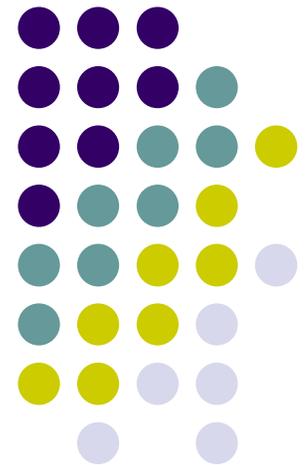


Συμπύεση και Μετάδοση Πολυμέσων: Πολυμέσα, Ψηφιοποίηση Σήματος



Πανεπιστήμιο Πελοποννήσου
Τμήμα Επιστήμης & Τεχνολογίας
Τηλεπικοινωνιών

Ευάγγελος Α. Κοσμάτος





Ορισμός πολυμέσων

- Πολυμέσα στον χώρο της τεχνολογίας πληροφορίας (information field) σημαίνει “πολλαπλά μέσα” μέσω των οποίων η πληροφορία αποθηκεύεται, μεταδίδεται, παρουσιάζεται ή γίνεται αντιληπτή.
 - π.χ. η εφημερίδα, που συνδυάζει κείμενο και εικόνα, και η τηλεόραση, που συνδυάζει ήχο και κινούμενη εικόνα.
- Στο μάθημα περιοριζόμαστε σε αυτά στα οποία η πληροφορία είναι ψηφιακή (ή ψηφιοποιημένα - digitized) και ελέγχεται από υπολογιστή, δηλαδή ψηφιακά πολυμέσα

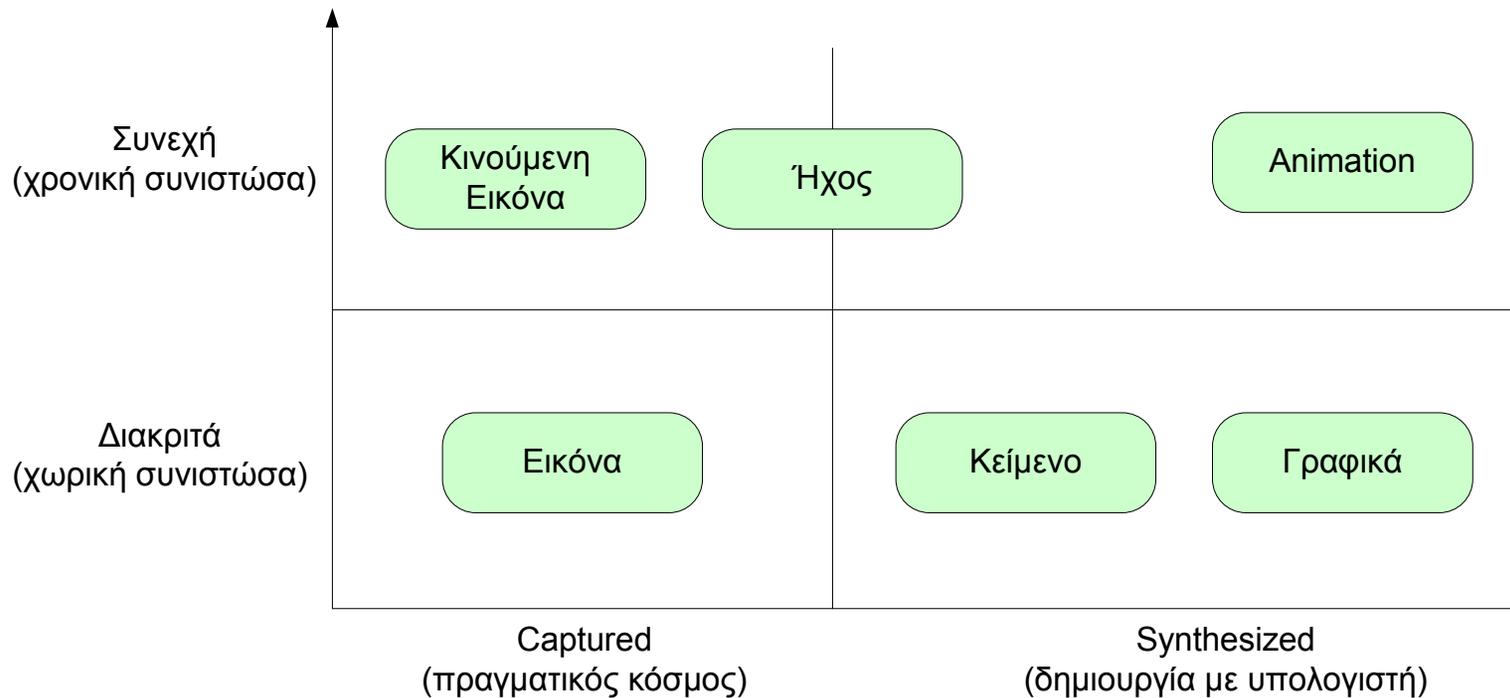


Ορισμός πολυμέσων

- Ψηφιακά πολυμέσα είναι ο τομέας που ασχολείται με την ελεγχόμενη από υπολογιστή ολοκλήρωση κειμένου, γραφικών, ακίνητης και κινούμενης εικόνας, animation, ήχου, και οποιουδήποτε άλλου μέσου ψηφιακής αναπαραστάσης, αποθήκευσης, μετάδοσης και επεξεργασίας της πληροφορίας
- Τα πολυμέσα αποτελούν το σημείο συνάντησης πέντε μεγάλων βιομηχανιών
 - πληροφορικής
 - τηλεπικοινωνιών
 - ηλεκτρονικών εκδόσεων
 - βιομηχανίας audio και video
 - βιομηχανίας της τηλεόρασης και του κινηματογράφου.

