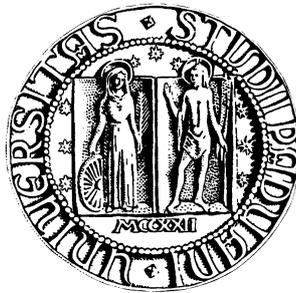


UNIVERSITÀ DEGLI STUDI DI PADOVA

FACOLTÀ DI SCIENZE STATISTICHE

CORSO DI PIANO DEGLI ESPERIMENTI I



ALGORITMI MPEG-4 PER LA COMPRESSIONE DI
FILE VIDEO: XVID CONTRO X264

Alberto Cavalin

ANNO ACCADEMICO 2005–2006

Indice

1	Lo studio	3
1.1	Lo scopo	3
1.1.1	Scopo primario	3
1.1.2	Scopo secondario	3
1.2	Il tipo	3
1.2.1	I trattamenti	4
1.2.2	I blocchi	4
1.3	La popolazione	5
1.3.1	Criteri di inclusione	5
1.3.2	Criteri di esclusione	5
2	Programmazione e disegno	7
2.1	Impostazioni generali	7
2.2	Metodo di randomizzazione	8
2.3	Gestione dei file	8
2.4	Valutazioni	9
2.4.1	Indice per lo scopo primario	9
2.4.2	Indice per lo scopo secondario	9
3	Metodi statistici	11

3.1	La dimensione campionaria	11
3.2	Il modello	12
3.3	La randomizzazione	13
3.4	Metodi di analisi	13
3.4.1	Considerazioni generali	13
3.4.2	Verifica delle assunzioni	13
3.4.3	L'analisi	14
3.5	Conclusioni	15
	Bibliografia	17

Introduzione

I codec XviD e x264 sono due implementazioni diverse dell'algoritmo MPEG-4, creato appositamente per offrire una compressione elevata di contenuti visivi.

Un codec (enCOder/DECoder) è un programma che, dato un flusso di dati in ingresso, lo restituisce codificato/decodificato secondo un algoritmo predefinito.

MPEG-4 è uno standard creato dal gruppo "Moving Picture Experts Group" (MPEG in breve) all'insegna di un forte risparmio sul bitrate¹, ed è anche diventato uno standard ISO (#14496) nel 2001. XviD segue lo standard MPEG-4 Part-2, conosciuto anche come MPEG-4 ASP (Advanced Simple Profile), mentre l'x264 implementa il più recente MPEG-4 Part-10, conosciuto anche come MPEG-4 AVC (Advanced Video Coding); ambedue permettono un gran risparmio di spazio mantenendo allo stesso tempo un'ottima qualità per il prodotto finale.

I dispositivi multimediali portatili attuali (CD, DVD, HD, PEP², ecc...) hanno ancora capacità poco elevate, ed è per questo che risulta di cruciale importanza la scelta del codec, in modo tale da minimizzare lo spazio occupato, immagazzinare più contenuti e quindi trarre più profitto/efficienza in termini di fruizione/durata del bene.

¹Quantità di bit usati per unità di tempo. L'unità di misura comune è il "kbps", e indica le migliaia di bit per secondo

²Personal Entertainment Player

Capitolo 1

Lo studio

1.1 Lo scopo

1.1.1 Scopo primario

Valutare la superiorità del codec x264 rispetto a XviD nella quantità di spazio occupato su disco.

1.1.2 Scopo secondario

Valutare la superiorità del codec x264 rispetto a XviD nella quantità di tempo utilizzato per la compressione.

1.2 Il tipo

Il tipo di studio che verrà effettuato è una piano fattoriale 2^k a blocchi, con $k=2^1$.

¹k fattori, 2 livelli per fattore

1.2.1 I trattamenti

Dato il formato debolmente compresso del DVD (MPEG-2)², i trattamenti che verranno utilizzati durante lo studio sono l'applicazione dei seguenti encoder:

- XviD
- x264

utilizzando i seguenti valori consigliati (vedere [1, Manuale Transcode]) di bitrate:

- 1000 kbps (1h:45' video \simeq 700MB)
- 1800 kbps (1h:45' video \simeq 1350MB)

NB: Il valore di bitrate fornito all'encoder è da intendersi come una sorta di riferimento per l'applicazione dell'algoritmo stesso, infatti il bitrate risultante non è costante ma variabile (se fosse costante l'intero algoritmo non sarebbe efficiente ed il confronto non avrebbe senso).

