

4. ΟΠΤΙΚΑ ΣΗΜΑΤΑ

4.1 Φυσιολογικά Χαρακτηριστικά της Ανθρώπινης Όρασης

Σε αντίθεση με το αυτί, το ανθρώπινο μάτι δεν έχει σημαντική διακριτική ικανότητα. Τρία παραδείγματα μπορούν να μας βοηθήσουν να κατανοήσουμε τους περιορισμούς της όρασης μας:

- Είναι εξαιρετικά δύσκολο να αντιληφθούμε ως διαφορετικές δύο ή περισσότερες εικόνες που επικαλύπτονται.
- Περιοδικές παραμορφώσεις στη συχνότητα ή παύσεις μετάδοσης εικόνας που μπορεί να φτάσουν μέχρι και δευτερόλεπτα περνούν απαρατήρητες κατά τη διάρκεια παρακολούθησης ενός βίντεο.
- Σε αντίθεση με τη δυνατότητα ανίχνευσης ανεπαίσθητων αλλά ασυνήθιστων ήχων στο σπίτι μας πολλές φορές αλλαγές στη διαρύθμιση του σπιτιού (π.χ. ξεκρέμασμα ενός κάδρου) περνούν απαρατήρητες για αρκετό καιρό.

Κατά συνέπεια η σχεδίαση μιας εφαρμογής πολυμέσων θα πρέπει να λαβει υπ' όψη της το γεγονός ότι οι άνθρωποι είναι πιο ευαίσθητοι σε παραμορφώσεις ηχητικών παρά οπτικών σημάτων. Κατά συνέπεια εφαρμογές που συμπιέζουν ή μεταδίδουν σε δίκτυα ηχητικά ή/και οπτικά σήματα θα πρέπει να είναι περισσότερο προσεκτικοί ώστε οι παραμορφώσεις και καθυστερήσεις που εισάγονται από τη συμπίεση και μετάδοση ηχητικών σημάτων να είναι μέσα σε ανεκτά όρια.

4.2 Η αίσθηση του βάθους

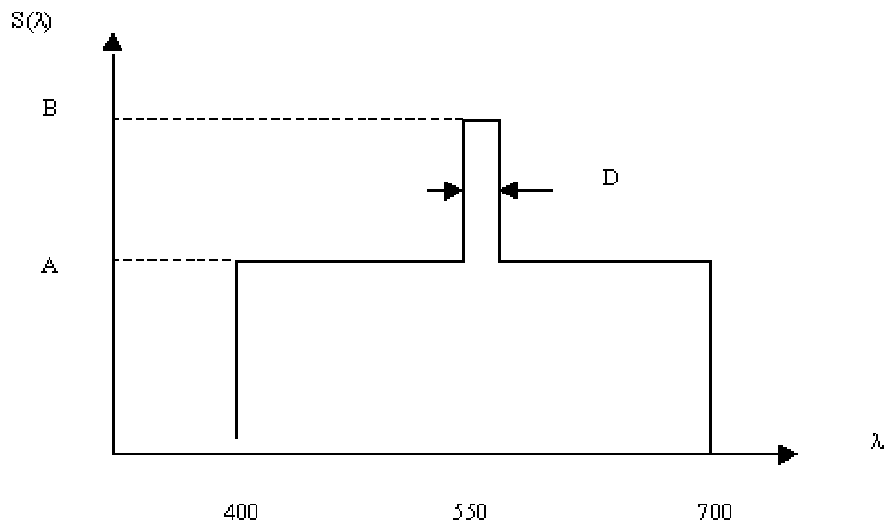
Η ανθρώπινη όραση χρησιμοποιεί δύο μηχανισμούς για να υπολογίζει τη σχετική απόσταση των αντικειμένων στο χώρο. Ο πρώτος μηχανισμός βασίζεται στη σύγκλιση μεταξύ των ευθειών παρατήρησης για το καθένα από τα δύο μάτια. Πιο συγκεκριμένα όταν παρατηρούμε ένα αντικείμενο από κοντά οι ευθείες παρατήρησης του αντικειμένου για το κάθε μάτι μας δεν είναι παράλληλες αλλά συγκλίνουν πάνω στο αντικείμενο. Όσο πιο κοντά βρίσκεται το αντικείμενο τόσο μεγαλύτερος βαθμός σύγκλισης είναι απαραίτητος για την παρατήρηση του. Όσο μεγαλύτερος όμως είναι και ο βαθμός σύγκλισης τόσο και μεγαλύτερη είναι η πίεση που εξασκούν οι μύες του ματιού για να το υποχρεώσουν να κινηθεί προς την επιθυμητή κατεύθυνση. Ο εγκέφαλος μετρά την πίεση των μυών του ματιού και έτσι λαμβάνει μια ένδειξη της απόστασης του αντικειμένου από το μάτι μας. Η δεύτερη ένδειξη που χρησιμοποιεί ο εγκέφαλος μας προέρχεται από την απόκλιση μεταξύ της εικόνας του κόσμου

που προέρχεται από το αριστερό και το δεξί μας μάτι. Πιο συγκεκριμένα όταν τα μάτια μας εστιάζουν στο ίδιο σημείο του χώρου η εικόνα που μεταδίδουν στον εγκέφαλο διαφέρει καθώς η εικόνα του δεξιού ματιού είναι μετατοπισμένη προς τα δεξιά και η εικόνα του αριστερού ματιού προς τα αριστερά. Όταν εστιάζουμε σε ένα μακρινό αντικείμενο η διαφορά μεταξύ των δύο αυτών εικόνων είναι αμελητέα ενώ όσο το αντικείμενο πλησιάζει προς το μάτι μας η διαφορά αυτή αυξάνεται. Ο εγκέφαλος συνδυάζει τη πληροφορία αυτή με την πίεση των μυών του ματιού για να δημιουργήσει την αίσθηση του βάθους.

4.3 Στοιχεία Θεωρίας Χρωμάτων

Το φως είναι μια μορφή ηλεκτρομαγνητικής ακτινοβολίας η οποία εκτείνεται σε ένα περιορισμένο εύρος μηκών κύματος. Οι ακτινοβολίες που γίνονται ορατές από το ανθρώπινο μάτι έχουν μήκος κύματος που κυμαίνεται μεταξύ 400 και 700 nm. Όταν ακτινοβολίες με μήκος κύματος μεταξύ αυτών των ορίων προσπίπτουν στο ανθρώπινο μάτι προκαλούν χρωματικές αισθήσεις ανάλογες με το μίγμα συχνοτήτων που περιέχουν. Πιο συγκεκριμένα το ανθρώπινο μάτι περιέχει δύο κατηγορίες κυττάρων: τους *κώνους* (*cones*) και τις *ράβδους* (*rods*). Το πλήθος των κώνων στο ανθρώπινο μάτι είναι περίπου 6-7 εκατομμύρια και είναι συγκεντρωμένα στο κέντρο του αμφιβληστροειδούς χιτώνα. Καθένας από τους κώνους αυτούς αντιδρά σε ένα περιορισμένο εύρος συχνοτήτων επιτρέποντας στο μάτι να διακρίνει μικρές χρωματικές λεπτομέρειες. Οι ράβδοι αντίθετα είναι πολύ περισσότεροι (75-100 εκατομμύρια) και είναι διασκορπισμένοι στην περιφέρεια του αμφιβληστροειδούς. Τα κύτταρα αυτά είναι ιδιαίτερα ευαίσθητα σε μικρά επίπεδα έντασης του φωτός επιτρέποντας στην περιφερειακή ανθρώπινη όραση να εντοπίζει κίνηση ή μορφές σε αρκετά χαμηλό φωτισμό, διευκολύνοντας έτσι την προστασία του ανθρώπινου είδους από πιθανούς εχθρούς.

Το φως που εκπέμπεται από διάφορες φωτεινές πηγές περιέχει ένα μίγμα από μήκη κύματος τα οποία περιγράφονται με την απεικόνιση της *φασματικής πυκνότητας* τους (*spectral density* $S(\lambda)$) σε κατάλληλα διαγράμματα. Τα διαγράμματα αυτά απεικονίζουν την ισχύ που έχει κάθε μήκος κύματος στο εκπεμπόμενο φως. Για παράδειγμα η Εικόνα 4.1 περιγράφει τη φασματική πυκνότητα μιάς πρασινωπής πηγής φωτός.



Εικόνα 4.1: Παράδειγμα φασματικής πυκνότητας φωτός.

Η φασματική πυκνότητα του συγκεκριμένου παραδείγματος παρατηρούμε ότι εμφανίζει ένα μέγιστο για μήκη κύματος περίπου 550 nm. Το μέγιστο αυτό ονομάζεται *δεσπόζων μήκος κύματος (dominant wavelength)*. Το μήκος κύματος στο οποίο παρουσιάζεται το μέγιστο καθορίζει και τη *χροιά (hue)* του χρώματος που εκπέμπεται από μια πηγή. Για παράδειγμα η συγκεκριμένη πηγή θα έχει πράσινο χρώμα. Επιπρόσθετα, στο συγκεκριμένο παράδειγμα εμφανίζεται και μια ποσότητα λευκού φωτός η οποία μειώνει τον κορεσμό του πράσινου και το ξανοίγει.

Η συνολική ισχύς του εκπεμπομένου φωτός μας δίνεται από το εμβαδό της καμπύλης της φασματικής πυκνότητας του. Η ισχύς αυτή ονομάζεται *φωτεινότητα (luminance)* και για το συγκεκριμένο παράδειγμα δίνεται από τον τύπο:

$$L = (B-A)D+AW$$

όπου $W = 700-400$.

Ο *κορεσμός (saturation)* ή *καθαρότητα (purity)* του φωτός μετρά το ποσοστό της φωτεινότητας που περιέχεται στο κύριο μήκος κύματος. Στην περίπτωση μας υπολογίζεται από τον τύπο:

$$S = 100\% \times (B-A)D/L$$

Όταν δύο πηγές φωτός διαφέρουν μόνο στη χροιά τους τότε το ανθρώπινο μάτι μπορεί να ξεχωρίσει μεταξύ περίπου 128 διαφορετικών τέτοιων χρωμάτων. Όταν δύο πηγές φωτός διαφέρουν μόνο στον κορεσμό τους τότε το μάτι μπορεί να ξεχωρίσει περίπου 20 διαφορετικά επίπεδα κορεσμού και αυτό εξαρτάται από τη συγκεκριμένη χροιά των χρωμάτων.

