

Συστήματα και Αλγόριθμοι Πολυμέσων

Ιωάννης Χαρ. Κατσαβουνίδης

Ομιλία #2: Αρχές Επεξεργασίας
Σημάτων Πολυμέσων

4 Οκτωβρίου 2005

Πανεπιστήμιο Θεσσαλίας
Τμήμα Μηχ. Η/Υ, Τηλεπ. & Δικτύων

Επανάληψη (1)

- Πολυμέσα είναι
 - Ήχος (1-D σήμα)
 - Εικόνα (2-D σήμα)
 - Βίντεο (3-D σήμα)
- Πολυμεσικό σύστημα αποτελείται από
 - Λογισμικό (επιμέρους συστατικά, δομή, GUI)
 - Υλικό (CPU, μνήμη, συσκευές απόδοσης ήχου και εικόνας)
 - Τηλεπικοινωνιακό δίκτυο

Επανάληψη (2)

- Πολυμεσικοί αλγόριθμοι
 - Τεχνικές ψηφιακής επεξεργασίας σήματος
 - Τεχνικές κωδικοποίησης και θεωρίας πληροφοριών
 - Συμπίεση και αποσυμπίεση δεδομένων (“codec”)
- Πολυμεσικές αρχιτεκτονικές υπολογιστών
 - PC (Intel – MMX/SSE/SSE2)
 - Embedded (TI – DSP, ARM – RISC)
 - ASIC/SOC (full hard-wired design, GP-CPU + hard-wired block acceleration, configurable & extensible – Tensilics)

Βασικές τεχνικές (αρχές) επεξεργασίας πολυμέσων

- Δειγματοληψία/Κβαντοποίηση
- Μετασχηματισμοί
- Θεωρία Πληροφορίας
- Προ- και Μετά-επεξεργασία πολυμεσικών σημάτων

Δειγματοληψία

- Βασίζεται στο θεώρημα του Nyquist:
 - $F_s \geq 2 * F_{MAX}$
 - Π.χ. Audio CD, $F_{MAX} = 20\text{kHz}$, $F_s = 44.1\text{kHz}$
 - Τηλέφωνο - ανθρώπινη φωνή $F_{MAX} = 4\text{kHz}$, $F_s = 8\text{kHz}$

Κβαντοποίηση

- Μη-αναστρέψιμη διαδικασία
- Αποτελείται από δύο τμήματα:
 - Μετατροπή συνεχούς εύρους τιμών σε διακριτές τιμές (=σύμβολα) – quantization
 - Μετατροπή των συμβόλων σε τιμές (centroids) που αντιπροσωπεύουν το αρχικό εύρος τιμών – inverse quantization
- Ο αριθμός του bits καθορίζεται από την αποδεκτή ποιότητα αναπαραγωγής του αρχικού σήματος

Κβαντοποίηση

- Π.χ. Στο Audio CD, N (# of bits) = 16
 - Άσκηση #1: επιβεβαιώστε ότι ο ρυθμός δεδομένων του Audio CD είναι 1411200 bits/sec
 - Άσκηση #2: αποδείξτε τη σηματοθορυβική σχέση ομοιόμορφου κβαντιστή $PSNR = 6 \cdot N + 11$ (dB) και κατά συνέπεια τη σηματοθορυβική σχέση του Audio CD ~ 107 dB

Επανα-δειγματοληψία

- Αύξηση του αριθμού των δειγμάτων (up-sampling)
- Μείωση του αριθμού των δειγμάτων (down-sampling)

Τεχνικές φίλτρων

- Βασίζεται στην χρήση της συνάρτησης $\text{sinc}(x) = \sin(x)/x$ για τον υπολογισμό των ενδιάμεσων τιμών, σε συνδιασμό με τη χρήση παραθύρων για τη μείωση της κυμάτωσης (ripples) και της επικάλυψης φάσματος (aliasing)
- Για τον υπολογισμό των τιμών εξόδου, γίνεται συνέλιξη (convolution) των τιμών εισόδου με τους συντελεστές του φίλτρου
- Παράδειγμα: $[1, -5, 20, 20, -5, 1]$

Τεχνικές φίλτρων (2)

