

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

SCUOLA DI INGEGNERIA E ARCHITETTURA

Corso di Laurea in Ingegneria Informatica

Porting di un algoritmo per la stima del flusso ottico su smartphone Android

Relatore:

Prof. Stefano Mattoccia

Candidato:

Guglielmo Palaferri

Appello II

Anno Accademico 2020-2021

Introduzione

Il monitoraggio costante della velocità di fiumi e correnti d'acqua può assumere notevole importanza sia nello studio di fenomeni idrologici puramente naturali, sia nella progettazione di opere ingegneristiche strettamente legate ad un particolare flusso d'acqua. Ad esempio, può aiutare ad analizzare e rilevare fenomeni come le inondazioni (specie gli avvenimenti improvvisi, che destano particolare attenzione), così come anche il trasporto di sedimenti o l'erosione delle rocce.

Molte delle tecniche tradizionali utilizzate per l'osservazione di un flusso idrico, tuttavia, non garantiscono grande efficienza e presentano costi elevati: spesso è richiesta la presenza di personale specializzato per la manutenzione di dispositivi complessi.[1] Una soluzione che preveda invece l'installazione di apparecchi ottici, e basi quindi il monitoraggio sull'elaborazione di immagini, può consentire di abbattere notevolmente i costi e di distribuire il sistema di osservazione ottenendo quindi maggiore resistenza ai guasti.

immagine esempio applicazione OTV

È proprio questo un caso di utilizzo di **OTV** (*Optical Tracking Velocimetry*), una tecnica che fa uso di particolari algoritmi di computer vision (in particolare l'algoritmo di Lucas-Kanade, utilizzato per la stima del flusso ottico) per tracciare le traiettorie e le velocità del flusso d'acqua a partire da una serie di immagini. Il tracciamento viene svolto grazie al riconoscimento di particelle quali detriti e altri residui e al confronto di fotogrammi consecutivi.

Il metodo OTV è pensato per essere applicato a dispositivi di elaborazione a basso costo e di dimensioni contenute: questi sarebbero posizionati lungo corsi d'acqua in aree geografiche remote. I dati poi raccolti da questi dispositivi potranno essere

spediti (tramite meccanismi semplici come l'invio di SMS) ad un sistema di raccolta dati centralizzato. Va da sé dunque che l'ottimizzazione dei consumi energetici dei dispositivi costituisca un punto cruciale per la realizzabilità di un tale sistema di monitoraggio. Questo tema verrà preso in considerazione e rappresenta uno dei punti principali degli studi finora condotti sull'argomento.

L'algoritmo è stato inizialmente testato su dispositivi della famiglia *Raspberry*, per via delle loro dimensioni molto contenute e in generale per le funzionalità da essi offerte, molto coerenti con i requisiti del progetto. Le analisi hanno riportato ottimi risultati dal punto di vista dei consumi energetici in particolare dei modelli Raspberry Pi 3B e 4.[2]

Altri dispositivi con buone potenzialità e con un profilo che si presti bene al contesto di utilizzo sono gli **smartphone**, con particolare riferimento a quelli basati su sistema operativo **Android**. L'utilizzo di tali dispositivi richiede ovviamente una seppur minima quantità di modifiche rispetto al deployment effettuato su Raspberry, ed è proprio questo il tema centrale del seguente documento.

Nei prossimi capitoli si procede quindi — dopo aver introdotto qualche informazione necessaria su OTV — a descrivere la realizzazione di un'applicazione per smartphone Android che adatti l'implementazione in C++ di OTV (disponibile su GitHub) e i risultati in termini di prestazioni e consumi energetici che ne sono conseguiti.

Indice

Introduzione	i
1 OTV	1
1.1 Contesto di utilizzo	1
1.2 Ciclo di funzionamento	1
1.3 Ottimizzazioni	3
1.3.1 Ottimizzazioni hardware	4
2 Processo di sviluppo	7
2.1 OpenCV	7
2.1.1 Cross-compilazione	8
2.1.2 Installazione	9
2.2 Sviluppo dell'applicazione Android	10

Capitolo 1

OTV

1.1 Contesto di utilizzo

Come già brevemente descritto, OTV prevede un deployment su dispositivi di dimensioni ridotte e autosufficienti dal punto di vista energetico. In particolare, la configurazione testata su Raspberry introduceva i seguenti componenti:

- Raspberry Pi 3B/4 per l'elaborazione
- Pannello solare 6 W (PiJuice Solar Panel) per sostenere i consumi energetici
- Batteria esterna (PiJuice Hat) per fornire alimentazione

[2] Una simile configurazione verrebbe usata con smartphone, salvo ovviamente l'utilizzo di una batteria aggiuntiva.

