14. Soglia di attraversamento stazione centrale di Napoli

E' stato calcolato l'index activity COUNT con cadenza di 20 secondi su un filmato di circa 160 secondi. I risultati sono mostrati nelle seguenti tabelle.

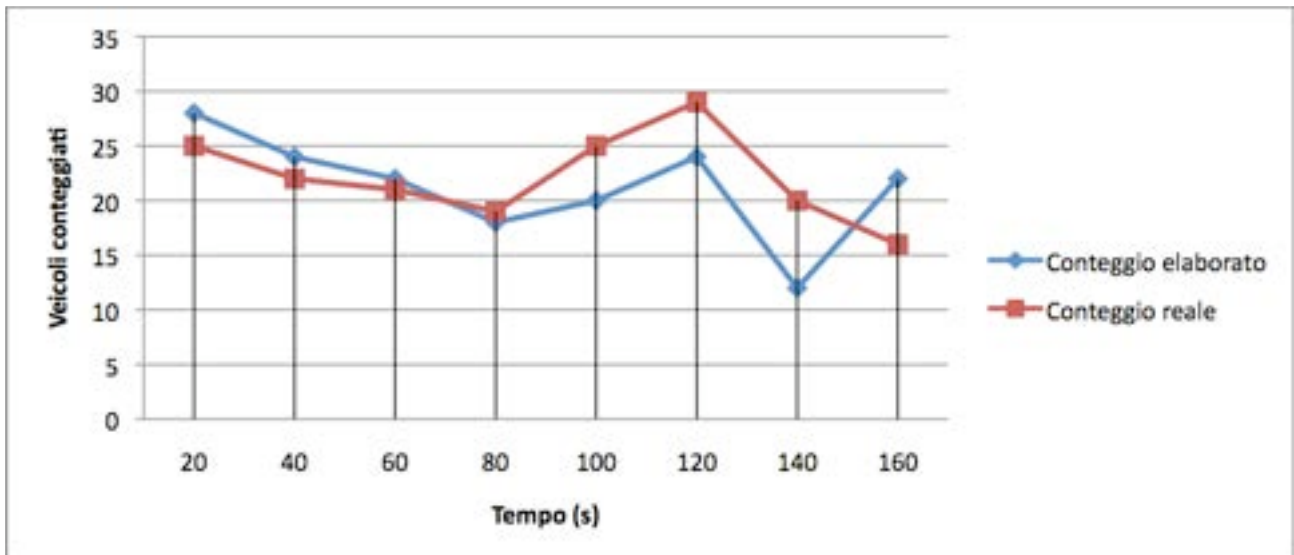
Timestamp(s)	Conteggio elaborato	Conteggio reale
20	28	25
40	24	22
60	22	21
80	18	19
100	20	25
120	24	29
140	12	20
160	22	16

Tabella 4. Test conteggio stazione centrale di Napoli

Conteggio totale elaborato	Conteggio totale reale	% Errore
96	109	11,93%

Tabella 5. Conteggio totale stazione centrale di Napoli

Rispetto al caso di studio precedente i risultati sono lievemente peggiori, per via della tipologia di traffico, molto caotico, intenso ed irregolare. Tuttavia riteniamo di poter essere ragionevolmente soddisfatti del risultato ottenuto, tenendo in considerazione il fatto che tali risultati siano ulteriormente migliorabili.



4.3 Roma Torre Vecchia

Nel caso di studio preso in esame in questa sezione la zona monitorata è via Torre Vecchia a Roma. Il traffico è di tipo urbano intenso in condizioni di illuminazione notturna. E' un caso di studio differente rispetto ai precedenti in quanto le condizioni di visibilità sono molto ridotte.

Le simulazioni effettuate hanno evidenziato che l'analisi effettuata con la configurazione dei parametri usata per i casi precedenti non produce risultati soddisfacenti. La motivazione principale è dovuta al fatto che il background non è facilmente distinguibile dai veicoli, soprattutto quando quest'ultimi sono di colore scuro. Per cui è stato necessario abbassare notevolmente la soglia sui pixel del background.