1.2.2 I blocchi

Il genere del file video non è un parametro di interesse ma è di disturbo perchè influisce sul risultato finale; i file verranno quindi randomizzati all'interno dei seguenti blocchi³:

- azione
- film-opera
- documentario

²MPEG-2 bitrate \simeq 6000 kbps

³Si veda: http://it.wikipedia.org/wiki/Generi_cinematografici

1.3 La popolazione

Un totale di 96 file verranno suddivisi in base al genere di appartenenza, in 3 gruppi da 32 unità ciascuno, ed ogni gruppo verrà ripartito equamente tra i propri trattamenti.

1.3.1 Criteri di inclusione

- file video in formato MPEG-2 (formato standard per i DVD)
- file video che appartengono ai generi:
 - azione
 - film-opera
 - documentario
- file video di durata compresa tra 1h:30' e 2h:00'

1.3.2 Criteri di esclusione

- file video in formato MPEG-2 codificati con bitrate inferiori a 5900 kbps
- file video non integri/parziali (detti anche zero filled/padded)
- file video che contengono lunghe sequenze di immagini fisse per una percentuale maggiore al 15% della durata complessiva

Capitolo 2

Programmazione e disegno

2.1 Impostazioni generali

I file obiettivo dello studio verranno reperiti e selezionati dagli archivi della S.I.A.E. previa gentile concessione della stessa.

L'utilizzo degli encoder verrà effettuato su di un personal computer con processore non inferiore ai 3GHz, con almeno 1GB di RAM, con due dischi da 200GB¹ in configurazione Raid 0² per contenere i file MPEG-2, ed un disco da 80GB³ per i file MPEG-4. Il sistema operativo in uso sarà una Linux box con installato il programma `transcode` ed i relativi codec `xvid4` ed `x264`.

La quantità di tempo richiesta per l'elaborazione dei files e la successiva raccolta ed elaborazione dei risultati è stata fissata pari a 30 giorni⁴ (utilizzando il computer 24h/24h).

Eventuali crash⁵ degli encoder implicheranno l'esclusione dallo studio dei file

¹4.4GB * 90 = 396 \implies 2x 200GB

²I due dischi figurano come un disco unico

³400GB * 0.14 = 56 \implies 80GB

⁴Dalla letteratura: i tempi di compressione non superano mai le 7 ore, perciò' $7*90/24 \simeq 27$ gg

⁵Interruzione inaspettata/Malfunzionamento del programma

generanti tali situazioni.

2.2 Metodo di randomizzazione

La randomizzazione va effettuata all'interno dei blocchi e si desidera avere pari numerosità per ogni trattamento, si userà quindi la tipologia *a blocchi permutati*. In questo caso si hanno 4 trattamenti e 8 file per trattamento, considerando le tavole di permutazione di 8 numeri, le assegnazioni verranno effettuate sulla base delle seguenti regole di associazione:

- {0,1} → T1 (XviD, 1000 kbps)
- {2,3} → T2 (XviD, 1800 kbps)
- {4,5} → T3 (x264, 1000 kbps)
- {6,7} → T4 (x264, 1800 kbps)

I file così scelti verranno prima suddivisi in 3 cartelle indicanti il genere, e poi in altre 4 sotto cartelle a seconda del trattamento.

2.3 Gestione dei file

Dopo aver organizzato i file secondo i precedenti criteri, verranno create le stesse cartelle nel disco di destinazione, e si comprimeranno i file con i seguenti comandi a seconda del bitrate e dell'encoder:

```
transcode -x mpeg2,null -y xvid4,null -w 1000
transcode -x mpeg2,null -y xvid4,null -w 1800
transcode -x mpeg2,null -y x264,null -w 1000
transcode -x mpeg2,null -y x264,null -w 1800
```

il comando `transcode` accetta inoltre i seguenti parametri per selezionare il file in ingresso e quello in uscita: `-i file.mp2 -o file.avi`.

