

Guida a FFmatch

Frank Sinapsi

5 gennaio 2012

Questo documento è una guida all'uso di FFmatch mediante la descrizione abbastanza dettagliata dell'applicazione del metodo a un esempio reale che ho scelto appositamente per scopi didattici.

Ho scelto un film americano che è stato trasmesso da una emittente televisiva italiana (doppiato in italiano). A quanto ne so, di tale film non esiste il dvd in italiano, ma ho potuto recuperarne una registrazione televisiva in AVI/Xvid (d'ora in avanti la indicherò con il termine *tvrip*) e un dvdrip in lingua originale (d'ora in avanti lo indicherò con *dvdrip*), anch'esso in AVI/Xvid.

Il tvrip ha una qualità video abbastanza scadente, mentre la qualità del dvdrip è nettamente superiore. Si tratta quindi di un caso tipico in cui si vorrebbe “mettere l'audio italiano nel dvdrip in inglese”.

Inoltre il tvrip ha delle interruzioni pubblicitarie (che sono state tolte da chi lo ha codificato), e il dvdrip ha un framerate di 23.976 fps.

Il primo problema che si presenta è che il tvrip ha delle grosse bande nere sopra e sotto. Ecco un fotogramma:



Le bande nere possono avere effetti negativi sul riconoscimento, e al momento FFmatch non ha un filtro di crop, quindi conviene preparare un nuovo filmato senza quelle grosse bande.

Per fare questo io procedo in questo modo, ma naturalmente ognuno può usare altri metodi equivalenti che gli sono più congeniali. Batto questa riga di comando:

```
mplayer Un\ ragazzo\ come\ gli\ altri.avi -vf cropdetect
```

Il filtro `cropdetect` di `mplayer` ha l'effetto di mostrare i valori di crop necessari per togliere le bande nere, e in questo caso ottengo le seguenti informazioni:

```
[CROP] Crop area: X: 0..479 Y: 92..482 (-vf crop=480:384:0:96).0 0
```

In genere per ottenere un crop più preciso è consigliabile controllare cosa viene croppato copiando i valori forniti da `mplayer` in un comando come questo:

```
mplayer Un\ ragazzo\ come\ gli\ altri.avi -vf rectangle=480:384:0:96
```

Ma in questo caso non c'è bisogno di andare tanto per il sottile, in quanto il filmato servirà esclusivamente per essere analizzato da `FFmatch`.

Quindi passo a generare la copia croppata, con questo comando:

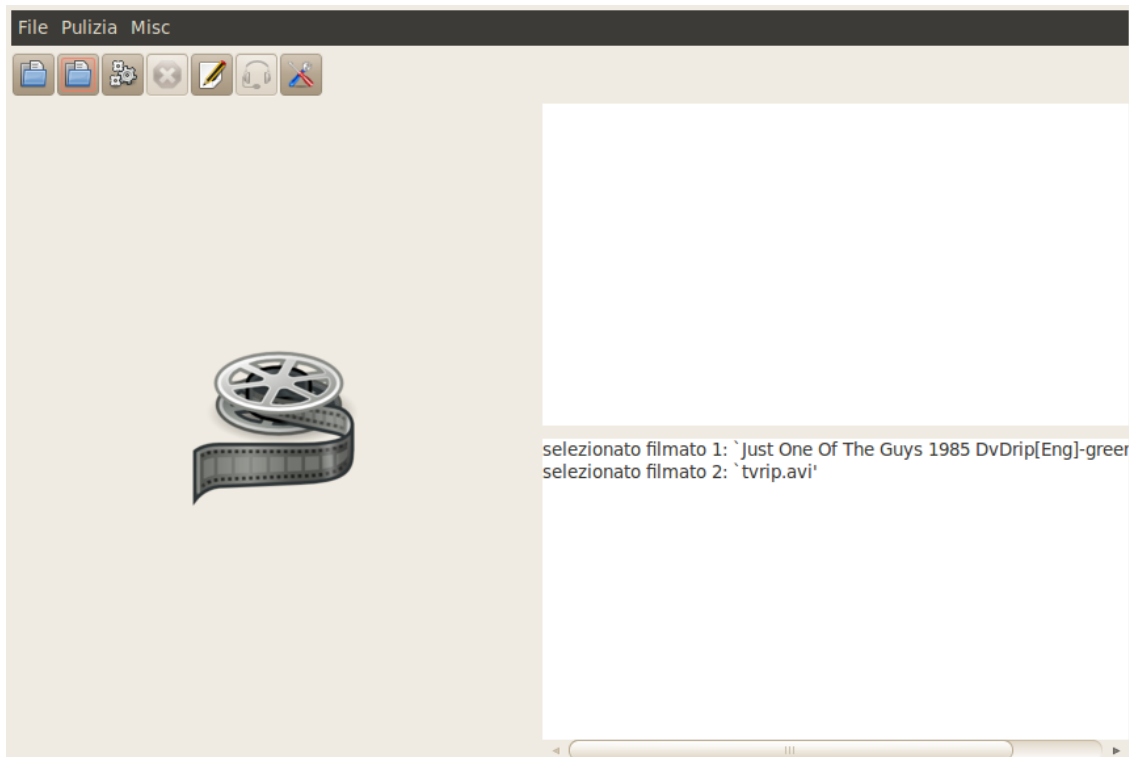
```
mencoder Un\ ragazzo\ come\ gli\ altri.avi -vf crop=480:384:0:96,harddup -ovc xvid -xvidencopts fixed_quant=2 -nosound -mc 0 -o tvrip.avi
```