4.3.1 Σύνθεση Χρωμάτων

Ένας από τους πιο διαδεδομένους πειραματικούς τρόπους σύνθεσης χρωμάτων στηρίζεται στην υπέρθεση ενός συνόλου από χρώματα αναφοράς για την παραγωγή του επιθυμητού χρώματος. Στην περίπτωση αυτή το επιθυμητό χρώμα το οποίο διαθέτει μια ορισμένη φασματική πυκνότητα $S(\lambda)$ προβάλλεται στη μία πλευρά μίας οθόνης ενώ στο υπόλοιπο μέρος της οθόνης προβάλλεται το χρώμα που προκύπτει από την υπέρθεση τριών φωτεινών πηγών με φασματικές πυκνότητες $A(\lambda)$, $B(\lambda)$ και $C(\lambda)$. Ο παρατηρητής μπορεί να ρυθμίσει τη στάθμη (α , β ή ϵ) της έντασης του φωτός που παράγεται από κάθε μία από τις πηγές ώστε το χρώμα $T(\lambda)$ που προκύπτει από την υπέρθεση των πηγών και που δίνεται από τον τύπο:

$$T(\lambda) = \alpha A(\lambda) + \beta B(\lambda) + \epsilon C(\lambda)$$

να είναι το ίδιο με το χρώμα που προβάλλεται στην άλλη άκρη της οθόνης παρόλο που οι φασματικές πυκνότητες των δύο χρωμάτων είναι διαφορετικές ($S(\lambda) \neq T(\lambda)$). Στην περίπτωση αυτή μπορούμε να πούμε ότι το χρώμα με φασματική πυκνότητα $S(\lambda)$ μπορεί να παραχθεί από την υπέρθεση των τριών χρωμάτων $A(\lambda)$, $B(\lambda)$ και $C(\lambda)$ με τις εντάσεις που βρέθηκαν από το πείραμα.

Δύο χρώματα τα οποία όταν προστίθενται σε ίσες ποσότητες προκύπτει το άσπρο χρώμα (δηλ. το χρώμα με τον ελάχιστο κορεσμό) ονομάζονται *συμπληρωματικά* (*complementary*). Παραδείγματα τέτοιων ζευγών χρωμάτων είναι το μπλε-κίτρινο, κόκκινο-κυανό, πράσινο-μωβ κλπ.

Πειράματα όπως το παραπάνω αποδεικνύουν την ύπαρξη μιας άλγεβρας χρωμάτων με συγκεκριμένες ιδιότητες. Πιο συγκεκριμένα, η θεωρία χρωμάτων καθορίζει ότι αν σε δυο χρώματα που προκαλούν την ίδια χρωματική αίσθηση παρόλο που μπορεί να έχουν διαφορετικές φασματικές πυκνότητες προσθέσουμε την ίδια ποσότητα του ίδιου χρώματος τότε τα χρώματα που θα παραχθούν θα συνεχίζουν να μας προκαλούν ταυτόσημες

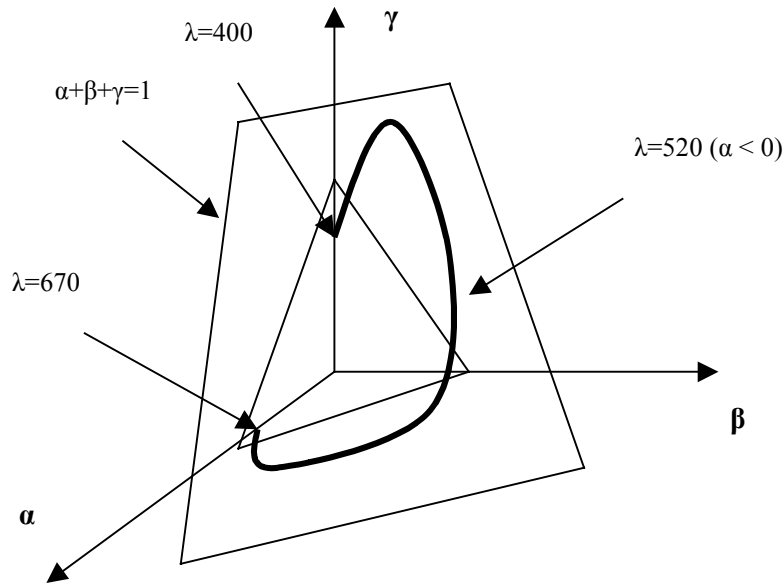
χρωματικές αισθήσεις. Επίσης η άλγεβρα καθορίζει ότι η αίσθηση οποιουδήποτε χρώματος είναι τρισδιάστατη, δηλαδή κάθε χρώμα μπορεί να αναπαραχθεί με την ανάμιξη σε κατάλληλες αναλογίες τριών *πρωταρχικών* (*primary*) πηγών φωτός, πηγών δηλαδή που η φασματική τους πυκνότητα λαμβάνει μη μηδενική τιμή για ένα μόνο μήκος κύματος. Κατά σύμβαση στην πλειονότητα των συστημάτων πολυμέσων η τριάδα των χρωμάτων που χρησιμοποιείται για το σκοπό αυτό αποτελείται από το κόκκινο (R) ($\lambda = 700 \text{ nm}$) το πράσινο (G) ($\lambda = 546,1 \text{ nm}$) και το μπλε (B) ($\lambda = 435,8 \text{ nm}$) αφού σε αυτά τα χρώματα είναι περισσότεροι ευαίσθητοι οι κώνοι που περιέχονται στο μάτι μας. Με άλλα λόγια οποιδήποτε χρώμα C μοναδιαίας λαμπρότητας μπορεί να προκύψει σαν ένας κυρτός συνδυασμός των παραπάνω χρωμάτων, δηλ:

$$C = \alpha R + \beta G + \gamma B \quad (1)$$

όπου

$$\alpha + \beta + \gamma = 1 \quad (2)$$

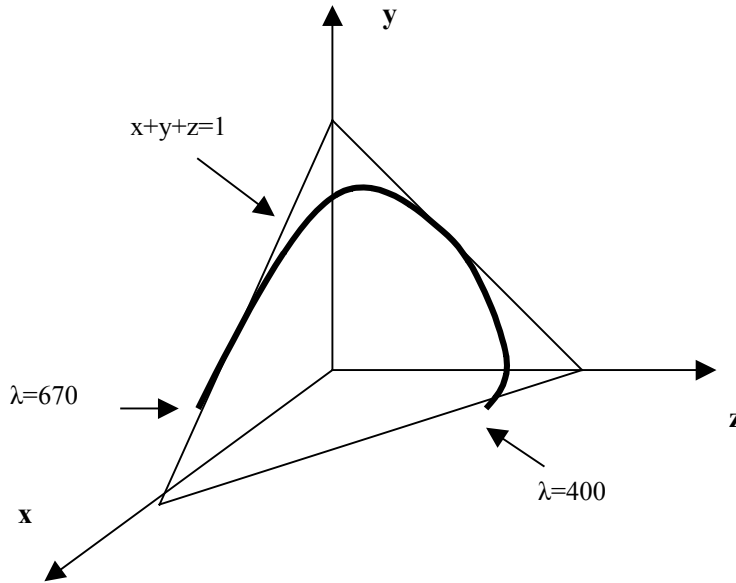
Αμα διπλασιάσουμε τους συντελεστές αυτούς θα προκύψει το ίδιο χρώμα με την προηγούμενη περίπτωση με διπλάσια όμως λαμπρότητα κ.ο.κ. Επομένως για την περιγραφή των χρωμάτων που μπορούν να προκύψουν από ένα τέτοιο γραμμικό συνδυασμό χρωμάτων μπορούμε να περιοριστούμε σε χρώματα μοναδιαίας λαμπρότητας. Στην περίπτωση αυτή, όλα τα χρώματα που μπορούν να παραχθούν θα παριστάνονται ως σημεία στο επίπεδο που ορίζεται από τη συνθήκη (2). Η Εικόνα 4.1 περιγράφει τη μορφή που θα έχει ένα τέτοιο επίπεδο και την καμπύλη που θα συνδέει τα καθαρά χρώματα στην περίπτωση αυτή. Κάθε σημείο στην καμπύλη αυτή θα παριστάνει τους συντελεστές α , β , γ που έχουν υπολογιστεί για καθαρά χρώματα μοναδιαίας έντασης χρησιμοποιώντας την πειραματική διαδικασία που περιγράψαμε στην αρχή της ενότητας.



Εικόνα 4.1: Καμπύλη παραγωγής καθαρών χρωμάτων από τρεις πρωταρχικές πηγές φωτός (R, G, B).

Το αξιοπερίεργο στην περίπτωση μας είναι ότι σε ορισμένες περιπτώσεις κάποιο από τα α , β ή γ μπορεί να είναι και αρνητικό (βλ. Στην Εικόνα 4.1 το σημείο που αντιστοιχεί σε καθαρό χρώμα μήκους κύματος 500 nm). Αυτό κυρίως συμβαίνει στην περίπτωση χρωμάτων που είναι υπερβολικά καθαρά καθώς η ανάμιξη δύο πρωταρχικών χρωμάτων σε οποιαδήποτε αναλογία παράγει ένα χρώμα που είναι λιγότερο κορεσμένο από τα αρχικά. Επειδή η υλοποίηση της αφαίρεσης ενός χρώματος δεν είναι δυνατή στα συστήματα προβολής που χρησιμοποιούν καθοδικούς σωλήνες (π.χ. τηλεοράσεις ή οθόνες υπολογιστών) κατά συνέπεια οι συσκευές αυτές δεν είναι σε θέση να αποδώσουν χρώματα για τα οποία τουλάχιστον ένας από τους συντελεστές α , β ή γ είναι αρνητικός.

Για την παράσταση όλων των χρωμάτων με κοινό τρόπο χρησιμοποιείται το πρότυπο CIE το οποίο εισήχθη το 1931 από την Commission Internationale de l' Eclairage (CIE). Τα πρωταρχικά χρώματα στο πρότυπο αυτό εκφράζονται σε σχέση με ένα σύνολο τριών πρωταρχικών χρωμάτων X, Y, Z τα οποία δεν υπάρχουν στη φύση αλλά επιλέχτηκαν έτσι ώστε οι συντελεστές που καθορίζουν την ένταση των υπολοίπων χρωμάτων να είναι πάντα θετικοί. Η Εικόνα 4.2 περιγράφει τη μορφή της καμπύλης καθαρών χρωμάτων στο πρότυπο CIE.



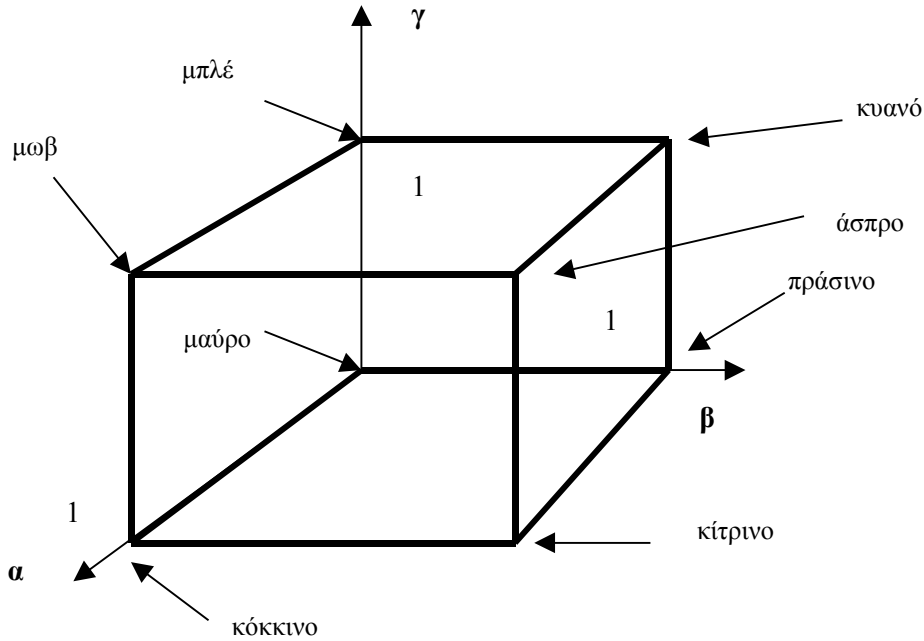
Εικόνα 4.2: Το πρότυπο CIE.