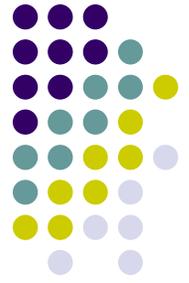


Μέσα πληροφορίας



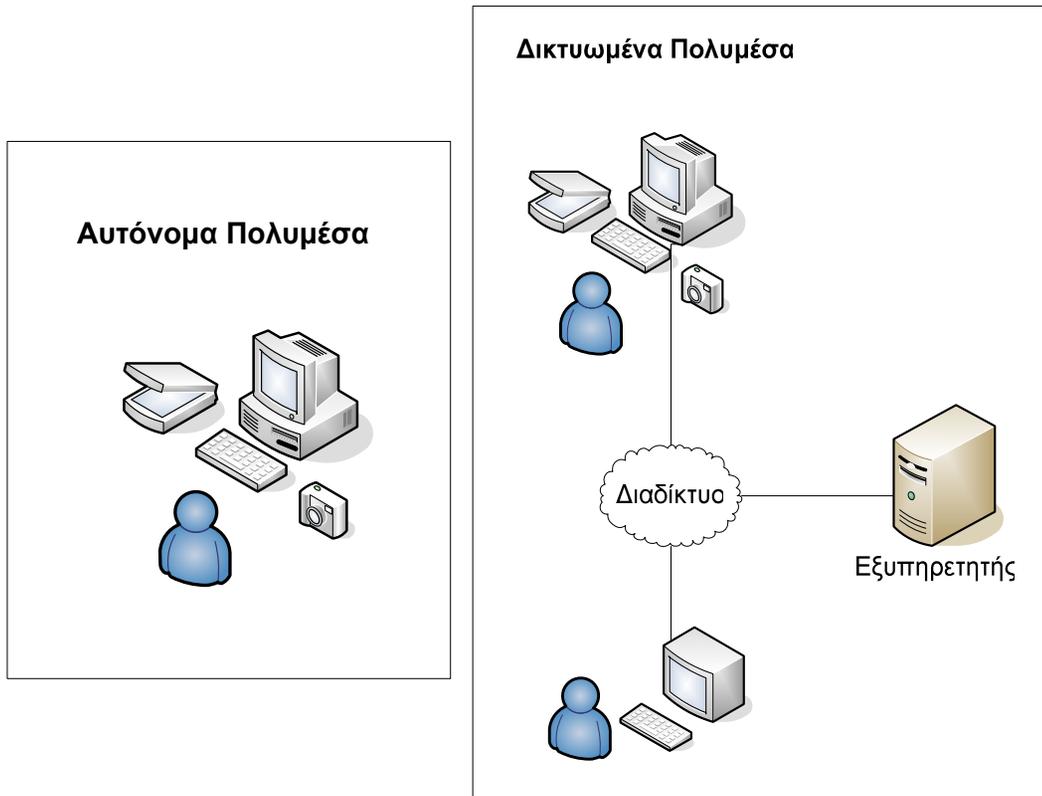
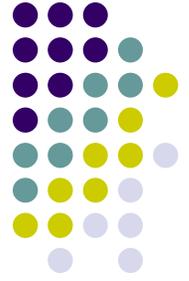
- πολυμέσο = ολοκλήρωση ενός τουλάχιστον “διακριτού” τύπου πληροφορίας και ενός “συνεχούς”.

Χαρακτηριστικά συστήματος πολυμέσων



- Έλεγχος από υπολογιστή.
 - αποθήκευση, επεξεργασία, παρουσίαση μέσω του υπολογιστή
- Ολοκλήρωση (Integration)
 - ελάχιστος αριθμός υποσυστημάτων
 - π.χ. οθόνη υπολογιστή
 - ενοποιημένη μεταφορά δεδομένων
- Ψηφιακή αναπαράσταση
 - Η πληροφορία ανεξάρτητα από τον τύπο και το περιεχόμενό της αναπαράσσεται σε ψηφιακή μορφή (σειρές 0 και 1).
- Αλληλεπίδραση (interaction).
 - ο χρήστης έχει την δυνατότητα να επεμβαίνει στο σενάριο ροής του συστήματος

Αυτόνομα και Δικτυωμένα Πολυμέσα



- **Αυτόνομα συστήματα πολυμέσων**
 - Επεξεργαστής
 - Κάρτα ήχου και κάρτα γραφικών
 - Ηχεία και μικρόφωνο
 - Επεξεργαστής και μνήμη αποθήκευσης
 - Οπτικός δίσκος (π.χ. CD-ROM)
- **Δικτυωμένα συστήματα πολυμέσων**
 - εγγενώς δικτυακές εφαρμογές
 - ηλεκτρονικό ταχυδρομείο
 - τηλεδιάσκεψη
 - Διανομή πολυμεσικών δεδομένων σε πολλαπλούς χρήστες
 - αρχιτεκτονική Client-server



Εφαρμογές πολυμέσων

- Τηλεδιάσκεψη
- Ψηφιακές βιβλιοθήκες
- Δικτυακές Ιστοσελίδες
- Δικτυακές αγορές
- Διαδραστική τηλεόραση (Interactive TV)
 - Video On Demand
- Παιχνίδια
- Εικονική πραγματικότητα (Virtual Reality)
- Εφαρμογές σε CD-ROM
 - Εκπαιδευτικές
 - Διαφημιστικές
 - Ψυχαγωγικές



Βασικές αρχές σημάτων

- Τι είναι Σήμα;
 - Πληροφορίες που αντιλαμβανόμαστε μέσω των αισθήσεων μας και μπορούν να περιγραφούν ως μια ή περισσότερες φυσικές μεταβλητές η τιμή των οποίων είναι μια συνάρτηση του χρόνου (και / ή του χώρου)
 - Σήματα: ήχος, εικόνα, βίντεο
- Μορφές Σημάτων
 - Ηλεκτρικά σήματα, οπτικά σήματα, ακουστικά σήματα, μαγνητικά σήματα



Βασικές αρχές σημάτων

- Αναπαραστάσεις Σημάτων
 - Ένας αναγνώστης οπτικών δίσκων (CD player) αποτελεί ένα σύστημα με το οποίο μετατρέπουμε την ακολουθία αριθμών που έχει αποθηκευτεί σε ένα οπτικό δίσκο (δηλ. την αριθμητική αναπαράσταση ενός σήματος) σε ένα ακουστικό σήμα.
- Μονοδιάστατα και Πολυδιάστατα σήματα
 - Μονοδιάστατα είναι τα σήματα που περιγράφονται μαθηματικά ως μια συνάρτηση μίας ανεξάρτητης μεταβλητής
 - Ένα ηχητικό σήμα είναι μια μεταβολή της πίεσης του αέρα ως συνάρτηση του χρόνου
 - Δυσδιάστατα σήματα:
 - Εικόνα: μεταβολή της φωτεινότητας ή του χρώματος ως συνάρτηση της θέσης (x,y)
 - Τρισδιάστατα σήματα:
 - Βίντεο: μεταβολή της φωτεινότητας ή του χρώματος ως συνάρτηση της θέσης και του χρόνου (x,y,t)



Βασικές αρχές σημάτων

- Αναλογικά Σήματα
 - Η πληροφορία που συλλέγεται από τα ανθρώπινα αισθητήρια όργανα είναι αναλογική, δηλαδή το σήμα που περιγράφει το φυσικό μέγεθος μεταβάλλεται ανάλογα με τον τρόπο που μεταβάλλεται και το φυσικό μέγεθος. Αναπαριστά αυτό το μέγεθος και παρακολουθεί τις όποιες μεταβολές του μέσα από μια σχέση αναλογίας.
- Συνεχή Σήματα
 - Ένα σήμα για το οποίο για κάθε τιμή της ανεξάρτητης (ή των ανεξάρτητων) μεταβλητής/ων μπορεί να υπολογιστεί μια τιμή του σήματος ονομάζεται συνεχές σήμα