Il comando precedente mi ha creato il file `tvrip.avi`, che è una versione AVI/Xvid croppata e senza audio del `tvrip` di partenza.

Se uno dei filmati di partenza fosse in un formato contenitore non supportato da `FFmatch` (attualmente sono supportati AVI e MP4), si procede in modo analogo, cioè si crea una copia temporanea “usa e getta” in AVI/Xvid (eventualmente croppando e scalando a dimensioni inferiori, se ciò dovesse essere necessario). Tale copia servirà solo per essere analizzata da `FFmatch`, dopodiché potrà essere cancellata.

Ciò che conta è che i frame dei filmati dati in pasto a `FFmatch` siano nelle stesse posizioni dei filmati originali.

Dopo questo lavoro preparatorio, si può lanciare `FFmatch` e selezionare i due filmati usando i primi due pulsanti in alto a sinistra, nella barra dei pulsanti.



La selezione di entrambi i filmati ha come effetto l'abilitazione del terzo pulsante (quello con le rotelline), che serve per far partire l'analisi; non resta che cliccarci sopra e aspettare che il programma si fermi.

Nel mio caso, quando il processo si arresta, nel riquadro di sopra ottengo queste righe:

```
(1) Just One Of The Guys 1985 DvDrip[Eng]-greenbud1969.avi
(2) tvrip.avi
```

```
free (1): 5680/144327
free (2): 1664/139753
```

```
controlli: 1522 (passati: 1494)
rapporto orizzontale: 0.985
rapporto verticale: 1.033
```

```
[000548-025545] ----> [000001-024998] {024998} (diff=-547)
[025546-025591] -/-> ??? {46-41}
[025592-027504] ----> [025040-026952] {001913} (diff=-552)
[027505-027834] -/-> ??? {330-327}
[027835-049744] ----> [027280-049189] {021910} (diff=-555)
[049748-054522] ----> [049190-053964] {004775} (diff=-558)
[054525-055658] ----> [053965-055098] {001134} (diff=-560)
[055662-072744] ----> [055099-072181] {017083} (diff=-563)
[072745-072757] -/-> ??? {13-7}
[072758-077570] ----> [072189-077001] {004813} (diff=-569)
[077571-077702] -/-> ??? {132-128}
[077703-078653] ----> [077130-078080] {000951} (diff=-573)
[078654-078676] -/-> ??? {23-36}
[078677-085230] ----> [078117-084670] {006554} (diff=-560)
[085231-085413] -/-> ??? {183-179}
[085414-109621] ----> [084850-109057] {024208} (diff=-564)
[109622-109688] -/-> ??? {67-42}
[109689-116406] ----> [109100-115817] {006718} (diff=-589)
[116410-139441] ----> [115818-138849] {023032} (diff=-592)
[139442-144327] -/-> ??? {4886-904}
```

frame(s) del primo filmato che non sono presenti nel secondo:

```
[000001-000547] ----> [000000] {547}
[049745-049747] ----> [049189] {3}
[054523-054524] ----> [053964] {2}
[055659-055661] ----> [055098] {3}
[116407-116409] ----> [115817] {3}
```

Le prime due righe riportano i nomi dei filmati analizzati.