4.3.2 Χρωματικά Μοντέλα

Το πρότυπο CIE παρόλο που είναι ακριβές δεν είναι αρκετά χρήσιμο για την αναπαραγωγή χρωμάτων σε υπολογιστικά συστήματα όπου τα πρωταρχικά χρώματα είναι το κόκκινο, το πράσινο και το μπλε αλλά και σε ορισμένες γραφικές τέχνες όπου η περιγραφή των χρωμάτων γίνεται με βάση την απόχρωση (hue), τον κορεσμό (saturation) και τη φωτεινότητα (lightness). Για να εξυπηρετήσουν εναλλακτικά συστήματα περιγραφών χρησιμοποιούνται μια σειρά από *χρωματικά μοντέλα (color models)* τέσσερα από τα οποία (RGB, CMYK, HLS και YIQ) θα εξετάσουμε στη συνέχεια.

4.3.2.1 RGB & CMYK

Το μοντέλο RGB παριστάνει κάθε χρώμα ως γραμμικό συνδυασμό τριών κατάλληλα επιλεγμένων πρωταρχικών πηγών κόκκινου, πράσινου και μπλε χρώματος όπως στη σχέση (1). Αν οι συντελεστές α , β και γ λαμβάνουν τιμές από το 0 έως και το 1 τότε όλα τα χρώματα που ορίζονται στο σύστημα αυτό θα περιγράφονται από τον κύβο της Εικόνας 4.3.



Εικόνα 4.3: Το χρωματικό μοντέλο RGB.

Το μοντέλο RGB αποτελεί ένα παράδειγμα *αρθροιστικού* (*additive*) χρωματικού μοντέλου. Κάθε χρώμα στο μοντέλο αυτό παράγεται από την υπέρθεση τριών πρωταρχικών χρωμάτων.

Αντίθετα, σε ένα *αφαιρετικό* (*subtractive*) χρωματικό μοντέλο κάθε χρώμα (έστω B) εκφράζεται και πάλι ως ένας γραμμικός συνδυασμός τριών χρωμάτων αλλά κάθε ένας από τους τρεις συντελεστές καθορίζει την ποσότητα ενός συγκεκριμένου χρώματος (του συμπληρωματικού του αντίστοιχου πρωταρχικού χρώματος) που θα πρέπει να αφαιρεθεί από το άσπρο χρώμα ώστε να προκύψει το χρώμα B. Για παράδειγμα το αφαιρετικό χρωματικό μοντέλο CMY χρησιμοποιεί τρία χρώματα (το κυανό (cyan), το μωβ (magenta) και το κίτρινο (yellow)) και οι τρεις συντελεστές α , β , γ καθορίζουν πόσο από το συμπληρωματικό του καθενός από αυτά τα χρώματα (κόκκινο, πράσινο και μπλε αντίστοιχα) πρέπει να αφαιρέσουμε από το άσπρο ώστε να προκύψει το επιθυμητό χρώμα. Προφανώς μεταξύ των συντελεστών στα χρωματικά μοντέλα RGB και CMY ισχύει η σχέση:

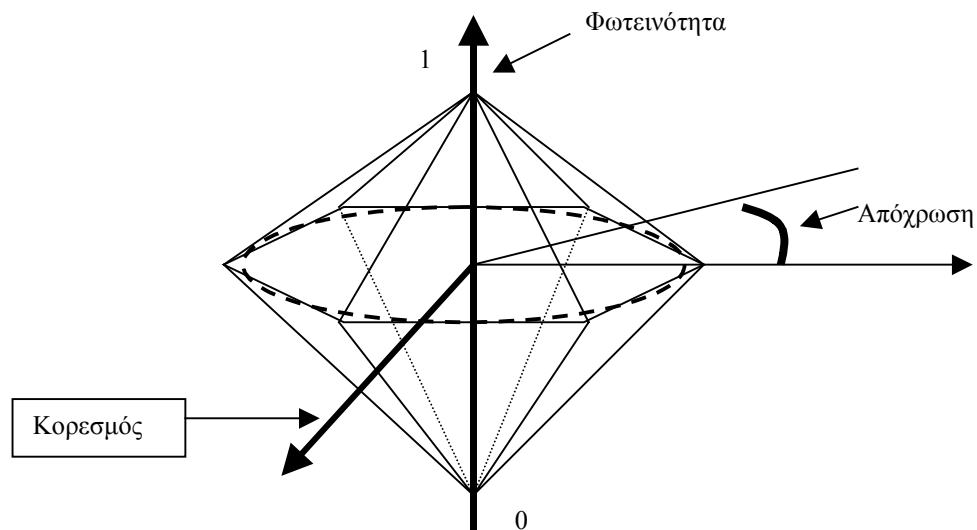
$$(\alpha, \beta, \gamma)_{\text{RGB}} = (1, 1, 1) - (\alpha, \beta, \gamma)_{\text{CMY}}$$

Ενώ το χρωματικό μοντέλο RGB χρησιμοποιείται για την παραγωγή χρωμάτων στη οθόνη του υπολογιστή, το σύστημα CMY χρησιμοποιείται κυρίως για την εκτύπωση χρωμάτων. Μια παραλλαγή του CMY αποτελεί και το CMYK στο οποίο εκτός από τα χρώματα που προαναφέραμε χρησιμοποιείται και το μαύρο. Στο σύστημα αυτό η τιμή του μαύρου

υπολογίζεται ως ο ελάχιστος συντελεστής για τα υπόλοιπα τρία χρώματα, ενώ οι τελικές τιμές των άλλων τριών χρωμάτων προκύπτουν από την αφαίρεση του μαύρου από τις αρχικές τους τιμές. Σε μία έγχρωμη εκτύπωση το χρώμα προέρχεται από φως το οποίο ανακλάται από μελάνι που αποτελείται από ένα συνδυασμό τριών χρωμάτων: κυανού (cyan), μωβ (magenta) και κίτρινου (yellow). Όταν αναμιχθούν και τα 3 χρώματα του κώδικα μας δίνουν μαύρο. Επειδή όμως δεν είναι δυνατό να παραχθεί ένα 'αληθινό' μαύρο από την ανάμιξη αυτή, γι' αυτό και ο χρωματικός αυτός κώδικας περιέχει και ένα μαύρο μελάνι (το K στο CMYK). Τα διάφορα χρώματα σε μια εκτύπωση παράγονται από την ανάμιξη ορισμένων ή όλων από τα χρώματα του κώδικα. Σε μια εκτύπωση δεν είναι δυνατό να μεταβληθεί η ένταση των χρωμάτων. Ο κορεσμός επιτυγχάνεται με τη μεταβολή του μεγέθους κάθε κουκίδας, έτσι ώστε να επιτραπεί σε λιγότερο ή περισσότερο άσπρο χαρτί να φαίνεται.

4.3.2.2 HLS ή HSV

Στο μοντέλο HLS ή HSV κάθε χρώμα εκφράζεται συναρτήσει τριών συντελεστών που καθορίζουν την απόχρωση (hue), τη φωτεινότητα (luminance) και τον κορεσμό (saturation) που διαθέτει. Το μοντέλο προκύπτει αν παραμορφώσουμε τον κύβο του RGB σε ένα κώνο που διαθέτει. Το μοντέλο προκύπτει αν παραμορφώσουμε τον κύβο του RGB σε ένα κώνο και προκύπτει να κοιτάξουμε κατά μήκος της διαγωνίου του κύβου του RGB από το σημείο (1, 1, 1) προς το σημείο (0, 0, 0) (βλ. Εικόνα 4.4).



Εικόνα 4.4: Το χρωματικό μοντέλο HLS.

Σε ένα τέτοιο σύστημα η απόχρωση ενός χρώματος καθορίζεται από τη γωνία που σχηματίζει με την ευθεία που συνδέει το σημείο (0.5, 0.5, 0.5) με το κόκκινο στον κύβο του RGB, η

φωτεινότητα καθορίζεται από τη θέση ενός χρώματος στη διαγώνιο ενώ ο κορεσμός καθορίζεται από την απόσταση του σημείου από τον άξονα της φωτεινότητας.

4.3.2.3 YUV & YIQ

Ο κύριος σκοπός της χρησιμοποίησης του χρωματικού μοντέλου YUV και των διαφόρων παραλλαγών του (YIQ, YCbCr-601, YCbCr-709, PhotoYCC) προέρχεται από την ανάγκη εξοικονόμησης εύρους ζώνης κατά τη μετάδοση εικόνων τηλεόρασης ή βίντεο. Η συγκεκριμένη οικογένεια χρωματικών μοντέλων εκμεταλλεύεται το γεγονός ότι τι ανθρώπινο μάτι είναι περισσότερο ευαίσθητο σε μεταβολές λαμπρότητας από ότι σε χρωματικές μεταβολές σε μια ακολουθία εικόνων. Τα συγκεκριμένα χρωματικά μοντέλα αποθηκεύουν την πληροφορία που περιγράφει τη λαμπρότητα (Y) ενός σημείου ξεχωριστά από τα χρωματικά του χαρακτηριστικά (UV ή IQ ανάλογα με το συγκεκριμένο μοντέλο). Επιπλέον ο χώρος που διατίθεται για την αποθήκευση της πληροφορίας που αφορά τη φωτεινότητα είναι μεγαλύτερος από το χώρο που διατίθεται για την αποθήκευση των χρωματικών χαρακτηριστικών κάθε σημείου. Με αυτό τον τρόπο μειώνονται οι απαιτήσεις σε εύρος ζώνης που είναι απαραίτητο για τη μετάδοση των χαρακτηριστικών μίας εικόνας. Η σχέση μεταξύ των συντελεστών YUV και διορθωμένου κατά γάμμα RGB περιγράφεται από τους παρακάτω τύπους:

$$Y = 0.299R + 0.587G + 0.114B$$

$$I = 0.596R - 0.274G - 0.322B$$

$$Q = 0.211R - 0.523G + 0.312B$$

4.3.3 Αρχές λειτουργίας οθονών υπολογιστών

Στην τρέχουσα ενότητα θα εξετάσουμε τα λειτουργικά χαρακτηριστικά των οθονών *καθοδικού σωλήνα* (*Cathode Ray Tube* ή *CRT*), οι οποίες αποτελούν και την πλειοψηφία των οθονών που χρησιμοποιούνται στα σύγχρονα υπολογιστικά συστήματα. Οι συσκευές αυτές αποτελούνται από επιστρώσεις ενώσεων του φωσφόρου οι οποίες ακτινοβολούν όταν βομβαρδιστούν με δέσμες ηλεκτρονίων. Η ένταση του φωτός που παράγουν οι επιστρώσεις φωσφόρου είναι ανάλογη με την ένταση της δέσμης με την οποία βομβαρδίζονται. Συνήθως οι έγχρωμες οθόνες αποτελούνται από επιστρώσεις τριών ειδών φωσφόρου που ακτινοβολούν όταν πέσει πάνω τους κόκκινο, μπλέ ή πράσινο φως αντίστοιχα. Οι περισσότερες έγχρωμες CRT οθόνες χρησιμοποιούν τρεις δέσμες ηλεκτρονίων τις οποίες και κατευθύνουν στο σημείο της οθόνης το οποίο θέλουν να ενεργοποιήσουν με τη χρήση κατάλληλων ηλεκτρομαγνητικών πεδίων. Κάθε μία από τις ακτίνες αυτές διατρέχει την επιφάνεια της

