



Università degli Studi di Salerno

Dipartimento di Ingegneria dell'Informazione ed Elettrica e
Matematica Applicata/DIEM

Dottorato di Ricerca in Informatica e Ingegneria dell'Informazione
XXIX Ciclo

TESI DI DOTTORATO

A Data-Flow Middleware Platform for Real-Time Video Analysis

CANDIDATO: **ROSARIO DI LASCIO**

COORDINATORE: **PROF. ALFREDO DE SANTIS**

TUTOR: **PROF. MARIO VENTO**

Anno Accademico 2016 -2017

Abstract (English)

In this thesis we introduce a new software platform for the development of real-time video analysis applications, that has been designed to simplify the realization and the deployment of intelligent video-surveillance systems.

The platform has been developed following the Plugin Design Pattern: there is an application-independent middleware, providing general purpose services, and a collection of dynamically loaded modules (plugins) carrying out domain-specific tasks.

Each plugin defines a set of node types, that can be instantiated to form a processing network, according to the data-flow paradigm: the control of the execution flow is not wired in the application-specific code but is demanded to the middleware, which activates each node as soon as its inputs are available and a processor is ready. A first benefit of this architecture is its impact on the software development process: the plugins are loosely coupled components that are easier to develop and test, and easier to reuse in a different project. A second benefit, due to the shift of the execution control to the middleware, is the performance improvement, since the middleware can automatically parallelize the processing using the available processors or cores, as well as using the same information or data for different thread of execution.

In order to validate the proposed software architecture, in terms of both performance and services provided by the middleware, we have undertaken the porting to the new middleware of two novel intelligent surveillance applications, by implementing all the nodes required by the algorithms.

The first application is an intelligent video surveillance system based on people tracking algorithm. The application uses a single, fixed camera; on the video stream produced by the camera, background subtraction is performed (with a dynamically updated background) to detect foreground objects. These objects are tracked, and their trajectories are used to detect events of interest, like entering a forbidden area, transiting on a one-way passage in the wrong direction, abandoning objects and so on.

The second application integrated is a fire detection algorithm, which combines information based on color, shape and movement in order to detect the flame.

Two main novelties have been introduced: first, complementary information, respectively based on color, shape variation and motion analysis, are combined by a multi expert system. The main advantage deriving from this approach lies in the fact that the overall performance of the system significantly increases with a relatively small effort made by designer. Second, a novel descriptor based on a bag-of-words approach has been proposed for representing motion.

The proposed method has been tested on a very large dataset of fire videos acquired both in real environments and from the web. The obtained results confirm a consistent reduction in the number of false positives, without paying in terms of accuracy or renouncing the possibility to run the system on embedded platforms.

Abstract (Italiano)

In questa tesi introduciamo una nuova piattaforma software per lo sviluppo di applicazioni di video analisi, progettato per semplificare lo sviluppo e la messa in opera di un sistema di video analisi intelligente.

La piattaforma è stata sviluppata seguendo il Design Pattern Plugin: c'è un middleware indipendente dalla piattaforma che mette a disposizione servizi per vari scopi, ed una collezione di moduli caricati dinamicamente (plugin) per la risoluzione di specifici task.

Ogni plugin definisce un set di tipi di nodi, che possono essere istanziati per formare una rete di elaborazione, in accordo al paradigma data-flow: Il controllo del flusso di esecuzione non è cablato nel codice specifico dell'applicazione ma viene richiesto al middleware che attiva ogni nodo non appena i suoi ingressi sono disponibili e un processore è pronto. Un primo vantaggio di questa architettura è il suo impatto sul processo di sviluppo del software: i plugin sono componenti poco accoppiati che sono più facili da sviluppare e testare e più facilmente riutilizzabili in un altro progetto. Un secondo beneficio, dovuto allo spostamento del controllo di esecuzione al middleware, è il miglioramento delle prestazioni, dal momento che il middleware può automaticamente parallelizzare l'elaborazione utilizzando i processori o i core disponibili, nonché utilizzando le stesse informazioni o dati per differenti thread di esecuzione .

Al fine di convalidare l'architettura software proposta, sia in termini di prestazioni che di servizi forniti dal middleware, è stato effettuato il porting all'interno del middleware di due applicazioni di sorveglianza intelligenti, implementando tutti i nodi richiesti dagli algoritmi.

La prima applicazione è un sistema di videosorveglianza intelligente basato su un algoritmo di tracking delle persone.

L'applicazione utilizza una singola telecamera fissa; sul flusso video prodotto dalla telecamera viene eseguita una sottrazione del background (con un aggiornamento dinamicamente del background) per rilevare oggetti di foreground. Questi oggetti vengono tracciati e le loro traiettorie vengono utilizzate per rilevare eventi di interesse, come accesso in una zona proibita, oggetti abbandonati e così via.

La seconda applicazione integrata è un algoritmo di rilevazione del fuoco che combina informazioni basate su colore, forma e movimento per rilevare le fiamme.

Sono state introdotte due novità principali: in primo luogo, informazioni complementari, rispettivamente basate sul colore, sulla variazione di forma e sull'analisi del movimento, sono combinate tra loro da un sistema multi-esperto. Il vantaggio principale derivante da questo approccio risiede nel fatto che le prestazioni complessive del sistema aumentano significativamente con uno sforzo relativamente piccolo. In secondo luogo, un innovativo descrittore basato su un approccio "bag-of-words" per rappresentare il movimento.

Il metodo proposto è stato testato su un grande dataset di video acquisiti sia in ambienti reali che dal web. I risultati ottenuti confermano una consistente riduzione del numero di falsi positivi, senza pagare in termini di precisione o rinunciare alla possibilità di eseguire il sistema su piattaforme embedded.