Nelle successive due righe sono riportati il numero di frame che sono rimasti liberi (non collegati) sul totale, per entrambi i filmati.

Segue il numero totale di controlli effettuati, e quanti di questi hanno dato esito positivo. Un controllo è un tentativo di mettere in corrispondenza un certo frame di un filmato con un certo frame dell'altro filmato. FFmatch non effettua controlli alla cieca, ma è guidato da uno schema di pattern matching applicato alle lunghezze delle scene. Quindi la maggior parte dei controlli effettuati ha un'alta probabilità di dare esito positivo. Per questo il numero di controlli falliti è di solito relativamente basso.

I rapporti orizzontale e verticale sono i fattori di espansione (se il rapporto è maggiore di 1) o di contrazione (se minore di 1) tra il primo e il secondo filmato. Tali rapporti non coincidono

necessariamente con i rapporti tra larghezze e altezze dei due filmati, perché questi potrebbero avere subito un cropping differente.

L'elenco successivo è la *mappatura* tra i due filmati, cioè la sequenza degli intervalli¹ nei quali il programma ha ripartito i filmati in base alle differenze interne rilevate.

Nel riquadro gli intervalli vengono mostrati in due colori diversi.

Gli intervalli in verde sono quelli che il programma ha collegato a intervalli dell'altro filmato; si noti che due intervalli collegati hanno lo stesso numero di frames, indicato tra parentesi graffe.

Gli intervalli in rosso sono intervalli del primo filmato che per qualche motivo il programma ha rinunciato a collegare. Sono quindi una parte dei frame riportati come liberi.

Possono esserci altri frame liberi, non riportati nelle righe rosse. Gli intervalli in rosso che vengono mostrati sono solo quelli in cui per entrambi i filmati c'è almeno un frame libero; le ampiezze di tali intervalli sono indicate tra parentesi graffe, nella forma $\{n1-n2\}$, dove $n1$ e $n2$ sono entrambi strettamente positivi.

Per rendere la mappatura più facile da leggere, gli intervalli liberi di un filmato che corrispondono a un intervallo vuoto dell'altro non vengono mostrati come righe rosse (dove una delle due lunghezze sarebbe stata 0), ma nelle sezioni successive del rapporto, come "frame di un filmato che non sono presenti nell'altro".

Per ottenere buoni risultati, è indispensabile capire bene le notazioni usate nel rapporto.

Un'altra cosa importante da capire è che la nuova traccia audio verrà costruita partendo dalla traccia audio in inglese, sostituendo a questa tutte le zone verdi con le corrispondenti zone della traccia audio in italiano.

Questo vuol dire, per fare un esempio, che se nel rapporto non c'è nemmeno una riga verde, la traccia ricostruita sarà una copia esatta della traccia in inglese.

Tornando all'esempio, possiamo notare che il programma ha mappato un alto numero di intervalli, ma ne ha anche lasciati alcuni non risolti (per l'esattezza 7 interni, e uno finale).

A questo punto si entra nella fase più impegnativa per l'utente: l'eliminazione di eventuali errori di mappatura e il completamento delle parti non risolte.

Per fare ciò dobbiamo passare nella *modalità supervisore*, cliccando sull'apposito pulsante con l'icona della matita. Così facendo compariranno alcuni strumenti tra i due riquadri nella parte destra della finestra, e in quella sinistra verranno mostrati i frame dei due filmati.

Gli strumenti della modalità supervisore sono questi:



Nella prima riga abbiamo, partendo da sinistra:

- un combo box (che d'ora in poi chiamerò semplicemente *combo*);
- uno spin button numerico;
- sei frecce blu, tre rivolte verso sinistra e tre verso destra;
- un secondo spin button numerico.

Nella seconda riga abbiamo:

- un primo gruppo costituito da uno spin button numerico e due frecce verdi;
- un secondo gruppo identico al primo eccetto il fatto che è preceduto da un pulsante, che in seguito chiamerò *pulsante di match*;

¹Un *intervallo*, o *zona*, è una sequenza di frame consecutivi di un filmato.

- un pulsante finale, che in seguito chiamerò *pulsante finale*.

Ci sono alcune caratteristiche che ogni utente di FFmatch dovrebbe subito imparare a memoria:

- tutti questi strumenti, con la sola eccezione del pulsante finale, permettono soltanto di fare spostamenti sui filmati;
- il pulsante finale è l'unico strumento che permette di modificare la mappatura, e soltanto dopo aver chiesto conferma all'utente;
- le frecce blu hanno la caratteristica di lasciare sempre inalterata la differenza tra le posizioni dei due filmati.