Immagine applicazione OTV 2 (quella con lo sketch del fiume usata nel paper nuovo)

1.2 Ciclo di funzionamento

Il dispositivo Android così composto, una volta accuratamente posizionato ed avviato, dovrebbe eseguire *quattro* misurazioni della velocità dell'acqua ogni ora, risultando quindi a regime in un ciclo di funzionamento periodico della durata di 15 minuti.

Sebbene la misurazione mediante l'algoritmo OTV sia svolta sul momento, non viene effettuata sulle immagini direttamente ricevute e lette in input dalla telecamera: il video acquisito necessita di una fase preliminare che prepari le immagini per essere elaborate. Questo viene fatto, tra le altre cose, per consentire di scegliere un settaggio particolare (ad esempio, selezionare una risoluzione diversa rispetto al video originale), utile successivamente al fine di ottimizzare l'elaborazione.

Il ciclo di funzionamento si articola quindi in questo modo:

1. Fase di **acquisizione**: le immagini vengono acquisite dalla telecamera. Questa fase ha una durata fissa e dipende dalla lunghezza del video che si vuole analizzare: tipicamente 20 secondi.
2. Fase di **estrazione** dei frame: a partire dal video acquisito, si estraggono i fotogrammi che lo compongono a seconda della configurazione scelta, in particolare è possibile specificare la risoluzione desiderata tra:
 - Full Resolution (**F**): Risoluzione originale
 - Half Resolution (**H**): Risoluzione dimezzata
 - Quarter Resolution (**Q**): Risoluzione 1/4 dell'originale
3. Fase di **elaborazione** (OTV): a questo punto le immagini estratte vengono effettivamente elaborate utilizzando OTV. Questa fase è cruciale dal punto di vista dei consumi in quanto è quella che può variare maggiormente a seconda della configurazione usata e delle ottimizzazioni implementate. È bene quindi analizzarla di conseguenza.
4. Fase di **idle**: una volta conclusa l'elaborazione (ed eventualmente spediti i dati rilevati) segue un periodo di stand-by, in cui si attende il tempo necessario prima della prossima rilevazione. Anche questa fase è molto importante per determinare i consumi energetici del processo: se il dispositivo dovesse disporre di una modalità di risparmio energetico, l'energia utilizzata potrebbe diminuire drasticamente.

Le fasi su cui è possibile effettivamente lavorare per ottenere risultati migliori sono quelle di elaborazione (in modo particolare) e di idle.

Prima di introdurre i vari livelli di ottimizzazione, va intanto fatto notare che la specifica implementazione di OTV presa in caso è basata sulla libreria open-source di computer vision **OpenCV**.

OpenCV fornisce un framework per la creazione di applicazioni legate alla computer vision e implementa una vasta gamma di algoritmi, tra cui l'algoritmo di Lucas-Kanade utilizzato da OTV già menzionato.

L'utilizzo di OpenCV prescrive una serie di passaggi di installazione che variano in base all'ambiente di sviluppo e che — nel caso specifico di Android — verranno analizzati nel successivo capitolo.

1.3 Ottimizzazioni

Le ottimizzazioni applicabili ad OTV analizzate in [3] consistono in una serie di tecniche e meccanismi che possano contribuire ad aumentare l'efficienza energetica del sistema. Tra queste, possiamo distinguere quelle legate al **software** e quelle invece a livello **hardware** (ad esempio, l'utilizzo di istruzioni particolari).

Le ottimizzazioni software consistono essenzialmente nella configurazione del dispositivo in modo ad esempio da disattivare le opzioni software che risultino superflue (Wi-Fi, Bluetooth ecc.). Si parla di ottimizzazioni derivanti dal sistema operativo utilizzato e dunque dipendenti dal dispositivo in questione. Si vedranno ottimizzazioni di questo tipo esclusive al sistema Android.