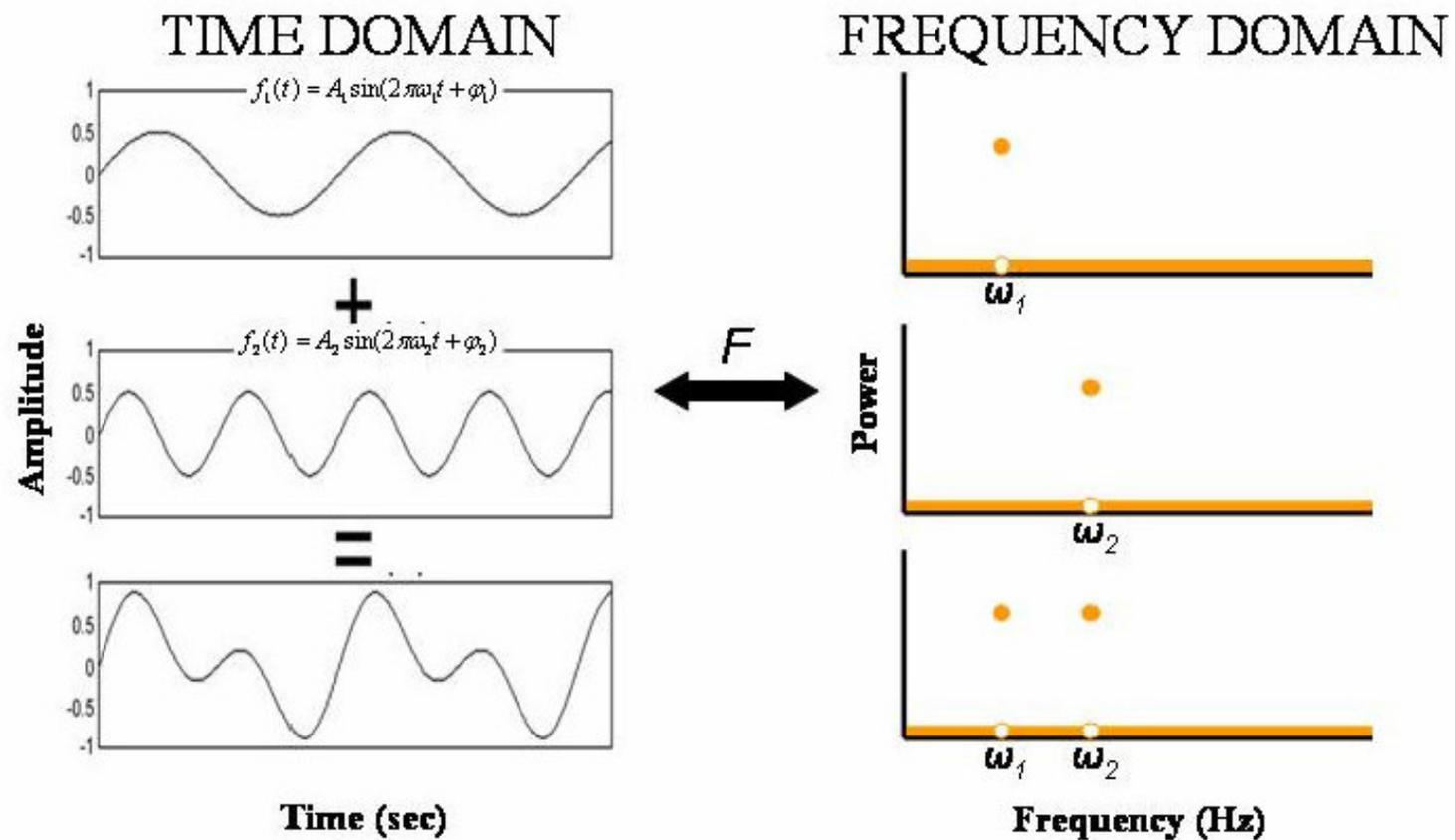


Βασικές αρχές σημάτων

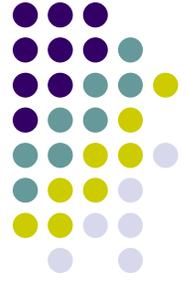
- Τα σήματα ως άθροισμα ημιτονικών συναρτήσεων
 - Κάθε συνεχές σήμα αναλύεται σε ένα άθροισμα ημιτονικών συναρτήσεων με διάφορα πλάτη, συχνότητες και φάσεις.
- Χρησιμοποιώντας την ανάλυση κάθε σήματος σε άθροισμα ημιτονικών συναρτήσεων μπορούμε να περιγράψουμε κάθε σήμα με ένα διάγραμμα που απεικονίζει το πλάτος των συχνοτήτων από τις οποίες αποτελείται.
 - Η αναπαράσταση αυτή αποτελεί το φάσμα συχνοτήτων (frequency spectrum) για το συγκεκριμένο σήμα.



