

Νικήτας Σγούρος

ΣΥΣΤΗΜΑΤΑ ΠΟΛΥΜΕΣΩΝ

Σημειώσεις

Πειραιάς, Ιανουάριος 2000

1. ΕΙΣΑΓΩΓΗ

Ο όρος Πολυμέσα αποτελεί έναν από τους πλέον χρησιμοποιημένους και ασαφείς τεχνικούς όρους των τελευταίων χρόνων. Το γνωστικό πεδίο που καλύπτει συνδυάζει δραστηριότητες όπως οι τηλεπικοινωνίες, η πληροφορική, η τέχνη, οι εκδοτικές δραστηριότητες, τα καταναλωτικά ηλεκτρονικά και τα μέσα μαζικής ενημέρωσης. Ο κοινός παρονομαστής όλων αυτών των δραστηριοτήτων βρίσκεται στην ανάγκη διαχείρισης ψηφιακής πληροφορίας που παρουσιάζεται με μια σειρά από διαφορετικούς τρόπους. Αν επιχειρήσουμε να δώσουμε ένα ορισμό του πεδίου θα λέγαμε ότι:

Τα Συστήματα Πολυμέσων είναι ο κλάδος της Πληροφορικής που ασχολείται με την ελεγχόμενη από τον υπολογιστή ενσωμάτωση κειμένου, γραφικών, ακίνητων και κινούμενων εικόνων, συνθετικής κίνησης, ήχων και κάθε άλλης μορφής πληροφορίας το περιεχόμενο της οποίας μπορεί να παρασταθεί να αποθηκευθεί, να μεταδοθεί και να τύχει επεξεργασίας με ψηφιακά μέσα.

Το παρόν μάθημα θα ασχοληθεί με την παρουσίαση των τρόπων παράστασης, αποθήκευσης, μετάδοσης και επεξεργασίας ψηφιακών πληροφοριών στα Συστήματα Πολυμέσων. Επίσης το μάθημα επιδιώκει να διδάξει βασικές αρχές σχεδίασης και προγραμματιστικά εργαλεία ανάπτυξης εφαρμογών πολυμέσων.

1.2. Μορφές πληροφορίας

Τα διάφορα είδη πληροφορίας που μπορούν να χειριστούν τα Συστήματα Πολυμέσων έχουν διαφορετικά χαρακτηριστικά. Η κατανόηση των χαρακτηριστικών αυτών θα μας βοηθήσει να καταλάβουμε τις απαιτήσεις που έχει η κάθε μορφή για την αποθήκευση, την επεξεργασία ή τη μετάδοση της.

Το κείμενο (*text*) αποτελεί την πιο παλιά μέθοδο επικοινωνίας του ανθρώπου με τον υπολογιστή. Κάθε πηγή κειμένου έχει δύο μορφές: την *αμορφοποίητη* (*unformatted* ή *plain text*) και τη *μορφοποιημένη* (*formatted* ή *rich text*). Σε αντίθεση με την πρώτη μορφή η μορφοποιημένη διαθέτει εκτεταμένα σύνολα χαρακτήρων και συμβολοσειρές διαφορετικού μεγέθους και

χαρακτήρα.

Ο όρος *γραφικά υπολογιστών (computer graphics)* περιγράφει εικόνες οι οποίες παράγονται από τον υπολογιστή και οι οποίες παριστώνται εσωτερικά σε μια συλλογή από πρωτογενή αντικείμενα, όπως είναι οι γραμμές, οι διάφοροι είδους καμπύλες, οι κύκλοι, τα πολύγωνα κλπ., από τα οποία και συντίθενται. Μια εικόνα γραφικών περιέχει τα πρωτογενή γεωμετρικά στοιχεία από τα οποία έχει δημιουργηθεί και επομένως διευκολύνει λειτουργίες όπως την ενημέρωση του περιεχομένου της με τη διαγραφή, προσθήκη, μετακίνηση ή οποιαδήποτε άλλου είδους μεταβολή των πρωτογενών αυτών στοιχείων.

Ένα υπολογιστικό σύστημα είναι σε θέση να δημιουργήσει εικόνες όχι μόνο από περιγραφές του γεωμετρικού (κυρίως) περιεχομένου τους αλλά και από πίνακες τιμών που περιγράφουν τη φωτεινότητα και το χρώμα του κάθε σημείου της εικόνας που θα παραχθεί. Οι τιμές ενός τέτοιου πίνακα τιμών ονομάζονται *ψηφίδες (pixels)*. Προφανώς εικόνες που συνθέτονται χρησιμοποιώντας πίνακες ψηφίδων δεν μπορούν να ενημερωθούν τόσο εύκολα όσο τα γραφικά, μιάς και αυτές δεν παριστούν εσωτερικά το σημασιολογικό τους περιεχόμενο. Πιο συγκεκριμένα, μια τέτοια παράσταση δε διευκολύνει την ενημέρωση μιας εικόνας σε επίπεδο αντικειμένων μιάς και αυτή απαιτεί τη χρησιμοποίηση εργαλείων ικανών να ανιχνεύσουν στον πίνακα ψηφίδων τα γεωμετρικά στοιχεία από τα οποία αποτελείται η εικόνα. Μια τέτοια ανίχνευση προφανώς δεν είναι μονοσήμαντη αφού υπάρχουν περισσότερα από ένα γεωμετρικά μοντέλα που μπορούν να απεικονιστούν σε μια διδιάστατη εικόνα. Το πεδίο της *υπολογιστικής όρασης (computer vision)* ασχολείται με την εύρεση αυτόματων μεθόδων περιγραφής του σημασιολογικού περιεχομένου μίας εικόνας που παριστάνεται σαν πίνακας ψηφίδων. Δυστυχώς η πρόοδος στο συγκεκριμένο πεδίο είναι σχετικά αργή παρόλο που έχουν αναπτυχθεί τεχνικές όπως η *οπτική αναγνώριση χαρακτήρων (optical character recognition)* που με σχετική επιτυχία μπορούν να διαβάσουν εικόνες κειμένων.

Σε σχέση με τα γραφικά, οι εικόνες που προέρχονται από πίνακες ψηφίδων έχουν δυο ή και τρεις τάξεις μεγέθους μεγαλύτερες απαιτήσεις αποθήκευσης. Το πλεονέκτημα τους βρίσκεται στο ότι δε χρειάζονται ιδιαίτερη επεξεργασία για την παρουσίαση τους σε αντίθεση με τα γραφικά στα οποία η εσωτερική αναπαράσταση τους θα πρέπει να μετασχηματιστεί σε ένα πίνακα ψηφίδων πριν από την προβολή τους. Ο μετασχηματισμός αυτός σε αρκετές περιπτώσεις έχει μεγάλες απαιτήσεις χώρου και χρόνου.

Όπως στο κινηματογράφο έτσι και στα συστήματα πολυμέσων η παρουσίαση μίας ακολουθίας εικόνων που περιγράφουν την κίνηση ενός ή περισσότερων αντικειμένων δημιουργεί τη ψευδαίσθηση της κίνησης όταν γίνεται με μία ορισμένη ταχύτητα. Κάθε εικόνα σε μια τέτοια ακολουθία εικόνων ονομάζεται *frame*. Η εντύπωση της συνεχούς κίνησης επιτυγχάνεται με μια συχνότητα προβολής που είναι μεγαλύτερη ή ίση από 15 ή 16 frames ανά δευτερόλεπτο (*frames per second* ή *fps*). Η τεχνική αυτή χρησιμοποιείται για τη προβολή εικόνων που έχουν παραχθεί είτε από βιντεοκάμερα (οπότε και μιλάμε για *captured motion video*) είτε από υπολογιστή με τη βοήθεια γραφικών (οπότε αναφερόμαστε σε *συνθετική κίνηση*). Πιο συγκεκριμένα, η *συνθετική κίνηση (computer animation)* ασχολείται με την αυτόματη σύνθεση προγραμμάτων που παράγουν ακολουθίες εικόνων. Οι ακολουθίες αυτές περιγράφουν την κίνηση ενός ή περισσότερων αντικειμένων.