Per estrarre le informazioni riguardo la dimensione dei file, basta semplicemente elaborare l'output del comando `ls -lR` tramite il comando `sed` al termine dello studio. Per quanto concerne i tempi di esecuzione, essi verranno misurati per ogni file col comando `time` utilizzando la stringa di formato `%e`⁶.

2.4 Valutazioni

2.4.1 Indice per lo scopo primario

Tale indice verrà espresso dal rapporto delle dimensioni del file prima e dopo la compressione:

$$\rho = \frac{\text{dimensione}_{MPEG-4}}{\text{dimensione}_{MPEG-2}} \in (0, 1)$$

2.4.2 Indice per lo scopo secondario

Tale indice verrà espresso dal numero di secondi trascorsi dall'inizio della compressione.

⁶numero di secondi trascorsi

Capitolo 3

Metodi statistici

3.1 La dimensione campionaria

La dimensione del campione è stata calcolata in base ai seguenti punti:

- La variabile risposta ρ è di tipo continuo
- Si vuole fare uno studio di superiorità
- Da studi precedenti si sa che:
 - $\rho_{XviD} = 0.14$ (valore atteso per l'encoder XviD)
 - $\sigma_{XviD}^2 = 0.0007392$ (varianza per l'encoder XviD)
 - il bitrate non influenza in modo significativo le precedenti stime
- Viene giudicata significativa una differenza di almeno il 15%, ciò implica $\rho_{x264} \leq 0.14 * 0.85 = 0.119$ e di conseguenza un $\Delta = |\rho_{x264} - \rho_{XviD}| = 0.021$
- I gradi di certezza ed incertezza sono rispettivamente:
 $\alpha = 0.05$ e $\beta = 4\alpha = 0.20$

nel seguente modo:

$$n = 2 * \lceil [2\sigma_{XviD}^2/\Delta^2] * [\Phi^{-1}(\alpha) + \Phi^{-1}(\beta)]^2 \rceil$$

In particolare: $2 * \lceil [2 * 0.0007392/0.021^2] * [1.96 + 1.64]^2 \rceil = 88$.

Considerando però alcuni bug che si potrebbero celare negli encoder si decide di aumentare il campione di un 5%¹: $n = \lceil 88 * 1.05 \rceil = 93$. Tenendo conto inoltre che si vuole suddividere il campione in parti uguali tra i blocchi ed i trattamenti, risulta conveniente utilizzare una numerosità pari a $n = 96$.

3.2 Il modello

Lo studio di superiorità da effettuare segue un piano fattoriale 2^k a blocchi, con $k=2$ indicante il numero dei fattori (encoder, bitrate); il numero di blocchi è pari a 3, da identificarsi coi generi dei file (azione, film-opera, documentario), i quali risultano essere un fattore di confondimento in quanto possono influire sul risultato finale.

Il modello da utilizzare sarà quindi il seguente:

$$Y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \xi_k + \varepsilon_{ijk} \quad \wedge \quad \varepsilon_{ijk} \sim N(0, \sigma^2)$$

con i seguenti significati:

- μ = effetto generale
- τ_i = effetto dell'encoder
- β_j = effetto del bitrate
- $(\tau\beta)_{ij}$ = effetto dell'interazione tra encoder e bitrate

¹vedere la documentazione sui relativi siti ufficiali per XviD e x264

- ξ_k = effetto del blocco
- ε_{ijk} = errore

3.3 La randomizzazione

Come descritto in precedenza verrà utilizzata la randomizzazione a blocchi permutati utilizzando le tavole di permutazione di 8 numeri o una lista generata dal calcolatore.

3.4 Metodi di analisi

3.4.1 Considerazioni generali

I dati degli esiti positivi della sperimentazione verranno raccolti ed analizzati tramite un'analisi della varianza a due fattori sia per verificare lo scopo primario che quello secondario.