Βασικές αρχές σημάτων



Ψηφιοποίηση πληροφορίας

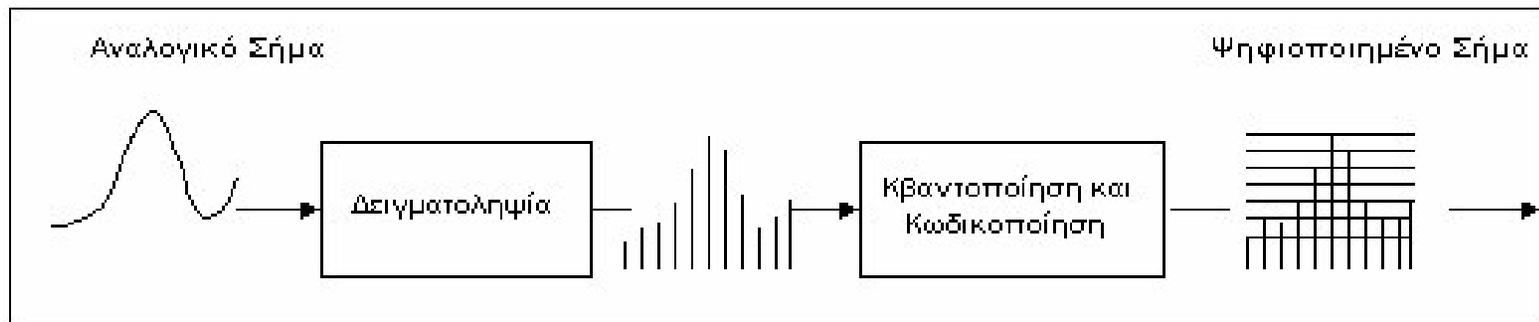


- Κάθε οπτικοακουστική πληροφορία που λαμβάνεται από τον εγκέφαλο αντιστοιχεί στη μεταβολή κάποιας συνεχούς μεταβλητής στο χώρο ή/και το χρόνο, το σήμα.
- Τα σήματα (αναλογικά) τα οποία είναι ουσιαστικά μια μορφή ενέργειας, μετατρέπονται σε άλλη μορφή ενέργειας (ηλεκτρικά αναλογικά σήματα), μέσω αισθητήρων (sensors).
- Στους Η/Υ η αναπαράσταση της κάθε πληροφορίας γίνεται σε δυαδική μορφή. Η δυαδική κωδικοποίηση χρησιμοποιεί δύο τιμές 1 και 0 οι οποίες αντιστοιχούν στις καταστάσεις θετικό και μη θετικό.
 - κάθε πληροφορία αποθηκεύεται σαν μια ακολουθία από 0 και 1
- Για να μπορέσουμε να αποθηκεύσουμε και να επεξεργαστούμε συνεχή σήματα στον Η/Υ πρέπει επομένως να τα μετατρέψουμε σε ακολουθίες αριθμών (ψηφιακό σήμα)



Ψηφιοποίηση πληροφορίας

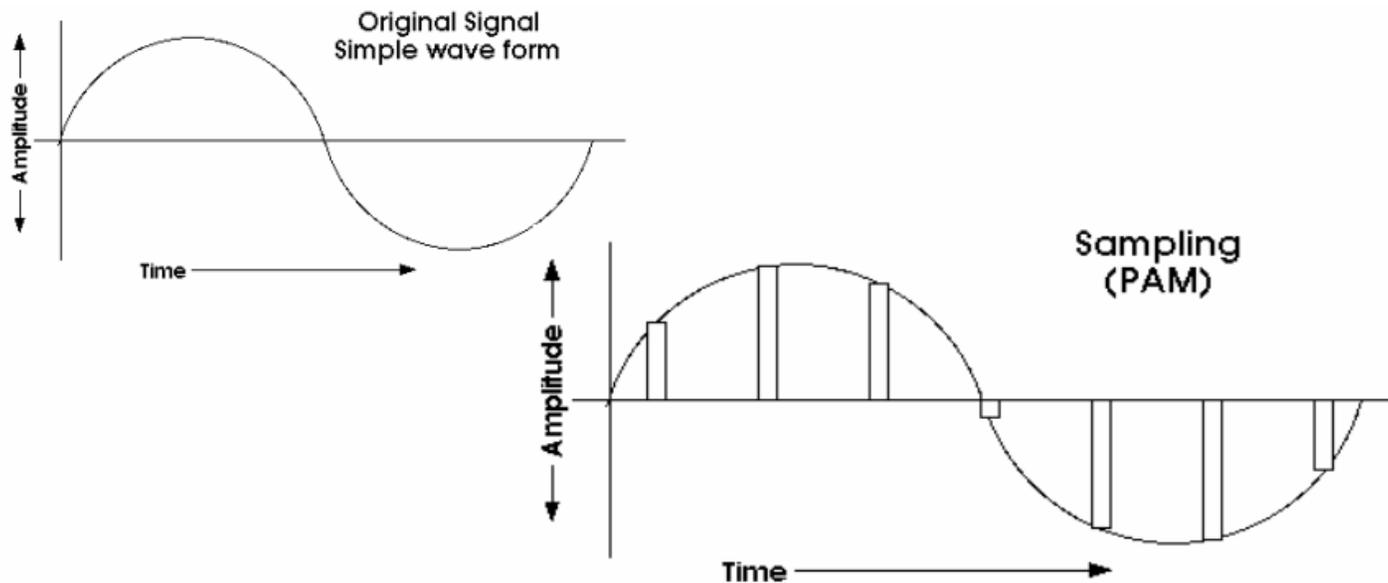
- Το πέρασμα από τον αναλογικό στον ψηφιακό κόσμο είναι η διαδικασία της ψηφιοποίησης (digitization).
 - Κατά την ψηφιοποίηση ένα σήμα αναλογικής μορφής μετατρέπεται σε ψηφιακό (συνήθως δυαδικό), δηλ. από συνεχή συνάρτηση του χρόνου μετατρέπεται σε μια σειρά διακριτών αριθμητικών τιμών.
- Η ψηφιοποίηση σημάτων περιλαμβάνει:
 - Δειγματοληψία (sampling)
 - Κβαντισμός (quantization)
 - Κωδικοποίηση (coding).





Δειγματοληψία

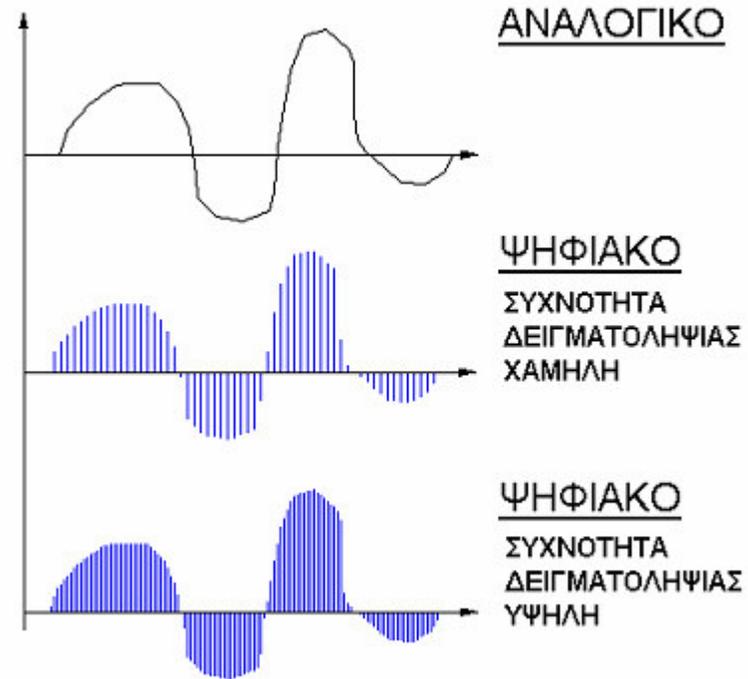
- Η διαδικασία με την οποία παράγουμε ένα διακριτό από ένα συνεχές σήμα ονομάζεται δειγματοληψία (sampling) και προκύπτει από την καταγραφή των τιμών του συνεχούς σήματος σε μια σειρά από διακριτά και ισαπέχοντα σημεία
 - στο χρόνο (ήχος)
 - στο χώρο (εικόνα)
 - στο χωροχρόνο (βίντεο).





