1. Εισαγωγή

Η ανάλυση εικόνων με υπολογιστή είναι μια τεχνολογία με σχετικά σύντομη ιστορία, αλλά με περισσότερο ή λιγότερο επιτυχημένη εφαρμογή σε κάθε τύπο εικόνας. Ως επιστημονικό πεδίο, συμπεριλαμβάνει ιδέες από την οπτική, την ηλεκτρονική, τα μαθηματικά, τη φωτογραφία και την επιστήμη των υπολογιστών. Τα τελευταία χρόνια, η ανάλυση εικόνων με υπολογιστή επιδεικνύει αλματώδη πρόοδο, γεγονός που οφείλεται τόσο στην εκθετική μείωση του κόστους των υπολογιστικών μονάδων και των μονάδων αποθήκευσης δεδομένων όσο και στη διαθεσιμότητα συσκευών ψηφιοποίησης και προβολής εικόνων.

Στην πιο βασική μορφή του, ένα σύστημα ανάλυσης εικόνων με υπολογιστή αποτελείται από έναν υπολογιστή για επεξεργασία εικόνων και από δύο είδη εξειδικευμένων μονάδων εισόδου / εξόδου : έναν **ψηφιοποιητή εικόνων (image digitizer)** και μια **συσκευή παρουσίασης εικόνων (image display device).** Η αναγκαιότητα ύπαρξης του ψηφιοποιητή εικόνων προκύπτει από το γεγονός ότι οι εικόνες δεν επιδέχονται ανάλυση με υπολογιστή απευθείας στη φυσική μορφή τους. Οι υπολογιστές επεξεργάζονται αριθμητικά δεδομένα και, επομένως, μια εικόνα πρέπει να ψηφιοποιηθεί και να μετατραπεί σε αριθμητική μορφή πριν αρχίσει οποιαδήποτε επεξεργασία της.

Η διαδικασία της ψηφιοποίησης υποδιαιρεί μια εικόνα σε μικρές περιοχές που ονομάζονται **εικονοστοιχεία (picture elements ή pixels).** Η πιο συνηθισμένη μέθοδος υποδιαίρεσης είναι η **ορθογώνια δειγματοληψία (rectangular sampling),** με την οποία η εικόνα διαιρείται σε οριζόντιες γραμμές διαδοχικών εικονοστοιχείων. Σε κάθε εικονοστοιχείο αντιστοιχίζεται ένας αριθμός που παριστάνει τη φωτεινότητα της εικόνας στο αντίστοιχο σημείο. Επομένως, η διαδικασία της ψηφιοποίησης δειγματίζει και κβαντίζει μια εικόνα και για κάθε εικονοστοιχείο παράγει έναν ακέραιο που παριστάνει τη φωτεινότητα της εικόνας στο σημείο αυτό. Τελικά, σε κάθε εικονοστοιχείο αντιστοιχίζεται μια ακέραια διεύθυνση (αριθμός γραμμής και αριθμός στήλης του εικονοστοιχείου) και μια ακέραια τιμή που ονομάζεται **επίπεδο γκρίζου (gray level)**. Δημιουργείται με αυτόν τον τρόπο ένας πίνακας ψηφιακών δεδομένων κατάλληλος για επεξεργασία με υπολογιστή.

Κατά την επεξεργασία, τα εικονοστοιχεία μεταβάλλονται σύμφωνα με τις εντολές του εκτελουμένου προγράμματος. Πέρα από τους συνήθεις προγραμματιστικούς περιορισμούς, τα στάδια επεξεργασίας είναι απεριόριστα. Μετά την επεξεργασία, τα αποτελέσματα παρουσιάζονται με μια διαδικασία αντίστροφη της ψηφιοποίησης : Το επίπεδο γκρίζου κάθε εικονοστοιχείου προσδιορίζει τη φωτεινότητα του αντιστοίχου σημείου στη συσκευή παρουσίασης εικόνων. Με τον τρόπο αυτό, η επεξεργασμένη εικόνα είναι και πάλι ορατή από ανθρώπινο χρήστη.

Εικόνες εμφανίζονται με διάφορες μορφές, ορατές ή μη ορατές, αφηρημένες ή φυσικές, κατάλληλες ή ακατάλληλες για επεξεργασία και ανάλυση με υπολογιστή. Το γεγονός αυτό σε συνδυασμό με τη συνεχή παρουσία τους στη ζωή του ανθρώπου κάνει αναγκαίο τον προσδιορισμό κάποιας τυποποιημένης συναφούς ορολογίας. Θα παρουσιαστεί στη συνέχεια το πιο βασικό κομμάτι αυτής της ορολογίας.

Στο λεξικό του Webster, ο όρος **εικόνα (image)** ορίζεται ως «μια αναπαράσταση, ομοιότητα ή απομίμηση ενός αντικειμένου ή πράγματος, μια ζωντανή γραφική περιγραφή, κάτι που εισάγεται για να περιγράψει κάτι άλλο». Για παράδειγμα, μια φωτογραφία προσώπου είναι η αναπαράσταση του προσώπου που στάθηκε μπροστά στη φωτογραφική μηχανή.

Μπορούμε να κατατάξουμε τις εικόνες σε διάφορες κατηγορίες με βάση τη μορφή τους ή τον τρόπο με τον οποίο δημιουργήθηκαν. Για παράδειγμα, **ορατές (visible)** είναι οι εικόνες που αντιλαμβάνεται το ανθρώπινο μάτι, όπως φωτογραφίες, σχέδια και πίνακες ζωγραφικής. Μια άλλη κατηγορία εικόνων είναι οι οπτικές **(optical)**, δηλαδή αυτές που δημιουργήθηκαν με τη χρήση οπτικών συσκευών όπως φακών και ολογραμμάτων. Οι φυσικές εικόνες είναι στην πραγματικότητα κατανομές ύλης και ενέργειας. Για παράδειγμα, οι οπτικές εικόνες είναι χωρικές κατανομές έντασης φωτός. Αυτές οι κατανομές γίνονται αντιληπτές από το ανθρώπινο μάτι και, επομένως, οι εικόνες αυτές είναι και ορατές. Παραδείγματα μη ορατών εικόνων αποτελούν χάρτες κατανομών θερμοκρασίας, πίεσης, υψομέτρου και πυκνότητας πληθυσμών. Μια ειδική κατηγορία φυσικών εικόνων αποτελούν οι **πολυφασματικές (multispectral**), αυτές, δηλαδή, που ορίζουν περισσότερες από μία ιδιότητες σε κάθε σημείο. Για παράδειγμα, οι εικόνες μιας έγχρωμης φωτογραφίας ή μιας έγχρωμης τηλεόρασης είναι **τριφασματικές (trispectral)** εικόνες του κόκκινου, του πράσινου και του μπλε χρώματος. Ενώ μια ασπρόμαυρη εικόνα περιέχει μόνο μια τιμή φωτεινότητας σε κάθε σημείο, μια έγχρωμη εικόνα περιέχει τρεις τιμές φωτεινότητας, μια για κάθε ένα από τα χρώματα κόκκινο, πράσινο και μπλε. Οι τρεις τιμές παριστάνουν ένταση φωτός σε διαφορετικές φασματικές ζώνες που το μάτι αντιλαμβάνεται ως διαφορετικά χρώματα.

Ένα άλλο σύνολο εικόνων αποτελείται από τις αφηρημένες εικόνες των μαθηματικών, δηλαδή από συνεχείς συναρτήσεις (συνεχείς εικόνες) και διακριτές συναρτήσεις (διακριτές εικόνες). Ένας υπολογιστής μπορεί να επεξεργαστεί μόνο διακριτές εικόνες.

Μια **ζωγραφιά (picture)** είναι ένας περιορισμένος τύπος εικόνας. Μπορούμε να ορίσουμε τη ζωγραφιά ως την κατανομή ύλης που είναι ορατή με κατάλληλο φωτισμό. Μπορούμε όμως και να θεωρήσουμε την έννοια της ζωγραφιάς ως ταυτόσημη με την έννοια της εικόνας.

Ο όρος *``***ψηφιακός (digital)***´´* αναφέρεται σε υπολογισμό με αριθμητικές μεθόδους ή με διακριτές μονάδες. Αν ορίσουμε μια **ψηφιακή εικόνα (digital image)** ως την αριθμητική αναπαράσταση ενός αντικειμένου (που μπορεί το ίδιο να αποτελεί μια εικόνα), τα εικονοστοιχεία αποτελούν τις διακριτές μονάδες και η κβαντισμένη (ακέραια) κλίμακα γκρίζου παρέχει την αριθμητική συνιστώσα.

Η επεξεργασία (processing) είναι η δράση που υποβάλλει κάτι σε μια ακολουθία λειτουργιών που οδηγούν σε επιθυμητό αποτέλεσμα. Επομένως, ορίζουμε την **ψηφιακή επεξεργασία εικόνας (digital image processing)** ως μια διαδικασία που υποβάλλει την αριθμητική αναπαράσταση ενός αντικειμένου σε μια ακολουθία λειτουργιών με σκοπό να επιτευχθεί κάποιο επιθυμητό αποτέλεσμα. Η ψηφιακή επεξεργασία εικόνας ξεκινάει με μια εικόνα και παράγει μια τροποποιημένη μορφή αυτής της εικόνας και, επομένως, απεικονίζει μια εικόνα σε μια άλλη εικόνα. Η **ψηφιακή ανάλυση εικόνας (digital image**

analysis) είναι μια διαδικασία που απεικονίζει μια ψηφιακή εικόνα σε κάτι διαφορετικό από ψηφιακή εικόνα, όπως ένα σύνολο δεδομένων από μέτρηση ή μια απόφαση. Ο όρος «ψηφιακή επεξεργασία εικόνας» χρησιμοποιείται συχνά για να υποδηλώσει τόσο επεξεργασία όσο και ανάλυση.

Τα **γραφικά με υπολογιστή (computer graphics)** ασχολούνται με την επεξεργασία και παρουσίαση εικόνων αντικειμένων που υπάρχουν μόνο νοητά ή ως μαθηματικές περιγραφές παρά ως φυσικά αντικείμενα. Η έμφαση είναι συχνά στη δημιουργία μιας εικόνας, δεδομένων ενός μοντέλου που περιγράφει ένα αντικείμενο, του φωτισμού του και της γεωμετρίας μιας υποθετικής φωτογραφικής μηχανής. Τα γραφικά με υπολογιστή περιλαμβάνουν και την **τέχνη με υπολογιστή (computer art)**, δηλαδή τη χρήση ψηφιακών συστημάτων απεικόνισης με μέσων καλλιτεχνικής έκφρασης.

Η **υπολογιστική όραση (computer vision)** ασχολείται με την ανάπτυξη συστημάτων που ερμηνεύουν το περιεχόμενο φυσικών σκηνών. Στην περιοχή της ρομποτικής, η υπολογιστική όραση παρέχει τεχνητά μάτια στο ρομπότ. Ο όρος **ψηφιακή απεικόνιση (digital imaging)** αναφέρεται σε κάθε μορφή επίδρασης σε δεδομένα εικόνας με υπολογιστή και συμπεριλαμβάνει τα γραφικά με υπολογιστή, την υπολογιστική όραση καθώς και την ψηφιακή επεξεργασία και ανάλυση εικόνων.

Στη συνέχεια, θα ορίζεται μια ψηφιακή εικόνα ως μια δειγματισμένη και κβαντισμένη συνάρτηση δύο διαστάσεων που έχει παραχθεί με οπτικά μέσα, ορθογώνια δειγματοληψία ίσου διαχωρισμού δειγμάτων και κβάντωση σε διαστήματα ίσου πλάτους. Επομένως, μια ψηφιακή εικόνα είναι ένας διδιάστατος ορθογώνιος πίνακας κβαντισμένων δειγμάτων.

Για την αποτελεσματική εφαρμογή μεθόδων ανάλυσης εικόνας με υπολογιστή, είναι απαραίτητο ο σχεδιασμός των υπολογιστικών προγραμμάτων να γίνεται με οργανωμένο και συστηματικό τρόπο. Μια τέτοια διαδικασία είναι χρήσιμη τόσο σε προγραμματιστές λογισμικού ανάλυσης εικόνας όσο και σε χρήστες του. Για την αποδοχή του λογισμικού, οι εύχρηστες διεπαφές προς το χρήστη και η καλή τεκμηρίωση είναι εξίσου σημαντικά με τη λειτουργικότητα και την ακρίβεια. Για τους περισσότερους χρήστες, οι **γραφικές διεπαφές με κατάλογο επιλογών (menu-driven graphical user interfaces)** είναι ευκολότερες στην εκμάθηση και τη λειτουργία από τις **διεπαφές κειμένου (textual interfaces)**. Παράλληλα, παρατηρείται η τάση μετάβασης από διεπαφές κειμένου σε **διεπαφές ρήσης (verbal interfaces)** και **οπτικές διεπαφές (visual interfaces)**.

2. Σημειακοί, αλγεβρικοί και γεωμετρικοί τελεστές

2.1. Ιστογράμματα επιπέδων γκρίζου

Ένα από τα απλούστερα και πλέον χρήσιμα εργαλεία επεξεργασίας εικόνας είναι το **ιστόγραμμα επιπέδων γκρίζου (gray level histogram)**. Αυτή η συνάρτηση συνοψίζει το περιεχόμενο επιπέδου γκρίζου μιας εικόνας, δείχνοντας, για κάθε επίπεδο γκρίζου (D), τον αριθμό των εικονοστοιχείων στην εικόνα που έχουν αυτό το επίπεδο γκρίζου. Με άλλα λόγια, το ιστόγραμμα επιπέδων γκρίζου είναι μια γραφική παράσταση στην οποία ο οριζόντιος άξονας παριστάνει το επίπεδο γκρίζου και ο κατακόρυφος τη συχνότητα εμφάνισής του.

Συχνά, είναι χρήσιμη η κατασκευή ιστογραμμάτων περισσοτέρων διαστάσεων. Για παράδειγμα, τέτοια ιστογράμματα είναι πολύ χρήσιμα στην επεξεργασία και ανάλυση εγχρώμων εικόνων. Ένα διδιάστατο ιστόγραμμα είναι μια συνάρτηση δύο μεταβλητών, του επιπέδου γκρίζου στην κόκκινη εικόνα και του επιπέδου γκρίζου στην μπλε εικόνα. Η τιμή του για κάποιο συνδυασμό (D_R,D_B) των δύο επιπέδων γκρίζου ισούται με τον αριθμό των αντιστοίχων εικονοστοιχείων με επίπεδο γκρίζου D_R στην κόκκινη εικόνα και D_B στην μπλε. Αν οι δύο εικόνες, κόκκινη και μπλε, είναι ταυτόσημες, το διδιάστατο ιστόγραμμα θα είχε παντού τιμή μηδέν, εκτός από τη διαγώνιο των 45^o. Εικονοστοιχεία με διαφορετικά κόκκινα και μπλε επίπεδα γκρίζου συνεισφέρουν στο ιστόγραμμα επάνω και κάτω από τη διαγώνιο των 45^o.