- Ρουτίνα συνέλιξης

```
for (j = 0; j < N; j++)  
{  
    sum = 0;  
    for (i = -2; i < 2; i++)  
        sum += h[i]*input[j+i];  
    output[j] = sum;  
}
```

Τεχνικές φίλτρων (3)

- Ιδιαίτερη αντιμετώπιση στα όρια του σήματος
 - Συνήθως, επανάληψη του οριακού δείγματος
- Πολυπλοκότητα
 - N MACs/output sample
 - Ιδανικός αλγόριθμος για DSPs

Τεχνικές παρεμβολής (1)

- Βασίζεται στην χρήση πολυώνυμων για τον υπολογισμό ενδιάμεσων τιμών
 - Παρεμβολή πρώτου βαθμού (γραμμική παρεμβολή)
 - Παρεμβολή δευτέρου & τρίτου βαθμού
 - Splines

Τεχνικές παρεμβολής (2)

- Γραμμική παρεμβολή

$$y(t) = x_1 t + x_0 (1 - t), 0 \leq t \leq 1$$

Μπορεί να γραφτεί και ως

$$y(t) = x_0 + (x_1 - x_0)t, 0 \leq t \leq 1$$

Τεχνικές παρεμβολής (3)

- Δεύτερου ή μεγαλύτερου βαθμού παρεμβολή απαιτεί την επίλυση συστήματος γραμμικών εξισώσεων
- Π.χ. Δευτεροβάθμια παρεμβολή

$$y(t) = \alpha t^2 + \beta t + \gamma, 0 \leq t \leq 1$$

$$y(0) = x_0$$

$$y(1) = x_1$$

Τεχνικές παρεμβολής (4)

- Μπορεί να γραφτεί ως

$$y(t) = (x_1 - x_0)\delta t^2 + (x_1 - x_0)(1 - \delta)t + x_0, 0 \leq t \leq 1, \delta \in R$$

όπου οι τιμές των δ (μία τιμή για κάθε διάστημα) υπολογίζονται απαιτώντας συνέχεια των παραγώγων στα όρια των διαστημάτων:

$$2(x_i - x_{i-1})\delta_i + (x_i - x_{i-1})(1 - \delta_i) = (x_{i+1} - x_i)(1 - \delta_{i+1})$$

$$(x_i - x_{i-1})(1 + \delta_i) = (x_{i+1} - x_i)(1 - \delta_{i+1})$$

Τεχνικές παρεμβολής (5)

- Αν έχουμε N σημεία $(x_0, x_1, \dots, x_{N-1})$, δηλαδή $(N-2)$ εσωτερικά σημεία, έχουμε $(N-2)$ γραμμικές εξισώσεις.
- 2 επιπρόσθετες εξισώσεις προκύπτουν με τη βοήθεια προέκτασης των οριακών σημείων, οπότε λύνουμε το σύστημα N εξισώσεων με N αγνώστους και αποκτούμε τους συντελεστές $\delta_0, \delta_1, \dots, \delta_{N-1}$

Τεχνικές παρεμβολής (6)

- Στην περίπτωση διδιάστατου σήματος (εικόνα), εφαρμόζουμε τις προηγούμενες τεχνικές με διαχωρίσιμο τρόπο – πρώτα κατά τη μία διάσταση και μετά κατά την άλλη.

Τεχνικές κβαντοποίησης (1)

- Ομοιόμορφη κβαντοποίηση
 - Για προσημασμένα σήματα, συνήθως κβαντοποιούμε το πρόσημο με 1-bit και το πλάτος (=απόλυτη τιμή) με τα υπόλοιπα $(N-1)$ bits “dead-zone quantization”
- Μη-ομοιόμορφη κβαντοποίηση
 - Οι υπο-περιοχές κβαντοποίησης (“bins”) δεν έχουν το ίδιο εύρος. Συνήθως επιλέγουμε το εύρος ούτως ώστε η πιθανότητα κάθε υπο-περιοχής να είναι σταθερή

Τεχνικές κβαντοποίησης (2)