οθόνης μία γραμμή τη φορά. Μια τέτοια γραμμή ονομάζεται *γραμμή σάρωσης (scan line)*. Κάθε φορά που η ακτίνα φτάνει στο τέλος μιάς γραμμής τότε μετακινείται στην αρχή της επόμενης. Η μετακίνηση αυτή ονομάζεται *horizontal retrace*. Κάθε φορά που η ακτίνα φτάσει στο τέλος της τελευταίας γραμμής της οθόνης τότε μετακινείται στην αρχή της πρώτης γραμμής. Η μετακίνηση αυτή ονομάζεται *vertical retrace*.

Η *διακριτικότητα (resolution)* μιάς οθόνης μετρά τον αριθμό των ψηφίδων (pixels) από τα οποία αποτελείται κάθε οριζόντια και κάθε κάθετη γραμμή μιάς οθόνης (π.χ. 1024 x 768). Ο χρόνος στον οποίο η δέσμη ηλεκτρονίων διατρέχει την οθόνη εκφράζεται με τον ρυθμό ανανέωσης της οθόνης (*refresh rate*) και μετριέται σε μονάδες συχνότητας. Π.χ. ένα ρυθμός ανανέωσης 70 Hz δηλώνει ότι η οθόνη επανασχεδιάζεται 70 φορές το δευτερόλεπτο. Ο ρυθμός σάρωσης μιάς γραμμής της οθόνης ονομάζεται *scan rate* και μετριέται και αυτός σε μονάδες συχνότητας. Προφανώς ισχύει ότι:

$$\text{Refresh rate} = \text{Scan rate} * n$$

όπου n ο αριθμός των γραμμών της οθόνης.

Η ικανότητα μιάς οθόνης να προβάλλει ένα σήμα βίντεο χωρίς να αλλοιώνει την ποιότητα του περιγράφεται από το *εύρος ζώνης βίντεο (video bandwidth)* της οθόνης το οποίο μετριέται (συνήθως) σε μονάδες συχνότητας. Ένα εύρος ζώνης 100 MHz σημαίνει ότι το σύστημα θα πρέπει να προβάλλει μια ψηφίδα (pixel) κάθε 10 ns. Κατά συνέπεια ο μετατροπέας ψηφιακού σε αναλογικό σήμα που συνδέεται με την οθόνη θα πρέπει να τροφοδοτεί την οθόνη με τιμές ψηφίδων με αυτή τη συχνότητα. Το εύρος ζώνης θα πρέπει να είναι τουλάχιστον ίσο με το γινόμενο της διακριτικότητας της οθόνης επί το ρυθμό ανανέωσης της οθόνης, δηλαδή:

$$\text{Video bandwidth} = \text{Refresh rate} * \text{Resolution}$$

όπου η διακριτικότητα της οθόνης δίνεται σαν το γινόμενο του αριθμού των ψηφίδων σε κάθε οριζόντια γραμμή της οθόνης επί τον αριθμό των ψηφίδων σε κάθε κάθετη γραμμή της οθόνης. Μια οθόνη που μπορεί να συγχρονίζει αυτόματα το ρυθμό σάρωσης με το εύρος ζώνης του σήματος βίντεο ονομάζεται *multiscanning* ή *multifrequency*. Αν το εύρος ζώνης της οθόνης είναι μικρότερο από το γινόμενο που αναφέραμε τότε η οθόνη θα φαίνεται θαμπή.

Ο λόγος μεταξύ του αριθμού των ψηφίδων σε κάθε οριζόντια γραμμή προς τον αριθμό των ψηφίδων σε κάθε κάθετη γραμμή της οθόνης (υποθέτοντας ότι χρησιμοποιούνται τετράγωνα ψηφίδες) ονομάζεται *aspect ratio*. Συνήθως στο βίντεο χρησιμοποιούνται διακριτικότητες

όπως 160 x 120, 320 x 240 ή 640 x 480 ή οποίες διατηρούν το aspect ratio 4:3 που χρησιμοποιείται στην τηλεόραση. Δύο διακριτικότητες που δεν ακολουθούν αυτόν τον λόγο είναι η διακριτικότητα *CIF* και *QCIF*. Η *CIF* (*Common Intermediate Format*) έχει μέγεθος 352 x 288 ενώ η *QCIF* (*Quarter Common Intermediate Format*) έχει μέγεθος 176 x 144. Παρόλο που κάθε μέγεθος είναι θεωρητικά πιθανό ορισμένες παραδοχές που γίνονται κυρίως από τα συστήματα συμπίεσης βίντεο μας οδηγούν να χρησιμοποιούμε μία από τις διακριτικότητες που αναφέραμε κατά τη διάρκεια της προβολής.

Η απόσταση μεταξύ δύο επιστρώσεων φωσφόρου σε μία οθόνη ονομάζεται *dot-pitch* και μετριέται σε χιλιοστά. Όσο πιο μικρή είναι αυτή η απόσταση τόσο βελτιώνεται η ποιότητα της εικόνας. Η εξάρτηση μεταξύ της καθαρότητας και της διακριτικότητας από το μέγεθος αυτό εξηγεί γιατί οι μονόχρωμες οθόνες έχουν καθαρότερη εικόνα από τις έγχρωμες αφού περιέχουν μία μόνο επίστρωση φωσφόρου.

Ανάλογα με τον τρόπο που γίνεται η ανανέωση της εικόνας υπάρχουν δύο τύποι οθονών: οι *συμπλεκόμενες* (*interlaced*) και οι *μη συμπλεκόμενες* (*noninterlaced*). Στις συμπλεκόμενες οθόνες κάθε σάρωση επισκέπτεται μόνο τις μισές γραμμές της οθόνης (εναλλακτικά τις γραμμές περιττής ή άρτιας τάξης). Αντίθετα στις μη συμπλεκόμενες οθόνες κάθε σάρωση αφορά όλες τις γραμμές της οθόνης.

Στις συμπλεκόμενες οθόνες υπάρχει η περίπτωση η δέσμη ηλεκτρονίων να σαρώνει την οθόνη αρκετά αργά, προκαλώντας έτσι την εξασθένιση της ακτινοβολίας που παράγεται από τις επιστρώσεις φωσφόρου. Σαν αποτέλεσμα η εικόνα στην οθόνη φαίνεται να αναβοσβήνει. Το φαινόμενο αυτό ονομάζεται *τρεμόσβημα* (*flickering*) και είναι αρκετά ενοχλητικό καθώς προκαλεί πονοκέφαλο και κουράζει τα μάτια. Το φαινόμενο γίνεται πιο έντονο όσο οι διαστάσεις της οθόνης μεγαλώνουν καθώς μια μεγάλη οθόνη καταλαμβάνει μεγαλύτερο μέρος από το οπτικό πεδίο του παρατηρητή και εμπλέκει και την περιφερειακή όραση του που είναι πιο ευαίσθητη στην αντίθεση κίνησης στις παρυφές του οπτικού πεδίου. Για την αποφυγή του τρεμοσβήματος οι συμπλεκόμενες οθόνες χρησιμοποιούν ενώσεις φωσφορου με μεγάλο βαθμό υστέρησης. Σαν αποτέλεσμα τα χρώματα στις οθόνες αυτές δεν είναι αρκετά φωτεινά (ιδίως το λευκό το οποίο μπορεί να φαίνεται σα γκρίζο ή μπλέ). Επίσης οι οθόνες αυτές δεν έχουν μεγάλη ένταση φωτός. Τέλος οι οθόνες αυτές έχουν προβλήματα στην εμφάνιση διαδοχικών λευκών γραμμών σε μαύρο φόντο. Στην περίπτωση αυτή επειδή οι δύο αυτές γραμμές σαρώνονται σε διαφορετικούς χρόνους φαίνονται σαν να μετακινούνται η μία στην άλλη. Από την άλλη οι συμπλεκόμενες οθόνες είναι αρκετά πιο φτηνές μιάς και έχουν μικρότερο ρυθμό ανανέωσης.

Η αναπαραγωγή έγχρωμων εικόνων σε μια οθόνη υπολογιστή θα πρέπει να λαμβάνει υπ' όψη της τα χαρακτηριστικά των οθονών. Τα πιο σημαντικά προβλήματα που μπορεί να παρουσιαστούν κατά τη συγκεκριμένη διαδικασία συνοψίζονται στα ακόλουθα τρία σημεία:

1. Το χρώμα σε μια οθόνη δεν παράγεται από την υπέρθεση τριών πηγών φωτός αλλά από την ακτινοβολία που παράγεται κατά τον φωσφορισμό τριών διαφορετικών σημείων στην οθόνη τα οποία βρίσκονται σε πολύ μικρή απόσταση μεταξύ τους. Η μίξη του χρώματος που εκπέμπουν οι τρεις διαφορετικές πηγές γίνεται από το ανθρώπινο μάτι. Κατά συνέπεια αρκετά καθαρά χρώματα δεν προβάλλονται στην φωτεινότητα που τους αναλογεί. Π.χ. μια περιοχή καθαρού κόκκινου αποτελείται από 1/3 κόκκινο και 2/3 μαύρο. Το πρόβλημα αυτό γίνεται περισσότερο αντιληπτό αν θελήσουμε να φωτογραφίσουμε την εικόνα που προβάλλεται σε μία οθόνη. Στην περίπτωση αυτή η απόδοση των καθαρών χρωμάτων στη φωτογραφία θα είναι αρκετά φτωγή.
2. Τα χαρακτηριστικά των επιστρώσεων φωσφόρου που χρησιμοποιούνται σε κάθε οθόνη διαφέρουν μεταξύ τους. Κατά συνέπεια το σύνολο των χρωμάτων (*color gamut*) που μπορεί να αναπαραχθεί διαφέρει από οθόνη σε οθόνη.
3. Η σχέση μεταξύ των RGB συντελεστών και της έντασης του φωτός που παράγεται από μία οθόνη είναι μη γραμμική. Η συγκεκριμένη αναντιστοιχία διορθώνεται με την εφαρμογή ενός μη γραμμικού μετασχηματισμού στους συντελεστές RGB ο οποίος ονομάζεται *διόρθωση γάμμα* (*gamma correction*) και αναλύεται στη συνέχεια.