3.4.2 Verifica delle assunzioni

1. Normalità ed indipendenza degli errori:

- QQ-plot: grafico dei quantili dei residui vs. i quantili della normale
- Test di Shapiro
- Normal probability plot: grafico dei residui vs. la loro probabilità

2. Omoschedasticità degli errori:

- boxplot per ogni trattamento / grafico valori stimati vs. residui

- Test di Levene (ANOVA sulle deviazioni $d_{ij} = |Y_{ij} - \tilde{Y}_i|$, $\tilde{Y}_i =$ trattamento mediano)

Nel caso in cui l'ipotesi di omoschedasticità non sia valida, è possibile applicare delle trasformazioni stabilizzanti ai dati di origine e rifare i test; se non si trova una trasformazione opportuna l'analisi non può procedere e lo studio termina.

3.4.3 L'analisi

Utilizzando la notazione standard si indica:

- “A” = fattore indicante la codifica (XviD o x264)
- “B” = fattore indicante il bitrate (1000 o 1800)
- “(1)” = $\sum \rho$ nel caso XviD @ 1000 kbps
- “a” = $\sum \rho$ nel caso x264 @ 1000 kbps
- “b” = $\sum \rho$ nel caso XviD @ 1800 kbps
- “ab” = $\sum \rho$ nel caso x264 @ 1800 kbps
- “n” = numerosità della singola cella = 8

si calcolano gli effetti principali e dell'interazione:

- $\text{Eff}(A) = \frac{1}{2n} * [ab + a - b - (1)]$
- $\text{Eff}(B) = \frac{1}{2n} * [ab + b - a - (1)]$
- $\text{Eff}(AB) = \frac{1}{2n} * [ab - b - a + (1)]$

e la tabella delle devianze per i successivi test di significatività:

devianza	g.l.	test $\sim F_{1,90}$
$SS_A = [ab + a - b - (1)]^2 / (4 * 24)$	1	$\frac{SS_A}{SS_{ERR}/90}$
$SS_B = [ab + b - a - (1)]^2 / (4 * 24)$	1	$\frac{SS_B}{SS_{ERR}/90}$
$SS_{AB} = [ab - b - a + (1)]^2 / (4 * 24)$	1	$\frac{SS_{AB}}{SS_{ERR}/90}$
$SS_{BLOCKS} = \sum_{i=1}^3 (Y_{j j \in Block_i} / 32) - Y_{\bullet\bullet}^2 / 96$	2	—
$SS_{ERR} = SS_{TOT} - SS_A - SS_B - SS_{AB} - SS_{BLOCKS}$	90	—
$SS_{TOT} = \sum_{i=1}^{96} (Y_i - \bar{Y}_{\bullet\bullet})^2$	95	—

Nel caso in cui l'interazione risultasse significativa, quest'ultima può oscurare il confronto tra le medie; una soluzione consiste nel fissare un livello di B ed effettuare il test di Turkey per il confronto delle medie di A, il quale afferma che due medie sono significativamente diverse se la loro differenza in valore assoluto è maggiore di: $T_\alpha = q_\alpha * \sqrt{\frac{MS_{ERR}}{n}}$.

3.5 Conclusioni

Utilizzando il seguente sistema di ipotesi:

$$\begin{cases} H_0 : |\delta| \geq \Delta & , \quad \delta = \text{differenza tra i trattamenti} \\ H_1 : -\Delta < \delta < \Delta & , \quad \Delta = \text{soglia di significatività} \end{cases}$$

si può affermare che l'encoder x264 è migliore rispetto ad Xvid se l'intervallo di confidenza per $\delta = \rho_{x264} - \rho_{Xvid}$ non interseca l'intervallo $(-\Delta, \Delta)$ e non comprende lo zero.

Bibliografia

- [1] Transcode
<http://www.transcoding.org>

- [2] The Unofficial XviD FAQ
<http://ronald.vslcatena.nl/docs/xvidfaq.html>

- [3] XviD (from Wikipedia)
<http://en.wikipedia.org/wiki/XviD>

- [4] H264 (from Wikipedia)
<http://en.wikipedia.org/wiki/H264>

- [5] Xvid.org
<http://www.xvid.org/>

- [6] x264 - a free h264/avc encoder
<http://developers.videolan.org/x264.html>

- [7] Bash Reference Manual
<http://www.gnu.org/software/bash/manual/bashref.html>