Gli spostamenti sono di due tipi: *sincroni* (o *vincolati*) e *svincolati*.

Uno spostamento è sincrono se lascia inalterata la differenza tra le posizioni. Quindi l'ultima caratteristica può essere riformulata così: "le frecce blu effettuano solo spostamenti sincroni".

Inoltre, le frecce blu sono gli unici strumenti che effettuano spostamenti sincroni.

Uno spostamento è svincolato se ha effetto sulla posizione di uno solo dei due filmati.

Ci sono tre tipi di frecce blu. Quelle al centro effettuano spostamenti di una singola posizione; poi ci sono quelle che spostano di 10 posizioni². Infine, le frecce blu laterali spostano sugli estremi dell'intervallo corrente (in riferimento al primo filmato).

I due spin della riga superiore permettono di fare spostamenti svincolati sui filmati, e mostrano sempre le posizioni correnti.

Le frecce verdi della riga inferiore permettono di saltare tra le scene dei filmati: cliccando sulla prima freccia verde rivolta a destra si salta al frame in cui inizia la prossima scena nel primo filmato; e così via. . .

Gli spin situati alla sinistra delle frecce verdi sono soglie che servono per regolare il rilevamento degli inizi di scena. Aumentando tali soglie si otterranno salti più ampi, perché verranno considerati "cambi di scena" solo quei frame che differiscono molto dal frame che li precede (in base al valore di soglia impostato); abbassando tali soglie fino al valore minimo (0), si otterranno salti sempre più piccoli fino a spostamenti di un singolo frame (con il valore minimo, cioè 0, ogni frame viene considerato un "cambio di scena").

Il pulsante di match può avere come effetto uno spostamento soltanto sul secondo filmato. Se il frame corrente del primo filmato è attualmente mappato (cioè, non libero), cliccando sul pulsante di match si otterrà uno spostamento del secondo filmato sul frame attualmente collegato a quello del primo filmato. In questo modo uno può controllare se il frame mostrato di sopra coincide con quello mostrato di sotto.

Ma la vera utilità del pulsante di match si ha quando lo si usa sui frame liberi; in tal caso, se possibile, il programma posizionerà il secondo filmato sui frame che ritiene più probabili, in base a considerazioni su allineamenti. Cliccando più volte sul pulsante di match è possibile che il programma si posizioni (a rotazione) su più frame probabili.

Il combo permette di spostarsi, su entrambi i filmati, sugli estremi degli intervalli dell'attuale mappatura.

È importante che l'utente prenda confidenza con il combo e con le notazioni usate per denotare le sue entries: gli estremi inferiori degli intervalli sono caratterizzati dalla parentesi quadra aperta, quelli superiori dalla parentesi quadra chiusa, e quelli non risolti dalle parentesi tonde.

Il pulsante finale, come ho già detto, è l'unico che permette di modificare la mappatura corrente.

Il suo funzionamento è il seguente: se si clicca sul pulsante finale quando il frame corrente del primo filmato è libero, il programma lo interpreterà come una richiesta di creazione del collegamento tra tale frame e il frame corrente del secondo filmato. Se la creazione di tale collegamento è possibile, il programma chiederà conferma all'utente, e in caso di risposta affermativa creerà il collegamento.

Se, invece, i due frame correnti sono attualmente già collegati tra loro, il programma chiederà all'utente se desidera che il collegamento venga rimosso.

²Questo valore può essere modificato nella finestra delle preferenze.

In quest'ultimo caso occorre fare molta attenzione, perché il collegamento che viene rimosso è in realtà un collegamento tra due intervalli che contengono i due frame correnti, e tipicamente tali intervalli comprenderanno altri frame oltre a quei due.

Questo è tutto ciò che c'è da sapere sugli strumenti della modalità supervisore.

La tecnica base per completare una mappatura consiste nel selezionare una dietro l'altra le entries del combo; su quelle che hanno almeno una parentesi quadra si controlla se il frame mostrato di sopra coincide con quello di sotto. Se sono anche leggermente diversi, si clicca sul pulsante finale per rimuovere il collegamento sbagliato.

Sulle entries con parentesi tonde occorre decidere, di volta in volta, come conviene “chiudere” la mappatura. In casi molto particolari potrebbe anche essere preferibile lasciare degli intervalli non risolti.