Δειγματοληψία

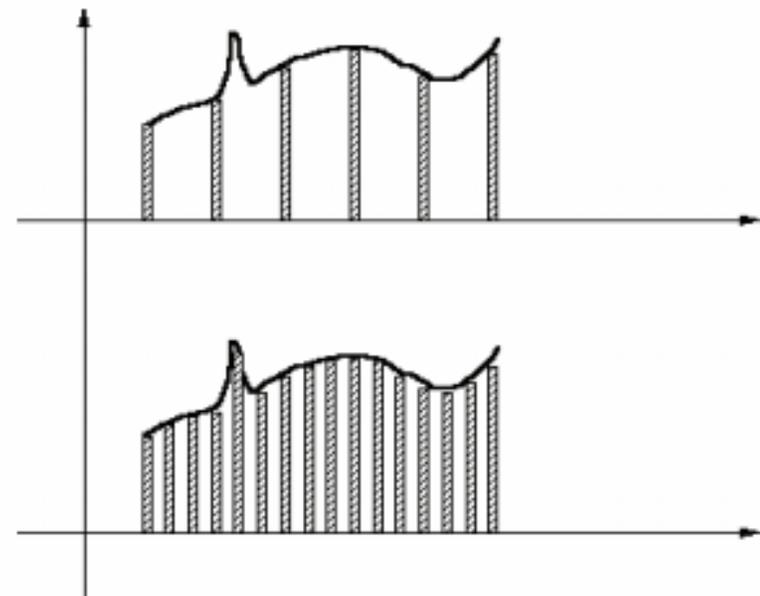
- Με τη δειγματοληψία το σήμα μετασχηματίζεται σε μια σειρά δειγμάτων (samples). Η μορφή αυτή του σήματος είναι γνωστή και ως PAM (Pulse Amplitude Modulation).
- Συχνότητα δειγματοληψίας: Ρυθμός καταγραφής των δειγμάτων
 - δείγματα ανά δευτερόλεπτο (ή Hz)
- Η τιμή της συχνότητας δειγματοληψίας πρέπει να ικανοποιεί δύο αντικρουόμενες απαιτήσεις:
 - ποιότητα ψηφιοποίησης (υψηλή συχνότητα)
 - μικρό μέγεθος ψηφιακού αρχείου (χαμηλή συχνότητα)





Δειγματοληψία

- Πόσο καλή μπορεί να είναι η προσέγγιση του αρχικού αναλογικού σήματος από το σήμα διακριτού χρόνου;
 - Στην περίπτωση χαμηλής συχνότητας δειγματοληψίας ο υψίσυχνος παλμός δεν καταγράφεται και η πληροφορία που μεταφέρει «χάνεται» αφού η διάρκειά του είναι μικρότερη από την περίοδο λήψης δειγμάτων.
 - Το πρόβλημα διορθώνεται αν αυξήσουμε τη συχνότητα δειγματοληψίας.
- Η τιμή της συχνότητας δειγματοληψίας πρέπει να ικανοποιεί δύο αντικρουόμενες απαιτήσεις:
 - ποιότητα ψηφιοποίησης (υψηλή συχνότητα)
 - μικρό μέγεθος ψηφιακού αρχείου (χαμηλή συχνότητα)





Θεώρημα Nyquist

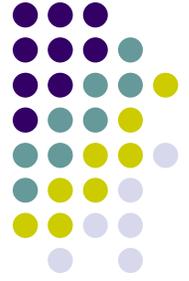
- Τα περισσότερα αναλογικά σήματα μπορεί να θεωρηθούν ότι είναι περιορισμένου εύρους ζώνης
 - υπάρχει μια μέγιστη συχνότητα f_n πάνω από την οποία η ενέργεια του σήματος είναι πρακτικά μηδέν
 - η f_n αναφέρεται και ως συχνότητα αποκοπής (cutoff frequency)
- Ένα συνεχές στο χρόνο σήμα $x(t)$ που περιέχει συχνότητες όχι μεγαλύτερες από f_n μπορεί να ανακατασκευαστεί ακριβώς από τα δείγματά του, αν τα δείγματα αυτά έχουν ληφθεί με συχνότητα $f_s = 1/T_s$ που είναι μεγαλύτερη από $2 \cdot f_n$.
 - Η $f_N = 2 \cdot f_n$ ονομάζεται συχνότητα Nyquist



Θεώρημα Nyquist

Παράδειγμα εφαρμογής του θεωρήματος

- Οι μουσικοί οπτικοί δίσκοι (Audio CD-ROM) χρησιμοποιούν συχνότητα δειγματοληψίας ίση με 44.1KHz (44100 δείγματα ανά δευτερόλεπτο) για την αποθήκευση του ηχητικού σήματος.
 - Η συχνότητα αυτή είναι ελαφρά μεγαλύτερη από 2 φορές τη συχνότητα των 20 KHz που κατά γενική παραδοχή αποτελεί το άνω όριο για τις συχνότητες που γίνονται αντιληπτές από το ανθρώπινο αυτί.
 - Με ένα τέτοιο ρυθμό δειγματοληψίας διασφαλίζεται ότι όλες οι συχνότητες που είναι μικρότερες ή ίσες με 20 kHz θα αποδοθούν σωστά στο διακριτό σήμα



Θεώρημα Nyquist

- Δειγματοληψία με συχνότητα χαμηλότερη από τη συχνότητα Nyquist παραμορφώνει το σήμα:
 - Χάνει τις υψηλότερες συχνότητες που τυχόν δεν καλύπτει η επιλεγμένη συχνότητα δειγματοληψίας f_s
 - Δημιουργεί συχνότητες οι οποίες δεν υπήρχαν στο αρχικό σήμα (το φαινόμενο ονομάζεται αναδίπλωση ή αλίαση “aliasing”).
- Συχνότητα δειγματοληψίας f_s υψηλότερη από τη συχνότητα Nyquist f_N αποτελεί σπατάλη αποθηκευτικού χώρου καθώς δημιουργεί πρόσθετα δείγματα χωρίς αυτά να είναι απαραίτητα.
- Στην πράξη επιλέγεται πάντα συχνότητα λίγο υψηλότερη από τη συχνότητα Nyquist.