Per quanto riguarda le ottimizzazioni software, una di queste è risultata particolarmente efficiente nei test effettuati su Raspberry: l'elaborazione in **scala di grigi** (o **monocromatica**)[2]. Non si tratta di una configurazione a livello di sistema operativo bensì di una piccola modifica al codice di OTV: i fotogrammi precedentemente estratti vengono acquisiti — mediante le API di OpenCV — come immagini monocromatiche invece che a colori. Questo passaggio non comporta risultati particolarmente diversi (in termini di velocità e numero di traiettorie rilevate) ma consente di ottenere un notevole guadagno in termini di prestazioni. Questo deriva dal fatto che le immagini in scala di grigi sono composte da un singolo canale, a differenza delle immagini a colori rappresentate invece da tre canali (Rosso, Verde, Giallo).

1.3.1 Ottimizzazioni hardware

Nel caso delle ottimizzazioni hardware, si parla di particolari metodi che introducono differenti modelli di esecuzione a livello di processore, facendo leva specialmente sulla parallelizzazione delle istruzioni. Questo, oltre agli ovvi vantaggi in termini di performance, può portare ad una maggiore efficienza in termini di consumi. Si delineano tre possibilità principali, eventualmente sovrapponibili, focalizzate su aspetti e modalità diverse di parallelizzazione:

- Esecuzione multi-core mediante **OpenMP** o **TBB**.
- Esecuzione di istruzioni **SIMD** tramite **NEON**
- Esecuzione su **GPU** mediante la libreria **OpenCL**

Immagine confronto tra varie ottimizzazioni nel paper 3

OpenMP e TBB

OpenMP e TBB sono due librerie che in fase di sperimentazione sono state utilizzate per testare il parallelismo *thread-level* in modo da sfruttare i multipli core disponibili nelle moderne CPU. Le due librerie — poiché forniscono lo stesso tipo di funzionalità — sono da utilizzare in modo mutuamente esclusivo. La scelta della libreria da utilizzare cadrà dunque su quella che garantisca le migliori prestazioni.

SIMD

SIMD (*Single Instruction Multiple Data*) è una classe di istruzioni che consente di ottenere parallelismo su un singolo core. Il modello prevede l'esecuzione della stessa operazione su una molteplicità di dati mediante l'utilizzo di una singola istruzione. All'interno di un'istruzione vengono quindi inglobati diversi dati e, ovviamente, l'operazione da svolgersi.

I processori ARM, montati sulla stragrande maggioranza di dispositivi Android così come anche sui modelli di Raspberry Pi già testati, dispongono di un'architettura SIMD avanzata chiamata NEON.

Poiché la libreria OpenCV fornisce nativamente supporto per le istruzioni NEON, risulta piuttosto immediato testare l'utilizzo di tali istruzioni nell'applicazione OTV.

L'utilizzo delle istruzioni SIMD può beneficiare specialmente i casi in cui una singola operazione debba essere ripetuta molte volte su un insieme di dati anche grande. È il caso dell'elaborazione di immagini, in cui spesso è richiesto operare su dati sotto forma di matrici ed eseguire operazioni anche semplici ma molto ripetitive.

GPU e OpenCL

L'utilizzo della potenza di calcolo di una GPU è giustamente considerato tra le ottimizzazioni da testare: qui il parallelismo è intrinseco al tipo di processore, che è progettato appositamente per svolgere operazioni semplici ma ripetitive su un grande insieme di dati, in particolare nei casi di elaborazione di immagini. Tuttavia, come evidenziato in [3] il potenziale guadagno in efficienza che questa soluzione garantirebbe non è facilmente stimabile e dipende da una varietà di fattori.

La libreria OpenCL permette di sfruttare le funzionalità della GPU, qualora supportata dal dispositivo in questione.

Capitolo 2

Processo di sviluppo

Parlando ora dell'effettiva operazione di porting di OTV su dispositivo Android, si espongono gli strumenti utilizzati e le complessità riscontrate durante il processo di sviluppo. Verranno trattati inoltre i dettagli tecnici dell'applicazione e i meccanismi utilizzati propri del sistema Android.

2.1 OpenCV

Come brevemente spiegato nel capitolo precedente, il codice disponibile di OTV è basato su OpenCV. Nel caso di Android, i moduli di OpenCV devono essere necessariamente incapsulati all'interno dell'applicazione, non essendo possibile installarli tra le librerie di sistema. Chiaramente, i moduli dovranno essere compilati con la toolchain di Android per poter ottenere binari effettivamente compatibili con il sistema operativo.