Giocando un po' con il programma, l'utente imparerà alcune strategie per svolgere velocemente questo compito. Un “trucco” consiste nello stabilire collegamenti in posizioni cruciali, e far partire nuovamente l'analisi. Se i collegamenti forzati manualmente vengono scelti in modo opportuno, gli algoritmi di analisi li sfrutteranno per portare avanti la creazione di nuovi collegamenti. E questa tecnica può essere ripetuta più volte. Occorre però fare attenzione al fatto che, se si sono rimossi dei collegamenti sbagliati e si fa partire nuovamente l'analisi, è abbastanza probabile che il programma crei nuovamente gli stessi collegamenti sbagliati. Quindi una buona strategia è quella di forzare collegamenti corretti e “cruciali” subito dopo aver rimosso un collegamento sbagliato; con “cruciali” intendo collegamenti che impediscano al programma di ricreare il collegamento sbagliato, nel caso in cui venga fatta ripartire l'analisi.

Applicando le tecniche accennate sopra all'esempio iniziale, mi sono portato in questa situazione:

```
(1) Just One Of The Guys 1985 DvDrip[Eng]-greenbud1969.avi
(2) tvrip.avi
```

```
free (1): 4886/144327
free (2): 904/139753
```

```
controlli: 1619 (passati: 1581)
rapporto orizzontale: 0.985
rapporto verticale: 1.033
```

```
[000548-025571] ----> [000001-025024] {025024} (diff=-547)
[025577-027573] ----> [025025-027021] {001997} (diff=-552)
[027576-027746] ----> [027022-027192] {000171} (diff=-554)
[027748-049649] ----> [027193-049094] {021902} (diff=-555)
[049653-054525] ----> [049095-053967] {004873} (diff=-558)
[054528-055658] ----> [053968-055098] {001131} (diff=-560)
[055662-072748] ----> [055099-072185] {017087} (diff=-563)
[072755-077636] ----> [072186-077067] {004882} (diff=-569)
[077641-078669] ----> [077068-078096] {001029} (diff=-573)
[078670-085252] ----> [078110-084692] {006583} (diff=-560)
[085257-109655] ----> [084693-109091] {024399} (diff=-564)
[109681-116406] ----> [109092-115817] {006726} (diff=-589)
[116410-139441] ----> [115818-138849] {023032} (diff=-592)
[139442-144327] -/-> ??? {4886-904}
```

```
frame(s) del primo filmato che non sono presenti nel secondo:
[000001-000547] ----> [000000] {547}
```

```

[025572-025576] ----> [025024] {5}
[027574-027575] ----> [027021] {2}
[027747-027747] ----> [027192] {1}
[049650-049652] ----> [049094] {3}
[054526-054527] ----> [053967] {2}
[055659-055661] ----> [055098] {3}
[072749-072754] ----> [072185] {6}
[077637-077640] ----> [077067] {4}
[085253-085256] ----> [084692] {4}
[109656-109680] ----> [109091] {25}
[116407-116409] ----> [115817] {3}

```

frame(s) del secondo filmato che non sono presenti nel primo:
[078097-078109] ----> [078669] {13}

Ho deciso di lasciare l'ultimo intervallo non risolto, perché pochi frame dopo l'estremo superiore dell'ultima riga verde iniziano i titoli di coda. Dato che questi sono troncati nel tvrip (si può notare che nel tvrip sono rimasti solo 904 frames contro i 4886 del dvdrip) e che da lì in poi non ci sono più dialoghi ma solo musica, lasciando l'ultima riga rossa si otterrà come risultato il passaggio dal doppiato italiano alla musica del dvdrip appena un po' prima dell'inizio dei titoli di coda.

A questo punto la mappatura può essere considerata completa. Non resta altro che generare la traccia audio. E qui sorge un altro problema, molto specifico: lo stream audio del tvrip è PCM non compresso con una frequenza di campionamento di 44100 Hz:

```

Audio tag: 0x1 (PCM)
Canali: 2
Freq. campionamento: 44100 Hz
Dimensione stream: 986,001,408 bytes (940.324219 MiB)

```

Estraendola dall'AVI ottengo quindi una traccia WAV non compressa di 986,001,452 bytes che chiamo `ita44100.wav`³. Tuttavia il dvdrip ha una traccia audio di tipo:

```

Audio tag: 0x55 (MP3)
Canali: 2
Freq. campionamento (bs): 48000 Hz

```

Inoltre devo tenere in considerazione che il dvdrip ha un framerate di 23.976 fps e il tvrip di 25 fps.