Πέρα από τις εικονικές μορφές πληροφορίας τα Συστήματα Πολυμέσων διαχειρίζονται και ηχητικές πηγές πληροφορίας. Τα διάφορα είδη ηχητικών πηγών περιλαμβάνουν την ανθρώπινη φωνή, διάφορα μουσικά κομμάτια, ή ένα μεγάλο εύρος φυσικών ή συνθετικών ήχων. Οι πιο ανεπτυγμένες εφαρμογές στον τομέα αυτό ασχολούνται με την επεξεργασία της ανθρώπινης φωνής, μιάς και η τελευταία περιέχει ένα σχετικά περιορισμένο φάσμα συχνοτήτων. Επιπλέον κατά τη διάρκεια της επικοινωνίας είναι σχετικά σπάνιο σε περισσότερους από έναν συμμετέχοντες να μιλάνε συγχρόνως. Τα δυο αυτά γεγονότα έχουν διευκολύνει αρκετά την

ανάπτυξη τεχνικών για την κωδικοποίηση, τη συμπίεση και τη μετάδοση ανθρώπινης φωνής. Για παράδειγμα οι απαιτήσεις σε αποθήκευση της ανθρώπινης φωνής είναι μέχρι και 10 τάξεις μεγέθους λιγότερες από ότι οι απαιτήσεις αποθήκευσης ενός μουσικού κομματιού με την ίδια διάρκεια. Οι κυριότερες εφαρμογές που ασχολούνται με την επεξεργασία της ανθρώπινης φωνής περιλαμβάνουν την *αναγνώριση φωνής (speech recognition)* που ασχολείται με την αναγνώριση των λέξεων από τις οποίες αποτελείται ο ανθρώπινος λόγος, την *κατανόηση του λόγου (speech understanding)* που ασχολείται με την αυτόματη σημασιολογική ερμηνεία του προφορικού λόγου και τη *σύνθεση φωνής (speech synthesis)* που ασχολείται με την παραγωγή λόγου από τον υπολογιστή. Από τους τομείς αυτούς τη μεγαλύτερη ανάπτυξη παρουσιάζει η σύνθεση φωνής ακολουθούμενη από την αναγνώριση φωνής ενώ η κατανόηση του λόγου απέχει αρκετά από το να οδηγήσει σε πρακτικές εφαρμογές μεγάλης κλίμακας.

Η σύνθεση φωνής όπως και η σύνθεση διαφόρων τεχνητών ήχων από τον υπολογιστή στα πλαίσια της *υπολογιστικής μουσικής (computer music)* διαφέρει από την απλή απόδοση ηχητικών μορφών πληροφορίας από τον υπολογιστή (*sound playback*) που είναι απλώς η δημιουργία ενός αναλογικού ηχητικού σήματος από μια ψηφιοποιημένη παράσταση του ήχου αυτού.

1.3. Εφαρμογές των Συστημάτων Πολυμέσων

Επειδή οι εφαρμογές των Συστημάτων Πολυμέσων εκτείνονται σε πολυάριθμους τομείς σκόπιμο είναι να προσπαθήσουμε να τις χαρτογραφήσουμε σε σχέση με τις ανάγκες των χρηστών που καλύπτουν. Μια τέτοια παρουσίαση θα μας επιτρέψει να περιγράψουμε διεξοδικά τις ανάγκες σε υπολογιστική ισχύ και σε δικτυακή υποδομή για καθεμία από αυτές. Η πρώτη κατηγοριοποίηση των εφαρμογών πολυμέσων τις διαχωρίζει σε αυτές που αποσκοπούν στην υποβοήθηση της επικοινωνίας μεταξύ των χρηστών τους και στις εφαρμογές που αποσκοπούν να βελτιώσουν την επικοινωνία μεταξύ χρηστών και πληροφοριακών συστημάτων. Η πρώτη κατηγορία είναι γνωστή ως εφαρμογές *ανθρώπου-προς-άνθρωπο (people-to-people)* ενώ ο όρος που χρησιμοποιείται για τις εφαρμογές που ανήκουν στη δεύτερη κατηγορία είναι *ανθρώπου-προς-συστήματα (people-to-systems)*.

Οι εφαρμογές ανθρώπου-προς-άνθρωπο διαχωρίζονται σε δύο κατηγορίες: τις *σύγχρονες (synchronous)* εφαρμογές όπως η *τηλεδιάσκεψη (teleconferencing)* και τις *ασύγχρονες (asynchronous)* εφαρμογές όπως το ηλεκτρονικό ταχυδρομείο πολυμέσων.

Οι σύγχρονες εφαρμογές μπορούν με τη σειρά τους να διαχωριστούν σε *διαπροσωπικές (interpersonal)*, σε εφαρμογές *ομαδικής διανομής (group distribution)* και σε εφαρμογές τηλεδιάσκεψης. Οι διαπροσωπικές εφαρμογές περιορίζουν τον αριθμό των συμμετεχόντων σε δύο όπως π.χ. η *βιντεοτηλεφωνία (videophony)*. Οι εφαρμογές ομαδικής διανομής αναφέρονται σε μετάδοση πληροφορίας σε πραγματικό χρόνο από ένα εκπομπό σε πολλούς δέκτες (*one-to-many*). Τέλος η τηλεδιάσκεψη αναφέρεται σε εφαρμογές στις οποίες μπορεί να υπάρχουν ταυτόχρονα πολλοί εκπομποί και πολλοί δέκτες (*many-to-many*).

Οι ασύγχρονες εφαρμογές περιλαμβάνουν υπηρεσίες ηλεκτρονικού ταχυδρομείου, ανταλλαγής πηγών πληροφοριών που περιέχουν πολυμέσα και υπηρεσίες ασύγχρονης τηλεδιάσκεψης που είναι μια μετεξέλιξη των ομάδων συζήτησης (*discussion groups*) σε δίκτυα υπολογιστών.

Μια μεγάλη ομάδα εφαρμογών που έχει στενή σχέση με τις εφαρμογές ανθρώπου-προς-άνθρωπο είναι και αυτή των *συνεργατικών υπολογιστικών συστημάτων υποβοήθησης εργασιών (computer-supported collaborative work ή CSCW)*. Το πεδίο αυτό ασχολείται με το σχεδιασμό και την ανάπτυξη υπολογιστικών περιβαλλόντων που διευκολύνουν και αυξάνουν την παραγωγικότητα ομάδων εργασίας ή ατόμων που δουλεύουν στο ίδιο αντικείμενο. Τα υπολογιστικά συστήματα

και τα προϊόντα που υλοποιούν τις μεθόδους των CSCW είναι γνωστά σαν *groupware*. Το χαρακτηριστικό των συστημάτων αυτών είναι ότι ασχολούνται αποκλειστικά με την υποβοήθηση περιβαλλόντων εργασίας και δεν επεκτείνονται στην υλοποίηση ψυχαγωγικών ή οποιουδήποτε άλλου είδους εφαρμογών. Παραδείγματα τέτοιων συστημάτων αποτελούν τα συστήματα *συνεργατικής σχεδίασης (collaborative design)*, τα συστήματα υποστήριξης αποφάσεων από ομάδες (*group decision support*), τα συστήματα ψηφοφορίας μέσω υπολογιστή ή τα συστήματα τηλεδιδασκαλίας.