I binari di OpenCV sono reperibili sul sito ufficiale oppure compilabili manualmente tramite una **cross-compilazione**: questa seconda opzione è preferibile in quanto consente di configurare a proprio piacimento l'installazione di OpenCV, aggiungendo o rimuovendo selettivamente le ottimizzazioni introdotte nel capitolo 1.

Tutte le ottimizzazioni hardware discusse, infatti, sono disponibili su OpenCV. Non avrebbe senso invece implementarle nel codice OTV poiché tutte le operazioni che facciano uso intensivo della CPU e che richiedano tempi di esecuzione piuttosto

lunghi sono interne alla libreria OpenCV. L'unica ottimizzazione che effettivamente viene implementata nel codice di OTV è l'elaborazione delle immagini in scala di grigi (non essendo a livello hardware): è sufficiente aggiungere una riga di codice nel momento in cui i frame estratti vengono letti dall'applicazione.

2.1.1 Cross-compilazione

Dovendo testare numerose configurazioni diverse, per la cross-compilazione di OpenCV è stato preparato un semplice script in bash che automatizzasse in minima parte il processo e fornisse inoltre uno scheletro da seguire nel caso in cui si volesse modificare la configurazione di OpenCV, in particolare aggiungendo o rimuovendo ottimizzazioni hardware. Si riporta di seguito il passaggio principale dello script:

```
for ABI in "armeabi-v7a" "arm64-v8a"; do
    mkdir -p build_${ABI}
    cd build_${ABI}
    rm CMakeCache.txt
    cmake \
        -DCMAKE_TOOLCHAIN_FILE=$ANDROID_TOOLCHAIN_PATH \
        -DANDROID_ABI=$ABI \
        -DANDROID_NATIVE_API_LEVEL=$ANDROID_API_LEVEL \
        -DANDROID_STL=c++_shared \
        -GNinja -DCMAKE_BUILD_TYPE:STRING=Release \
        -DBUILD_FAT_JAVA_LIB=OFF \
        -DBUILD_JAVA=OFF \
        -DBUILD_SHARED_LIBS=ON \
        -DWITH_TBB=ON -DBUILD_TBB=ON \
        -DENABLE_NEON=OFF \
        -DWITH_OPENMP=OFF \
        -DWITH_OPENCL=OFF \
        -DWITH_CAROTENE=OFF \
        -DWITH_PTHREADS_PF=OFF \
        -DANDROID_SDK_ROOT=$ANDROID_SDK_ROOT \
        -DANDROID_NDK=$ANDROID_NDK \
        -DOPENCV_ENABLE_NONFREE=ON \
        -DBUILD_TESTS=FALSE -DBUILD_PERF_TESTS=FALSE \
        -DBUILD_ANDROID_EXAMPLES=FALSE \
        -DBUILD_DOCS:BOOL=OFF -DWITH_IPP=OFF -DWITH_MSMF=OFF \
    ..
```

```
ninja
ninja install
cd ..
done
```

In questo caso viene utilizzato **CMake** per configurare la compilazione (scelta obbligatoria, poiché il file CMake è fornito nella distribuzione di OpenCV) e **ninja** come *build system*. CMake infatti non esegue la vera e propria compilazione, ma prepara l'ambiente e crea i file di compilazione che verranno poi usati dal build system per effettivamente compilare il progetto. In alternativa a ninja era possibile utilizzare **Make**, la scelta fra i due non comporta nessuna differenza nel risultato ma può influire leggermente sui tempi di compilazione.

Tra le diverse opzioni di CMake si nota intanto **ANDROID_ABI**, che specifica l'ABI per il quale compilare il progetto. Un ABI (*Application Binary Interface*) rappresenta sostanzialmente l'insieme di set di istruzioni supportate da un dispositivo [4]. Si limita la scelta ad **armeabi-v7a** e **arm64-v8a**, rispettivamente le ABI delle architetture ARM a 32 e 64 bit. Architetture differenti da ARM non verrebbero comunque testate su Android.