La mia scelta è di creare un filmato con framerate 25 fps, usando ovviamente il video del dvdrip.

Guardando la mappatura è evidente che la traccia audio in italiano sarà costituita per la maggior parte dall'audio del tvrip, che quindi è già pronto per essere usato (file `ita44100.wav`).

Resta quindi soltanto da preparare una traccia WAV non compressa a 44100 Hz del dvdrip, tenendo conto del fatto che il video verrà accelerato da 23.976 a 25 fps.

Bene, la procedura consiste nell'estrarre dal dvdrip un file MP3, che chiamo `eng.mp3`, e decomprimerlo in formato WAV, ad esempio con questo comando:

```

madplay eng.mp3 -o wav:eng48000.wav

```

In questo modo ho ottenuto il file `eng48000.wav`, che ho chiamato così per ricordarmi che ha 48000 campioni al secondo.

³I 44 bytes in più, rispetto allo stream composto dai soli campioni, costituiscono l'header iniziale RIFF del WAV.

Adesso va fatto un resampling, tenendo conto di due vincoli:

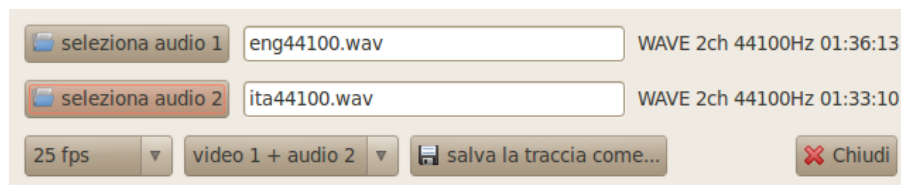
1. devo ottenere un WAV a 44100 Hz, per essere compatibile con `ita44100.wav`;
2. il video verrà accelerato di un fattore $25/23.976$.

Dato che moltiplicando 48000 per $25/(24000/1001)$ si ottiene 50050, il resampling che mi serve può essere ottenuto con questo comando:

```
sox -r 50050 eng48000.wav -r 44100 eng44100.wav
```

L'utility `sox` ha tante opzioni (filtri) che possono essere usate per effettuare il resampling, ma usando una versione non troppo vecchia del programma può anche essere sufficiente lo scarno comando che ho riportato sopra, considerando anche che la traccia in inglese verrà quasi interamente rimpiazzata dai campioni prelevati dalla traccia in italiano (dell'audio originale resteranno infatti soltanto la canzoncina finale sui titoli di coda, l'audio dei primi 548 frames del dvdrip, e qualche "buco" di pochi frames sparsi qua e là nel corso del film).

Adesso ho tutto ciò che serve per generare la traccia in italiano, quindi clicco sul pulsante con le cuffie e seleziono le due tracce WAV che ho preparato (`eng44100.wav` per il primo filmato e `ita44100.wav` per il secondo); inoltre seleziono l'opzione "video 1+audio 2", dato che voglio adattare l'audio del tvrip (filmato 2) al video del dvdrip (filmato 1):



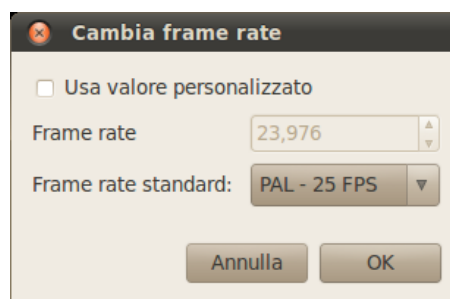
Infine clicco sul pulsante *salva la traccia come...* e salvo la nuova traccia WAV come `ffmatch.wav`.

Il file che ho ottenuto (`ffmatch.wav`) ha la stessa dimensione (in bytes) del file `eng44100.wav`. Non si tratta di un caso poiché, come ho già detto in precedenza, la traccia audio generata da FFmatch non è che una copia dell'audio in inglese in cui sono state sostituite le zone verdi con le corrispondenti zone in italiano.

L'ultima cosa che resta da fare è ricomprimere e muxare l'audio ottenuto al dvdrip, ricordando che quest'ultimo va accelerato da 23.976 a 25 fps.

Questa operazione può essere facilmente fatta usando uno tra tanti programmi esistenti.