- Μη-ομοιόμορφη κβαντοποίηση
 - Οι τιμές αντικατάστασης υπολογίζονται με τον αλγόριθμο Lloyd-Max
 - Εναλλακτικά, το σήμα μετατρέπεται με τη βοήθεια μιας μη-γραμμικής συνάρτησης σε ομοιόμορφα κατανεμημένο και ακολουθεί ομοιόμορφη κβαντοποίηση (“companding”)

Αλγόριθμος κβαντοποίησης

- Quantization:

```
for (i=0; i<8; i++)  
{  
    sign[i] = (x[i]>0) ? 0 : 1;  
    q[i] = abs(x[i])/(2*QP);  
}
```

Αλγόριθμος κβαντοποίησης

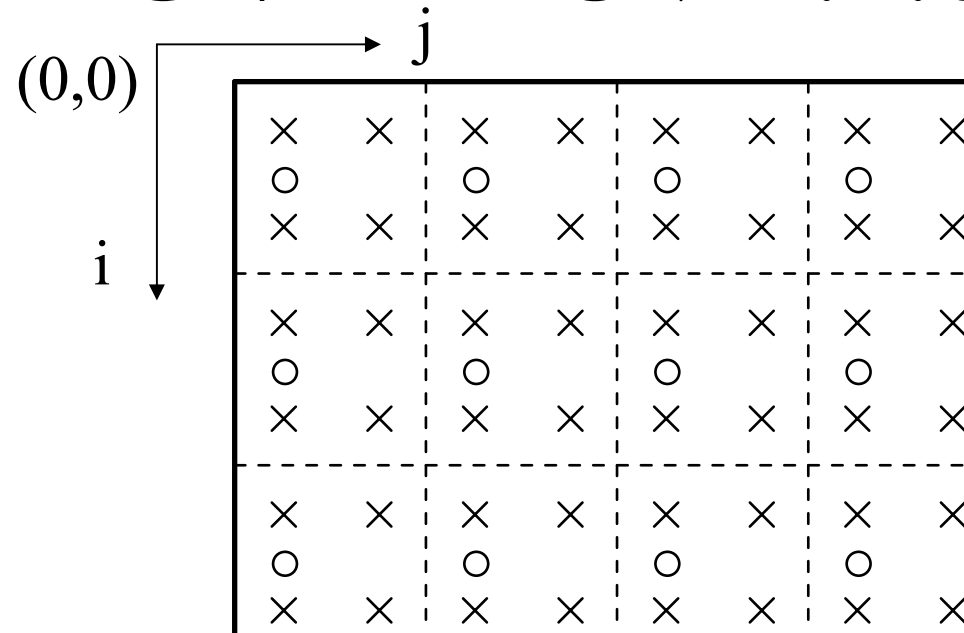
- Inverse Quantization:

```
for (i=0; i<8; i++)  
{  
    if(q[i]==0) r[i] = 0;  
    else if (sign[i]==0)  
        r[i] = (2*q[i]+1)*QP;  
    else  
        r[i] = -(2*q[i]+1)*QP;  
}
```

Technology overview of MPEG1/2

- Basic unit of processing is a **picture**
- Each picture consists of an array of **pixels**; for each pixel, we associate one integer value (usually 8-bit) called **luminance** and for each 2x2 block of pixels we associate two **chrominance** values, called Cb/Cr or U/V – also called YUV 4:2:0 format. YUV 4:2:2 and 4:4:4 are also supported (MPEG2 only).

MPEG1/2 – YUV 4:2:0 format



- × Represent luminance samples
- Represent chrominance samples

Ασκήσεις

- Να υπολογίσετε τις παρακάτω στατιστικές ποσότητες για κάθε ένα από τα τρία συστατικά (Y/U/V) του 1ου καρέ του βίντεο “src13”
 - Μέσος όρος (μ)
 - Διασπορά (σ)
 - Συσχέτιση 1ου βαθμού κατά τη διάσταση x και κατά τη διάσταση y
$$\rho_{\text{hor}} = [E(x(i,j)*x(i,j+1))-\mu^2]/\sigma^2$$
$$\rho_{\text{ver}} = [E(x(i,j)*x(i+1,j))-\mu^2]/\sigma^2$$
 - Συσχέτιση 2ου βαθμού κατά τη διάσταση x και κατά τη διάσταση y