Οι εφαρμογές ανθρώπου-προς-συστήματα μπορούν να διαχωριστούν σε κατηγορίες ανάλογα με τις μεθόδους προσπέλασης προς το πληροφοριακό σύστημα. Τα συστήματα *διαλογικής (interactive)* προσπέλασης επιτρέπουν στο χρήστη να πάρει την πρωτοβουλία για την προσπέλαση της πληροφορίας. Ένα παράδειγμα τέτοιου συστήματος αποτελούν τα συστήματα *video-on-demand* στα οποία ο χρήστης επιλέγει τον τίτλο της ταινίας και την ώρα που θέλει να την παρακολουθήσει. Αντίθετα στα συστήματα ομαδικής διανομής (*group distribution*) το σύστημα επιλέγει την ώρα και τη μορφή της πληροφορίας που θα παρουσιαστεί. Τα συστήματα αυτά προσομοιάζουν την εκπομπή ελεύθερων ή συνδρομητικών τηλεοπτικών προγραμμάτων.

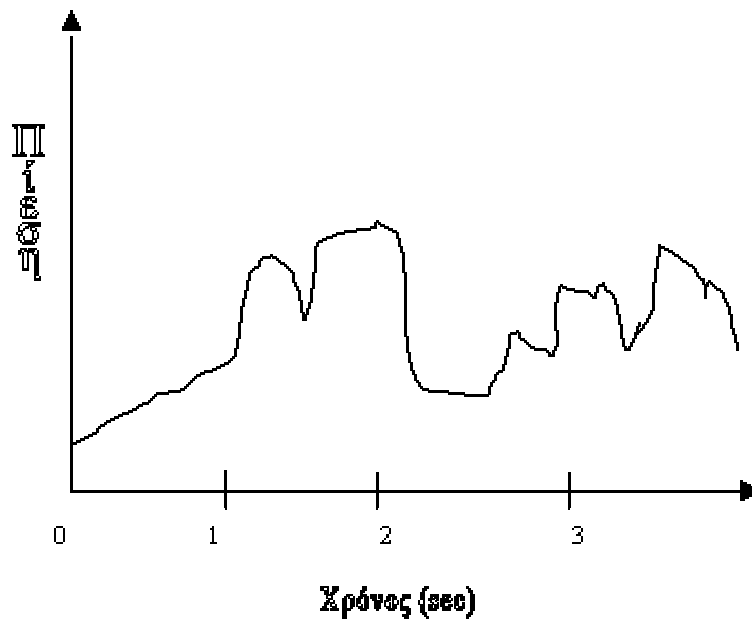
Τέλος, ένας διαχωρισμός των εφαρμογών πολυμέσων αφορά εφαρμογές στις οποίες ο χρήστης θα πρέπει να μεταβεί σε ένα ειδικά εξοπλισμένο χώρο για να τις χρησιμοποιήσει (*room-based mode*) και εφαρμογές τις οποίες ο χρήστης μπορεί να χρησιμοποιήσει μέσω του προσωπικού του υπολογιστικού συστήματος (*desktop mode*).

2. ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ ΕΠΕΞΕΡΓΑΣΙΑΣ ΣΗΜΑΤΩΝ

2.1 Εισαγωγή

Τα Συστήματα Πολυμέσων ασχολούνται με την ανάπτυξη υπολογιστικών μηχανισμών με τους οποίους μπορούμε να χειριστούμε, να μεταβάλλουμε, να καταγράψουμε ή να μεταδώσουμε συνδυασμούς σημάτων. Όλοι μας διαισθητικά έχουμε μια ιδέα της σημασίας των όρων *σήμα* και *σύστημα*. Για τις ανάγκες του συγκεκριμένου μαθήματος θα θεωρήσουμε ότι σήμα είναι ένα σύνολο μεταβολών που αναπαριστούν πληροφορίες. Παραδείγματα τέτοιων σημάτων αποτελούν ο ανθρώπινος λόγος, η μουσική, οι εικόνες ή τα σήματα βίντεο. Εδώ θα πρέπει να σημειώσουμε ότι τα σήματα μπορούν να πάρουν αρκετές ισοδύναμες μορφές ή αναπαραστάσεις. Για παράδειγμα, η ανθρώπινη φωνή αποτελεί ένα ακουστικό σήμα το οποίο μπορεί να μετατραπεί σε ηλεκτρικό σήμα με τη βοήθεια ενός μικροφώνου, σε ένα σύνολο από διαταράξεις του μαγνητικού πεδίου σε μια κασέτα ή ακόμη και σε μια ακολουθία από αριθμούς σε μια ψηφιακή ηχογράφηση.

Με τον όρο σύστημα θα αναφερόμαστε σε υπολογιστικά περιβάλλοντα με τα οποία μπορούμε να χειριστούμε να μεταβάλλουμε, να καταγράψουμε ή να μεταδώσουμε σήματα. Για παράδειγμα 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. Γενικά τα συστήματα εφαρμόζονται πάνω σε σήματα για να παράγουν καινούργιες μορφές ή αναπαραστάσεις σημάτων.



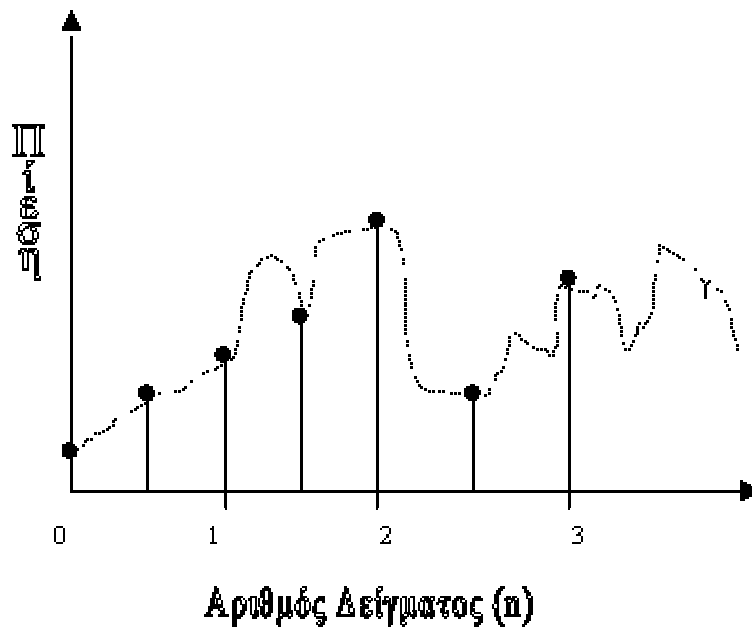
Εικόνα 1.1: Παράδειγμα ενός μονοδιάστατου, συνεχούς στο χρόνο σήματος.