Το ιστόγραμμα επιπέδων γκρίζου έχει πολλαπλές χρησιμότητες στην επεξεργασία εικόνας με υπολογιστή, μερικές από τις οποίες συνοψίζονται παρακάτω :

- Το ιστόγραμμα δίνει μια απλή οπτική ένδειξη για το κατά πόσο μια εικόνα κάνει καλή και πλήρη χρήση του διαθεσίμου διαστήματος επιπέδων γκρίζου. Κανονικά, μια ψηφιακή εικόνα θα πρέπει να χρησιμοποιεί όλα ή σχεδόν όλα τα διαθέσιμα επίπεδα γκρίζου. Στην αντίθετη περίπτωση, το αποτέλεσμα είναι ουσιαστικά κβάντωση σε **λιγότερα** επίπεδα, δηλαδή **αύξηση** του διαστήματος ανάμεσα στα επίπεδα κβάντωσης. Ομοίως, αν ο ψηφιοποιητής δεν καλύπτει όλο το εύρος φωτεινοτήτων της εικόνας, τότε τα επίπεδα γκρίζου θα αποκοπούν με αποτέλεσμα να εμφανίζονται απομονωμένες μεγάλες τιμές στα δύο άκρα του ιστογράμματος. Γενικά, είναι καλή τακτική να επιθεωρείται το ιστόγραμμα επιπέδων γκρίζου κατά τη διαδικασία της ψηφιοποίησης, ώστε να αποκαλύπτονται έγκαιρα προβλήματα στην ψηφιοποίηση.
- Οι υψομετρικές γραμμές παρέχουν μια αποτελεσματική μέθοδο για τον προσδιορισμό του συνόρου ενός απλού αντικειμένου σε μια εικόνα. Η τεχνική αυτή ονομάζεται **κατωφλίωση (thresholding)**. Ας θεωρήσουμε ότι μια εικόνα περιέχει ένα σκούρο αντικείμενο. Ας θεωρήσουμε ότι μια εικόνα περιέχει ένα σκούρο αντικείμενο. Τα σκούρα εικονοστοιχεία μέσα στο αντικείμενο παράγουν ένα μέγιστο (μια «κορυφή») στο δεξί άκρο του ιστογράμματος. Τα εικονοστοιχεία του παρασκηνίου παράγουν ένα άλλο μέγιστο (μια άλλη «κορυφή») στο αριστερό άκρο του ιστογράμματος. Τα σχετικά λίγα εικονοστοιχεία γύρω από το άκρο του αντικειμένου παράγουν ένα ελάχιστο (μια «κοιλάδα») ανάμεσα στα δύο μέγιστα του ιστογράμματος. Η επιλογή του επιπέδου γκρίζου στην περιοχή της κοιλάδας

ως κατωφλίου θα οδηγήσει τελικά σε μια λογική εκτίμηση του συνόρου του αντικειμένου.

- Το ιστόγραμμα μπορεί να οδηγήσει σε εκτίμηση του εμβαδού ενός αντικειμένου σε εικόνα χωρίς ενδιάμεση χρήση της ίδιας της εικόνας. Αυτό μπορεί να επιτευχθεί στην περίπτωση που η εικόνα περιέχει μόνο ένα ομοιόμορφο αντικείμενο και το παρασκήνιο και είναι γνωστή η τιμή του επιπέδου γκρίζου του αντικειμένου. Μια άλλη μέτρηση που μπορεί να υπολογιστεί κατευθείαν από το ιστόγραμμα μιας εικόνας χωρίς χρήση της ίδιας της εικόνας είναι η μέτρηση της **ολοκληρωμένης οπτικής πυκνότητας (integrated optical density)**, ενός μέτρου της «μάζας» μιας εικόνας. Για μια ψηφιακή εικόνα, η ολοκληρωμένη οπτική πυκνότητα ορίζεται ως

$$\mathrm{ООП} = \sum_{i=1}^{NR}\sum_{j=1}^{NC} D(i,j),$$

όπου D(i,j) είναι το επίπεδο γκρίζου του εικονοστοιχείου στη γραμμή i και στη στήλη j και NR(NC) είναι ο αριθμός των γραμμών (στηλών) της εικόνας. Αν N(k) είναι ο αριθμός των εικονοστοιχείων στην εικόνα με επίπεδο γκρίζου ίσο με k, έχουμε

$$\mathrm{ООП} = \sum_{k=1}^{NL} k\,N(k),$$

όπου NL είναι ο αριθμός των επιπέδων γκρίζου.

2.2. Ιστογράμματα επιπέδων γκρίζου

Οι σημειακοί τελεστές αποτελούν μια οικογένεια απλών αλλά σημαντικών τεχνικών επεξεργασίας εικόνας και επιτρέπουν στο χρήστη την τροποποίηση της κατανομής των δεδομένων εικόνας στο διαθέσιμο διάστημα επιπέδων γκρίζου. Επομένως, επηρεάζουν τη μορφή με την οποία εμφανίζεται η εικόνα.

Ένας **σημειακός τελεστής (point operator)** δέχεται ως όρισμα μια (εικόνα εισόδου) και παράγει μια νέα εικόνα (εικόνα εξόδου) σε τρόπο ώστε το επίπεδο γκρίζου κάθε εικονοστοιχείου της εικόνας εξόδου να προσδιορίζεται αποκλειστικά από το επίπεδο γκρίζου του αντίστοιχου εικονοστοιχείου της εικόνας εισόδου. Σε αντίθεση με τους σημειακούς τελεστές, ένας **τοπικός τελεστής (local operator)** προσδιορίζει το επίπεδο γκρίζου κάθε εικονοστοιχείου της εικόνας εξόδου χρησιμοποιώντας μια περιοχή εικονοστοιχείων γύρω από το αντίστοιχο εικονοστοιχείο της εικόνας εισόδου. Τα παραπάνω σημαίνουν ότι ένας σημειακός τελεστής δεν μπορεί να μεταβάλλει τις χωρικές σχέσεις μεταξύ εικονοστοιχείων σε μια εικόνα.

Οι σημειακοί τελεστές συχνά αναφέρονται και ως **ενίσχυση αντίθεσης (contrast enhancement) ή μετασχηματισμοί κλίμακας γκρίζου (gray scale transformations)**. Συχνά ενσωματώνονται στο λογισμικό ψηφιοποίησης και παρουσίας εικόνων και αποσκοπούν στην τροποποίηση του ιστογράμματος επιπέδου γκρίζου της εικόνας με έναν προβλέψιμο τρόπο. Ενας σημειακός τελεστής δρα σε μια εικόνα εισόδου A(m,n) και παράγει μια εικόνα εξόδου B(m,n) σύμφωνα με μια μαθηματική σχέση

$$B\ (m,\ n) = f\ [\ A\ (m,\ n)\ ].$$

Η συνάρτηση f(˙) ονομάζεται **συνάρτηση μετασχηματισμού κλίμακας γκρίζου (gray scale transformation function)** και προσδιορίζει πλήρως το σημειακό τελεστή. Οι συναρτήσεις μετασχηματισμού κλίμακας γκρίζου συνηθισμένων σημειακών τελεστών περιγράφονται στη συνέχεια :

- **Γραμμικοί σημειακοί τελεστές** : Στην περίπτωση αυτή, το επίπεδο γκρίζου στην εικόνα εξόδου είναι γραμμική συνάρτηση του επιπέδου γκρίζου της εικόνας εισόδου. Επομένως, η συνάρτηση μετασχηματισμού κλίμακας γκρίζου έχει τη μορφή

$$D_B = f\ (D_A) = a\ D_A + b\ ,$$

όπου D_B είναι το επίπεδο γκρίζου του σημείου εξόδου που αντιστοιχεί σε σημείο εισόδου με επίπεδο γκρίζου D_A. Προφανώς, αν a=1 και b=0, έχουμε τον ταυτοτικό σημειακό τελεστή που απλώς αντιγράφει την εικόνα A(m,n) στην εικόνα B(m,n). Αν η παράμετρος a είναι μεγαλύτερη της μονάδας (a>1), η αντίθεση (contrast) θα αυξηθεί στην εικόνα εξόδου. Για a<1, η αντίθεση θα μειωθεί, ενώ για a=1 και μη μηδενικό b, ο τελεστής απλώς ολισθαίνει τις τιμές των επιπέδων γκρίζου όλων των εικονοστοιχείων επάνω ή κάτω αφήνοντας την αντίθεση αμετάβλητη. Το αποτέλεσμα του τελευταίου τελεστή είναι να γίνει ολόκληρη η εικόνα περισσότερο ή λιγότερο σκούρα. Αν η παράμετρος a είναι αρνητική, οι σκούρες περιοχές γίνονται περισσότερο φωτεινές και οι φωτεινές περισσότερο σκούρες σε τρόπο ώστε δημιουργείται το **συμπλήρωμα (complement)** της εικόνας εισόδου.

- **Μη γραμμικοί μονοτονικοί σημειακοί τελεστές** : Συναρτήσεις μετασχηματισμού κλίμακας γκρίζου με πεπερασμένη θετική κλίση παντού διατηρούν τη βασική εμφάνιση μιας εικόνας, αλλά δεν είναι τόσο περιοριστικές όσο οι γραμμικοί τελεστές. Οι μη γραμμικοί σημειακοί τελεστές ταξινομούνται με βάση την επίδραση τους στο κέντρο του διαστήματος επιπέδων γκρίζου.

Η συνάρτηση μετασχηματισμού κλίμακας γκρίζου

$$f\ (x) = x + Cx\ (D_m\text{-}x)$$

επηρεάζει το επίπεδο γκρίζου των εικονοστοιχείων του κέντρου του διαστήματος επιπέδων γκρίζου αλλά αφήνει τα σκοτεινά και τα φωτεινά εικονοστοιχεία αναλλοίωτα. Πιο συγκεκριμένα, η παράμετρος D_m προσδιορίζει το μέγιστο επίπεδο γκρίζου, ενώ η παράμετρος C καθορίζει το μέγεθος της αύξησης (C>0) ή της μείωσης (C<0) των κεντρικών επιπέδων γκρίζου.

Άλλες συναρτήσεις μετασχηματισμού κλίμακας γκρίζου αυξάνουν την αντίθεση σε αντικείμενα του κέντρου του διαστήματος τιμών σε βάρος των φωτεινών και των σκοτεινών αντικειμένων. Τέτοιες συναρτήσεις έχουν σιγμοειδές σχήμα με κλίση μεγαλύτερη της μονάδας στο κέντρο και μικρότερη της μονάδας προς τα άκρα του διαστήματος επιπέδων γκρίζου. Ένα παράδειγμα αποτελεί η ημιτονοειδής συνάρτηση

$$f(x)=\frac{D_m}{2}\left\{1+\frac{1}{\sin(\alpha\frac{\pi}{2})}\sin\left[\alpha\pi\left(\frac{x}{D_m}-\frac{1}{2}\right)\right]\right\}, \qquad o<\alpha<1,$$

όπου [0, D_m] είναι το διάστημα επιπέδων γκρίζου στο οποίο το ιστόγραμμα είναι μη μηδενικό. Η παράμετρος α καθορίζει την επίδραση του τελεστή στο κέντρο του διαστήματος επιπέδων γκρίζου με μεγαλύτερη τιμή να υποδηλώνει ισχυρότερη επίδραση.

Συναρτήσεις μετασχηματισμού κλίμακας γκρίζου που έχουν κλίση μικρότερη της μονάδας στο κέντρο και μεγαλύτερη της μονάδας κοντά στα άκρα του διαστήματος επιπέδων γκρίζου μειώνουν την αντίθεση στο κέντρο και την αυξάνουν στα φωτεινά και τα σκοτεινά αντικείμενα. Ένα παράδειγμα αποτελεί η συνάρτηση

$$f(x)=\frac{D_m}{2}\left\{1+\frac{1}{\tan(\alpha\frac{\pi}{2})}\tan\left[\alpha\pi\left(\frac{x}{D_m}-\frac{1}{2}\right)\right]\right\}, \qquad o<\alpha<1,$$

όπου οι παράμετροι D_m και α παίζουν τον ίδιο ρόλο όπως προηγουμένως.

Οι παραπάνω σημειακοί τελεστές χρησιμοποιούνται συχνά για τη μείωση των περιορισμών τόσο κατά την ψηφιοποίηση εικόνων όσο και κατά την παρουσίαση τους. Πιο συγκεκριμένα, οι σημειακοί τελεστές χρησιμοποιούνται στις παρακάτω διαδικασίες :

- **Φωτομετρική βαθμονόμηση (photometric calibration) :** Είναι συχνά επιθυμητό τα επίπεδα γκρίζου μιας ψηφιακής εικόνας να αντανακλούν κάποια φυσική ιδιότητα, όπως της έντασης του φωτός ή της οπτικής πυκνότητας. Οι σημειακοί τελεστές το επιτυγχάνουν αυτό αφαιρώντας την επίδραση μη γραμμικοτήτων του αισθητηρίου. Αν, για παράδειγμα, μια εικόνα έχει ψηφιοποιηθεί από συσκευή με μη γραμμική απόκριση στην ένταση του φωτός, τότε ένας σημειακός τελεστής μπορεί να χρησιμοποιηθεί για να μετασχηματίσει την κλίμακα γκρίζου σε τρόπο ώστε τα επίπεδα γκρίζου να αναπαριστάνουν ίσες αυξήσεις έντασης του φωτός.
- **Ενίσχυση αντίθεσης (contrast enhancement) :** Σε μερικές ψηφιακές εικόνες, τα ενδιαφέροντα χαρακτηριστικά καταλαμβάνουν μόνο ένα σχετικά περιορισμένο διάστημα της κλίμακας γκρίζου. Ένας σημειακός τελεστής θα μπορούσε να χρησιμοποιηθεί για να ενισχύσει την αντίθεση των ενδιαφερόντων

χαρακτηριστικών ώστε να καταλαμβάνουν μεγαλύτερο μέρος του διαστήματος επιπέδων γκρίζου.