Altre opzioni interessanti sono quelle riguardanti le ottimizzazioni hardware: ad esempio **WITH_TBB**, **ENABLE_NEON** o **WITH_OPENCL**. Per abilitare le ottimizzazioni è sufficiente settare a **ON** l'opzione di interesse.

2.1.2 Installazione

Una volta ottenuti i binari di tutti i moduli necessari di OpenCV (tramite cross-compilazione o download), è possibile procedere con l'inserimento all'interno del progetto di Android Studio: come già detto, queste librerie dovranno essere parte integrante dell'applicazione.

L'inserimento è molto meccanico e avviene con la creazione di due cartelle nel progetto: una per contenere gli header di OpenCV, l'altra per contenere le librerie dinamiche contenenti il codice (file **.so**). Sia gli header che le librerie vengono inclusi nel progetto aggiungendo istruzioni ad hoc all'interno del file **CMakeLists.txt** generato automaticamente da Android Studio per compilare la componente nati-

va del progetto che comunichi con gli entry point Java. Segue un esempio con le principali istruzioni CMake utilizzati per aggiungere le librerie di OpenCV:

```
includedirectories("../jniIncludes");
```

2.2 Sviluppo dell'applicazione Android

Il deployment e l'installazione di applicazioni su dispositivi Android avviene mediante l'uso di file .APK. I file APK sono sostanzialmente degli archivi compressi in formato .ZIP il cui contenuto è quindi visualizzabile con un qualsiasi gestore di archivi. Contengono tutti i file necessari per una singola applicazione Android: librerie native, classi Java compilate, ma anche risorse e immagini.

La creazione dei file APK a partire dai file sorgente e dal progetto iniziale spetta solitamente a sistemi di compilazione. Per facilitare questa operazione, viene di norma utilizzato un ambiente di sviluppo e quindi un IDE specializzato.

Nel caso di Android l'IDE consigliato è Android Studio, che fornisce un bundle di installazione contenente tutti gli strumenti e le toolchain necessarie, facilmente configurabili.

Il linguaggio principale supportato da Android è **Java**: ogni applicazione richiede anche solo una minima parte scritta in Java per poter utilizzare le API di sistema e gestire i cosiddetti *Entry Point* dell'applicazione, ovvero i meccanismi tramite i quali l'utente può avviare ed interagire con l'applicazione, l'interfaccia utente è infatti manipolabile esclusivamente con l'utilizzo di classi Java.

È comunque disponibile il supporto per codice nativo C/C++ grazie all'NDK (*Native Development Kit*): un insieme di strumenti per la compilazione e il debug integrabili con Android Studio. Lo strumento di configurazione di default per codice nativo è CMake, tra i più diffusi per progetti C e C++, il cui utilizzo è stato brevemente dimostrato nella sezione di cross-compilazione di OpenCV.

Per configurare infine il progetto nella sua interezza, compresi i dettagli di deployment, Android Studio richiede l'utilizzo del plugin **Gradle**, che può essere integrato

con CMake (in modo molto automatizzato) qualora sia prevista la presenza di codice nativo.

Bibliografia

- [1] F. Tauro, F. Tosi, S. Mattoccia, E. Toth, R. Piscopia, and S. Grimaldi, “Optical tracking velocimetry (OTV): Leveraging optical flow and trajectory-based filtering for surface streamflow observations,” *Remote Sensing*, vol. 10, no. 12, 2018. [Online]. Available: <https://www.mdpi.com/2072-4292/10/12/2010>
- [2] A.-H. Livoroi, A. Conti, L. Foianesi, F. Tosi, F. Aleotti, M. Poggi, F. Tauro, E. Toth, S. Grimaldi, and S. Mattoccia, “On the deployment of out-of-the-box embedded devices for self-powered river surface flow velocity monitoring at the edge,” *Applied Sciences*, vol. 11, no. 15, 2021. [Online]. Available: <https://www.mdpi.com/2076-3417/11/15/7027>
- [3] F. Tosi, M. Rocca, F. Aleotti, M. Poggi, S. Mattoccia, F. Tauro, E. Toth, and S. Grimaldi, “Enabling image-based streamflow monitoring at the edge,” *Remote Sensing*, vol. 12, no. 12, 2020. [Online]. Available: <https://www.mdpi.com/2072-4292/12/12/2047>
- [4] “Android ABIs.” [Online]. Available: <https://developer.android.com/ndk/guides/abis>