Αρκετά σήματα μπορούν να θεωρηθούν σαν ένα σύνολο μεταβολών στο χρόνο. Για παράδειγμα η ανθρώπινη φωνή προκύπτει από τις μεταβολές της πίεσης του αέρα λόγω της ταλάντωσης των φωνητικών χορδών στο λάρυγγα. Οι μεταβολές αυτές εξελίσσονται στο χρόνο δημιουργώντας μια *κυματομορφή στο χρόνο (time waveform)*. Για παράδειγμα, η Εικόνα 1.1 περιγράφει την κυματομορφή που αντιστοιχεί σε ένα κομμάτι ανθρώπινης ομιλίας. Στην εικόνα αυτή ο οριζόντιος άξονας περιγράφει το χρόνο ενώ ο κάθετος άξονας περιγράφει την πίεση του αέρα. Το σήμα της Εικόνας 1.1 αποτελεί ένα παράδειγμα ενός *μονοδιάστατου (one-dimensional), συνεχούς (continuous)* σήματος. Το σήμα είναι μονοδιάστατο γιατί μπορεί να περιγραφεί μαθηματικά σαν μια συνάρτηση μίας ανεξάρτητης μεταβλητής που συνήθως ονομάζεται t . Το σήμα είναι συνεχές γιατί για κάθε τιμή του t ($t > 0$) μπορεί να υπολογιστεί μια τιμή του σήματος.

Το γεγονός ότι οι υπολογιστές είναι σε θέση να χειριστούν μόνο διακριτά δεδομένα μας αναγκάζει να μετατρέψουμε τα σήματα που μας ενδιαφέρουν από τη συνεχή σε μία *διακριτή στο χρόνο (discrete-time)* μορφή. 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. Το αποτέλεσμα της δειγματοληψίας είναι μια ακολουθία αριθμών που μπορεί να παρασταθεί σα μια συνάρτηση μιας μεταβλητής δείκτη που παίρνει μόνο διακριτές τιμές. Μαθηματικά θα έχουμε ότι:

$$s[n] = s(nT_s)$$

όπου $s()$ η συνεχής συνάρτηση του σήματος, $s[]$ η διακριτή συνάρτηση του σήματος, n είναι ένας ακέραιος ($n = \dots, -2, -1, 0, 1, 2, \dots$) και T_s η περίοδος της δειγματοληψίας (δηλ. η απόσταση μεταξύ διαδοχικών δειγμάτων στο χρόνο). Για παράδειγμα η Εικόνα 1.2 περιγράφει το αποτέλεσμα της δειγματοληψίας του σήματος της Εικόνας 1.1 με περίοδο δειγματοληψίας ίση με 0.5 sec.



Εικόνα 1.2: Παράδειγμα ενός μονοδιάστατου διακριτού στο χρόνο σήματος.

Εκτός από τα σήματα που μεταβάλλονται στο χρόνο, όπως τα ηχητικά σήματα, υπάρχουν και κατηγορίες σημάτων που μεταβάλλονται μόνο στο χώρο ή και στο χώρο και στο χρόνο. Για παράδειγμα, μια φωτογραφία αποτελεί παράδειγμα ενός σήματος που δε μεταβάλλεται στο χρόνο. Ένα τέτοιο σήμα μπορεί να περιγραφεί σα μια συνάρτηση (π.χ. $f(x,y)$) δύο μεταβλητών κάθε μία από τις οποίες περιγράφει τις συντεταγμένες x και y κάθε σημείου της εικόνας σε δύο κάθετους άξονες συντεταγμένων. Στην περίπτωση αυτή, η τιμή της συνάρτησης f σε κάθε σημείο (x, y) περιγράφει το χρώμα στο αντίστοιχο σημείο της φωτογραφίας. Η συγκεκριμένη συνάρτηση $f(x, y)$ αποτελεί παράδειγμα μια *διδιάστατης, συνεχούς στο χώρο* συνάρτησης. Η διακριτή μορφή της f θα προέλθει από τη δημιουργία ενός πίνακα από αριθμούς, οι γραμμές και οι στήλες του οποίου θα περιγράφουν τα σημεία στο χώρο στα οποία γίνεται η δειγματοληψία στον οριζόντιο και κάθετο άξονα αντίστοιχα και η τιμή σε κάθε σημείο του πίνακα θα μας δίνει το χρώμα στο αντίστοιχο σημείο. Μαθηματικά θα έχουμε ότι:

$$f[m, n] = f(m\Delta_x, n\Delta_y)$$

όπου $f()$ η συνεχής συνάρτηση του σήματος, $f[]$ η διακριτή συνάρτηση του σήματος, n, m είναι ακέραιοι ($m = \dots, -2, -1, 0, 1, 2, \dots$, $n = \dots, -2, -1, 0, 1, 2, \dots$) και Δ_x, Δ_y οι περίοδοι δειγματοληψίας στην οριζόντια και κατακόρυφη διεύθυνση αντίστοιχα. Οι περίοδοι δειγματοληψίας αντιστοιχούν στην απόσταση στο χώρο μεταξύ διαδοχικών δειγμάτων της εικόνας στον οριζόντιο και κάθετο άξονα.

Σήματα που μεταβάλλονται και στο χώρο και στο χρόνο όπως λ.χ. οι εικόνες ενός βίντεο, περιγράφονται από τρισδιάστατες συναρτήσεις (π.χ. $v(x, y, t)$) όπου x, y οι συντεταγμένες ενός σημείου της οθόνης προβολής και t ο χρόνος στον οποίο προβάλλεται η εικόνα στη συγκεκριμένη οθόνη.

Όπως αναφέραμε και προηγουμένως ένα σύστημα αποτελεί ένα μηχανισμό ο οποίος λαμβάνει

σαν εισόδους σήματα τα οποία μετασχηματίζει σε καινούργιες μορφές ή αναπαραστάσεις σημάτων. Κατά αναλογία με τα σήματα και τα συστήματα μπορούν να είναι μονοδιάστατα ή πολυδιάστατα, διακριτά ή συνεχή. Για παράδειγμα ένα μονοδιάστατο, συνεχές στο χρόνο σύστημα λαμβάνει σαν είσοδο ένα σήμα $x(t)$ και παράγει στην έξοδο ένα σήμα $y(t)$. Μαθηματικά μπορούμε να παραστήσουμε τη λειτουργία αυτή σαν:

$$y(t) = \mathcal{A} \{ x(t) \}$$

που σημαίνει ότι το σύστημα επενεργεί μέσω της διαδικασίας \square στο σήμα εισόδου $x(t)$ για να παράγει το σήμα εξόδου $y(t)$. Για παράδειγμα το ακόλουθο σύστημα:

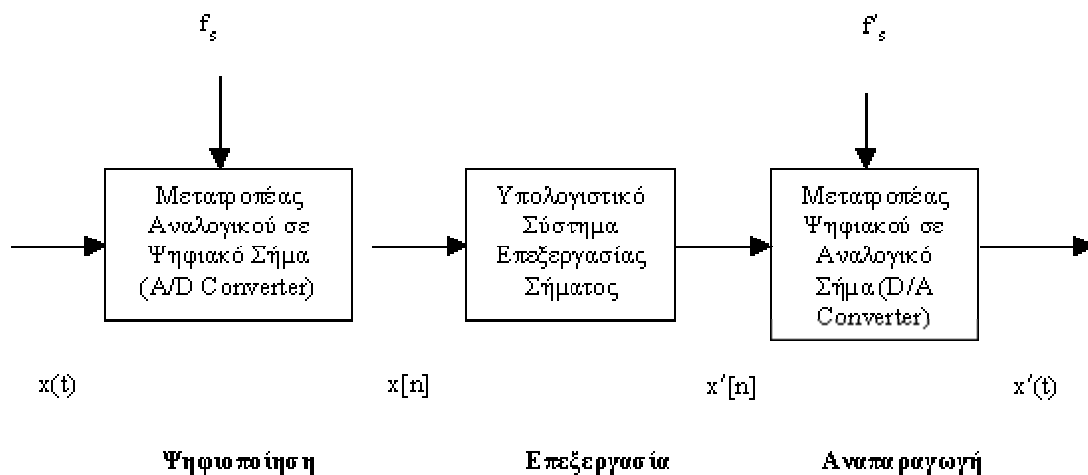
$$y(t) = [x(t)]^3$$

αποτελεί ένα παράδειγμα συνεχούς συστήματος στο οποίο σε κάθε χρονική στιγμή το σήμα εξόδου είναι ίσο με τον κύβο του σήματος εισόδου. Η διακριτή μορφή του συγκεκριμένου συστήματος θα είναι:

$$y[n] = (x[n])^3$$

2.2 Καταγραφή, Επεξεργασία και Αναπαραγωγή Σημάτων

Η λειτουργία της πλειονότητας των Συστημάτων Πολυμέσων μπορεί να περιγραφεί από ένα διάγραμμα που αποτελείται από τρεις φάσεις: τη φάση της ψηφιοποίησης, της επεξεργασίας και της αναπαραγωγής ενός σήματος (βλ. Εικόνα 2.1).