- **Βαθμονόμηση παρουσίασης (display calibrations) :** Μερικές συσκευές παρουσίασης έχουν ένα διάστημα επιπέδων γκρίζου στο οποίο τα χαρακτηριστικά εικόνων παρουσιάζονται πιο ευκρινή. Πιο σκοτεινά και πιο φωτεινά χαρακτηριστικά με την ίδια αντίθεση στην εικόνα δεν εμφανίζονται τόσο ευκρινώς σε τέτοιες συσκευές. Με τη χρήση σημειακού τελεστή, μπορεί να εξασφαλιστεί ότι τα ενδιαφέροντα χαρακτηριστικά εμπίπτουν στο μέγιστο ορατό διάστημα της συσκευής. Η χρήση ενός τέτοιου σημειακού τελεστή εμπίπτει στα πλαίσια διαδικασιών βαθμονόμησης παρουσίασης.
- **Υψομετρικές γραμμές (contour lines) :** Με τη χρήση ενός σημειακού τελεστή, μπορούν να προστεθούν υψομετρικές γραμμές σε μια εικόνα. Μια τέτοια διαδικασία μπορεί να είναι χρήσιμη στον προσδιορισμό των συνόρων μεταξύ αντικειμένων ή στη δημιουργία μασκών.
- **Αποκοπή (clipping) :** Οι ψηφιακές εικόνες αποθηκεύονται σε ακέραια μορφή, συνήθως του ενός byte. Επομένως, το διάστημα των διαθέσιμων επιπέδων γκρίζου είναι περιορισμένο. Για παράδειγμα, για εικόνες των 8 bits, το επίπεδο γκρίζου πρέπει να περιοριστεί στο διάστημα 0-255 προτού αποθηκευτεί κάθε εικονοστοιχείο. Υποθέτουμε ότι κάθε σημειακός τελεστής ακολουθείται από ένα βήμα που θέτει αρνητικές τιμές σε μηδέν και περιορίζει θετικές τιμές στο μέγιστο επίπεδο γκρίζου.

2.3. Αλγεβρικοί τελεστές

Ένας **αλγεβρικός τελεστής (algebraic operator)** παράγει μια εικόνα εξόδου υπολογίζοντας το εικονοστοιχείο – προς - εικονοστοιχείο άθροισμα, διαφορά, γινόμενο ή πηλίκο δύο εικόνων. Στην περίπτωση αθροισμάτων ή γινομένων, είναι δυνατό να εμπλέκονται περισσότερες από δύο εικόνες.

Οι τέσσερεις αλγεβρικοί τελεστές επεξεργασίας εικόνας εκφράζονται μαθηματικά ως

- $C(m,n)=A(m,n)+B(m,n)$ (άθροισμα)
- $C(m,n)=A(m,n)-B(m,n)$ (διαφορά)
- $C(m,n)=A(m,n)\ B(m,n)$ (γινόμενο)
- $C(m,n)=A(m,n)/B(m,n)$ (πηλίκο),

όπου A(m,n) και B(m,n) είναι οι εικόνες εισόδου και C(m,n) είναι η εικόνα εξόδου. Οι παραπάνω αλγεβρικοί τελεστές μπορούν να συνδυαστούν σχηματίζοντας περισσότερο πολύπλοκες αλγεβρικές εκφράσεις.

Οι βασικοί αλγεβρικοί τελεστές βρίσκουν ποικίλες εφαρμογές στην επεξεργασία και ανάλυση εικόνων με υπολογιστή. Μερικές από αυτές είναι οι επόμενες :

- **Ελάττωση θορύβου (noise reduction) :** Ο αλγεβρικός τελεστής της άθροισης μπορεί να χρησιμοποιηθεί για την ελάττωση τυχαίου προσθετικού θορύβου σε πολλαπλές εικόνες της ίδιας στάσιμης σκηνής. Με την υπέρθεση όλων των εικόνων, τα στάσιμα χαρακτηριστικά τους παραμένουν αναλλοίωτα, ενώ τα χαρακτηριστικά του θορύβου, τυχαία και ασυσχέτιστα από εικόνα σε εικόνα, αλληλοαναιρούνται.

- **Διπλή έκθεση (double exposure) :** Η άθροιση εικόνων μπορεί επίσης να χρησιμοποιηθεί για να υπερθέσει το περιεχόμενο μιας εικόνας στο περιεχόμενο μιας άλλης εικόνας, δημιουργώντας έτσι ένα αποτέλεσμα διπλής έκθεσης.
- **Αφαίρεση παρασκηνίου (background subtraction) :** Ο αλγεβρικός τελεστής της αφαίρεσης εικόνων μπορεί να χρησιμοποιηθεί για την απομάκρυνση ανεπιθύμητων προσθετικών χαρακτηριστικών από μια εικόνα. Τέτοια χαρακτηριστικά αποτελούν αργά μεταβαλλόμενες σκιάσεις στο παρασκήνιο, περιοδικός θόρυβος ή οποιαδήποτε άλλη μορφή θορύβου με γνωστή τιμή σε κάθε εικονοστοιχείο.
- **Ανίχνευση κίνησης (motion detection) :** Η αφαίρεση εικόνων είναι επίσης χρήσιμη στην ανίχνευση διαφόρων μεταξύ εικόνων της ίδιας σκηνής. Για παράδειγμα, ανίχνευση κίνησης μπορεί να γίνει με αφαίρεση διαδοχικών εικόνων μιας σκηνής.
- **Ανίχνευση ακμών (edge detection) :** Η αφαίρεση εικόνων μπορεί να χρησιμοποιηθεί για τον υπολογισμό μιας σημαντικής παραγώγου της εικόνας, της **βαθμίδας (gradient)**. Η βαθμίδα ορίζεται ως εξής : Έστω μια δεδομένη βαθμωτή συνάρτηση f(x,y) και ένα καρτεσιανό σύστημα συντεταγμένων με μοναδιαία διανύσματα $\hat{x}$ και $\hat{y}$ κατά τους άξονες x και y, αντίστοιχα. Η βαθμίδα της συνάρτησης f(x,y) ορίζεται ως η διανυσματική συνάρτηση :

$$\nabla f(x,y) = \hat{x}\frac{\theta f(x,y)}{\theta x} + \hat{y}\frac{\theta f(x,y)}{\theta y} \;,$$

όπου το σύμβολο ∇ υποδηλώνει τον τελεστή της βαθμίδας. Το διάνυσμα ∇f(x,y) «δείχνει» προς την κατεύθυνση της μέγιστης ανοδικής κλίσης και το μήκος του ισούται με την τιμή της κλίσης. Δηλαδή, η βαθμωτή συνάρτηση

$$|\nabla f| = \sqrt{\left(\frac{\theta f}{ux}\right)^2 + \left(\frac{\theta f}{\theta u}\right)^2}$$

υποδηλώνει πόσο «απότομη» είναι η κλίση της συνάρτησης f σε κάθε σημείο. Η βαθμίδα λαμβάνει μεγάλες απόλυτες τιμές σε περιοχές μεγάλης κλίσης όπως οι ακμές αντικειμένων, και επομένως ο υπολογισμός της αποτελεί μια δυνητική μέθοδο ανίχνευσης και εντοπισμού ακμών αντικειμένων σε εικόνες.

- **Επεξεργασία εικόνων με μάσκες :** Μάσκες μπορούν να χρησιμοποιηθούν για να καταπιέσουν επιλεκτικά το περιεχόμενο εικόνων. Μια μάσκα έχει τιμή «ένα» σε περιοχές της εικόνας που θα παραμείνουν αναλλοίωτες κα «μηδέν» εκεί όπου το περιεχόμενο της εικόνας θα καταπιεστεί. Πολλαπλασιασμός της εικόνας με τη μάσκα θα μηδενίσει το περιεχόμενο της αντίστοιχης περιοχής.
- **Διόρθωση χωρικά μεταβλητών ψηφιοποιητών :** Ο αλγεβρικός τελεστής της διαίρεσης χρησιμοποιείται για να διορθώσει τα αποτελέσματα ψηφιοποιητών των οποίων η ευαισθησία στο φως μεταβάλλεται από σημείο σε σημείο στην εικόνα.

2.4. Γεωμετρικοί τελεστές

Οι **γεωμετρικοί τελεστές (geometric operators)** επηρεάζουν τις χωρικές σχέσεις μεταξύ των αντικειμένων μιας εικόνας. Κατά κάποιο τρόπο, αυτοί οι τελεστές αλλάζουν τις θέσεις εικονοστοιχείων μέσα στην εικόνα. Το αποτέλεσμα τους είναι παρόμοιο με εκτύπωση της εικόνας σε ελαστικό φύλλο και ελαστική παραμόρφωση του φύλλου. Στην πραγματικότητα, οι γεωμετρικοί τελεστές μπορούν να έχουν πολύ γενικότερη μορφή αν κάθε σημείο της εικόνας εισόδου επιτρέπεται να μετακινηθεί εντελώς ελεύθερα. Επομένως, κάποιοι περιορισμοί πρέπει να επιβληθούν ώστε το περιεχόμενο της εικόνας να μην περιπλέκεται σε βαθμό καταστροφικό.

Ένας γεωμετρικός τελεστής απαιτεί δύο διαδικασίες για τον προσδιορισμό του, πρώτα, πρέπει να προσδιοριστεί ο αλγόριθμος του ίδιου του χωρικού μετασχηματισμού, δηλαδή η κίνηση του κάθε εικονοστοιχείου από την αρχική στην τελική του θέση στην εικόνα. Στη συνέχεια, πρέπει να προσδιοριστεί ένας αλγόριθμος για παρεμβολή επιπέδων γκρίζου. Η δεύτερη αυτή αναγκαιότητα προκύπτει από το γεγονός ότι ακέραιες θέσεις m,n στην εικόνα εισόδου απεικονίζονται γενικά σε κλασματικές θέσεις στην εικόνα εξόδου.

Στις περισσότερες εφαρμογές, είναι επιθυμητό να διατηρηθεί η συνέχεια καμπυλόγραμμων χαρακτηριστικών και η συνεκτικότητα των αντικειμένων στην εικόνα. Στην αντίθετη περίπτωση, ένας λιγότερο περιορισμένος χωρικός μετασχηματισμός θα έτεινε να σπάσει γραμμές και αντικείμενα και να σκορπίσει το περιεχόμενο της εικόνας.

Για τον προσδιορισμό του χωρικού μετασχηματισμού, καθορίζεται μια μαθηματική σχέση που συνδέει σημεία της εικόνας εισόδου με σημεία της εικόνας εξόδου. Στη γενική περίπτωση, ένας γεωμετρικός τελεστής ικανοποιεί μια σχέση της μορφής

$$g(x,y) = f[\alpha(x,y),\ b(x,y)],$$

όπου f(x,y) και g(x,y) είναι οι εικόνες εισόδου και εξόδου, αντίστοιχα, και οι συναρτήσεις α(x,y) και b(x,y) προσδιορίζουν μονοσήμαντα το χωρικό μετασχηματισμό. Αν οι συναρτήσεις α(x,y) και b(x,y) είναι συνεχείς, τότε διατηρείται η συνεκτικότητα της εικόνας κατά το μετασχηματισμό.

Από την παραπάνω εξίσωση, γίνεται σαφές ότι για τον υπολογισμό των τιμών επιπέδων γκρίζου στην εικόνα εξόδου g απαιτούνται γενικά και τιμές της εικόνας εισόδου f σε κλασματικές (μη ακέραιες) θέσεις. Ελλείψει αυτών των τιμών, χρησιμοποιείται κάποιος αλγόριθμος παρεμβολής για την εκτίμηση τους με βάση τις γνωστές τιμές της εικόνας εισόδου f σε κλασματικές θέσεις. Μερικές από τις πιο κοινές μεθόδους παρεμβολής είναι οι επόμενες :

- **Παρεμβολή πλησιέστερου γείτονα (nearest neighbor intepolation) :** Η παρεμβολή πλησιέστερου γείτονα είναι γνωστή και ως **παρεμβολή μηδενικής τάξης (zeroth-order interpolation)** και αποτελεί το απλούστερο σχήμα παρεμβολής. Το επίπεδο γκρίζου εικονοστοιχείου εξόδου υπολογίζεται ως το επίπεδο γκρίζου του εικονοστοιχείου εισόδου που βρίσκεται πλησιέστερα στη θέση του εικονοστοιχείου εξόδου. Αυτή η διαδικασία είναι υπολογιστικά απλή και σε πολλές περιπτώσεις επαρκής. Είναι, όμως, δυνατό να εισάγει **τεχνητές δομές (artifacts)** σε εικόνες με εκλεπτυσμένη δομή στις οποίες τα επίπεδα γκρίζου αλλάζουν σημαντικά από εικονοστοιχείο σε εικονοστοιχείο.

- **Διγραμμική παρεμβολή :** Η διγραμμική παρεμβολή ή **παρεμβολή πρώτης τάξης (first-order interpolation)** έχει ελαφρώς αυξημένη προγραμματιστική πολυπλοκότητα και χρόνο εκτέλεσης σε σχέση με την παρεμβολή μηδενικής τάξης, αλλά παράγει καλύτερα αποτελέσματα. Η βασική ιδέα της διγραμμικής παρεμβολής έγκειται στο εξής. Έστω f(x,y) η συνάρτηση που πρέπει να παρεμβληθεί. Θέτουμε (διγραμμική εξίσωση παρεμβολής)

$$f(x, y) \approx ax + by + cxy + d$$

και επιλέγουμε τους συντελεστές a, b, c και d έτσι ώστε η παραπάνω προσεγγιστική ισότητα να γίνεται ακριβής ισότητα σε τέσσερεις διαφορετικές θέσεις (εικονοστοιχεία). Όταν διαδοχικές περιοχές τεσσάρων εικονοστοιχείων παρεμβάλλονται με τη διγραμμική εξίσωση, οι επιφάνειες που προκύπτουν ταιριάζουν ως προς το πλάτος στα σύνορα των περιοχών, αλλά όχι και ως προς την κλίση. Ετσι, μια επιφάνεια που παράγεται από διγραμμική παρεμβολή κατά τμήματα είναι συνεχής, αλλά με ασυνεχείς παραγώγους στα σύνορα των περιοχών παρεμβολής.

- **Παρεμβολή υψηλότερης τάξης (higher-order interpolation) :** Το εξαμβλυντικό αποτέλεσμα της διγραμμικής παρεμβολής μπορεί να επηρεάσει αρνητικά τις λεπτομέρειες μιας εικόνας, ειδικά αν συμπεριλαμβάνεται μεγέθυνση. Σε άλλες εφαρμογές, οι ασυνέχειες στην κλίση που προκαλεί η διγραμμική παρεμβολή παράγει ανεπιθύμητα αποτελέσματα. Μια λύση αυξημένου υπολογιστικού κόστους αποτελούν οι παρεμβολές υψηλότερης τάξης. Η γενική ιδέα είναι να χρησιμοποιηθεί κάποια εξίσωση παρεμβολής και οι άγνωστοι συντελεστές της να υπολογιστούν από ίσο αριθμό θέσεων (εικονοστοιχείων). Παραδείγματα τύπων παρεμβολής αποτελούν κυβικές splines, συναρτήσεις Legendre κλπ.