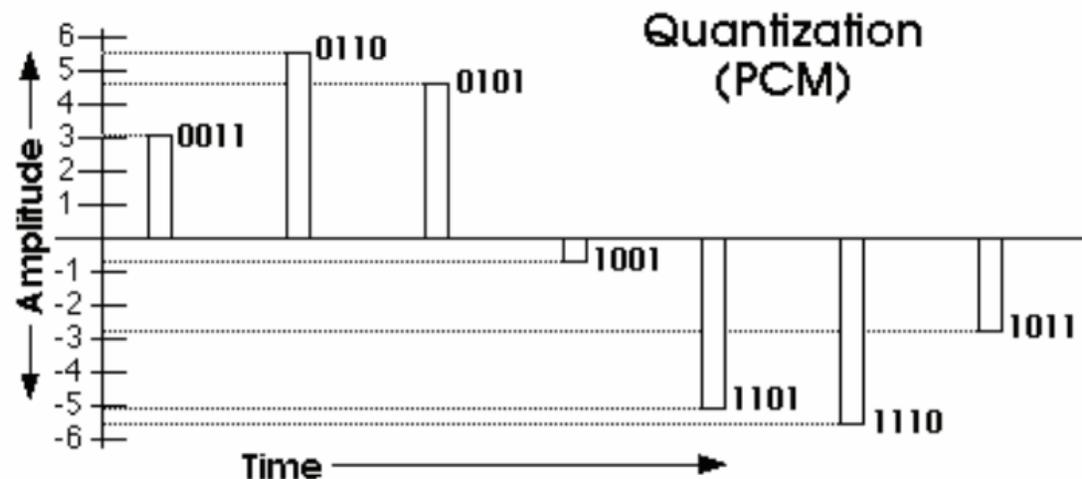
Κβαντισμός

- Η δειγματοληψία δημιουργεί διακριτοποίηση του σήματος στο χρόνο. Οι τιμές όμως των δειγμάτων (πλάτος σήματος) μπορούν να έχουν οποιαδήποτε τιμή.
 - Για να αναπαραστήσουμε τις τιμές αυτές στον υπολογιστή απαιτείται διακριτοποίηση και των τιμών πλάτους
 - προσέγγιση των πραγματικών τιμών με κάποιες προκαθορισμένες
- Η διακριτοποίηση των τιμών πλάτους των δειγμάτων ονομάζεται κβαντισμός.



Κβαντισμός

- Κατά τη διαδικασία του κβαντισμού υποδιαιρείται το εύρος τιμών του συνεχούς σήματος σε συγκεκριμένα επίπεδα και σε κάθε τέτοιο επίπεδο εκχωρείται ένας ψηφιακός κωδικός από τους διαθέσιμους της κλίμακα
- Κάθε δείγμα παίρνει εκείνη την ψηφιακή τιμή πλησιέστερα στην οποία βρίσκεται η αρχική του τιμή (στρογγυλοποίηση στον κοντινότερο διαθέσιμο κωδικό).





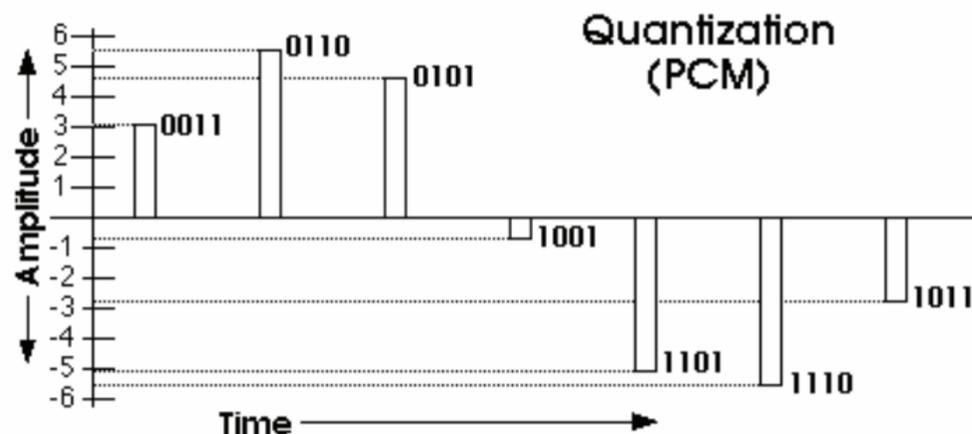
Κβαντισμός

Υπόθεση:

- Γνωρίζουμε την μέγιστη V_{max} και την ελάχιστη V_{min} τιμή που μπορεί να πάρει ένα σήμα
- Έχουμε διαθέσιμα n bits για την κωδικοποίηση κάθε δείγματος

Πώς θα κβαντίσουμε ένα σήμα:

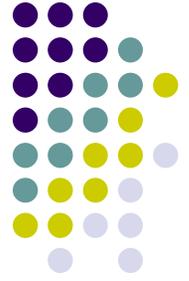
- Διαχωρισμός του εύρους του σήματος σε διαστήματα (βήμα κβαντισμού) που το καθένα έχει εύρος $q=(V_{max}-V_{min})/2^n$
- Αναπαράσταση των τιμών στο μέσο των διαστημάτων κβαντοποίησης
- Εκχώρηση σε κάθε δείγμα την ψηφιακή τιμή που είναι πλησιέστερα στην πραγματική τιμή