4.3.3.1 Διόρθωση Γάμμα

Σε αρκετές περιπτώσεις η εισαγωγή και επεξεργασία ψηφιακών εικόνων υποθέτει ότι υπάρχει μια γραμμική σχέση μεταξύ της τιμής των ψηφίδων μίας εικόνας και της έντασης του φωτός με την οποία οι ψηφίδες αυτές προβάλλονται στην οθόνη. Δυστυχώς στην πλειονότητα των περιπτώσεων η υπόθεση αυτή δεν ισχύει καθώς υπάρχει μία μη γραμμική σχέση μεταξύ των RGB συντελεστών μίας ψηφίδας και της έντασης του φωτός στην οποία αυτοί μεταφράζονται στην οθόνη. Η *διόρθωση γάμμα* (*gamma correction*) εφαρμόζεται στις τιμές των ψηφίδων μίας εικόνας ώστε να αποκαταστήσει μια γραμμική σχέση μεταξύ της τιμής μιας ψηφίδας και της έντασης του φωτός στην οποία αυτή αντιστοιχεί.

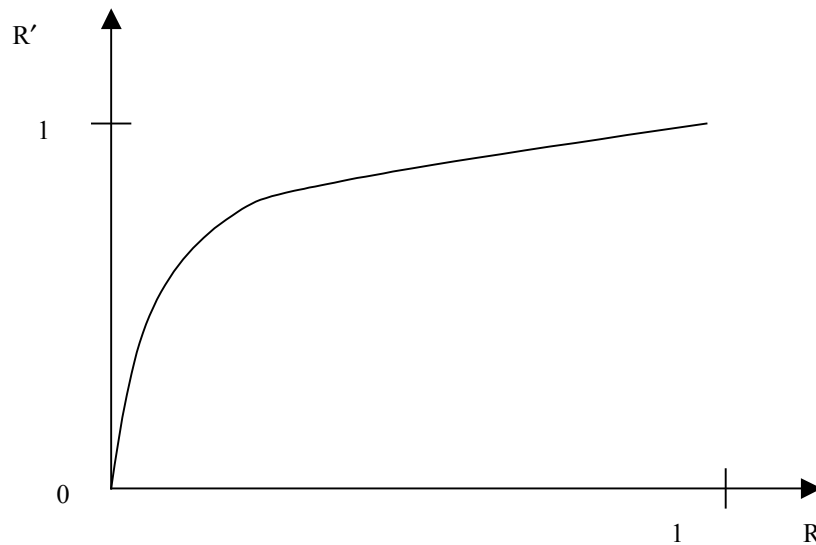
Πιο συγκεκριμένα, η ένταση R_o φωτός που παράγεται από μια οθόνη υπολογιστή όταν τροφοδοτείται με ένα RGB συντελεστή που περιγράφει την ένταση R λ.χ. της κόκκινης πρωταρχικής συνιστώσας δίνεται από τον τύπο:

$$R_o = KR^{\gamma}$$

όπου η σταθερά γ συνήθως λαμβάνει τιμές από 1.6 ως 2.6 και εξαρτάται από τον ιδιαίτερο τρόπο λειτουργίας κάθε οθόνης. Για να μπορέσουμε να μετατρέψουμε σε γραμμική τη συγκεκριμένη σχέση θα πρέπει να αντιστοιχίσουμε τη τιμή R σε μια καινούργια τιμή R' η οποία θα δίνεται από τον τύπο:

$$R' = kR^{1/\gamma}$$

Η αντιστοίχιση αυτή ονομάζεται διόρθωση γάμμα (βλ. Εικόνα 4.5) και θα πρέπει να γίνει και για τις υπόλοιπες δύο χρωματικές συνιστώσεις του RGB. Το αντίτιμο που πληρώνουμε από την εφαρμογή ενός τέτοιου μετασχηματισμού είναι ο περιορισμός του δυναμικού εύρους της εικόνας που προκύπτει. Για παράδειγμα αν η διόρθωση γάμμα που θα εφαρμόσουμε αντιστοιχίζει τη τιμή 0 στο 0 και την τιμή 255 στο 255 τότε το δυναμικό εύρος της εικόνας που θα προκύψει περιορίζεται από το 255 στο 167.



Εικόνα 4.4: Διόρθωση Γάμμα.

4.3 Μέθοδοι Επεξεργασίας Εικόνας

Οι μέθοδοι επεξεργασίας εικόνων που θα μας απασχολήσουν υποδιαιρούνται σε τέσσερις κύριες κατηγορίες: στις *σημειακές μεθόδους*, στα *φίλτρα γειτνίασης*, στα *φασματικά φίλτρα* και στις *γεωμετρικές μεθόδους*. Οι σημειακές μέθοδοι υπολογίζουν μια καινούργια τιμή για κάθε ψηφίδα στην εικόνα με βάση αποκλειστικά την τρέχουσα τιμή της ψηφίδα. Τα φίλτρα γειτνίασης υπολογίζουν μια καινούργια τιμή για κάθε ψηφίδα στην εικόνα με βάση την

τρέχουσα τιμή της ψηφίδας και την τιμή των «γειτονικών» σε αυτή ψηφίδων. Τα φασματικά φίλτρα υπολογίζουν και μεταβάλλουν τις συχνότητες των περιοδικών συναρτήσεων από τις οποίες απαρτίζεται μία εικόνα. Τέλος οι γεωμετρικές μέθοδοι εφαρμόζουν μια σειρά από γεωμετρικούς μετασχηματισμούς σε κάθε ψηφίδα με σκοπό να μεταβάλλουν τη θέση της στην εικόνα.

Οι εικόνες στις οποίες αναφερόμαστε στο υπόλοιπο της ενότητας υποθέτουμε ότι περιέχουν μόνο τόνους του γκρι (*gray-scale images*). Σε αυτό το είδος εικόνων η τιμή κάθε ψηφίδας περιγράφει την απόχρωση του γκρι στην αντίστοιχη θέση της εικόνας. Η τιμή αυτή περιέχεται σε ένα εύρος τιμών από 0 (μαύρο) έως M (άσπρο) όπου το M συνήθως ισούται με 255.

4.3.1 Σημειακές Μέθοδοι

Οι σημειακές μέθοδοι ταξινομούνται στις ακόλουθες κατηγορίες:

Εκτέλεση αριθμητικών και λογικών πράξεων. Οι συγκεκριμένες πράξεις μπορεί να είναι μοναδικές ή δυαδικές ανάλογα με το αν εφαρμόζονται στις ψηφίδες μίας ή δύο εικόνων. Οι Πίνακες 4.1 και 4.2 περιγράφουν τις πιθανές μοναδικές και δυαδικές πράξεις μεταξύ εικόνων.. Στον Πίνακα 4.1 τα σύμβολα $s(x, y)$ και $d(x, y)$ αναφέρονται στις ψηφίδες των εικόνων εισόδου και εξόδου αντίστοιχα στη θέση (x, y) , ενώ K είναι μία σταθερά.

Περιγραφή	Πράξη
Πρόσθεση/Αφαίρεση με σταθερά	$d(x, y) = s(x, y) \pm K$
Πολλαπλασιασμός/Διαίρεση με σταθερά	$d(x, y) = s(x, y) * K$ ή $d(x, y) = s(x, y) / K$ ή $d(x, y) = K / s(x, y)$
Διάζευξη	$d(x, y) = s(x, y) \vee K$
Σύζευξη	$d(x, y) = s(x, y) \wedge K$
Αποκλειστική Διάζευξη	$d(x, y) = s(x, y) \underline{\vee} K$
Απόλυτη τιμή	$d(x, y) = s(x, y) $

Πίνακας 4.1: Μοναδιαίες Πράξεις Επεξεργασίας Εικόνων.

Στον Πίνακα 4.2 τα σύμβολα $s1(x, y)$, $s2(x, y)$ και $d(x, y)$ αναφέρονται στις ψηφίδες των δύο εικόνων εισόδου και εξόδου αντίστοιχα στη θέση (x, y) .

Περιγραφή	Πράξη
Πρόσθεση/Αφαίρεση	$d(x, y) = s1(x, y) \pm s2(x, y)$
Πολλαπλασιασμός/Διαίρεση	$d(x, y) = s1(x, y) * s2(x, y)$ ή $d(x, y) = s1(x, y) / s2(x, y)$
Διάζευξη	$d(x, y) = s1(x, y) \vee s2(x, y)$
Σύζευξη	$d(x, y) = s1(x, y) \wedge s2(x, y)$
Αποκλειστική Διάζευξη	$d(x, y) = s1(x, y) \underline{\vee} s2(x, y)$
Ελάχιστο	$d(x, y) = \min(s1(x, y), s2(x, y))$
Μέγιστο	$d(x, y) = \max(s1(x, y), s2(x, y))$

Πίνακας 4.2: Δυαδικές Πράξεις Επεξεργασίας Εικόνων.

Αντιστοίχιση Τιμών. Οι μέθοδοι που ανήκουν στη συγκεκριμένη κατηγορία χρησιμοποιούν την τρέχουσα τιμή της ψηφίδας μίας εικόνας σα δείκτη σε ένα πίνακα τιμών από τον οποίο εξάγεται η τελική τιμή της ψηφίδας. Παράδειγμα εφαρμογής μίας τέτοιας μεθόδου αποτελεί η *ενίσχυση αντίθεσης (contrast enhancement)*. Αν γνωρίζουμε ότι το εύρος τιμών A των ψηφίδων μίας εικόνας είναι αρκετά μικρό σε σχέση με το εύρος τιμών B που μπορεί να προβληθεί από το υπολογιστικό σύστημα τότε μπορούμε να αντιστοιχήσουμε κάθε τιμή του τρέχοντος εύρους A σε μια τιμή του εύρους B με την εφαρμογή της ακόλουθης συνάρτησης:

$$d(x, y) = \frac{(d1 - d0)}{(s1 - s0)}(x - s1) + d1$$

όπου $s0, s1$ το άνω και κάτω όριο αντίστοιχα του εύρους τιμών της αρχικής εικόνας και $d1, d0$ το άνω και κάτω όριο αντίστοιχα του δυναμικού εύρους της τελικής εικόνας.