Ad esempio, lanciando `Avidemux` e aprendo il dvdrip, per cambiare il framerate nominale (senza ricodifica del bitstream video) da 23.976 a 25 fps basta cliccare su Video->FrameRate e impostare i dati in questo modo:



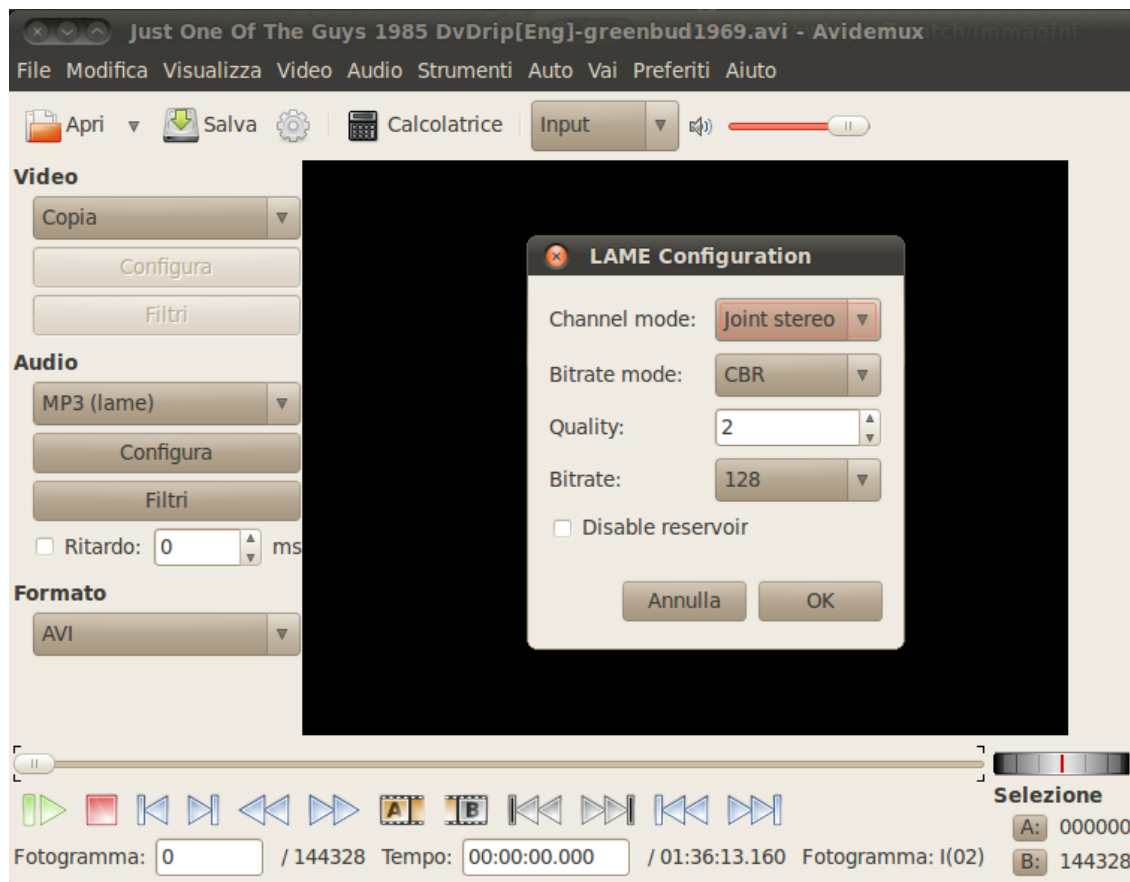
Poi si conferma il cambiamento cliccando su *OK*.

Per sostituire la traccia audio con quella generata da FFmpeg occorre cliccare su Audio->TracciaPrincipale, selezionare “Sorgente Audio: WAV esterno”, cliccare su *Browse...* e selezionare il file `ffmpeg.wav` salvato in precedenza con FFmpeg:



Poi naturalmente si conferma l'impostazione cliccando su *OK*.

Adesso occorre selezionare l'algoritmo di compressione per la traccia WAV non compressa, usando le opzioni per l'audio sulla sinistra della finestra:



Nella schermata sopra si può notare che ho optato per una compressione in MP3 128 kbps CBR joint stereo, usando la libreria libmp3lame.

Non resta che salvare il nuovo AVI, cliccando su *Salva*. Durante il salvataggio Avidemux provvederà a comprimere l'audio in MP3; il video, invece, non subirà alcuna modifica.