Οι γεωμετρικοί τελεστές βρίσκουν σημαντικές εφαρμογές στην επεξεργασία και ανάλυση εικόνων με υπολογιστή. Μερικές από τις εφαρμογές αυτές είναι οι επόμενες :

- **Γεωμετρική βαθμονόμηση (geometric calibration) :** Μια σημαντική εφαρμογή των γεωμετρικών τελεστών είναι η αφαίρεση γεωμετρικών παραμορφώσεων σε ψηφιακές εικόνες που έχουν προκληθεί από γεωμετρικές ατέλειες του αισθητηρίου. Τέτοιες παραμορφώσεις είναι πολύ συνηθισμένες σε εικόνες που έχουν ληφθεί από δορυφόρους και εναέρια πλαγιοσκοπικά συστήματα ραντάρ.
- **Επανόρθωση εικόνας (image rectification) :** Ορισμένα συστήματα απεικόνισης χρησιμοποιούν καμπυλόγραμμα συστήματα συντεταγμένων εικονοστοιχείων. Για τη σωστή παρουσίαση εικόνων ψηφιοποιημένων με τέτοια συστήματα είναι αναγκαία η επανόρθωσή τους, δηλαδή ο μετασχηματισμός τους σε ορθογώνιες συντεταγμένες εικονοστοιχείων.
- **Καταγραφή εικόνας (image registration) :** Μια άλλη εφαρμογή γεωμετρικών τελεστών αποτελεί η καταγραφή παρομοίων εικόνων με σκοπό τη

σύγκριση. Τυπικό παράδειγμα αποτελεί η αφαίρεση εικόνων για ανίχνευση κίνησης ή αλλαγής. Αν όμοιες εικόνες παρουσιαστούν ελαφρώς μετατοπισμένες, στη διαφορά τους θα είναι εμφανής μια συνιστώσα μερικής παραγώγου. Ως αποτέλεσμα, εύκολα θα μπορούσαν να καλυφθούν σημαντικές διαφορές στις εικόνες. Αν εικόνες ενός στασίμου αντικειμένου ψηφιοποιηθούν από ένα σταθερό αισθητήριο, τότε λαμβάνονται με **αντιστοιχία καταγραφής (registration correspondence)**. Αλλιώς, πριν την αφαίρεση πρέπει να εφαρμοστεί μια λειτουργία καταγραφής. Η διαδικασία καταγραφής αποτελείται συνήθως από τελεστές μετατόπισης αλλά και περιστροφής.

- **Μετατροπή διαμόρφωσης εικόνας (image format conversion) :** Οι γεωμετρικοί τελεστές χρησιμοποιούνται πολλές φορές απλώς και μόνο για να μετατρέψουν τη διαμόρφωση μιας εικόνας σε άλλη περισσότερο βολική για ερμηνεία.
- **Προβολή χαρτών (map projection) :** Μια άλλη πολύ σημαντική εφαρμογή των γεωμετρικών τελεστών είναι στην προβολή εικόνων με σκοπό τη χαρτογράφηση. Στην περίπτωση αυτή, οι γεωμετρικοί τελεστές εφαρμόζονται ώστε εικόνες που έχουν ληφθεί από δορυφόρο, για παράδειγμα, να μετασχηματιστούν στη μορφή με την οποία θα εμφανίζονται στο χάρτη.
- **Μορφοποίηση (morphing) :** Οι γεωμετρικοί τελεστές χρησιμοποιούνται τα τελευταία χρόνια για την επίτευξη ειδικών εφφέ που έγιναν δημοφιλή στον κινηματογράφο και στην τηλεόραση. Η μορφοποίηση είναι μια τεχνική που επιτρέπει το σταδιακό μετασχηματισμό ενός αντικειμένου σε ένα άλλο.

3. Κατάτμηση και ανάλυση εικόνων σε περιοχές

3.1. Γενικά

Στο προηγούμενο κεφάλαιο, εξετάστηκαν κυρίως μέθοδοι για τη βελτίωση της παρουσίασης εικόνων. Τέτοιες μέθοδοι αποσκοπούν στην απλοποίηση εικόνων, με ανίχνευση ή ενίσχυση ακμών τους για παράδειγμα, και εμπίπτουν στο αντικείμενο της επεξεργασίας εικόνων. Στη συνέχεια του βιβλίου, θα εξεταστούν τεχνικές συμπληρωματικές της επεξεργασίας εικόνων που αποσκοπούν στην **αντίληψη (perception)** και **κατανόηση (understanding)** του περιεχομένου εικόνων κα αναφέρονται ως ανάλυση εικόνων.

Στο κεφάλαιο αυτό, εξετάζονται τεχνικές για απλοποίηση εικόνων με διαμερισμό τους σε ένα σύνολο από ασύνδετες περιοχές. Οι τεχνικές αυτές αναφέρονται συνολικά ως **κατάτμηση εικόνων (image segmentation)** ή **ανάλυση σε περιοχές (region analysis)**. Στην απλούστερη περίπτωση, κάθε περιοχή αποτελείται από ένα σύνολο συνδεδεμένων εικονοστοιχείων του ιδίου επιπέδου γκρίζου. Για την παρουσίαση των τεχνικών κατάτμησης εικόνων, απαιτείται αυστηρός ορισμός της έννοιας των **συνδεδεμένων (connected)** εικονοστοιχείων. Πιο συγκεκριμένα, χρειάζεται να οριστεί αν ένα εικονοστοιχείο είναι συνδεδεμένο με όλα τα οκτώ εικονοστοιχεία που το περιβάλλουν (**οκταπλή σύνδεση, eight-connectedness**) ή μόνο με τα τέσσερα εικονοστοιχεία με τα οποία έχει κοινή ακμή (**τετραπλή σύνδεση, four-connectedness**).

Για να γίνει κατανοητή η σημασία του παραπάνω ερωτήματος, θεωρούμε την επόμενη απλή δυαδική εικόνα, όπου τα μη σημειωμένα εικονοστοιχεία αποτελούν εικονοστοιχεία του παρασκηνίου με τιμή μηδέν. Στην αρχή, εξετάζουμε την περίπτωση οκταπλής σύνδεσης. Εφόσον δύο εικονοστοιχεία με μια μόνο κοινή κορυφή θεωρούνται συνδεδεμένα, το σχήμα είναι συνδεδεμένο. Το πρόβλημα έγκειται στο ότι, για τον ίδιο λόγο, και το παρασκήνιο είναι συνδεδεμένο. Η «οπή» στο κέντρο της εικόνας δεν διαχωρίζεται από το παρασκήνιο που την «περιβάλλει». Επομένως, αντιμετωπίζουμε ένα σχήμα με δομή κελύφους χωρίς εσωτερικό. Ας εξετάσουμε τώρα την περίπτωση της τετραπλής σύνδεσης. Το σχήμα δεν είναι πλέον συνδεδεμένο. Επίσης, ούτε το παρασκήνιο είναι συνδεδεμένο. Επομένως αντιμετωπίζουμε ένα μη συνδεδεμένο σχήμα με διακριτό εσωτερικό.

		I	I		
	I			I	
	I			I	
		I	I		

Το παραπάνω παράδειγμα παρουσιάζει μια βασική ιδιότητα του τετραγωνικού πλέγματος, δηλαδή τη δυνατότητα ύπαρξης συνδεδεμένων εικονοστοιχείων που έχουν κοινή μόνο μια κορυφή. Αντίθετα προς το τετραγωνικό πλέγμα, το εξαγωνικό πλέγμα δεν έχει αυτή την ιδιότητα, αφού κάθε ζευγάρι εξαγωνικών εικονοστοιχείων με μια κοινή κορυφή αναγκαστικά μοιράζεται και μια κοινή ακμή. Πάντως, στο βιβλίο αυτό η σύνδεση εικονοστοιχείων θα θεωρείται πάντα ως τετραπλή.

Με βάση τον προηγούμενο ορισμό, μπορούμε τώρα να επιστρέψουμε στο ερώτημα της ανάλυσης εικόνων με διαμερισμό τους σε περιοχές. Ένα σύνολο εικονοστοιχείων θα ονομάζεται **στοιχειώδης συνδεδεμένη περιοχή (elementary connected region)** αν

1. όλα τα εικονοστοιχεία του συνόλου έχουν την ίδια τιμή επιπέδου γκρίζου,
2. δύο οποιαδήποτε εικονοστοιχεία του συνόλου συνδέονται με μια αλυσίδα συνδεδεμένων εικονοστοιχείων κάθε ένα από τα οποία ανήκει στο σύνολο και
3. οποιοδήποτε γνήσιο υπερσύνολο εικονοστοιχείων δεν ικανοποιεί τις προηγούμενες δύο συνθήκες.

Από μόνη της η δεύτερη συνθήκη αποτελεί τον ορισμό **της συνδεδεμένης περιοχής (connected region)**. Η τρίτη συνθήκη εξασφαλίζει το να είναι οι στοιχειώδεις συνδεδεμένες περιοχές όσο το δυνατό μεγαλύτερες.

Ο διαμερισμός μιας εικόνας σε στοιχειώδεις συνδεδεμένες περιοχές που ικανοποιούν και τις τρεις παραπάνω συνθήκες αναδεικνύει τόσο τις «σημαντικές» περιοχές που ορίζονται από πλευρές γεωμετρικών αντικειμένων όσο και ανεπιθύμητες περιοχές που οφείλονται σε ασήμαντες αλλαγές της σκίασης. Για να περιοριστεί η ανάδειξη ανεπιθύμητων περιοχών, μπορεί κανείς είτε να χαλαρώσει την πρώτη συνθήκη είτε να προχωρήσει σε συγχώνευση περιοχών. Στην πρώτη περίπτωση, συνδεδεμένα εικονοστοιχεία θα τοποθετηθούν στην ίδια στοιχειώδη συνδεδεμένη περιοχή αν η διαφορά των επιπέδων γκρίζου τους είναι μικρότερη από κάποιο προκαθορισμένο κατώφλι. Γενικά, αυτή η διαδικασία επιτρέπει σε συνδεδεμένα εικονοστοιχεία αυθαιρέτων τιμών επιπέδου γκρίζου να βρεθούν στην ίδια στοιχειώδη συνδεδεμένη περιοχή. Εναλλακτικά, μπορούμε να σχηματίσουμε πρώτα στοιχειώδεις συνδεδεμένες περιοχές και, χρησιμοποιώντας κάποιο κριτήριο, να αποφασίσουμε πότε να συγχωνεύσουμε δύο περιοχές με κοινό σύνορο. Το γενικό αποτέλεσμα αυτής της διαδικασίας εξαρτάται από τη συγκεκριμένη υλοποίηση του αλγορίθμου.

Πέρα από τα προηγούμενα, μπορούν να κατασκευαστούν και άλλα κριτήρια για τον προσδιορισμό και τη συγχώνευση περιοχών. Ένα παράδειγμα αποτελεί ο επόμενος αλγόριθμος. Εστωσαν R_1 και R_2 δύο περιοχές με κοινό σύνορο. Έστω L το μήκος του κομματιού εκείνου του κοινού συνόρου που διαχωρίζει τα εικονοστοιχεία των οποίων τα επίπεδα γκρίζου διαφέρουν λιγότερο από d. Θα συγχωνεύουμε δύο περιοχές αν, για δεδομένο κατώφλι t,

$$\max\left\{\frac{L}{P_1}, \frac{L}{P_2}\right\} \geq t,$$

όπου P_1 και P_2 είναι το μήκος της περιμέτρου των περιοχών R_1 και R_2, αντίστοιχα. Η ερμηνεία του αλγορίθμου είναι η εξής : Θεωρούμε ότι L είναι το αμφίβολο κομμάτι του κοινού συνόρου. Αν μια από τις περιοχές, η R_2 για παράδειγμα, περιβάλλεται κατά κάποιο τρόπο από την R_1 τότε το κοινό σύνορο θα είναι ένα μεγάλο κλάσμα του P_2. Επιπλέον, αν αυτό το κοινό σύνορο είναι αμφίβολο, τότε το μήκος L θα είναι ένα μεγάλο κλάσμα του P_2 και οι δύο περιοχές θα συγχωνευθούν. Ως αποτέλεσμα, ο αλγόριθμος θα τείνει να κάνει την περιοχή R_1 πιο ομαλή.

Οι βασικές έννοιες της κατάτμησης εικόνων μπορούν να επεκταθούν έτσι ώστε να συμπεριλάβουν και άλλες ιδιότητες εικόνων επιπρόσθετα της έντασης. Μια από τις πιο ενδιαφέρουσες δυνατότητες αποτελεί ο διαμερισμός εικόνων σε περιοχές σύμφωνα με το χρώμα. Στην περίπτωση αυτή, η ανάλυση σε περιοχές ξεκινάει με το σχηματισμό στοιχειωδών συνδεδεμένων περιοχών εικονοστοιχείων με ταυτόσημες τιμές κόκκινου, πράσινου και μπλε χρώματος. Στη συνέχεια, μπορεί να γίνει συγχώνευση περιοχών σύμφωνα με κάποιο κριτήριο ανάλογο προς αυτά που χρησιμοποιούνται στην ανάλυση ασπρόμαυρων εικόνων.

Σε σύγκριση με την ανθρώπινη εμπειρία, δεν μπορεί να θεωρηθεί ότι η διαδικασία της κατάτμησης εικόνων είναι συναφής με κάποια ανθρώπινη λειτουργία. Όταν ένας ανθρώπινος παρατηρητής αντικρίζει μια σκηνή, η επεξεργασία που λαμβάνει χώρα στο σύστημα όρασης ουσιαστικά εκτελεί αυτόματα κατάτμηση γι' αυτόν. Η επεξεργασία αυτή γίνεται τόσο αποτελεσματικά που κανείς δεν βλέπει μια περίπλοκη σκηνή, αλλά κάτι που το αντιλαμβάνεται ως μια συλλογή αντικειμένων. Με την ψηφιακή επεξεργασία, όμως, πρέπει να απομονωθούν τα αντικείμενα σε μια εικόνα με επίπονη διαδικασία διαμερισμού της εικόνας σε σύνολα εικονοστοιχείων, κάθε ένα από τα οποία αποτελεί την εικόνα ενός αντικειμένου. Η διαδικασία αυτή είναι μη τετριμμένη και πολύ σημαντική στην ανάλυση εικόνων με υπολογιστή.