Ενα άλλο παράδειγμα μεθόδου αντιστοίχισης τιμών αποτελεί ο υπολογισμός του αρνητικού μίας εικόνας. Η αρνητική τιμή κάθε ψηφίδας δίνεται από τον τύπο:

$$d(x, y) = M - s(x, y)$$

όπου M η μέγιστη τιμή που μπορεί να έχει μία ψηφίδα στο υπολογιστικό σύστημα (συνήθως $M = 255$).

Μεταβολή Ιστογράμματος. Το *ιστόγραμμα* μίας εικόνας στην οποία οι ψηφίδες μπορούν να λάβουν N πιθανές τιμές (από 0 ως N-1) ορίζεται ως ο μονοδιάστατος πίνακας $h(n)$, $n = 0, \dots, N-1$ όπου $h(k)$ είναι ο αριθμός των ψηφίδων της εικόνας που έχει τη τιμή k. Με άλλα λόγια το

ιστόγραμμα περιγράφει τη συχνότητα με την οποία συναντάται κάθε πιθανή τιμή ψηφίδας σε μία εικόνα.

Η πιο συνηθισμένη μέθοδος επεξεργασίας που βασίζεται στο ιστόγραμμα μίας εικόνας είναι η *ισοστάθμιση* (*histogram equalization*). Η ισοστάθμιση αποσκοπεί στη μετατροπή του τρέχοντος ιστογράμματος h μίας εικόνας σε ένα καινούργιο ιστόγραμμα h' στο οποίο όλες οι δυνατές τιμές για μία ψηφίδα έχουν την ίδια πιθανότητα εμφάνισης. Μία συνάρτηση που είναι ικανή να μετασχηματίσει το τρέχον ιστόγραμμα με τον τρόπο που αναφέραμε περιγράφεται από τον τύπο:

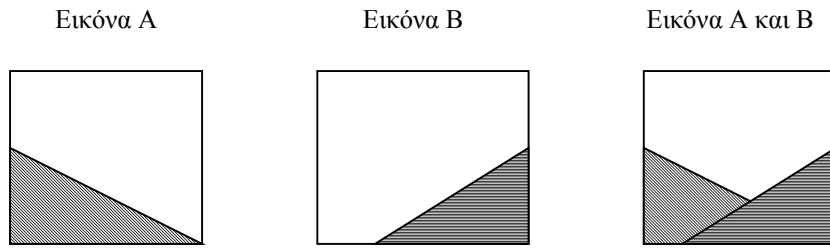
$$h'[k] = \sum_{n=0}^{k-1} h(n)$$

όπου $k = 0, \dots, N-1$. Η ισοστάθμιση σε μία εικόνα συνήθως χρησιμοποιείται για την ανάδειξη λεπτομερειών που είναι δύσκολα ανιχνεύσιμες λόγω της ανώμαλης φύσης του ιστογράμματος της αρχικής εικόνας.

Μίξη Εικόνων. Η μίξη (*image compositing*) αποτελεί μία από τις περισσότερο συνηθισμένες μεθόδους επεξεργασίας εικόνων. Ένας απλοϊκός τρόπος υλοποίησης της μίξης θα μπορούσε να είναι η πρόσθεση των τιμών των ψηφίδων που έχουν την ίδια θέση και στις δύο εικόνες. Δυστυχώς η εικόνα που θα παρήγαγε ένας τέτοιος συνδυασμός δεν θα απέδιδε σωστά τις ακμές που υπάρχουν σε κάθε μία εικόνα. Ο γενικός αλγόριθμος με τον οποίο γίνεται η μίξη δύο εικόνων A, B διακρίνει μια σειρά από 4 περιπτώσεις στις οποίες υποδιαιρείται ο χώρος που καλύπτει καθεμία από τις ψηφίδες της εικόνας που προκύπτει:

1. Καμμία κάλυψη από τις A και B
2. Κάλυψη μόνο από την A
3. Κάλυψη μόνο από τη B
4. Κάλυψη από την A και B .

Η εικόνα 4.5 περιγράφει τις περιπτώσεις που αναφέραμε.



Εικόνα 4.5: Παράδειγμα επικάλυψης κατά τη μίξη δύο εικόνων A και B.

Επομένως για να υλοποιηθεί στη γενική της μορφή η μίξη δύο εικόνων θα πρέπει να αντιστοιχήσουμε για κάθε εικόνα N που συμμετέχει στη μίξη μια συνάρτηση $\alpha_N(x, y)$ η οποία περιγράφει την επικάλυψη ης ψηφίδας της εικόνας που προκύπτει στη θέση (x, y) από την εικόνα N. Για παράδειγμα, αν έχουμε δύο εικόνες S1 και S2 τότε οι συναρτήσεις $\alpha_{S1}(x, y)$ και $\alpha_{S2}(x, y)$ περιγράφουν σε τι ποσοστό οι S1 και S2 καλύπτουν το χώρο της ψηφίδας (x, y) στην εικόνα D που προκύπτει. Στηριζόμενοι στις δύο συναρτήσεις α_{S1} και α_{S2} και στην υποδιαίρεση του χώρου κάθε ψηφίδας που περιγράψαμε παραπάνω μπορούμε να περιγράψουμε τη μέθοδο με την οποία θα υπολογίζουμε την τιμή κάθε ψηφίδας στην D σαν μια τετράδα (a, b, c, d) κάθε στοιχείο της οποίας περιγράφει αν το 0, η εικόνα S1 ή η εικόνα S2 χρησιμοποιείται για τον υπολογισμό της τιμής της ψηφίδας. Για παράδειγμα, η τετράδα $(0, S1, S2, S1)$ καθορίζει ότι σε κάθε ψηφίδα της εικόνας D που προκύπτει από τη μίξη των S1 και S2, ο χώρος στον οποίο δεν επικαλύπτονται οι S1 και S2 θα έχει την παλιά του τιμή, ο χώρος ο οποίος καλύπτεται από την S1 λαμβάνει την τιμή που έχει η S1, ο χώρος που καλύπτει η S2 λαμβάνει την τιμή που έχει η S2 και ο χώρος που αλληλοκαλύπτεται από την S1 και την S2 λαμβάνει την τιμή της S1.

Στην πράξη ο συνηθέστερος τρόπος μίξης δύο εικόνων προσεγγίζει την τετράδα $(0, S1, S2, S1)$ που περιγράψαμε παραπάνω και ονομάζεται *alpha blending*. Η συγκεκριμένη μίξη χρησιμοποιεί τρεις εικόνες, τις S1 και S2 και μία ξεχωριστή εικόνα α (alpha image). Η τιμή των ψηφίδων της εικόνας D που προκύπτει δίνεται από τον τύπο:

$$d(x, y) = (1-\alpha(x, y))S1(x, y) + \alpha(x, y)S2(x, y)$$

4.3.2 Φίλτρα Γειτνίασης

Οι κύριες εφαρμογές των φίλτρων γειτνίασης στα Συστήματα Πολυμέσων αφορούν την εξάλειψη του θορύβου από μία εικόνα όπως και την άμβλυνση (blurring) και την όξυνση (sharpening) των χαρακτηριστικών μίας εικόνας. Τα φίλτρα γειτνίασης υλοποιούνται ως διδιάστατοι πίνακες συντελεστών και η εφαρμογή τους συνίσταται στον υπολογισμό της διδιάστατης διακριτής συνέλιξης τους με τις τιμές των ψηφίδων της εικόνας στην οποία εφαρμόζονται. Πιο συγκεκριμένα, η διδιάστατη συνέλιξη μεταξύ μίας εικόνας f με διαστάσεις $M \times N$ και των συντελεστών ενός φίλτρου h διάστασης $K \times L$ ορίζεται από τον τύπο:

$$d(x, y) = \sum_{x=0}^{K-1} \sum_{y=0}^{L-1} f\left(m + \frac{K-1}{2} - x, n + \frac{L-1}{2} - y\right) h(x, y) \quad (1)$$

Στην εφαρμογή του παραπάνω τύπου υποθέτουμε ότι το σημείο $(0, 0)$ και στους δύο πίνακες f και h βρίσκεται στο άνω αριστερό άκρο τους. Οι ψηφίδες της εικόνας που προκύπτει από την εφαρμογή της (1) έχουν αποδεκτές τιμές για:

$$\frac{(K-1)}{2} \leq m \leq M - \frac{(K-1)}{2}$$

και

$$\frac{(L-1)}{2} \leq n \leq N - \frac{(L-1)}{2}$$

Κατά συνέπεια η εικόνα που προκύπτει από την εφαρμογή ενός τέτοιου φίλτρου θα έχει μικρότερη «ωφέλιμη» επιφάνεια από την αρχική. Το φαινόμενο αυτό προξενεί προβλήματα σε περιπτώσεις κατά τις οποίες μια εικόνα περνά από πολλαπλά στάδια φιλτραρίσματος ή όταν ένα τμήμα μόνο μίας εικόνας φιλτράρεται από ένα πρόγραμμα επεξεργασίας εικόνας. Στην περίπτωση αυτή θα πρέπει να προσεχθεί η απόδοση των ψηφίδων στα σύνορα της εικόνας ή του τμήματος της εικόνας που προκύπτει από το φιλτράρισμα.

Η υλοποίηση μεθόδων άμβλυνσης των χαρακτηριστικών μίας εικόνας στηρίζεται στη σύνθεση φίλτρων τα οποία υπολογίζουν ως τιμή της κάθε ψηφίδα στην εικόνα που προκύπτει το μέσο όρο των τιμών των ψηφίδων στην αρχική εικόνα σε ένα παράθυρο με κέντρο την τρέχουσα ψηφίδα και διάσταση ίση με τη διάσταση του πίνακα που υλοποιεί το φίλτρο. Για παράδειγμα, τα παρακάτω φίλτρα $H1$, $H2$, $H3$ όταν συνελιχθούν με μία εικόνα υλοποιούν την άμβλυνση των χαρακτηριστικών της:

$$H1 = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, H2 = \frac{1}{10} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}, H3 = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

Το φίλτρο H1 υπολογίζει το μέσο όρων των τιμών των ψηφίδων σε ένα παράθυρο 3x3 γύρω από την τρέχουσα ψηφίδα. Τα υπόλοιπα δύο φίλτρα αναθέτουν στην τρέχουσα τιμή της ψηφίδας μεγαλύτερη βαρύτητα. από ότι στις τιμές των γειτονικών της ψηφίδων.