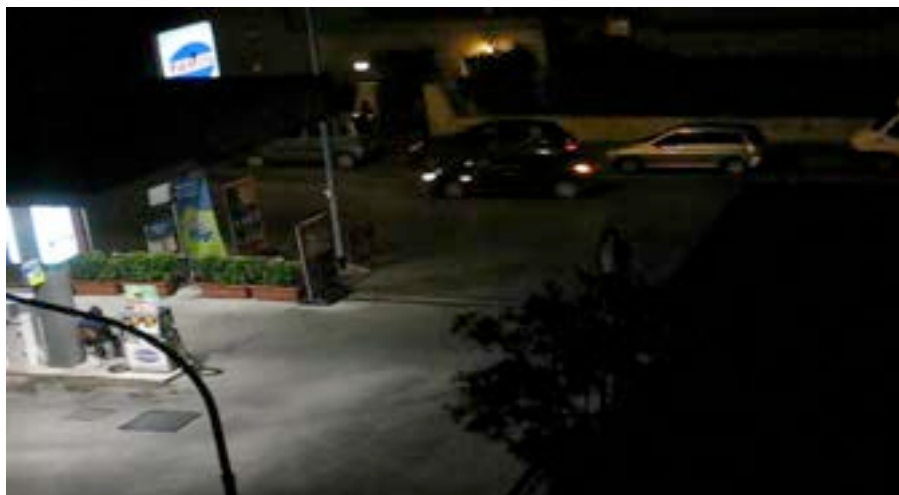


Figura 15. Roma, via Torre Vecchia

Anche in questo caso l'area di interesse è stata impostata in modo tale da costituire una soglia di attraversamento per il conteggio dei veicoli, come mostrato nella Figura 16.



Figura 16. Soglia di attraversamento via Torre Vecchia

E' stato calcolato l'index activity COUNT con cadenza di 30 secondi su un filmato di circa 120 secondi. I risultati sono mostrati nelle seguenti tabelle.

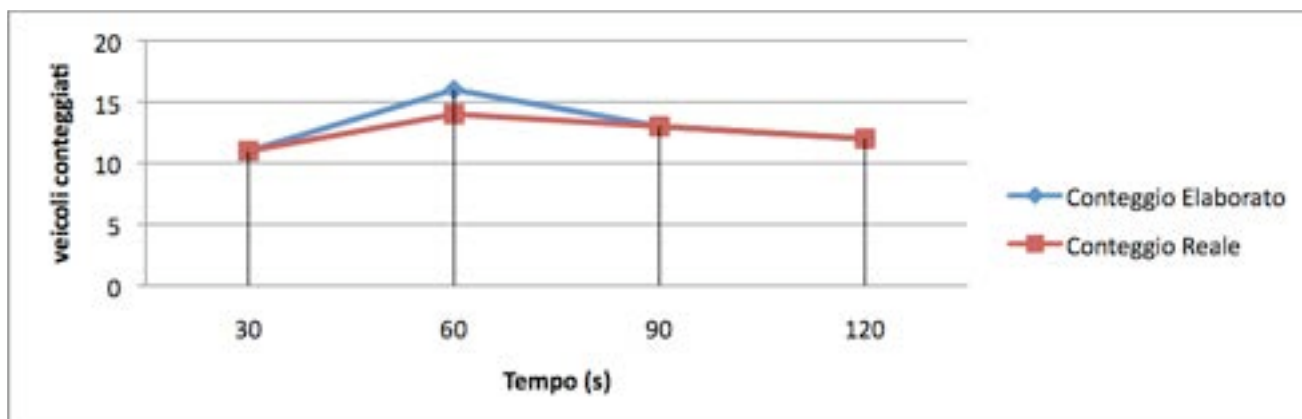
Timestamp(s)	Conteggio elaborato	Conteggio reale
30	11	11
60	16	14
90	13	13
120	12	12

Tabella 6. Test conteggio Roma zona Torre Vecchia

Conteggio totale elaborato	Conteggio totale reale	% Errore
52	50	4,00%

Tabella 7. Conteggio totale Roma zona torre Vecchia

I risultati ottenuti sono stati molto soddisfacenti in quanto nella maggior parte dei casi il conteggio è risultato esatto ed il conteggio totale elaborato ha un margine di errore molto esiguo.