Εικόνα 2.1: Στάδια Επεξεργασίας Σήματος σε Συστήματα Πολυμέσων.

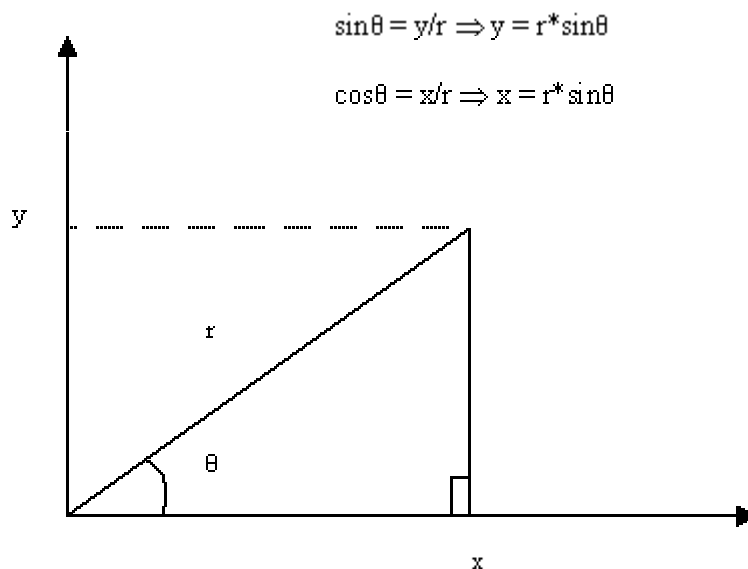
Η φάση της ψηφιοποίησης λαμβάνει ως είσοδο ένα αναλογικό (δηλ. συνεχές) σήμα $x(t)$ το οποίο και μετασχηματίζει σε ένα ψηφιακό (δηλ. διακριτό) σήμα $x[n]$. Το ψηφιακό αυτό σήμα τροφοδοτείται στη συνέχεια σε ένα σύστημα επεξεργασίας σήματος το οποίο και δημιουργεί ένα άλλο ψηφιακό σήμα $x'[n]$. Το $x'[n]$ τροφοδοτείται σε ένα μετατροπέα ψηφιακού σε αναλογικό σήμα ο οποίος και παράγει στην έξοδο του ένα αναλογικό σήμα $x'(t)$. Το $x'(t)$ αποτελεί και την έξοδο του συστήματος που γίνεται αντιληπτή από το χρήστη. Τόσο η ψηφιοποίηση όσο και η μετατροπή του σήματος από ψηφιακό σε αναλογικό γίνονται με τη χρησιμοποίηση αντίστοιχων συχνοτήτων δειγματοληψίας (f_s και f'_s , αντίστοιχα, στην Εικόνα 2.1).

Για παράδειγμα, η αλυσίδα ηχογράφησης και ανάγνωσης μουσικών οπτικών δίσκων (CD) δέχεται σαν είσοδο ένα αναλογικό σήμα (τον ήχο της ηχογράφησης) ο οποίος ψηφιοποιείται και

αποθηκεύεται σε διακριτή μορφή σε έναν οπτικό δίσκο. Για να είναι δυνατή η ακρόαση του συγκεκριμένου δίσκου το σήμα που περιέχει εισάγεται με τη βοήθεια του αναγνώστη οπτικών δίσκων (CD player) σε ένα μετατροπέα ψηφιακού σε αναλογικό σήμα ο οποίος μετατρέπει το αποθηκευμένο διακριτό σήμα σε αναλογικό σήμα. Το αναλογικό σήμα που παράγεται διεγείρει τη μεμβράνη ενός ηχείου και δημιουργεί ένα αναλογικό ακουστικό σήμα το οποίο και φτάνει στα αυτιά μας..

2.3 Μαθηματικές Αναπαραστάσεις Σημάτων

Η πλειονότητα των σημάτων στα Συστήματα Πολυμέσων μπορούν να περιγραφούν σαν τριγωνομετρικές συναρτήσεις και πιο συγκεκριμένα σαν συναρτήσεις συνημιτόνων ή ημιτόνων. Στην ενότητα αυτή θα περιγράψουμε συνοπτικά τις ιδιότητες των συγκεκριμένων συναρτήσεων στις οποίες θα αναφερόμαστε συνοπτικά σαν τριγωνομετρικές συναρτήσεις και θα παρουσιάσουμε την ορολογία που χρησιμοποιείται στην περιγραφή των συγκεκριμένων σημάτων.



Εικόνα 2.1: Ορισμός των συναρτήσεων συνημιτόνου και ημιτόνου μιας γωνίας θ σε ένα ορθογώνιο τρίγωνο.

Οι ορισμοί των συναρτήσεων του *συνημιτόνου* (*cos*) και του *ημιτόνου* (*sin*) δίνονται συνήθως με τη βοήθεια ενός διαγράμματος παρόμοιου με αυτό της Εικόνας 2.1. Οι συγκεκριμένες συναρτήσεις λαμβάνουν σαν όρισμα μια γωνία θ η οποία περιγράφεται σε *ακτίνια* (*rad*). Όταν η γωνία θ βρίσκεται στο πρώτο τεταρτημόριο ($0 \leq \theta \leq \pi/2$) τότε το ημίτονο της θ ισούται με το μήκος της πλευράς y του τριγώνου που βρίσκεται απέναντι από τη γωνία θ δια το μήκος της υποτείνουσας του ορθογωνίου τριγώνου. Παρομοίως, το συνημίτονο της θ ορίζεται από το λόγο του μήκους της παρακείμενης πλευράς x προς το μήκος της υποτείνουσας. Όταν η θ αυξάνεται από το 0 προς το $\pi/2$, το συνημίτονο της μειώνεται από το 1 στο 0 και το ημίτονο της αυξάνεται από το 0 στο 1. Όταν η γωνία θ είναι μεγαλύτερη του $\pi/2$ τότε τα πρόσημα των x και y λαμβάνονται υπ' όψη στον υπολογισμό των συναρτήσεων. Το x είναι αρνητικό στο δεύτερο και τρίτο τεταρτημόριο ενώ το y είναι αρνητικό στο τρίτο και τέταρτο τεταρτημόριο.