Τα φίλτρα όξυνσης τονίζουν τις ακμές των αντικειμένων μίας εικόνας. Στα περισσότερα από αυτά η τιμή των ψηφίδων στην εικόνα που προκύπτει υπολογίζεται από την αφαίρεση των τιμών των γειτονικών ψηφίδων από την τιμή της τρέχουσας ψηφίδας πολλαπλασιασμένης με έναν κατάλληλο συντελεστή βαρύτητας. Τα παρακάτω φίλτρα F1, F2 και F3 αποτελούν εναλλακτικές υλοποιήσεις φίλτρων όξυνσης:

$$F1 = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}, F2 = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{bmatrix}, F3 = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 5 & -2 \\ 1 & -2 & 1 \end{bmatrix}$$

Μια αρκετά δημοφιλής μέθοδος εξάλειψης του θορύβου από μία εικόνα βασίζεται στη χρήση φίλτρων μέσης τιμής (*median filters*). Τα συγκεκριμένα φίλτρα είναι μη γραμμικά και επομένως η υλοποίησή τους δε στηρίζεται στη συνέλιξη τους με τις ψηφίδες μίας εικόνας. Για να υπολογιστεί η τιμή της κάθε ψηφίδας στην εικόνα D που προκύπτει από την εφαρμογή της μεθόδου ταξινομούνται αρχικά οι τιμές των ψηφίδων της εικόνας που βρίσκονται γύρω από την τρέχουσα ψηφίδα. Η τιμή της ψηφίδας στην εικόνα D είναι ίση με την τιμή που βρίσκεται στο μέσον της ταξινομημένης ακολουθίας τιμών που υπολογίστηκε στο προηγούμενο βήμα. Η Εικόνα 4.6 περιγράφει την εφαρμογή του φίλτρου μέσης τιμής στην κεντρική ψηφίδα μίας εικόνας.

$$\begin{array}{ccc} \text{Ψηφίδες Αρχικής Εικόνας} & & \text{Τελική Εικόνα} \\ \begin{bmatrix} 5 & 6 & 6 \\ 4 & 32 & 5 \\ 3 & 5 & 2 \end{bmatrix} & \xrightarrow{\text{Μέση Τιμή}} (2,3,4,5,6,32) \rightarrow & \begin{bmatrix} 5 \end{bmatrix} \end{array}$$

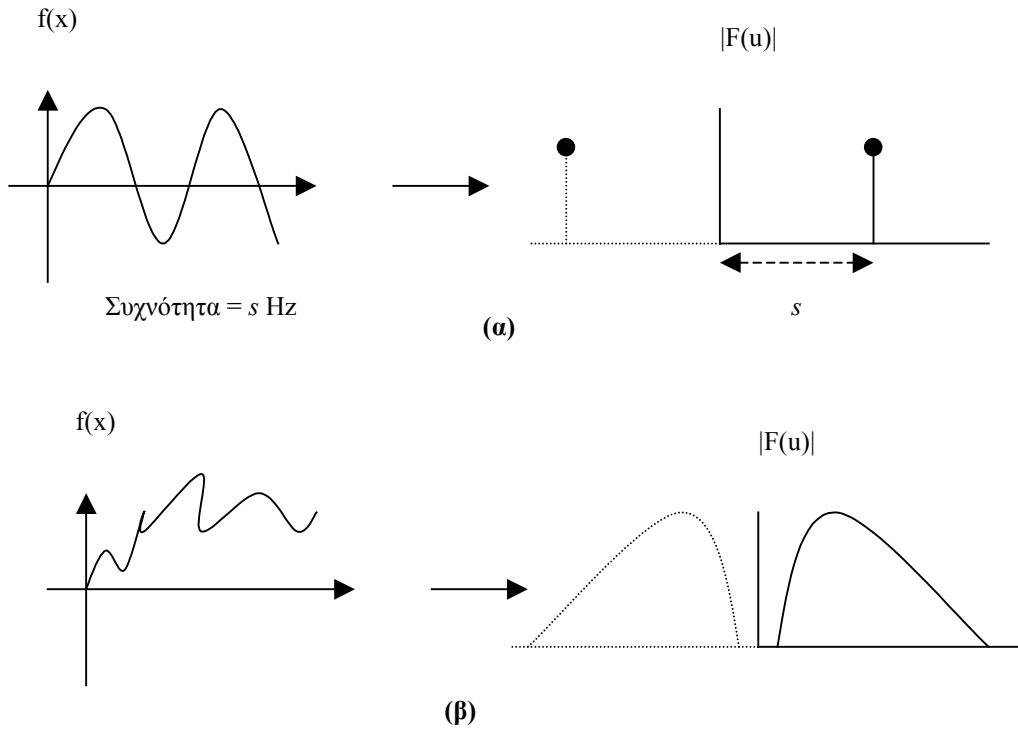
Εικόνα 4.6: Εφαρμογή φίλτρου μέσης τιμής.

4.3.3 Φασματικά Φίλτρα

Η λειτουργία των φασματικών φίλτρων βασίζονται στον μετασχηματισμό της πληροφορίας που περιέχει μία εικόνα από το πεδίο του διδιάστατου χώρου στο πεδίο της συχνότητας στο οποίο μπορούμε να μεταβάλλουμε με μεγαλύτερη ευκολία ορισμένα χαρακτηριστικά της εικόνας. Η εφαρμογή των φασματικών φίλτρων σε μία εικόνα προκύπτει από την εκτέλεση των παρακάτω τριών βημάτων:

1. Μετασχημάτισε την εικόνα από πεδίο του χώρου στο πεδίο της συχνότητας. Το βήμα αυτό πραγματοποιείται με τον μετασχηματισμό της εικόνας κατά Fourier.
2. Μετέβαλλε ορισμένα από τα χαρακτηριστικά της εικόνας στο πεδίο της συχνότητας.
3. Εφάρμοσε τον αντίστροφο μετασχηματισμό Fourier για το μετασχηματισμό της εικόνας στο πεδίο του χώρου.

Πιο συγκεκριμένα, ο μετασχηματισμός Fourier βασίζεται στο γεγονός ότι ένα σήμα μπορεί να παρασταθεί σαν ένα άθροισμα συνημιτονικών συναρτήσεων με διάφορες συχνότητες, πλάτη και φάσεις. Για παράδειγμα, αν θεωρήσουμε ότι το σήμα μας περιγράφεται από τη μονοδιάστατη συνάρτηση $I(x, y)$. Ο μετασχηματισμός Fourier της συνάρτησης αυτής θα δημιουργήσει μια καινούργια συνάρτηση $F(u)$, όπου u ένας μιγαδικός αριθμός. Η απόλυτη τιμή $|F(u)|$ της $F(u)$ ονομάζεται *φάσμα πλάτους* (*amplitude spectrum*) και καθορίζει τα πλάτη των συνημιτονικών συναρτήσεων που θα πρέπει να προστεθούν για να προκύψει η $I(x)$. Η φάση της $F(u)$ καθορίζει την αρχική φάση κάθε συνημιτονικής συνάρτησης (δηλ. την τιμή της $I(x)$ για $x = 0$) που περιέχεται στο φάσμα πλάτους και ονομάζεται *φάσμα φάσης* (*phase spectrum*). Η Εικόνα 4.5 περιγράφει το φάσμα πλάτους που αντιστοιχεί σε δύο μονοδιάστατες συναρτήσεις.



Εικόνα 4.6: Φάσμα πλάτους: (α) περιοδικού συνημιτονικού σήματος (β) τυχαίου σήματος.

Ο μετασχηματισμός Fourier ενός διδιάστατου σήματος όπως είναι μια εικόνα παράγει ένα σύνολο $|F(u, v)|$ που περιέχει το πλάτος κάθε χωρικής συχνότητας από την οποία συντίθεται το σήμα. Μια χωρική συχνότητα αντιστοιχεί σε μια επιφάνεια (μια συνημιτονική κυμάτωση) της οποίας η συχνότητα (ή ο ρυθμός κυματισμού) δίνεται από την απόσταση του σημείου (u, v) από την αρχή των αξόνων ($\sqrt{u^2 + v^2}$) ενώ η φάση της ϕ (δηλ. η γωνία που σχηματίζει ο κυματισμός της με τον οριζόντιο άξονα συντεταγμένων της $I(x,y)$) δίνεται από τον τύπο $\phi = \tan^{-1}(v/u)$. Η τιμή ενός σημείου (u, v) στο φάσμα πλάτους μίας τέτοιας συνάρτησης αντιστοιχεί στο πλάτος της χωρικής κυμάτωσης την οποία περιγράφει. Μια σημαντική ιδιότητα του διδιάστατου φάσματος πλάτους για την επεξεργασία εικόνας είναι ότι ο γεωμετρικός τόπος όλων των συνημιτονικών κυματώσεων που συνθέτουν την εικόνα και έχουν την ίδια συχνότητα f είναι η περιφέρεια ενός κύκλου με κέντρο το κέντρο του φάσματος πλάτους και ακτίνα ίση με f . Η ιδιότητα αυτή διευκολύνει σημαντικά την κατασκευή υπιπερατών, βαθυπερατών ή ζωνοπερατών φίλτρων καθώς η λειτουργία τους ανάγεται στην μεταβολή των τιμών του φάσματος πλάτους που βρίσκονται εντός, εκτός ή μεταξύ, αντίστοιχα, κύκλων που αντιστοιχούν στις συχνότητες που μας ενδιαφέρουν.

4.3.4 Γεωμετρικές Μέθοδοι

Οι γεωμετρικές μέθοδοι ασχολούνται με τη μεταβολή της θέσης, δηλ. των συντεταγμένων, των ψηφίδων μίας εικόνας μέσω της εφαρμογής ενός ή περισσότερων γεωμετρικών μετασχηματισμών. Η μεταβολή του περιεχομένου μίας εικόνας με την εφαρμογή γεωμετρικών μεθόδων είναι γνωστή ως *image warping*. Οι πιο συνηθισμένες γεωμετρικές μέθοδοι είναι η *μεγέθυνση/σμίκρυνση (scaling)*, η *περιστροφή (rotation)*, οι *ομοπαράλληλοι μετασχηματισμοί (affine transformations)* και οι *μεταθέσεις (transpositions)*.

Ειδικότερα, με τον όρο γεωμετρικοί μετασχηματισμοί αναφερόμαστε σε μεθόδους μεταβολής των συντεταγμένων κάθε ψηφίδας μίας εικόνας χωρίς μεταβολή του συστήματος συντεταγμένων στο οποίο αναφέρονται τα σημεία αυτά. Στη συγκεκριμένη ενότητα θα ασχοληθούμε με διδιάστατους γεωμετρικούς μετασχηματισμούς που είναι και οι συνηθέστεροι.