4.4 Riepilogo sperimentazioni

In questa sezione viene presentata una tabella riepilogativa dei test effettuati sui casi di studio presentati precedentemente. Tale tabella mostra che i risultati ottenuti sono stati soddisfacenti sia nei casi diurni che notturni, il picco di errore si rileva nel caso di Napoli, per via dell'elevata complessità della scena urbana, ma che comunque è una soglia di errore che riteniamo sia possibile abbattere con l'introduzione di ulteriori funzionalità. L'errore medio ottenuto è ragionevolmente basso

Caso Studio	di	Conteggio totale elaborato	Conteggio totale reale	% Errore
Aurelia		17	18	5,56%
Napoli		96	109	11,93%
Roma		52	50	4,00%
Errore medio				7,16%

Tabella 8. Riepilogo dei risultati dei test

5 Conclusioni

Il rapporto realizzato ha inizialmente introdotto il contesto in cui si inserisce lo studio sull'analisi delle scene urbane e le finalità cui tale studio è indirizzato. In particolare è stato mostrato come le metodologie realizzate a seguito di tale studio possano essere applicate in primo luogo per l'attuazione di controlli adattivi dei punti luce dedicati all'illuminazione pubblica ed in secondo luogo per offrire servizi utili al cittadino riguardanti l'infomobilità.

Sono state mostrate successivamente le varie fasi di sviluppo dell'applicativo, dallo studio di fattibilità alle successive migliorie apportate per finire alla versione reale ed alla realizzazione di algoritmi per la valutazione degli indici di attività di traffico evidenziando le problematiche riscontrate. In particolare è stato necessario trovare un giusto compromesso tra qualità del processo di analisi del traffico e risorse computazionali e di memoria utilizzate. L'ottica è infatti quella di realizzare un applicativo autonomo che sia in grado di effettuare rilevamenti real time lungo archi temporali piuttosto elevati (nell'ordine delle settimane), ed è stato quindi importante ragionare nell'ottica della scalabilità e della robustezza.

Infine sono stati effettuati dei benchmark su differenti casi di studio per ricoprire una variegata gamma di situazioni in cui il rilevamento del traffico può essere applicato. I risultati ottenuti sono stati molto soddisfacenti sia in casi di traffico diurno che notturno. L'errore medio ottenuto nel conteggio dei veicoli è ampiamente al di sotto della soglia del 10%, per cui riteniamo di poter essere soddisfatti dei risultati ottenuti.

Si ritiene inoltre che tali risultati siano ulteriormente migliorabili attraverso l'uso di tecniche di pattern recognition e metodi di segmentazione dei veicoli, come proposto da Pang et Al. in [1] e [2].

6 Bibliografia

- [1] C. C. Pang, W. Leung Lam, N. Ching Yung. *A Method for Vehicle Count in the Presence of Multiple-Vehicle Occlusions in Traffic Images*.
- [2] C. C. Pang, W. Leung Lam, N. Ching Yung. *Highly accurate texture-based vehicle segmentation method*.
- [3] OpenCV Wikipedia. <http://opencv.willowgarage.com/wiki/>
- [4] G. Bradsky, A. Kaehler. *Learning OpenCV*. O'Reilly.
- [5] Qt references. <http://www.qt-italia.org/>
- [6] Vlc references. <http://www.videolan.org/vlc/>
- [7] R. Cucchiara, M. Piccardi, A. Prati, N. Scarabottolo. *Real-time detection of moving vehicle*. 10th International Conference on Image Analysis and Processing (ICIAP 99).