Ο πιο γενικός τύπος περιγραφής της συνημιτονικής συνάρτησης προκύπτει από την έκφραση του

ορίσματος της συνάρτησης (δηλ. της γωνίας θ) ως συνάρτησης του χρόνου. Δύο ισοδύναμες μορφές του τύπου αυτού είναι οι:

$$x(t) = A\cos(\omega_0 t + \phi) = A\cos(2\pi f_0 t + \phi) \quad (2.1)$$

όπου $\omega_0 = 2\pi f_0$. Τα σύμβολα στην (2.1) έχουν τα ακόλουθα ονόματα και φυσική σημασία:

1. Το A ονομάζεται *πλάτος (amplitude)* της συνημιτονικής συνάρτησης. Το πλάτος καθορίζει το εύρος των τιμών που λαμβάνει η συνημιτονική συνάρτηση. Επειδή η συγκεκριμένη συνάρτηση ταλαντώνεται μεταξύ $+1$ και -1 , το σήμα $x(t)$ στην (2.1) λαμβάνει τιμές μεταξύ $+A$ και $-A$.
2. Το ϕ ονομάζεται *φάση (phase shift)* της συνάρτησης και μετριέται σε ακτίνια. Το ω_0 ονομάζεται *ακτινική συχνότητα (radian frequency)* και μετριέται σε rad/sec. Κατ' αναλογία το $f_0 (= \omega_0/2\pi)$ ονομάζεται *κυκλική συχνότητα (cyclic frequency)* και μετριέται σε Hz (Hertz). Στο υπόλοιπο της ενότητας με τον όρο συχνότητα θα αναφερόμαστε στην κυκλική συχνότητα των συναρτήσεων που εξετάζουμε.

Για παράδειγμα, η έκφραση:

$$x(t) = 20\cos(2\pi 40t - 0.4\pi) \quad (2.2)$$

Περιγράφει ένα συνημιτονικό σήμα με πλάτος 20, κυκλική συχνότητα 40 Hz και φάση ίση με -0.4π rad.

2.3.1 Συχνότητα & Περίοδος

Τόσο η συνημιτονική όσο και η ημιτονική συνάρτηση είναι *περιοδικές*, δηλαδή παράγουν την ίδια ακολουθία τιμών σε τακτά χρονικά διαστήματα (κύκλους). Η περίοδος των δύο συναρτήσεων συμβολίζεται σαν T_0 και είναι ίση με το μήκος ενός κύκλου της συνάρτησης. Γενικά η συχνότητα της συνάρτησης καθορίζει την περίοδο της με βάση τις ακόλουθες εξισώσεις:

$$T_0 = 2\pi/\omega_0 \text{ ή } T_0 = 1/f_0$$

Η κυκλική συχνότητα των συναρτήσεων που εξετάζουμε μετρά τον αριθμό των περιόδων (κύκλων) ανά δευτερόλεπτο.

Για παράδειγμα, η περίοδος του σήματος της (2.2) είναι:

$$T_0 = 1/f_0 = 1/40 = 0.025 \text{ sec}$$

Κατά συνέπεια, ένας πλήρης κύκλος του σήματος αυτού διαρκεί 0.025 sec.

2.3.2 Φάση & Μετατόπιση στο Χρόνο

Η φάση των συναρτήσεων που εξετάζουμε σε συνδυασμό με τη κυκλική συχνότητα καθορίζει τις θέσεις στο χρόνο των μεγίστων και ελαχίστων τιμών της συνάρτησης. Όταν $\phi \neq 0$ η φάση καθορίζει πόσο οι τιμές της συνάρτησης είναι μετατοπισμένες σε σχέση με τις τιμές μίας συνάρτησης με το ίδιο πλάτος και συχνότητα αλλά με φάση ίση με το 0. Η μετατόπιση δίνεται από τον τύπο:

$$\Delta t = -\phi/2\pi f_0$$

Για παράδειγμα, το σήμα της (2.2) θα είναι μετατοπισμένο σε σχέση με το σήμα $x'(t) = 20\cos(2\pi 40t)$ κατά $\Delta t = -\phi/2\pi f_0 = -1/2 \times 3.14 \times 40 \approx 0.004 \text{ sec}$.

Όταν δύο περιοδικά σήματα της ίδιας συχνότητας και πλάτους ξεκινούν από το ίδιο πλάτος τότε λέμε ότι βρίσκονται *σε φάση*. Αν το ένα από τα δύο σήματα αποτελεί μία καθυστερημένη έκδοση του άλλου τότε λέμε ότι τα δύο σήματα βρίσκονται *εκτός φάσης* ή ότι *έχουν διαφορά φάσης*. Για παράδειγμα ένας ημιτονικός και ένας συνημιτονικός παλμός έχουν διαφορά φάσης 90° . Σήματα που έχουν μεταξύ τους διαφορά φάσης 180° αλληλοαναιρούνται όταν προστεθούν. Στην περίπτωση αυτή λέμε ότι τα σήματα *έχουν αντίθετη πολικότητα*.

2.4 Φάσμα Συχνοτήτων

Υπάρχουν δύο κύριοι λόγοι για τους οποίους οι τριγωνομετρικές συναρτήσεις έχουν εξέχουσα σημασία στα Συστήματα Πολυμέσων. Ο πρώτος λόγος προέρχεται από το γεγονός ότι οι συγκεκριμένες συναρτήσεις περιγράφουν τις λύσεις των διαφορικών εξισώσεων με τις οποίες μοντελοποιούνται ταλαντώσεις στη φύση. Για παράδειγμα, η διέγερση ενός διαπασών μπορεί να περιγραφεί με τη χρήση μίας τέτοιας συνάρτησης. Επειδή οι ταλαντώσεις αποτελούν τη φυσική αιτία της παραγωγής ηχητικών αλλά και πολλών άλλων ειδών σημάτων, μπορούμε να αναπαραστήσουμε τα σήματα αυτά χρησιμοποιώντας κατάλληλους συνδυασμούς τριγωνομετρικών συναρτήσεων.

Ο δεύτερος και πλέον σημαντικός λόγος για τη χρήση των συγκεκριμένων συναρτήσεων προέρχεται από το γεγονός ότι μπορούμε να θεωρήσουμε ότι κάθε σήμα αναλύεται σε ένα άθροισμα συνημιτονικών συναρτήσεων με διάφορα πλάτη, συχνότητες και φάσεις. Χρησιμοποιώντας την ανάλυση αυτή μπορούμε να περιγράψουμε ένα οποιοδήποτε 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. Στις επόμενες ενότητες θα περιγράψουμε τη μορφή που έχει το φάσμα συχνοτήτων για ηχητικά ή οπτικά σήματα.

2.5 Δειγματοληψία

Στην § 2.2 αναφέραμε ότι η μετατροπή ενός αναλογικού σε ψηφιακό σήμα και αντίστροφα γίνεται με μία ορισμένη συχνότητα δειγματοληψίας. Το ερώτημα που τίθεται είναι πόσο μικρή ή μεγάλη μπορεί να γίνει η συχνότητα δειγματοληψίας έτσι ώστε όλες οι συχνότητες που περιέχονται στο σήμα εισόδου να περιέχονται αναλλοίωτες και στο σήμα εξόδου. Για παράδειγμα, ποιό εύρος συχνοτήτων δειγματοληψίας μπορούμε να χρησιμοποιήσουμε έτσι ώστε το ψηφιακό σήμα που προκύπτει από τη δειγματοληψία του σήματος της σχέσης (2.2) να είναι περιοδικό με συχνότητα ίση με τη συχνότητα της (2.2);

Η απάντηση στο συγκεκριμένο ερώτημα δίνεται από ένα από τα θεμελιώδη θεωρήματα της επεξεργασίας σήματος, το *θεώρημα του Shannon*, το οποίο αποδεικνύει ότι:

Ένα συνεχές στο χρόνο σήμα $x(t)$ που περιέχει συχνότητες όχι μεγαλύτερες από f_{max} μπορεί να ανακατασκευαστεί ακριβώς από τα δείγματα του $x[n] = x[nT_s]$, αν τα δείγματα αυτά έχουν ληφθεί με συχνότητα $f_s = 1/T_s$ που είναι μεγαλύτερη από $2f_{max}$.