Η κατάτμηση εικόνων μπορεί να προσεγγιστεί από τρεις διαφορετικές φιλοσοφικές πλευρές. Η πρώτη προσέγγιση ονομάζεται **προσέγγιση περιοχών (region approach)** και αναθέτει κάθε εικονοστοιχείο σε ένα συγκεκριμμένο αντικείμενο ή περιοχή. Η δεύτερη προσέγγιση ονομάζεται **προσέγγιση συνόρων (boundary approach)** και προσπαθεί να εντοπίσει τα σύνορα μεταξύ περιοχών. Τέλος, η τρίτη προσέγγιση ονομάζεται **προσέγγιση ακμών (edge approach)** και αποσκοπεί στην ταυτοποίηση εικονοστοιχείων ακμών και στη σύνδεσή τους για σχηματισμό συνόρων. Και οι τρεις προσεγγίσεις είναι χρήσιμες στην επίλυση του προβλήματος. Εδώ, θα εξετάσουμε διάφορες τεχνικές για απομόνωση αντικειμένων σε μια ψηφιακή εικόνα. Αν απομονωθούν τα αντικείμενα, μπορούν στη συνέχεια να μετρηθούν και να ταξινομηθούν. Τεχνικές για μέτρηση και ταξινόμηση θα εξεταστούν στο επόμενο κεφάλαιο.

3.2. Κατάτμηση εικόνων με κατωφλίωση

Η **κατωφλίωση (thresholding)** είναι μια ιδιαίτερα χρήσιμη τεχνική για την κατάτμηση εικόνων σκηνών που περιέχουν στερεά αντικείμενα πάνω σε ένα παρασκήνιο. Είναι υπολογιστικά απλή και ποτέ δεν αποτυγχάνει στον προσδιορισμό διακριτών περιοχών με κλειστά συνδεδεμένα σύνορα. Βασίζεται στο διαμερισμό εικόνων σε σύνολο εσωτερικών και εξωτερικών σημείων.

Κατά τη χρήση ενός κανόνα κατωφλίωσης για κατάτμηση εικόνων, όλα τα εικονοστοιχεία με επίπεδο γκρίζου πάνω από το κατώφλι ανατίθενται στο αντικείμενο, ενώ όλα τα εικονοστοιχεία με επίπεδο γκρίζου κάτω από το κάτωφλι εμπίπτουν έξω από το αντικείμενο. Το σύνορο είναι εκείνο το σύνολο εσωτερικών σημείων, κάθε ένα από τα οποία έχει τουλάχιστον ένα γείτονα έξω από το αντικείμενο.

Η κατωφλίωση λειτουργεί καλά όταν όλα τα ενδιαφέροντα αντικείμενα έχουν ομοιόμορφο επίπεδο γκρίζου στο εσωτερικό τους και βρίσκονται πάνω σε ένα παρασκήνιο διαφορετικού, αλλά ομοιόμορφου επιπέδου γκρίζου. Αν το αντικείμενο διαφέρει από το παρασκήνιο ως προς μια ιδιότητα άλλη από το επίπεδο γκρίζου (όπως, για παράδειγμα, την υφή), μπορεί πρώτα να χρησιμοποιηθεί ένας τελεστής που μετατρέπει αυτή την ιδιότητα σε επίπεδο γκρίζου και στη συνέχεια να χρησιμοποιηθεί κατωφλίωση του επιπέδου γκρίζου για να κατατμηθεί η εικόνα.

Για την επιλογή του κατωφλίου μπορούν να χρησιμοποιηθούν διάφορες μέθοδοι :

- **Ολική κατωφλίωση (global thresholding) :**Στην απλούστερη μορφή κατωφλίωσης, η τιμή του κατωφλίου διατηρείται σταθερή σε όλη την έκταση της εικόνας. Αν το επίπεδο γκρίζου του παρασκηνίου είναι σχετικά σταθερό και αν τα αντικείμενα έχουν όλα κατά προσέγγιση ίση αντίθεση με το παρασκήνιο, τότε ένα σταθερό ολικό κατώφλι θα δουλέψει καλά εφόσον επιλεγεί σωστά.
- **Προσαρμοστική κατωφλίωση (adaptive thresholding) :** Σε πολλές περιπτώσεις, το επίπεδο γκρίζου του παρασκηνίου και η αντίθεση των αντικειμένων μεταβάλλονται μέσα στην εικόνα. Σε τέτοιες περιπτώσεις, ένα κατώφλι που λειτουργεί καλά σε μια περιοχή της εικόνας μπορεί να λειτουργεί πολύ άσχημα σε άλλες περιοχές. Σε τέτοιες περιπτώσεις, είναι βολικό να χρησιμοποιείται ένα κατώφλι που αποτελεί μια αργά μεταβαλλόμενη συνάρτηση της θέσης στην εικόνα.
- **Βέλτιστη κατωφλίωση (optimal thresholding) :** Εκτός από την περίπτωση που το αντικείμενο στην εικόνα έχει εξαιρετικά απότομες πλευρές, η ακριβής τιμή του κατωφλίου έχει σημαντική επίδραση στη θέση του συνόρου και του συνολικού μεγέθους του αντικειμένου. Αυτό σημαίνει ότι επόμενες μετρήσεις μεγέθους, κυρίως του εμβαδού, είναι ευαίσθητες στην τιμή του κατωφλίου. Για το λόγο αυτό, χρειάζεται μια βέλτιστη μέθοδος επιλογής του κατωφλίου. Στο κεφάλαιο 2.1 είδαμε ότι αν το ιστόγραμμα επιπέδων γκρίζου μιας εικόνας εμφανίζει μια «κοιλάδα» στο κέντρο του και δύο «κορυφές» στα δύο άκρα του, τότε η επιλογή του κατωφλίου στο βαθύτερο σημείο της κοιλάδας είναι βέλτιστη. Εναλλακτικά, μια εικόνα μπορεί να διαμεριστεί σε ίσους τομείς σε πρώτη φάση. Στη συνέχεια, κάθε τομέας κατωφλιώνεται με κατώφλι που προσδιορίζεται από το βαθύτερο σημείο της κοιλάδας του ιστογράμματός του. Αν το ιστόγραμμα κάποιου τομέα δεν εμφανίζει κοιλάδα, ο τομέας αυτός αγνοείται. Τα αντικείμενα που αναγνωρίζονται με τη διαδικασία αυτή κατωφλιώνονται εκ νέου με κατώφλι που προσδιορίζεται στο μέσο της απόστασης του επιπέδου γκρίζου του παρασκηνίου του τομέα και του μέσου επιπέδου γκρίζου του αντικειμένου. Μια τέτοια διαδικασία προσπαθεί προσαρμοστικά να επιτύχει βέλτιστη κατωφλίωση.

- **Ο αλγόριθμος της μεσοποτάμιας κοιλάδας (watershed algorithm) :** Μια παραλλαγή της προσαρμοστικής κατωφλίωσης αποτελεί ο αλγόριθμος της παραποτάμιας κοιλάδας. Υποθέτουμε ότι τα αντικείμενα της εικόνας είναι χαμηλού επιπέδου γκρίζου σε αντίθεση με το παρασκήνιο που αποτελείται από εικονοστοιχεία με υψηλό επίπεδο γκρίζου. Ας υποθέσουμε ότι μια τυπική γραμμή σάρωσης διέρχεται διαμέσου δύο αντικειμένων που βρίσκονται σε κοντινή απόσταση μεταξύ τους.

Στην αρχή, η εικόνα κατωφλιώνεται σε χαμηλό επίπεδο γκρίζου, έτσι ώστε να κατατμηθεί στο σωστό αριθμό αντικειμένων αλλά με μικρά σύνορα. Στη συνέχεια, το κατώφλι ανυψώνεται σταδιακά, ένα επίπεδο κάθε φορά με αποτέλεσμα να μεγαλώνουν τα σύνορα των αντικειμένων. Η διαδικασία τερματίζεται όταν τα σύνορα των αντικειμένων αγγίξουν το ένα το άλλο χωρίς όμως να επιτευχθεί συγχώνευση. Με τον τρόπο αυτό, τα σημεία πρώτης επαφής αποτελούν και τα τελικά σύνορα μεταξύ των αντικειμένων. Ο αλγόριθμος της μεσοποτάμιας κοιλάδας μπορεί να επιλύσει το πρόβλημα της ορθής κατάτμησης για αντικείμενα που βρίσκονται πολύ κοντά μεταξύ τους και δεν διαχωρίζονται με μέθοδο ολικού κατωφλίου. Η ορθή έκβαση της κατωφλίωσης εξαρτάται, όμως, από την ορθή επιλογή τόσο του αρχικού όσο και του τελικού κατωφλίου.

3.3. Κατάτμηση εικόνων με βάση τη βαθμίδα

Οι μέθοδοι κατάτμησης εικόνων με κατωφλίωση βασίζεται στο διαμερισμό των εικόνων σε σύνολα εσωτερικών και εξωτερικών εικονοστοιχείων. Αντίθετα, οι μέθοδοι κατωφλίωσης με βάση τη βαθμίδα προσπαθούν να εντοπίσουν τις ακμές των συνόρων αντικειμένων με βάση τις υψηλές τιμές του μέτρου της βαθμίδας της εικόνας.

- **Ιχνηλάτηση συνόρου (boundary tracking) ή παρακολούθηση περιγράμματος (contour following)**. Ας θεωρήσουμεμε την εικόνα του μέτρου της βαθμίδας μιας εικόνας που περιέχει ένα μόνο αντικείμενο τοποθετημένο σε κάποιο παρασκήνιο. Η διαδικασία ιχνηλάτησης του συνόρου ξεκινάει με αναγνώριση του εικονοστοιχείου με το υψηλότερο επίπεδο γκρίζου, δηλαδή του εικονοστοιχείου με την υψηλότερη τιμή βαθμίδας στην αρχική εικόνα. Το σημείο αυτό αποτελεί το πρώτο συνοριακό σημείο, αφού πρέπει αναγκαστικά να ανήκει στο σύνορο του αντικειμένου. Αν υπάρχουν περισσότερα από ένα εικονοστοιχεία με τη μέγιστη τιμή επιπέδου γκρίζου, επιλέγεται αυθαίρετα ένα από αυτά. Στη συνέχεια, εξετάζουμε την 3Χ3 περιοχή εικονοστοιχείων με κέντρο το πρώτο συνοριακό σημείο.

Βρίσκουμε το εικονοστοιχείο της περιοχής με το μέγιστο επίπεδο γκρίζου και το θεωρούμε ως το δεύτερο συνοριακό σημείο. Αν υπάρχουν περισσότερα από ένα εικονοστοιχεία στην περιοχή με τη μέγιστη τιμή γκρίζου, τότε επιλέγουμε ένα από αυτά αυθαίρετα. Από εδώ και πέρα, επαναληπτικά βρίσκουμε επόμενα συνοριακά σημεία με βάση τα δύο τελευταία συνοριακά σημεία. Σχηματίζουμε την 3Χ3 περιοχή με κέντρο το πιο πρόσφατο συνοριακό σημείο και εξετάζουμε το εικονοστοιχείο που βρίσκεται διαμετρικά αντίθετα του προτελευταίου συνοριακού σημείου καθώς και τα δύο εικονοστοιχεία σε κάθε πλευρά του. Το επόμενο συνοριακό εικονοστοιχείο είναι εκείνο από τα τρία εικονοστοιχεία με το υψηλότερο επίπεδο γκρίζου. Αν και τα τρία ή δύο γειτονικά συνοριακά εικονοστοιχεία έχουν τη μέγιστη τιμή επιπέδου γκρίζου, διαλέγουμε το μεσαίο εικονοστοιχείο. Αν δύο μη συνεχόμενα εικονοστοιχεία έχουν τη μέγιστη τιμή επιπέδου γκρίζου, διαλέγουμε αυθαίρετα.

Στην περίπτωση εικόνων μιας κηλίδας χωρίς θόρυβο, ο παραπάνω αλγόριθμος θα ιχνηλατήσει το σύνορο που ορίζεται από τα μέγιστα της βαθμίδας. Ομως, η μέθοδος είναι ευαίσθητη ακόμα και σε μικρά ποσά θορύβου. Η επίδραση του θορύβου μπορεί να περιοριστεί με εξομάλυνση της εικόνας της βαθμίδας ή με τη χρήση ενός **εντόμου-ιχνηλάτη (tracking bug)**. Ενα έντομο-ιχνηλάτης είναι ένα ορθογώνιο παράθυρο με μεταβλητό προσανατολισμό που χρησιμοποιείται για τον υπολογισμό της μέσης τιμής της βαθμίδας των εικονοστοιχείων του. Ο αλγόριθμος του εντόμου-ιχνηλάτη είναι μια παραλλαγή της μεθόδου της ιχνηλάτησης συνόρου που περιγράφτηκε προηγουμένως. Όσο μεγαλύτερο το έντομο-ιχνηλάτης, τόσο μεγαλύτερη είναι η εξομάλυνση της εικόνας της βαθμίδας και τόσο μικρότερη η επίδραση του θορύβου. Όμως και η χρήση εντόμου-ιχνηλάτη δεν αποκλείει την απόκλιση του αλγορίθμου έξω από τα όρια της εικόνας ούτε εγγυάται τον προσδιορισμό ενός κλειστού τελικού συνόρου.