Μέγεθος Δείγματος

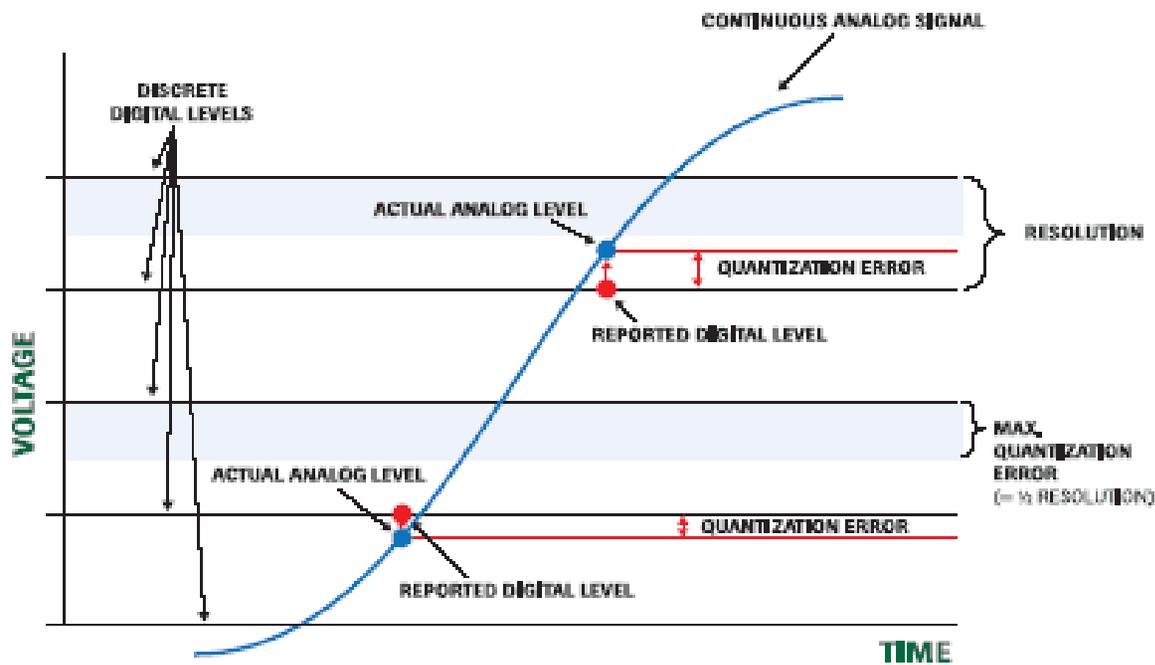
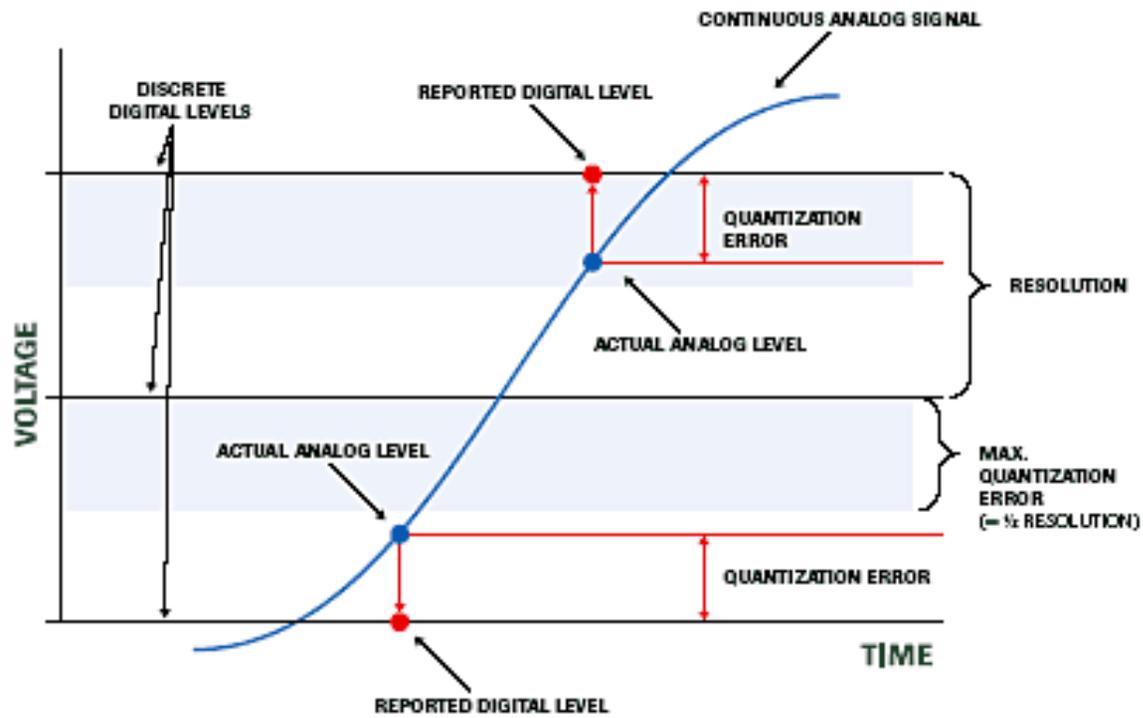
- Το πλήθος των διαθέσιμων τιμών (σταθμών) οι οποίες μπορούν να ανατεθούν στη τιμή ενός δείγματος εξαρτάται από το πλήθος ψηφίων του κωδικού (codeword) που χρησιμοποιείται για την αναπαράσταση των τιμών αυτών.
- Το μέγεθος αυτό (n) χαρακτηρίζεται ως «μέγεθος δείγματος» (sampling size).
 - Πλήθος σταθμών κωδικοποίησης = 2^n
- Αφού κάθε δείγμα παριστάνεται και αποθηκεύεται με χρήση ενός τέτοιου κωδικού το μέγεθος δείγματος δείχνει ταυτόχρονα και το ποσό της μνήμης που απαιτείται για την ψηφιακή αποθήκευση του κάθε δείγματος.
 - Αν για παράδειγμα χρησιμοποιήσουμε μήκος λέξης 16 bit τότε κάθε δείγμα του ψηφιοποιημένου σήματος θα καταλαμβάνει 2 bytes μνήμης
- Παραδείγματα μεγέθους δείγματος
 - 1 bit -> 2 στάθμες κβάντωσης
 - 2 bit -> 4 στάθμες κβάντωσης
 -
 - n bit -> 2^n στάθμες κβάντωσης



Σφάλμα Κβαντισμού

- Διαφορά της πραγματικής τιμής του δείγματος από την τιμή που τελικά κωδικοποιείται
 - Η διαφορά αυτή εισάγει παραμόρφωση στο κβαντισμένο σήμα. Η παραμόρφωση αυτή είναι τόσο πιο μικρή όσο περισσότερες στάθμες κβαντισμού υπάρχουν
- Μέγιστο σφάλμα κβαντισμού
 - Αν κωδικοποιήσουμε το μέσο των διαστημάτων κβαντισμού τότε το μέγιστο σφάλμα κβαντισμού είναι:

$$e_{\max} = q/2 = (V_{\max} - V_{\min})/2^{n+1}$$



Το σφάλμα κβαντισμού ως Σηματοθορυβικός λόγος

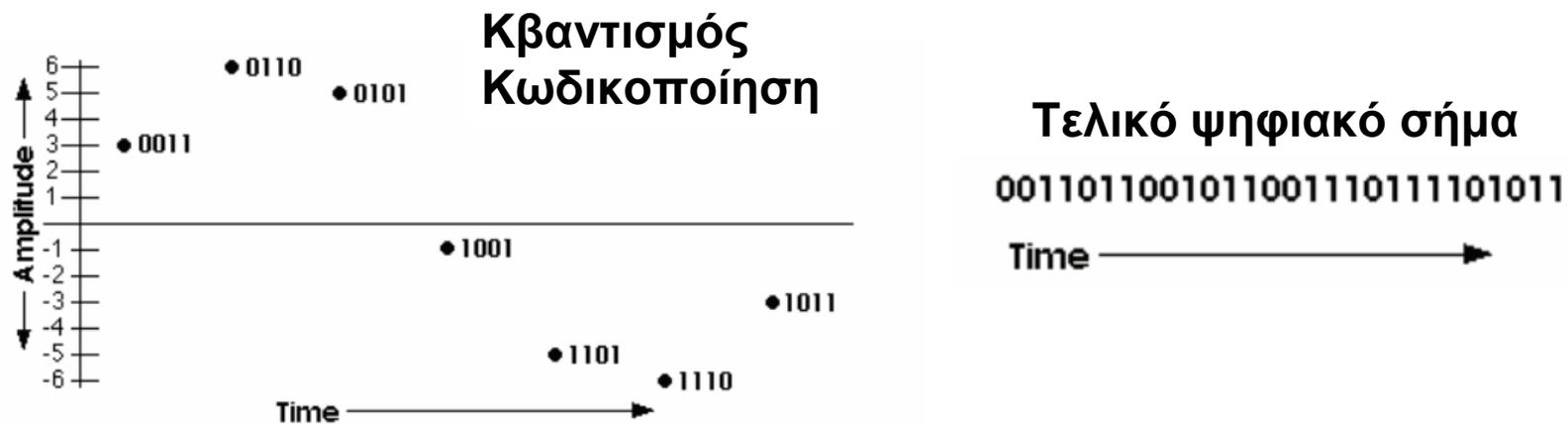


- Σηματοθορυβικό λόγο ονομάζουμε το λόγο της τιμής του σήματος προς τη τιμή των ανεπιθύμητων παρεμβολών που επενεργούν στο σήμα (θόρυβος)
- Το σφάλμα κβαντισμού θεωρείται μια τέτοια παρεμβολή
- Ο σηματοθορυβικός λόγος εκφράζεται συνήθως σε db (decibel)
 $SNR_{dB} = 20\log_{10}(\text{Πλάτος Σήματος} / \text{Πλάτος Θορύβου})$
- Αν χρησιμοποιήσουμε τη μέγιστη τιμή του σήματος και τη μέγιστη τιμή του θορύβου τότε έχουμε το λόγο PSNR (Peak Signal to Noise Ratio):
 $PSNR_{dB} = 20\log_{10}(V_{max} / e_{max})$
- Αν χρησιμοποιήσουμε την ελάχιστη τιμή του σήματος και τη μέγιστη τιμή του θορύβου τότε έχουμε τον ελάχιστο σηματοθορυβικό λόγο (Min Signal to Noise Ratio):
 $SNR_{min} = 20\log_{10}(V_{min} / e_{max})$



Κωδικοποίηση

- Κωδικοποίηση ονομάζουμε την αναπαράσταση των κβαντισμένων τιμών πλάτους των δειγμάτων του σήματος σε δυαδική ακολουθία (ακολουθία δυαδικών ψηφίων).
 - Δημιουργείται έτσι το τελικό ψηφιακό σήμα σαν μια σειρά από bits.





Κωδικοποίηση

- Ο κώδικας κάθε τιμής στάθμης ονομάζεται **κωδική λέξη (codeword)**
- Το σύνολο των διαφορετικών κωδικών λέξεων που χρησιμοποιούνται για την κωδικοποίηση ονομάζεται **κώδικας**
- Όταν το σύνολο των κωδικών λέξεων έχει το ίδιο μέγεθος (αριθμό bits) τότε ο κώδικας ονομάζεται **κώδικας σταθερού μήκους**.



Αναδημιουργία

- Στον δέκτη εκτελείται η αντίστροφη διαδικασία αποκωδικοποίηση (decoding) ώστε από τους ψηφιακούς κωδικούς να αναπαραχθούν οι τιμές των δειγμάτων
- Στη συνέχεια με τεχνικές παρεμβολής (interpolation) είναι δυνατόν να δημιουργηθούν περισσότερες ενδιάμεσες τιμές και να αναγεννηθεί το αρχικό αναλογικό σήμα κατά μεγάλη προσέγγιση.
- Η ποιότητα του παραγόμενου σήματος εξαρτάται
 - μέγεθος δείγματος
 - συχνότητα δειγματοληψίας



Ρυθμός μετάδοσης δεδομένων

- Ένα σήμα στο οποίο έχει γίνει δειγματοληψία, κβαντοποίηση και κωδικοποίηση αποκαλείται μερικές φορές και «σήμα PCM» (Pulse Code Modulation ή σήμα παλμοκωδικά διαμορφωμένο).

Υπολογισμός ρυθμού μετάδοσης του σήματος:

- Ρυθμός μετάδοσης:
 $fs * n$ bps (bits per second)

όπου:

fs : συχνότητα δειγματοληψίας του σήματος είναι

n : μέγεθος δείγματος (bits ανά δείγμα)

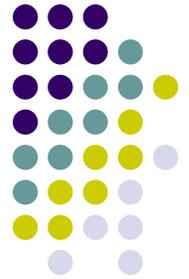


Ρυθμός μετάδοσης δεδομένων

Παράδειγμα:

- Στις εφαρμογές VoIP (Voice over IP)
 - πραγματοποιείται δειγματοληψία με συχνότητα 8KHz
 - μέγεθος δείγματος κβαντισμού 8 bits
- Αποτελέσματα
 - Δειγματοληψία: 8000 δείγματα/sec
 - Στάθμες κβαντισμού: $2^8 = 256$ στάθμες
 - Κωδικοποίηση: 8 bits/δείγμα
 - Ρυθμός μετάδοσης: $8000 * 8 = 64000$ bps = 64 Kbps

Ψηφιακή πληροφορία: πλεονεκτήματα & μειονεκτήματα



- Πλεονεκτήματα:
 - μικρότερη ευαισθησία στο θόρυβο
 - υψηλός βαθμός ολοκλήρωσης (integration)
 - Ο όρος «ολοκλήρωση» αναφέρεται στην τεχνολογική δυνατότητα να διαχειρίζεται με τον ίδιο τρόπο οποιαδήποτε ψηφιακή πληροφορία ανεξάρτητα από το τι αυτή παριστάνει (κείμενα, εικόνες, video, ήχο, κλπ)
 - Ο αντίλογος για την ολοκλήρωση: Σε πολλές περιπτώσεις οι δυνατότητες τεχνολογικής ολοκλήρωσης δεν φαίνονται να έχουν ανταπόκριση στο αγοραστικό κοινό,
 - Εύκολη υλοποίηση μεθόδων εντοπισμού και διόρθωσης λαθών, συμπίεσης και κρυπτογράφησης της πληροφορίας
- Μειονεκτήματα
 - Πιστότητα στην καταγραφή της πληροφορίας
 - Η δημιουργία ψηφιακής πληροφορίας ισορροπεί ανάμεσα στην απαίτηση για υψηλή ποιότητα αλλά και μικρό μέγεθος αρχείων