Ενας διδιάστατος γεωμετρικός μετασχηματισμός απεικονίζει κάθε σημείο A (x_A, y_A) του επιπέδου σε ένα άλλο σημείο B (x_B, y_B) μέσω μίας συνάρτησης T έτσι ώστε:

$$T(x_A, y_A) = (x_B, y_B)$$

ή πιο συνοπτικά: $T(A) = B$

Από όλους τους πιθανούς διδιάστατους γεωμετρικούς μετασχηματισμούς οι πλέον χρησιμοποιούμενοι στα Συστήματα Πολυμέσων είναι οι *ομοπαράλληλοι (affine)* μετασχηματισμοί. Οι μετασχηματισμοί αυτοί έχουν μια αρκετά απλή μορφή. Πιο συγκεκριμένα, αν ένας τέτοιος μετασχηματισμός απεικονίζει το σημείο A που αναφέραμε προηγουμένως σε ένα σημείο B τότε οι συντεταγμένες των δύο σημείων θα συνδέονται με τους τύπους:

$$\begin{aligned}x_B &= a \cdot x_A + c \cdot y_A + l_x \\y_B &= b \cdot x_A + d \cdot y_A + l_y\end{aligned}$$

όπου a, b, c, d, l_x, l_y σταθερές και ad διάφορο του bc . Η μορφή που γράψαμε μπορεί να εκφραστεί σε μορφή πινάκων ως:

$$(x_B, y_B) = (x_A, y_A) \cdot M + (l_x, l_y) \quad \mathbf{(1)}$$

όπου ο M είναι ένας 2x2 πίνακας με τη μορφή:

$$\begin{matrix} a & b \\ c & d \end{matrix}$$

Ο μηχανισμός που περιγράψαμε χρησιμοποιείται για την υλοποίηση τεσσάρων πρωτογενών μετασχηματισμών που από τη σύνθεση τους μπορεί να προκύψει οποιοσδήποτε άλλος μετασχηματισμός. Οι πρωτογενείς μετασχηματισμοί είναι:

- Η μετατόπιση (*translation*)
- Η μεγέθυνση/σμίκρυνση (*scaling*)
- Η περιστροφή (*rotation*)
- Η κύρτωση (*shear*)

Στη συνέχεια εξετάζουμε καθέναν από αυτούς τους μετασχηματισμούς.

4.3.4.1 Μετατόπιση

Η μετατόπιση ενός σημείου σε ένα γεωμετρικό μετασχηματισμό περιγράφεται από τις συντεταγμένες του σημείου I (I_x, I_y). Στον συγκεκριμένο μετασχηματισμό ο πίνακας M έχει τη μορφή:

$$\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}$$

Το αποτέλεσμα της εφαρμογής ενός τέτοιου μετασχηματισμού σε ένα σημείο A είναι η μετατόπιση του A κατά I_x και κατά I_y αντίστοιχα στους άξονες x και y.

4.3.4.2 Μεγέθυνση/Σμίκρυνση

Η μεγέθυνση/σμίκρυνση ενός σχήματος κατά S_x και S_y αντίστοιχα στους άξονες x και y επιτυγχάνεται με τον πολλαπλασιασμό των αντίστοιχων συντεταγμένων κάθε σημείου του με τους δύο αυτά ποσοστά μεγέθυνσης/σμίκρυνσης. Για την υλοποίηση της παραπάνω λειτουργίας ο πίνακας M έχει τη μορφή:

$$\begin{matrix} S_x & 0 \\ 0 & S_y \end{matrix}$$

και το l έχει τη μοφή $(0, 0)$.

Εδώ θα πρέπει να σημειώσουμε ότι αν κάποιο από τα S_x, S_y είναι αρνητικό τότε ο συγκεκριμένος μετασχηματισμός πέρα από τη μεταβολή των διαστάσεων του σχήματος το μετατοπίζει στο συμμετρικό του σχήματος κατά τους άξονες y και x αντίστοιχα.

4.3.4.3 Περιστροφή

Η περιστροφή ενός σημείου κατά γωνία θ ως προς το κέντρο των αξόνων του συστήματος συντεταγμένων περιγράφεται από τον τύπο (1) όταν ο πίνακας M έχει τη μορφή:

$$\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

και το l έχει τη μοφή $(0, 0)$.

4.3.4.4 Κύρτωση

Η κύρτωση περιλαμβάνει τη μεταβολή των συντεταγμένων στον άξονα των x ενός σημείου κατά ένα ποσό που είναι ανάλογο της συντεταγμένης του ίδιου σημείου κατά τον άξονα των y . Ένα παράδειγμα ενός τέτοιου μετασχηματισμού αποτελεί η μετατροπή ορθής γραφής σε πλάγια (*italics*). Κατά το μετασχηματισμό αυτό η γενική μορφή του πίνακα M είναι:

$$\begin{pmatrix} 1 & g \\ h & 1 \end{pmatrix}$$

ενώ το l έχει τη μοφή $(0, 0)$.

4.3.4.5 Αντίστροφος Ομοπαράλληλικός Μετασχηματισμός

Δοθέντος ενός ομοπαράλληλικού μετασχηματισμού με πίνακα M της μορφής:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

τότε ο αντίστροφος μετασχηματισμός προκύπτει να υπολογιστεί ο αντίστροφος πίνακας M^{-1} και ακολουθηθεί η ανάλογη σειρά υπολογισμού. Ο M^{-1} δίνεται από τον τύπο:

$$M^{-1} = k * A$$

όπου $k = 1/(ad-bc)$ και ο πίνακας A έχει τη μορφή:

$$\begin{array}{cc} d & -b \\ -c & a \end{array}$$

4.3.4.6 Ιδιότητες των ομοπαράλληλικών μετασχηματισμών

Οι ομοπαράλληλικοί μετασχηματισμοί έχουν μία σειρά από τρομερά χρήσιμες ιδιότητες που εξηγούν και την ευρεία χρησιμοποίησή τους για το μετασχηματισμό σχημάτων στα γραφικά υπολογιστών. Οι ιδιότητες αυτές μπορούν να συνοψιστούν ως εξής:

- Οι ομοπαράλληλικοί μετασχηματισμοί διατηρούν τις ευθείες. Μια ευθεία γραμμή πριν από ένα τέτοιο μετασχηματισμό συνεχίζει να είναι ευθεία και μετά το μετασχηματισμό αυτό.
- Οι ομοπαράλληλικοί μετασχηματισμοί διατηρούν την παράλληλία μεταξύ των ευθειών. Ευθείες που είναι παράλληλες πριν από την εφαρμογή του μετασχηματισμού συνεχίζουν να είναι παράλληλες και μετά την εφαρμογή του.
- Οι ομοπαράλληλικοί μετασχηματισμοί διατηρούν την αναλογία μεταξύ των αποστάσεων. Ένα σημείο P που βρίσκεται σε αναλογία αποστάσεων f μεταξύ δύο άλλων σημείων A και B θα συνεχίσει να βρίσκεται σε μια τέτοια αναλογία όταν τα P, A και B θα έχουν μετασχηματιστεί.
- Το λόγο του εμβαδού του σχήματος που προκύπτει από το μετασχηματισμό προς το λόγο του αρχικού σχήματος ισούται με την απόλυτη τιμή της ορίζουσας του πίνακα M.

4.3.4.7 Σύνθεση ομοπαράλληλικών μετασχηματισμών

Υποθέτουμε ότι $T1()$ και $T2()$ περιγράφουν δύο ομοπαράλληλους μετασχηματισμούς με πίνακες $M1$ και $M2$ και μετατόπιση $L1$ και $L2$ αντίστοιχα. Πως μπορούμε να εκφράσουμε το μετασχηματισμό που προκύπτει αν σε ένα σημείο A εφαρμόσουμε διαδοχικά τον $T1$ και $T2$;

Για να απαντήσουμε στο ερώτημα αυτό υποθέτουμε ότι η εφαρμογή του μετασχηματισμού $T1$ αντιστοιχίζει το σημείο A σε ένα άλλο σημείο B και η εφαρμογή του μετασχηματισμού $T2$ αντιστοιχίζει το B σε ένα άλλο σημείο Γ . Στην περίπτωση αυτή θα ισχύει ότι:

$$B = A * M1 + L1$$

$$\Gamma = B * M2 + L2 = (A * M1 + L1) * M2 + L2 = A * (M1 * M2) + L1 * M2 + L2$$

Επομένως ο μετασχηματισμός που προκύπτει από τη σύνθεση των δύο μετασχηματισμών που αναφέραμε θα είναι ένας ομοπαράλληλος μετασχηματισμός με πίνακα M ίσο με το γινόμενο των πινάκων των επιμέρους μετασχηματισμών με τη σειρά που εφαρμόζονται και με μετατόπιση L που θα δίνεται από τον τύπο:

$$L = L1 * M2 + L2$$

Με την ίδια λογική μπορούμε να γενικεύσουμε τον παραπάνω τύπο για τη σύνθεση περισσότερων των δύο μετασχηματισμών.

Η αντικατάσταση περισσότερων του ενός διαδοχικών μετασχηματισμών από τη σύνθεση τους αποτελεί μια διαδικασία με σημαντικό υπολογιστικό όφελος. Για παράδειγμα, αν υποθέσουμε ότι θέλουμε να εφαρμόσουμε δύο διαδοχικούς ομοπαράλληλους μετασχηματισμούς κάθε μία από τις κορυφές ενός N -γώνου. Για τη σύνθεση ενός μετασχηματισμού που περιγράφει και τους δύο μετασχηματισμούς θα χρειαστούμε 12 πολλαπλασιασμούς και για την εφαρμογή του τελικού μετασχηματισμού σε καθένα από τα σημεία του N -γώνου θα χρειαστούμε άλλους 4 πολλαπλασιασμούς. Συνολικά επομένως θα χρειαστούμε $4N+12$ πολλαπλασιασμούς. Αν εφαρμόζαμε σε κάθε σημείο κάθε μετασχηματισμό ξεχωριστά θα χρειαζόμασταν συνολικά $8N$ μετασχηματισμούς. Για τα N -γωνα για τα οποία το $8N$ είναι μεγαλύτερο από το $4N+12$ συμφέρει να συνθέσουμε τους μετασχηματισμούς.

Ως παράδειγμα σύνθεσης ενός μετασχηματισμού από μια σειρά άλλων απλότερων μετασχηματισμών μπορούμε να θεωρήσουμε την περιστροφή κατά γωνία θ ενός σημείου A (x_A, y_A) γύρω από ένα τυχαίο κέντρο περιστροφής B (x_B, y_B). Ένας τέτοιος μετασχηματισμός

μπορούμε να θεωρήσουμε ότι υλοποιείται με τη σύνθεση τριών απλούστερων μετασχηματισμών στο σημείο A:

- Μίας μετατόπισης των A και B κατά $(-x_B, -y_B)$
- Μίας περιστροφής του A που προκύπτει από τον παραπάνω μετασχηματισμό κατά γωνία θ γύρω από το σημείο B που προκύπτει από τον παραπάνω μετασχηματισμό
- Μίας μετατόπισης των δύο σημείων κατά (x_B, y_B)

4.3.4.8 Περιγραφή Μετασχηματισμών σε Ομογενείς Συντεταγμένες

Ενας από τους τρόπους απλοποίησης των τύπων που χρησιμοποιούμε για την περιγραφή των ομοπαράλληλικών μετασχηματισμών στηρίζεται στη χρησιμοποίηση ομογενών (*homogeneous*) συντεταγμένων. Πιο συγκεκριμένα, στον τύπο (1) η μετατόπιση και ο πίνακας M μπορούν να συμπυκνωθούν σε έναν πίνακα με την ακόλουθη μετατροπή:

$$(x_B, y_B, 1) = (x_A, y_A, 1) * M$$

όπου ο πίνακας M έχει την ακόλουθη μορφή:

a	b	0
c	d	0
L_x	l_y	1