- **Κατωφλίωση εικόνων βαθμίδας (gradient image thresholding) :** Η κατωφλίωση μιας εικόνας βαθμίδας σε κάποιο ενδιάμεσο επίπεδο γκρίζου θέτει τόσο τα αντικείμενα όσο και το παρασκήνιο κάτω από το κατώφλι και τα περισσότερα εικονοστοιχεία των ακμών πάνω από το κατώφλι. Στη συνέχεια, επαναλαμβάνεται η κατωφλίωση με σταδιακά αυξανόμενο κατώφλι. Το αποτέλεσμα είναι τόσο τα αντικείμενα όσο και το παρασκήνιο να μεγαλώσουν. Όταν αγγίξουν το ένα τα άλλα, η διαδικασία τερματίζεται χωρίς συγχώνευση και τα σημεία επαφής ορίζουν το σύνορο. Η μέθοδος αυτή είναι εφαρμογή του αλγορίθμου της μεσοποτάμιας κοιλάδας στην εικόνα της βαθμίδας και αποφεύγει πολλά από τα προβλήματα των μεθόδων ιχνηλάτησης συνόρου.
- **Λαπλασιανή ανίχνευση ακμών (Laplacian edge detection) :** Η Λαπλασιανή είναι ένας διαφορικός τελεστής δεύτερης τάξης που δρα σε συναρτήσεις δύο ανεξαρτήτων μεταβλητών. Αν θεωρήσουμε τη συνάρτηση δύο μεταβλητών f(x,y), η Λαπλασιανή της ορίζεται ως

$$\nabla^2 f(x,y) = \frac{\partial^2}{\partial x^2} f(x,y) + \frac{\partial^2}{\partial y^2} f(x,y).$$

Η ψηφιακή υλοποίησή της γίνεται συνήθως με έναν από τους επόμενους συνελικτικούς πυρήνες (μάσκες) :

0	-1	0
-1	4	-1
0	-1	0

-1	-1	-1
-1	8	-1
-1	-1	-1

Η Λαπλασιανή παράγει μια απότομη διέλευση από το μηδέν σε μια ακμή. Ως τελεστής, η Λαπλασιανή είναι γραμμικός, χρονικά αναλλοίωτος και η συνάρτηση μεταφοράς του μηδενίζεται σε μηδενική συχνότητα. Επομένως, μια εικόνα φιλτραρισμένη με τον τελεστή της Λαπλασιανής θα έχει μηδενικό μέσο επίπεδο γκρίζου. Η Λαπλασιανή ανιχνεύει τις απότομες ακμές μιας εικόνας χωρίς θόρυβο. Η δυαδική εικόνα που παράγεται με κατωφλίωση μιας εικόνας φιλτραρισμένης με το Λαπλασιανό τελεστή στο μηδενικό επίπεδο γκρίζου θα δώσει κλειστά, συνεκτικά περιγράμματα όταν εξαλειφθούν τα εσωτερικά σημεία.

Η Λαπλασιανή, ως διαφορικός τελεστής, είναι ευαίσθητη στο θόρυβο και, επομένως, πριν την εφαρμογή της απαιτείται εξομάλυνση της εικόνας. Εξομάλυνση μπορεί να γίνει, για παράδειγμα, με ένα Γκαουσσιανό φίλτρο. Η Λαπλασιανή μπορεί να συνδυαστεί με το Γκαουσσιανό φίλτρο σε ένα **πυρήνα Γκαουσσιανής Λαπλασιανής (Laplacian of Gaussian Kernel)** :

$$-\nabla^2 \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} = \frac{1}{\pi\sigma^4}\left(1-\frac{x^2+y^2}{2\sigma^2}\right) e^{-\frac{x^2+y^2}{2\sigma^2}} .$$

Ο πυρήνας αυτός αποτελεί την κρουστική απόκρουση ενός γραμμικού, χρονικά αναλλοίωτου συστήματος (φίλτρου). Επειδή η κρουστική αυτή απόκριση είναι **διαχωρίσιμη (separable)**, η υλοποίηση του φίλτρου γίνεται πολύ αποτελεσματικά.

3.4.Ανίχνευση και σύνδεση ακμών (edge detection and linking)

Η τρίτη προσέγγιση για τον προσδιορισμό των συνόρων αντικειμένων σε μια εικόνα έγκειται στην εξέταση κάθε εικονοστοιχείου και της άμεσης περιοχής του για να αποφασιστεί αν το εικονοστοιχείο ανήκει πραγματικά στο σύνορο ενός αντικειμένου. Τέτοια εικονοστοιχεία προσδιορίζονται ως **εικονοστοιχεία ακμών (edge pixels)**. Μια εικόνα στην οποία τα επίπεδα γκρίζου υποδηλώνουν πόσο πληροί τις απαιτήσεις ενός εικονοστοιχείου ακμών το αντίστοιχο εικονοστοιχείο ονομάζεται **εικόνα ακμών (edge image)** ή **χάρτης ακμών (edge map)**. Μια εικόνα που καταγράφει μόνο τη θέση (όχι το μέγεθος) των εικονοστοιχείων ακμών ονομάζεται **δυαδική εικόνα ακμών (binary edge image)**. Μια εικόνα που καταγράφει την κατεύθυνση των ακμών είναι μια **κατευθυντική εικόνα ακμών (directional edge image)** .

Μια εικόνα ακμών συνήθως δείχνει το περίγραμμα κάθε αντικειμένου, αλλά σπάνια οδηγεί αυτό στα κλειστά, συνδεδεμένα σύνορα που απαιτεί η κατάτμηση εικόνων. Επομένως, απαιτείται άλλο ένα επεξεργαστικό στάδιο που ονομάζεται σύνδεση **σημείων ακμών (edge point linking)** και συσχετίζεται παραπλήσια σημεία ακμών σε τρόπο ώστε να σχηματίζεται ένα κλειστό, συνδεδεμένο σύνορο.

Η ανίχνευση ακμών γίνεται με εξέταση της περιοχής κάθε εικονοστοιχείου και ποσοτικοποίηση της κλίσης και της κατεύθυνσης της αλλαγής επιπέδων γκρίζου. Για το σκοπό αυτό, χρησιμοποιούνται διάφοροι συνελικτικοί τελεστές, μερικοί από τους οποίους είναι οι επόμενοι :

- **Ο ανιχνευτής ακμών του Roberts :** Είναι ένας τοπικός διαφορικός τελεστής που δέχεται μια εικόνα εισόδου A(m,n) και παράγει την εικόνα εξόδου

$$B(m,n) = \left\{\left[\sqrt{f(m,n)} - \sqrt{f(m+1,n+1)}\right]^2 + \left[\sqrt{f(m+1,n)} - \sqrt{f(m,n+1}\right]^2\right\}^{1/2}$$

Οι εσωτερικές τετραγωνικές ρίζες κάνουν τη διαδικασία παρόμοια με την επεξεργασία που λαμβάνει χώρα στο ανθρώπινο σύστημα όρασης

- **Ο ανιχνευτής ακμών του Sobel :** Χρησιμοποιεί τους συνελικτικούς πυρήνες των επομένων πινάκων. Κάθε εικονοστοιχείο συνελίσσεται και με τους δύο πυρήνες. Ο ένας πυρήνας έχει μέγιστη απόκριση σε μια κατακόρυφη ακμή, ενώ ο άλλος σε μια οριζόντια. Τελικά, χρησιμοποιείται η μέγιστη τιμή μεταξύ των δύο συνελίξεων και το αποτέλεσμα είναι μια εικόνα του μεγέθους των ακμών.

-1	-2	-1
0	0	0
1	2	1

-1	0	1
-2	0	2
-1	0	1

- **Ο ανιχνευτής ακμών του Prewitt :** Η διαδικασία είναι ακριβώς η ίδια όπως και με τον ανιχνευτή ακμών του Sobel, αλλά χρησιμοποιούνται οι επόμενοι δύο πυρήνες.

-1	-1	-1
0	0	0
1	1	1

1	0	-1
1	0	-1
1	0	-1

- **Ο ανιχνευτής ακμών του Kirsch :** Χρησιμοποιεί τους επόμενους οκτώ συνελικτικούς πυρήνες. Κάθε εικονοστοιχείο συνελίσσεται και με τις οκτώ μάσκες. Κάθε μάσκα έχει μέγιστη απόκριση σε ακμές κάποιου συγκεκριμένου προσανατολισμού. Το ολικό μέγιστο επιλέγεται και, τελικά, δημιουργείται μια εικόνα του μεγέθους των ακμών.

+5	+5	+5
-3	0	-3
-3	-3	-3

-3	+5	+5
-3	0	+5
-3	-3	-3

-3	-3	+5
-3	0	+5
-3	-3	+5

-3	-3	-3
-3	0	+5
-3	+5	+5

-3	-3	-3
-3	0	-3
+5	+5	+5

-3	-3	-3
+5	0	-3
+5	+5	-3

+5	-3	-3
+5	0	-3
+5	-3	-3

+5	+5	-3
+5	0	-3
-3	-3	-3

Μετά τη δημιουργία μιας εικόνας ακμών, οι ακμές θα είναι σχετικά εμφανείς και το επίπεδο του θορύβου χαμηλό κάτω από ιδανικές συνθήκες. Τότε, με κατωφλίωση και λέπτυνση μπορεί να δημιουργηθεί μια εικόνα συνδεδεμένων συνόρων πάχους ενός μόνο εικονοστοιχείου. Κάτω, όμως, από λιγότερο ιδανικές συνθήκες, η εικόνα ακμών θα παρουσιάζει κενά που θα πρέπει να καλυφθούν. Η διαδικασία της κάλυψης των κενών των ακμών ονομάζεται **σύνδεση ακμών (edge linking)**.

Μικρά κενά μπορούν να καλυφθούν απλά με αναζήτηση μιας 5Χ5 ή μεγαλύτερης περιοχής με κέντρο ένα ανοικτό άκρο συνόρου για άλλα ανοικτά άκρα συνόρων και συμπλήρωση των συνοριακών εικονοστοιχείων που λείπουν. Σε περίπλοκες σκηνές με πολλά σημεία ακμών, αυτή η προσέγγιση μπορεί να οδηγήσει σε υπερβολική κατάτμηση. Για να αποφευχθεί η υπερβολική κατάτμηση, απαιτείται τα ακραία εικονοστοιχεία να συμφωνούν ως προς το επίπεδο γκρίζου και τον προσανατολισμό των ακμών στα πλαίσια κάποιων ανοχών.

3.5. Μεγέθυνση περιοχών

Η **μεγέθυνση περιοχών (region growing)** είναι μια προσέγγιση στην κατάτμηση εικόνων που έχει τύχει της προσοχής των ερευνητών της τεχνητής όρασης και της τεχνητής νοημοσύνης. Κανείς ξεκινάει τη διαδικασία διαμερίζοντας την εικόνα σε πολλές μικροσκοπικές περιοχές. Αυτές οι αρχικές περιοχές μπορεί να είναι μικρές περιοχές εικονοστοιχείων ή ακόμα και μεμονωμένα εικονοστοιχεία. Σε κάθε περιοχή, υπολογίζονται κατάλληλα ορισμένες ιδιότητες που αντανακλούν συμμετοχή σε ένα αντικείμενο. Τέτοιες ιδιότητες μπορεί να είναι το μέσο επίπεδο γκρίζου, η υφή ή το χρώμα. Έτσι, το πρώτο βήμα αναθέτει σε κάθε περιοχή ένα σύνολο παραμέτρων των οποίων οι τιμές αντανακλούν το αντικείμενο στο οποίο ανήκουν.

Στη συνέχεια, εξετάζονται τα σύνορα μεταξύ γειτονικών περιοχών και υπολογίζεται το σθένος τους με βάση τις διαφορές των μέσων ιδιοτήτων των γειτονικών περιοχών. Ένα σύνορο είναι **σθεναρό (strong)** αν οι ιδιότητες διαφέρουν σημαντικά μεταξύ των δύο πλευρών του συνόρου, αλλιώς είναι **ασθενές (weak)**. Τα σθεναρά σύνορα διατηρούνται. Τα ασθενή διαλύονται και οι περιοχές που διαχωρίζουν συγχωνεύονται.

Η διαδικασία επαναλαμβάνεται υπολογίζοντας εκ νέου τις ιδιότητες για τις μεγαλύτερες περιοχές και διαλύοντας τα ασθενή σύνορα. Η συγχώνευση περιοχών συνεχίζεται μέχρι τα σύνορα να μην είναι αρκετά ασθενή για να διαλυθούν. Τότε, έχει τερματιστεί η διαδικασία της κατάτμησης. Η όλη διαδικασία δίνει την εντύπωση ότι στο εσωτερικό αντικειμένων αναπτύσσονται περιοχές που μεγαλώνουν μέχρι τα σύνορα τους να ταυτιστούν με τις ακμές των αντικειμένων.

Οι αλγόριθμοι μεγέθυνσης περιοχών είναι υπολογιστικά πιο επίπονοι από άλλες απλούστερες τεχνικές, αλλά χρησιμοποιούν άμεσα και ταυτόχρονα πολλές ιδιότητες της εικόνας για τον προσδιορισμό του τελικού συνόρου. Υπόσχονται επιτυχία σε εικόνες φυσικών σκηνών, όπου δεν διατίθεται εκ των προτέρων γνώση για την υποβοήθηση της διαδικασίας της κατάτμησης.

3.6. Μορφολογική επεξεργασία εικόνων

Ο όρος **μορφολογία (morphology)** αναφέρεται στη μελέτη της μορφής και της δομής και με την έννοια αυτή χρησιμοποιείται στη βιολογία, τη γεωγραφία και τη γλωσσολογία. Η εφαρμογή της μορφολογικής ανάλυσης στην επεξεργασία (ασπρόμαυρης) εικόνας προτάθηκε στα τέλη της δεκαετίας του 1960 από τον Georges Matheron και βασίστηκε στην ανάλυση της γεωμετρικής δομής που είναι έμφυτη σε μια εικόνα. Η εργασία του Matheron, στην πρώιμη μορφή της, δημοσιεύτηκε το 1975 ως μονογραφία με τίτλο **Τυχαία Σύνολα και Ολοκληρωτική Γεωμετρία (Random Sets and Integral Geometry)**. Η αρχική προσέγγιση του ήταν, βασικά, στατιστικής φύσεως και αποτελούσε σύνθεση της γεωμετρικής πιθανότητας που χρησιμοποιείται στη στερεολογία και μιας άλγεβρας Minkowski για την ανάλυση του σχήματος που είχε προτείνει ο Hans Hadwiger. Μετά το 1975, όμως, η έρευνα και η εφαρμογή της μορφολογικής ανάλυσης επικεντρώθηκε στη χρήση των θεμελιωδών μορφολογικών τελεστών χωρίς κάποια στατιστική ερμηνεία τους.

Σε γενικές γραμμές, στη μορφολογική ανάλυση μελετάμε τον τρόπο με τον οποίο ένα προκαθορισμένο γεωμετρικό σχήμα (**το δομικό στοιχείο (structuring element)**) ταιριάζει μέσα σε μια **δίτιμη (binary)** εικόνα. Η μελέτη αυτή χρησιμοποιεί μορφολογικές λειτουργίες όπως **ανίχνευση ακμών (edge detection)**, **κατάτμηση (segmentation)** και **ενίσχυση (enhancement)** εικόνων. Η κατασκευή **μορφολογικών φίλτρων (morphological filters)** με τη σύνθεση διαφόρων μορφολογικών λειτουργιών προσφέρει πλεονεκτήματα έναντι των γραμμικών φίλτρων: (i) σε αντίθεση με τα γραμμικά φίλτρα, τα μορφολογικά φίλτρα τείνουν να αφήσουν απαραμόρφωτη τη γεωμετρική δομή μιας εικόνας, (ii) η μορφολογική ανάλυση μπορεί να δημιουργήσει χρήσιμα (μορφολογικά) χαρακτηριστικά και (iii) αυστηρά θεωρήματα αναπαράστασης επιτρέπουν την έκφραση δημοφιλών κλάσεων γραμμικών και μη γραμμικών φίλτρων ως συνθέσεων των θεμελιωδών μορφολογικών λειτουργιών. Τελικά, η μορφολογική επεξεργασία επιτρέπει τη συστηματική αλλοίωση του γεωμετρικού περιεχομένου μιας εικόνας, ενώ ταυτόχρονα διατηρεί σημαντικά γεωμετρικά χαρακτηριστικά της αναλλοίωτα. Σημειώνουμε ότι οι αλγόριθμοι μορφολογικής ανάλυσης και επεξεργασίας επιδέχονται υλοποιήσεις σε μορφή **πίνακα αναζήτησης (look-up table)** σε σχετικά απλό υλικό για ταχεία **σωληνωτή επεξεργασία (pipeline processing)**.