Για παράδειγμα, η δειγματοληψία του σήματος (2.2) συνίσταται στον υπολογισμό της τιμής της $x(t)$ σε ένα διακριτό σύνολο χρόνων $t_n = nT_s$, όπου το n είναι ακέραιος. Το σύνολο τιμών του σήματος που προκύπτει δίνεται από τον τύπο:

$$x(nT_s) = 20\cos(80\pi nT_s - 0.4\pi)$$

όπου το T_s ονομάζεται *περίοδος δειγματοληψίας* και το n είναι ακέραιος. Επειδή η μέγιστη συχνότητα στο σήμα (2.2) είναι 40 Hz κατά το θεώρημα του Shannon η ελάχιστη συχνότητα δειγματοληψίας (f_{sample}) που θα πρέπει να χρησιμοποιηθεί για τη ψηφιοποίηση του (2.2) είναι $2 \times 40 = 80$ Hz άρα η μέγιστη περίοδος δειγματοληψίας του συγκεκριμένου σήματος είναι $1/80 =$

0.0125 sec.

Υπάρχουν δύο κύριες παρατηρήσεις που μπορούμε να κάνουμε για το θεώρημα του Shannon. Ενώ το θεώρημα αποδεικνύει ότι μια πλήρης ανακατασκευή του σήματος $x(t)$ είναι δυνατή με τις συχνότητες δειγματοληψίας που περιγράφονται, δε μας δίνει ένα αλγόριθμο με τον οποίο μπορούμε να ανακατασκευάσουμε το σήμα μας. Επίσης το θεώρημα καθορίζει μια ελάχιστη συχνότητα δειγματοληψίας κάτω από την οποία η ανακατασκευή του σήματος μας δεν είναι δυνατή. Η ελάχιστη αυτή συχνότητα είναι γνωστή ως *συχνότητα Nyquist (Nyquist rate)*.

Υπάρχουν πάρα πολλές εφαρμογές του θεωρήματος του Shannon. Για παράδειγμα, οι μουσικοί οπτικοί δίσκοι χρησιμοποιούν συχνότητες δειγματοληψίας ίσες με 44.1 KHz για την αποθήκευση του ηχητικού σήματος. Η συγκεκριμένη συχνότητα 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 εξασφαλίζεται ότι όλες οι συχνότητες που είναι μικρότερες ή ίσες των 20 kHz θα αποδοθούν σωστά στο διακριτό σήμα.

Όταν η συχνότητα δειγματοληψίας είναι μικρότερη της συχνότητας Nyquist τότε εμφανίζεται το φαινόμενο της *αναδίπλωσης (aliasing ή foldover)* για τις συχνότητες που περιέχονται στο σήμα μας και είναι μεγαλύτερες από το μισό της συχνότητας Nyquist. Πιο συγκεκριμένα, αν f_0 είναι η συχνότητα ενός σήματος το οποίο θέλουμε να δειγματοληψήσουμε και για την οποία ισχύει ότι:

$$f_{\text{Nyquist}}/2 < f_0 < f_{\text{Nyquist}}$$

τότε το συγκεκριμένο σήμα θα εμφανιστεί μετά τη δειγματοληψία με συχνότητα:

$$f_1 = f_{\text{sample}} - f_0$$

Για παράδειγμα αν ψηφιοποιούμε ένα σήμα συχνότητας 26 KHz με συχνότητα δειγματοληψίας $f_{\text{sample}} = 50$ KHz τότε θα έχουμε αναδίπλωση και το σήμα εξόδου θα έχει συχνότητα $50 - 26 = 24$ KHz.

2.6 Γραμμικά, Αμετάβλητα στο Χρόνο Συστήματα

Στην επεξεργασία σήματος ιδιαίτερο ενδιαφέρον παρουσιάζει μια κλάση συστημάτων που χαρακτηρίζονται ως *γραμμικά (linear)* και *αμετάβλητα στο χρόνο (time-invariant)*. Ας υποθέσουμε ότι ένα ψηφιακό σύστημα παράγει στη έξοδο τα σήματα $y_1[n]$ και $y_2[n]$ όταν δέχεται σαν είσοδο τα σήματα $x_1[n]$ και $x_2[n]$ αντίστοιχα. Το συγκεκριμένο σύστημα χαρακτηρίζεται ως γραμμικό όταν για οποιοδήποτε α και β η εφαρμογή του σήματος εισόδου $\alpha x_1[n] + \beta x_2[n]$ σε αυτό παράγει στην έξοδο το σήμα $\alpha y_1[n] + \beta y_2[n]$.

Ας υποθέσουμε ότι ένα ψηφιακό σύστημα παράγει στην έξοδο το σήμα $y[n]$ όταν δέχεται ως είσοδο το σήμα $x[n]$. Το συγκεκριμένο σύστημα χαρακτηρίζεται ως αμετάβλητο στο χρόνο όταν για οποιοδήποτε n_0 , η εφαρμογή του σήματος εισόδου $x[n-n_0]$ σε αυτό παράγει στην έξοδο το σήμα $y[n-n_0]$.

Τα ψηφιακά συστήματα που είναι γραμμικά και αμετάβλητα στο χρόνο παρουσιάζουν ιδιαίτερο ενδιαφέρον καθώς η κρουστική τους απόκριση περιγράφει πλήρως τη συμπεριφορά τους (βλ. § 3.7.2).

2.7 Συμπύεση

Ο μέγας όγκος πληροφορίας που περιέχεται στα σήματα που χρησιμοποιούνται στα Συστήματα Πολυμέσων επιβάλλει τη χρησιμοποίηση μεθόδων συμπίεσης που αποσκοπούν στη μείωση των απαιτήσεων μετάδοσης ή αποθήκευσης των συγκεκριμένων σημάτων. Τη συμπίεση

ενός τέτοιου σήματος ακολουθεί η αντίστροφη διαδικασία της αποσυμπίεσης στο δέκτη που σκοπό έχει να επαναφέρει το σήμα στην αρχική του μορφή.

Ανάλογα με τη σχέση που έχει το αρχικό σήμα με το αποτέλεσμα της αποσυμπίεσης υπάρχουν δύο κατηγορίες μεθόδων συμπίεσης: οι τεχνικές με απώλειες (*lossy compression schemes*) και οι τεχνικές χωρίς απώλειες (*lossless compression schemes*). Οι τεχνικές χωρίς απώλειες δε μεταβάλλουν τα χαρακτηριστικά του σήματος κατά τη διάρκεια της συμπίεσης. Κατά συνέπεια το σήμα που προκύπτει κατά την αποσυμπίεση είναι ακριβές αντίγραφο του αρχικού. Τεχνικές αυτής της μορφής χρησιμοποιούνται για τη συμπίεση σημάτων τα οποία δε θα πρέπει να αλλοιωθούν κατά τη διαδικασία της συμπίεσης/αποσυμπίεσης (π.χ. ιατρικά σήματα).