Βασικές μορφολογικές πράξεις

Στη μορφολογική ανάλυση, τόσο μια δίτιμη (ασπρόμαυρη) εικόνα B όσο και το δομικό στοιχείο S θεωρούνται ως σύνολα σημείων ενός χώρου IR^n. Μια μορφολογική πράξη είναι μια διαδικασία η οποία τελικά παράγει ένα νέο σημειοσύνολο και ορίζεται μέσω μιας συνολοθεωριτικής εξίσωσης. Οι βασικές μορφολογικές πράξεις είναι η **διάβρωση (erosion)** και η **διαστολή (dilation)**, όπως ορίζονται παρακάτω. Προηγουμένως, όμως, χρειάζονται οι ορισμοί των πράξεων της **μετατόπισης (translation)** σημειοσυνόλου κατά ένα σημείο στο χώρο IR^n, της **πρόσθεσης κατά Minkowski** και της **αφαίρεσης κατά Minkowski**.

Μετατόπιση : Δεδομένου ενός υποσυνόλου A του χώρου IR^n η μετατόπιση A+$\underline{x}$ του A κατά το σημείο $\underline{x}$ του IR^n ορίζεται ως

$$A + \underline{x} = \{\underline{\alpha} + \underline{x} : \underline{\alpha} \in A \}.$$

Ουσιαστικά, η μετατόπιση υποδηλώνει ότι το σημειοσύνολο A και το σχήμα που αυτό ορίζει μετατοπίζεται πάνω στο χώρο IR^n κατά ένα διάνυσμα $\underline{x}$.

Πρόσθεση κατά Minkowski : Δεδομένων δύο σημειοσυνόλων *A* και *B* στο χώρο IR^n ορίζουμε την κατά Minkowski πρόσθεση $A \oplus B$ ως

$$A \oplus B = \bigcup_{\underline{b} \in B} A + \underline{b}\,.$$

Η πρόσθεση κατά Minkowski δύο σημειοσυνόλων A και B συντελείται, επομένως, αν πρώτα υπολογιστεί η μετατόπιση του A κατά κάθε στοιχείο του B και στην συνέχεια υπολογιστεί η **ένωση (union)** των μετατοπίσεων.

Αφαίρεση κατά Minkowski : Δεδομένων δύο σημειοσυνόλων A και B στο χώρο IR^n, ορίζουμε την κατά Minkowski αφαίρεση A□B ως:

A□B = $\bigcap_{\underline{b} \in B} A + \underline{b}$.

Η αφαίρεση κατά Minkowski δύο σημειοσυνόλων A και B συντελείται, επομένως, αν πρώτα υπολογιστεί η μετατόπιση του A κατά κάθε στοιχείο του B και στη συνέχεια υπολογιστεί η **τομή (section)** των μετατοπίσεων.

Με βάση τα παραπάνω, μπορούμε τώρα να ορίσουμε τη διάβρωση και τη διαστολή σημειοσυνόλων από άλλα σημειοσύνολα.

Διάβρωση : Η διάβρωση $A \otimes B$ του σημειοσυνόλου A από το σημειοσύνολο B ορίζεται ως:

$A \otimes B$ =A□(-B)

όπου -$B = \{-\underline{b} : \underline{b} \in B\}$. Όταν το σημειοσύνολο A διαβρώνεται από το σημειοσύνολο B, τότε το δεύτερο ονομάζεται δομικό στοιχείο. Η διάβρωση ενός σημειοσυνόλου από ένα δομικό στοιχείο έχει ως αποτέλεσμα τη συρρίκνωση του σημειοσυνόλου με τρόπο που προσδιορίζεται από το δομικό στοιχείο.

Διαστολή : Η διαστολή του σημειοσυνόλου A από το σημειοσύνολο B ορίζεται απλά ως η πρόσθεση κατά Minkowski $A \oplus B$. Το σημειοσύνολο B ονομάζεται και πάλι δομικό στοιχείο και το αποτέλεσμα της διαστολής είναι επέκταση του σημειοσυνόλου A.

Αλλες μορφολογικές πράξεις που συναντώνται στη μορφολογική ανάλυση περιλαμβάνουν τις επόμενες:

Άνοιγμα (opening) : Το άνοιγμα ενός σημειοσυνόλου A από ένα δομικό στοιχείο B ορίζεται ως η διάβρωση του A από το B ακολουθούμενη από διαστολή του αποτελέσματος από το B. Δηλαδή:

$$A \circ B = (A \otimes B) \oplus B.$$

Το αποτέλεσμα του ανοίγματος ενός σημειοσυνόλου από ένα δομικό στοιχείο είναι να απαλειφθούν μικρά και λεπτά «αντικείμενα», να διασπαστούν «αντικείμενα» σε λεπτά σημεία τους και, γενικά, να εξομαλυνθούν τα σύνορα μεγαλυτέρων «αντικειμένων» χωρίς να αλλάξει σημαντικά ο όγκος τους.

Κλείσιμο (closing) : Το κλείσιμο ενός σημειοσυνόλου A από ένα δομικό στοιχείο B ορίζεται ως η διαστολή του A από το B ακολουθούμενη από διάβρωση του αποτελέσματος από το B. Δηλαδή:

$$A \bullet B = (A \oplus B) \otimes B.$$

Το αποτέλεσμα του κλεισίματος ενός σημειοσυνόλου από ένα δομικό στοιχείο είναι να «γεμίσουν» μικρές και λεπτές τρύπες σε «αντικείμενα», να συνδεθούν παραπλήσια «αντικείμενα» και, γενικά, να εξομαλυνθούν τα σύνορα μικροτέρων αντικειμένων.

Πέραν των προηγούμενων πράξεων, στη μορφολογική ανάλυση και επεξεργασία συναντάμε και άλλες πράξεις, όπως **συρρίκνωση (shrinking)**, **λέπτυνση (thinning)**, **σκελετοποίηση (skeletonization)**, **κλάδεμα (pruning)** και **πάχυνση (thickening)**. Και αυτές οι μορφολογικές πράξεις επιτυγχάνονται με συνδυασμούς των βασικών μορφολογικών πράξεων της διάβρωσης και της διαστολής.

Αναπαράσταση σημάτων με σημειοσύνολα

Μια δίτιμη (ασπρόμαυρη) εικόνα είναι ένα πολυδιάστατο ψηφιακό σήμα $f(\underline{x})$ με μόνες επιτρεπτές τιμές 1 ή 0. Μια δίτιμη εικόνα μπορεί να αναπαρασταθεί άμεσα ως σημειοσύνολο του χώρου IR^n. Πιο συγκεκριμένα, το **προσκήνιο (foreground)** της εικόνας ορίζεται ως το σύνολο S των σημείων $\underline{x}$ του χώρου IR^n για τα οποία $f(\underline{x})$=1, δηλαδή

$$S = \{\underline{x} \in IR^n : f(\underline{x}) = 1\}.$$

Αντίστοιχα, το συμπλήρωμα S^c του σημειοσυνόλου S αποτελείται από το σύνολο των σημείων $\underline{x}$ του χώρου IR^n για τα οποία $f(\underline{x})$= 0 και ονομάζεται **παρασκήνιο (background)** της εικόνας:

$$S^c = \underline{x} \in IR^n : f(\underline{x}) = 0.$$

Για την αναπαράσταση πολυδιάστατων αναλογικών σημάτων με σημειοσύνολα, μπορούν να χρησιμοποιηθούν δύο διαφορετικές, αλλά μαθηματικώς ισοδύναμες, προσεγγίσεις: η προσέγγιση των **εγκάρσιων τομών (cross sections)** και η προσέγγιση της **πλανητικής σκιάς (umbra)**. Όπως γίνεται φανερό στη συνέχεια, η χρήση εγκαρσίων τομών αναπαριστάνει ένα n-διάστατο αναλογικό σήμα με μια (άπειρη) συλλογή n-διαστάτων συνόλων, των εγκαρσίων τομών του σήματος. Αντίθετα η προσέγγιση της πλανητικής σκιάς αναπαριστάνει ένα n-διάστατο αναλογικό σήμα με ένα μόνο (n+1)-διάστατο σύνολο, την πλανητική σκιά του σήματος.

Η εγκάρσια τομή $T_{\xi}(f)$ του πολυδιάστατου αναλογικού σήματος $f(\underline{x})$ σε επίπεδο ξ ορίζεται ως:

$$\mathrm{T}_{\xi}(f)=\{\underline{x}: f(\underline{x}) \geq \xi\}, \xi \in IR.$$

Με άλλα λόγια η εγκάρσια τομή $T_{\xi}(f)$ αποτελείται από το σύνολο όλων των σημείων $\underline{x}$ του χώρου IR^{n} στα οποία η τιμή του σήματος $f(\underline{x})$ ξεπερνάει το **κατώφλι (threshold)** ξ. Ο Serra δείχνει ότι, αν το σήμα $f(\underline{x})$ είναι ημισυνεχές εκ των άνω, μπορεί να ανακατασκευασθεί μονοσήμαντα από τη συλλογή $\{T_{\xi}(f)$, $\xi \in IR\}$ των εγκάρσιων τομών του ως:

$$f(\underline{x})=\sup\{\xi \in R: \underline{x} \in T_{\xi}(f)\}.$$

Η πλανητική σκιά $U(f)$ του πολυδιάστατου αναλογικού σήματος $f(\underline{x})$ ορίζεται ως:

$$U(f)=\{(\underline{x}, \xi) \in IR^{n} \times IR: f(\underline{x}) \geq \xi\}.$$

Δηλαδή, η πλανητική σκιά $U(f)$ αποτελείται από το σύνολο των σημείων τα οποία βρίσκονται κάτω από το γράφημα του σήματος $f(\underline{x})$ μέχρι το $-\infty$. Μπορεί να αποδειχθεί ότι ένα ημισυνεχές εκ των άνω σήμα $f(\underline{x})$ ανακατασκευάζεται μονοσήμαντα από την πλανητική σκιά του ως:

$$f(\underline{x})=\sup\{\xi \in IR:(\underline{x}, \xi) \in U(f)\}.$$

Από τα παραπάνω παρατηρούμε ότι τόσο οι εγκάρσιες τομές όσο και η πλανητική σκιά αναπαριστάνουν ένα n-διάστατο αναλογικό σήμα ως ένα σημειοσύνολο του χώρου $IR^{n+1}=IR^{n} \times IR$. Επειδή ένα σημειοσύνολο του χώρου IR^{n+1} είναι, όπως είδαμε, ισοδύναμο με μια (n+1)-διάστατη δίτιμη (ασπρόμαυρη) εικόνα, συμπεραίνουμε ότι το «κόστος» της δίτιμης αναπαράστασης ενός αναλογικού σήματος είναι η αύξηση κατά μια των διαστάσεών του.

Για τη μορφολογική επεξεργασία και ανάλυση πολυδιάστατων αναλογικών σημάτων, οι προσεγγίσεις των εγκάρσιων τομών και της πλανητικής σκιάς είναι κατάλληλες. Πιο συγκεκριμένα, τα πολυδιάστατα σήματα αναπαριστάνονται με σημειοσύνολα, είτε με τη χρήση εγκαρσίων τομών είτε με τη χρήση της πλανητικής σκιάς, και οι μεθοδολογίες της μορφολογικής επεξεργασίας και ανάλυσης εφαρμόζονται στα σημειοσύνολα που προκύπτουν. Θα εξετάσουμε στη συνέχεια τη δυνατότητα αναπαράστασης γραμμικών, χρονικά αναλλοίωτων συστημάτων ως μορφολογικών φίλτρων[*]. Η θεωρία αυτή γενικεύεται για αναπαράσταση ευρέων κλάσεων μη γραμμικών συστημάτων[♥], αλλά η γενίκευση δεν θα παρουσιαστεί εδώ.

Θεωρήματα αναπαράστασης συστημάτων με μορφολογικά φίλτρα

Οι βασικές μορφολογικές πράξεις της διάβρωσης και της διαστολής μπορούν να χρησιμοποιηθούν για την αναπαράσταση γραμμικών, χρονικά αναλλοίωτων συστημάτων, αλλά και ευρέων κλάσεων μη γραμμικών συστημάτων. Ως εισαγωγή σ' αυτή τη δυνατότητα των μορφολογικών φίλτρων, ας δούμε την αναπαράσταση γραμμικών, χρονικά αναλλοίωτων συστημάτων.

Θεωρούμε ένα γραμμικό, χρονικά αναλλοίωτο σύστημα με κρουστική απόκριση $h(\underline{x})$, το οποίο δέχεται διεγέρσεις από μια κλάση ημισυνεχών εκ των άνω σημάτων εισόδου. Θεωρούμε ότι η συνάρτηση μεταφοράς του συστήματος ισούται με τη μονάδα για μηδενική συχνότητα, δηλαδή το σύστημα αφήνει αναλλοίωτα σταθερά σήματα εισόδου. Ορίζουμε τον **πυρήνα (kernel)** του συστήματος ως ένα σύνολο K των σημάτων εισόδου $g(\underline{x})$:

$$K = \{\rho(\underline{\chi}) : (h * g)(\underline{o}) \geq o \ \},$$

όπου $(h * g)(\underline{x})$ είναι η τιμή της συνέλιξης των συναρτήσεων h και g στο $\underline{x}$. Ο Μαραγκός δείχνει ότι ο πυρήνας K χαρακτηρίζει και ανακατασκευάζει μονοσήμαντα το γραμμικό, χρονικά αναλλοίωτα σύστημα. Πιο συγκεκριμένα, ο Μαραγκός παρουσιάζει το επόμενο θεώρημα:

Θεώρημα : Έστω ένα γραμμικό, χρονικά αναλλοίωτο σύστημα με μη αρνητική κρουστική απόκριση $h(\underline{x}) \geq 0 \quad \forall \underline{x}$ και διεγέρσεις από μια κλάση ημισυνεχών εκ των άνω σημάτων εισόδου. Έστω επίσης $H(\underline{o}) = 1$, όπου

$$H(\underline{s}) = \int h(\underline{x}) e^{i \underline{s}^T \bullet \underline{x}} d\underline{x}$$

η συνάρτηση μεταφοράς του συστήματος. Η απόκριση $(h * u)(\underline{x})$ του συστήματος σε διέγερση από ημισυνεχές εκ των άνω σήμα εισόδου $u(\underline{x})$ δίνεται από τη σχέση

$$(h * u)(\underline{x}) = \sup_{g \in IR^n} \left\{ \inf_{z \in IR^N} \{ u(\underline{z}) - \rho(\underline{z} - \underline{x}) \ \} \ \right\}$$

όπου K είναι ο πυρήνας του συστήματος, όπως ορίστηκε παραπάνω.