Αντίθετα, οι τεχνικές με απώλειες αλλοιώνουν τα χαρακτηριστικά του σήματος κατά τη διαδικασία της συμπίεσης. Τεχνικές αυτής της μορφής χρησιμοποιούνται κυρίως για τη συμπίεση ηχητικών σημάτων ή εικόνων. Οι περισσότερες από τις μεθόδους που ανήκουν σε αυτήν την κατηγορία εκμεταλλεύονται τα φυσιολογικά χαρακτηριστικά της ανθρώπινης όρασης και ακοής για να εισάγουν επιλεκτικά αλλοιώσεις στο σήμα οι οποίες γίνονται δύσκολα αντιληπτές από το δέκτη.

Ανάλογα με τη μέθοδο που ακολουθείται για τη συμπίεση μίας πηγής πληροφορίας υπάρχουν δύο κατηγορίες τεχνικών συμπίεσης: οι τεχνικές κωδικοποίησης εντροπίας (*entropy encoding*) και οι τεχνικές κωδικοποίησης πηγής (*source encoding*).

2.7.1 Κωδικοποίηση Εντροπίας

Ο όρος κωδικοποίηση εντροπίας αναφέρεται σε μία σειρά τεχνικών χωρίς απώλειες οι οποίες δε λαμβάνουν υπόψη τους τη φύση των σημάτων στα οποία εφαρμόζονται. Οι συγκεκριμένες μέθοδοι θεωρούν ότι το σήμα που συμπιέζεται δεν είναι τίποτα άλλο παρά μία σειρά από δυαδικά ψηφία.

Η συμπίεση σε αρκετές από τις μεθόδους αυτές στηρίζεται στην ανίχνευση σε ένα σήμα ακολουθιών που αποτελούνται από ένα και μόνο στοιχείο και στην αντικατάστασή τους με ένα αριθμό που δηλώνει το μήκος τους ακολουθούμενο από το στοιχείο από το οποίο συντίθενται. Για παράδειγμα, κατά τη συμπίεση ενός κειμένου μια ακολουθία της μορφής “αβββββαββββ” μπορεί να συμπιεσθεί στην ακολουθία “α6βα3β”. Στην κατηγορία αυτή ανήκουν τεχνικές όπως η *run-length encoding (RLE)* που χρησιμοποιείται για τη μετάδοση τηλεμοιοτυπιών (*fax*).

Μια εναλλακτική μέθοδος κωδικοποίησης εντροπίας, η *στατιστική κωδικοποίηση*, ανιχνεύει ακολουθίες στοιχείων με μεγάλη συχνότητα εμφάνισης σε ένα σήμα. Στη συνέχεια η μέθοδος κωδικοποιεί τις ακολουθίες που ανίχνευσε με τέτοιο τρόπο ώστε να καταλαμβάνουν το ελάχιστο δυνατό μήκος στο σήμα. Το μήκος της κωδικοποίησης που χρησιμοποιείται για κάθε ακολουθία είναι αντιστρόφως ανάλογο της συχνότητας εμφάνισης της στο σήμα. Στην περίπτωση αυτή η αποσυμπίεση στηρίζεται στην εφαρμογή μίας μεθόδου αποκωδικοποίησης των συμπιεσμένων ακολουθιών που περιέχονται στο σήμα. Μία αρκετά δημοφιλής μέθοδος στατιστικής κωδικοποίησης είναι η *κωδικοποίηση Huffman*.

2.7.2 Κωδικοποίηση Πηγής

Οι μέθοδοι κωδικοποίησης πηγής λαμβάνουν υπόψη τους τη φύση του σήματος που συμπιέζεται. Για παράδειγμα, μία τέτοια μέθοδος μπορεί να ανιχνεύει και να συμπιέζει δραστικά περιόδους σιωπής σε ένα ηχητικό σήμα μιάς και οι τελευταίες δεν περιέχουν καμμία χρήσιμη ακουστική πληροφορία πέραν της διάρκειάς τους. Συνήθως οι τεχνικές κωδικοποίησης πηγής πετυχαίνουν μεγαλύτερους βαθμούς συμπίεσης από τις κωδικοποιήσεις εντροπίας αν και ο βαθμός συμπίεσης είναι μεταβλητός και εξαρτάται από τη μορφή του συγκεκριμένου σήματος. Επιπλέον αρκετές από τις μεθόδους κωδικοποίησης πηγής υλοποιούν συμπίεση με απώλειες.

Υπάρχουν τρεις κύριες παραλλαγές της μεθόδου: η *κωδικοποίηση μετασχηματισμού (transform encoding)*, η *διαφορική κωδικοποίηση (differential encoding)* και η *διανυσματική κβαντοποίηση (vector quantization)*.

Η αρχή λειτουργίας της κωδικοποίησης μετασχηματισμού στηρίζεται στο μετασχηματισμό ενός σήματος σε μια αναπαράσταση που διευκολύνει τη συμπίεση του (π.χ. μετασχηματισμός Fourier). Η συμπίεση πραγματοποιείται στην εναλλακτική αυτή αναπαράσταση και το συμπεριεσμένο σήμα επανέρχεται μέσω του αντίστροφου μετασχηματισμού στην αρχική του αναπαράσταση.

Η διαφορική κωδικοποίηση στηρίζεται στην αποθήκευση των διαφορών μεταξύ διαδοχικών τιμών ενός σήματος. Ειδικότερα η μέθοδος στηρίζεται στην παρατήρηση ότι υψηλές συχνότητες δειγματοληψίας συνήθως οδηγούν σε μικρές διαφορές μεταξύ διαδοχικών δειγμάτων ενός σήματος. Κατά συνέπεια το μήκος του σήματος που προκύπτει από την αποθήκευση των διαφορών μεταξύ διαδοχικών δειγμάτων ενός σήματος είναι κατά κανόνα σημαντικά μικρότερο από το σήμα που περιέχει τις ονομαστικές τιμές όλων των δειγμάτων.

Τέλος, η διανυσματική κβαντοποίηση επιδιώκει να ανιχνεύσει και να κωδικοποιήσει τις στοιχειώδεις *μορφοκλασματικές (fractal)* μορφές από τις οποίες συντίθεται ένα σήμα. Δυστυχώς η υλοποίηση τεχνικών συμπίεσης που στηρίζονται στη συγκεκριμένη μέθοδο δεν υπήρξε ιδιαίτερα επιτυχής μέχρι τώρα, κυρίως λόγω της πολυπλοκότητας της ανάλυσης ενός τυχαίου σήματος σε στοιχειώδεις μορφοκλασματικές μορφές.

2.7.3 Συμμετρικές & Ασύμμετρες Μέθοδοι Συμπίεσης

Ανάλογα με την πολυπλοκότητα των διαδικασιών συμπίεσης και αποσυμπίεσης ενός σήματος οι μέθοδοι συμπίεσης χωρίζονται σε *συμμετρικές* (πολυπλοκότητα αποσυμπίεσης = πολυπλοκότητα συμπίεσης) και *ασύμμετρες* (πολυπλοκότητα αποσυμπίεσης \neq πολυπλοκότητα συμπίεσης). Για παράδειγμα οι τεχνικές διανυσματικής κωδικοποίησης είναι ασύμμετρες καθώς οι διαδικασίες αποσυμπίεσης που χρησιμοποιούν είναι σημαντικά απλούστερες από τις αντίστοιχες διαδικασίες συμπίεσης. Οι συμμετρικές διαδικασίες χρησιμοποιούνται κατά κόρο σε σύγχρονες εφαρμογές (π.χ. τηλεδιάσκεψη) ενώ οι ασύμμετρες χρησιμοποιούνται κυρίως σε ασύγχρονες εφαρμογές όπως το Video on Demand.