---

[*] P.Maragos and R.W. Schafer, “Morphological Filters- Part I: Their Spt-Theoretic Analysis and Relation to Linear, Shift-Invariant Filters”, IEEE Trans. Acoust., Speech, Sign. Proc., vol. ASSP-35, pp.1153-1169, 1987

[♥]P.Maragos and R.W. Schafer, “Morphological Filters-Part II: Their Relation to median, Order-Statistic, and Stack Filters ”, IEEE Trans. Acoust., Speech, Sign. Proc., vol. ASSP-35, pp. 1170-1184, 1987.

[♥]

Παρατήρηση : Μπορεί κανείς να δει εύκολα ότι η τομή των πλανητικών σκιών δύο σημάτων αντιστοιχεί στην πλανητική σκιά του ελαχίστου (ακριβέστερα, του κάτω ορίου (infinimum)) των δύο συναρτήσεων. Με βάση την παρατήρηση αυτή και τον ορισμό της μορφολογικής διάβρωσης, συμπεραίνουμε από το παραπάνω θεώρημα ότι *η απόκριση του συστήματος υπολογίζεται ως το μέγιστο (ακριβέστερα, το άνω όριο(supremum)) των διαβρώσεων του σήματος εισόδου με όλες τις συναρτήσεις του πυρήνα του συστήματος.*

6. Εισαγωγή στην τριδιάστατη τεχνητή όραση

6.1 Γενικά

Η ανακατασκευή επιφανειών αντικειμένων είναι ένα ιδιαίτερο πεδίο της τεχνητής όρασης το οποίο ασχολείται με την ανακατασκευή του σχήματος αντικειμένων ή τον υπολογισμό αποστάσεων ανάμεσα σε αισθητήριο (π.χ. κάμερα) και αντικείμενα σε μια σκηνή. Η αρχική πληροφορία αποτελείται από δεδομένα που μετριούνται από ένα ή περισσότερα αισθητήρια. Πιθανές εφαρμογές της περιλαμβάνουν την υπολογιστική μοντελοποίηση τριδιαστάτων αντικειμένων (π.χ. στην αρχιτεκτονική, στη μηχανολογία, στη χειρουργική, κλπ.), την εκτίμηση αποστάσεων ή θέσεων εμποδίων (π.χ. στον έλεγχο ροής οχημάτων, στη ρομποτική, κλπ.), την επιθεώρηση επιφανειών (π.χ. στον έλεγχο ποιότητας), την προσεγγιστική ή ακριβή εκτίμηση της θέσης τριδιαστάτων αντικειμένων (π.χ. στην αυτοματοποιημένη παραγωγή), ή την ταχεία εκτίμηση της θέσης εμποδίων χωρίς απαίτηση αναγνώρισης (π.χ. στην πλοήγηση).

Η διαδικασία της ανακατασκευής του σχήματος τριδιάστατων αντικειμένων συμπεριλαμβάνει διαδικασίες επεξεργασίας εικόνας (π.χ. φιλτράρισμα εικόνας, ανακατασκευή εικόνας, ενίσχυση εικόνας, κλπ.) ή ανάλυσης προτύπου (φώραση ακμών, κατάτμηση, αναγνώριση χαρακτηριστικών, κλπ.). Η χρήση των τεχνικών επεξεργασίας εικόνας και ανάλυσης προτύπου στην τεχνητή όραση εξετάζεται σε άλλα μέρη του μαθήματος. Σε αυτό το κομμάτι του μαθήματος, θα ασχοληθούμε με θέματα ανακατασκευής τριδιαστάτων αντικειμένων από μετρήσεις αισθητηρίων που αναπαράγουν μια **στατική (static)** ή **δυναμική σκηνή (dynamic scene)**. Σε μια στατική σκηνή δεν λαμβάνουν χώρα μετακινήσεις αντικειμένων ή αλλαγές φωτισμού κατά τη διάρκεια λήψης των εικόνων. Αν τέτοιες αλλαγές είναι επιτρεπτές, τότε η σκηνή είναι δυναμική. Σε προβλήματα τεχνητής όρασης, η προβαλλόμενη σκηνή μπορεί να είναι ένα φυσικό περιβάλλον, όπως η σκηνή ενός δρόμου ή μια αεροφωτογραφία από χαμηλό ύψος, ή ένα τεχνητό περιβάλλον, όπως ένα εργαστήριο. Στην πρώτη περίπτωση, η σκηνή ονομάζεται **εξωτερική σκηνή**, ενώ στη δεύτερη **εσωτερική**.

Η διαδικασία της **τρι**διάστατης ανάλυσης διαφέρει σημαντικά από αυτή της *δι*διάστατης ανάλυσης γιατί τα αντικείμενα που περιλαμβάνει μια τριδιάστατη σκηνή μπορεί να είναι στερεά ή παραμορφώσιμα, αδιάφανα ή διαφανή. Το πρόβλημα της **ανακατασκευής του σχήματος (shape reconstruction)** ενός αντικειμένου σε μια σκηνή συνήθως έγκειται στην **ανακατασκευή της επιφάνειας (surface reconstruction)** (δηλαδή του συνόλου των **πλευρών (object faces)**) του αντικειμένου. Σε γενικές γραμμές, πλευρά ενός αντικειμένου είναι ένα φραγμένο σύνολο επιφανειακών σημείων του που περιβάλλεται από **ακμές (edges)** του αντικειμένου και, προφανώς, το σύνολο των πλευρών ενός αντικειμένου αποτελεί την επιφανειά του. Το πρόβλημα της ανακατασκευής του σχήματος ενός αντικειμένου σε μια σκηνή μπορεί να θεωρηθεί και ως πρόβλημα εκτίμησης του βάθους (απόστασης) μεμονωμένων σημείων πάνω στην επιφάνεια του αντικειμένου και, κατά συνέπεια, αναφέρεται και ως **ανάλυση δεδομένων απόστασης (range data analysis)**.

Για την ανακατασκευή επιφάνειας και την ανάλυση απόστασης, απαιτείται τόσο τεχνική ανάλυση όσο κατανόηση της εφαρμογής. Η μέθοδος ανάλυσης που μπορεί να χρησιμοποιηθεί προσδιορίζεται από τα διαθέσιμα αισθητήρια για τη λήψη εικόνων (π.χ. κάμερες), τα διαθέσιμα εργαλεία για τη μεταβολή των κατανομών των τιμών της εικόνας (π.χ. εξοπλισμό φωτισμού), αλλά και τα αντικείμενα της σκηνής, τη διάταξη και τις κινήσεις τους. Ο σχεδιασμός του πλαισίου, δηλαδή η επιλογή της σκηνής, του φωτισμού, της κάμερας, του υλικού, κλπ., είναι συμπληρωματικός της επιλογής των τεχνικών ανάλυσης.

Παράλληλα, απαιτούνται και ορισμένες μαθηματικές υποθέσεις, επιφανειακά πολύ περιοριστικές, προκειμένου να οδηγηθούμε σε αυστηρές λύσεις και διαδικασίες. Γενικά, πρέπει να μοντελοποιηθούν οι επιρροές και οι δυνατότητες της τρίτης διάστασης για ανακατασκευή σχήματος. Η τρίτη διάσταση δεν είναι απλά μια νέα διάσταση που προστίθεται στο διδιάστατο πρόβλημα. Για παράδειγμα, ο φωτισμός και το βάθος επηρεάζουν την **υφή (texture)** και η υφή μπορεί να χρησιμοποιηθεί για ανάλυση του φωτισμού ή του βάθους. Η διαμόρφωση σκιών ή οι κινήσεις ενός αντικειμένου κάνουν δυνατή την άντληση συμπερασμάτων σχετικών με τη γεωμετρία του αντικειμένου. Για παράδειγμα, βιολόγοι κατάφεραν να ανακατασκευάσουν το φυσικό σχήμα δεινοσαύρων από αποτυπώματα των πελμάτων τους. Αντίστοιχα, για την ανακατασκευή σχήματος στην τεχνητή όραση χρειάζεται να εξαχθεί και να χρησιμοποιηθεί ανάλογη πληροφορία.

Δεδομένα σχετικά με απόσταση ή σχήμα μπορούν να εμφυτευφθούν κατευθείαν στο μηχανισμό σχηματισμού εικόνων με **τεχνικές ενεργού απόκτησης εικόνων (active image acquisition techniques)**. Για παράδειγμα, μπορεί να χρησιμοποιηθεί ελεγχόμενη ποσότητα προβαλλόμενης και ανακλώμενης ενέργειας (π.χ. υπέρηχος ή φως laser) και με μέτρηση του χρόνου μετάβασης από τον πομπό στο δέκτη να δημιουργηθούν εικόνες απόστασης. Αλλες τεχνικές ενεργού απόκτησης εικόνων βασίζονται στη χρήση **δομημένου φωτισμού (structured lighting), φωτομετρικής ανάλυσης (photometric stereo analysis), μεταβολών της εστίασης (shape from focus)**, και **ελεγχόμενων κινήσεων του αντικειμένου ή του αισθητηρίου (shape from motion** ή **shape from occluding boundaries)**.

Οι τεχνικές απόκτησης εικόνων χωρίζονται σε **μονόφθαλμες (monocular)** (ένα αισθητήριο), **διόφθαλμες (binocular)** (δύο αισθητήρια) ή **πολυόφθαλμες (polyocular)** (πολλά αισθητήρια). Επίσης, τόσο οι ενεργητικές όσο και οι παθητικές τεχνικές συνήθως απαιτούν **βαθμονόμηση (calibration)**, δηλαδή ακριβή καθορισμό των τιμών παραμέτρων για απόκτηση της εικόνας. Αυτό απαιτεί χρήση μοντέλων της διαδικασίας απόκτησης εικόνων που παρουσιάζουν τη μεγαλύτερη δυνατή ακρίβεια.

6.2 Μαθηματικός προσδιορισμός του βασικού προβλήματος

Για το μαθηματικό προσδιορισμό του προβλήματος της ανακατασκευής του σχήματος αντικειμένων, χρησιμοποιούμε την αναλογία με την προβολική οπτική και εισάγουμε μερικές βασικές έννοιες. Οι **σκηνές (scenes)** s ορίζονται σε συντεταγμένες της κάμερας ή σε παγκόσμιες συντεταγμένες *XYZ* για ένα τριδιάστατο χώρο R^3. Για **διάνυσμα θέσης a** = $(X, Y, Z)^T$ από την αρχή **O** = (0, 0, 0) στο σημείο $(X, Y, Z)^T$ υποθέτουμε ότι *s*(**a**) είναι η **τιμή σκηνής (scene value)** ορισμένη σε όλο το τριδιάστατο περιβάλλον του οπτικού αισθητηρίου σε κάποια χρονική στιγμή. Ενας **χώρος σκηνής (scene space)** αποτελείται από όλες τις σκηνές που υποτίθεται ότι έχουν τη δυνατότητα να περιβάλλουν το οπτικό αισθητήριο. Σε γενικές γραμμές, ένας χώρος σκηνής αντιστοιχεί σε μια περιοχή εφαρμογής της ανάλυσης σκηνής.

Στην περίπτωση χρήσης μιας κάμερας CCD (charged coupled device) ως οπτικού αισθητηρίου, οι τιμές σκηνής είναι οι μετρούμενες τιμές γκρίζας ή έγχρωμης απόχρωσης. Οι αποστάσεις που μετριούνται με τη χρήση παλμών φωτός laser μπορούν επίσης να θεωρηθούν ως τιμές σκηνής.

Οι προβολικές **εικόνες (images)** *E* σκηνών ορίζονται σε συντεταγμένες εικόνας *xy* ενός διδιάστατου επιπέδου R^2 και μόνο για ένα πεπερασμένο πίνακα του **επιπέδου εικόνας (image plane)**, π.χ. για $1 \leq x \leq M$ και $1 \leq y \leq N$. Οι **τιμές εικόνας (image values)** $E(x,y)$ αντιστοιχούν σε τιμές σκηνής *s*(**a**) σύμφωνα με κάποια προβολική απεικόνιση. Μια **οπτική απεικόνιση (optical mapping)** *A* είναι κάποια προβολή μιας σκηνής *s* σε μια εικόνα *E*,

$$E = A(s).$$

Η εικόνα *E* αναπαριστάνει τα φυσικά αντικείμενα στο επίπεδο εικόνας με τιμές **ακτινοβόλησης (irradiance)**.[1] Το γενικό πρόβλημα της προβολικής οπτικής συνίσταται στην υλοποίηση της απεικόνισης *A* και απαιτεί ορισμένες προϋποθέσεις όπως την έλλειψη παραμορφώσεων, σκίασης, κλπ. Εδώ θα ασχοληθούμε με το πρόβλημα της ανακατασκευής τριδιαστάτων αντικειμένων με βάση ένα σύνολο διαθέσιμων εικόνων *E* και θα παρουσιάσουμε διάφορες τεχνικές ανακατασκευής σχήματος. Κατά βάση, το πρόβλημα έγκειται σε αντιστροφή της απεικόνισης *A*, πλήρους στην περίπτωση της ανακατασκευής της ακριβούς δομής της επιφάνειας αντικειμένων και μερικής στην περίπτωση της ανάλυσης δεδομένων απόστασης. Επομένως, το ενδιαφέρον μας κατευθύνεται προς την εύρεση μιας απεικόνισης

$$s = A^{-1} (E)$$

και, στη συνέχεια, υλοποίησή της. Το πρόβλημα αυτό συνήθως λύνεται μόνο εν μέρει: είναι δυνατή μόνο η ανακατασκευή μερικών όψεων ενός αντικειμένου και μόνο μερικών αντικειμένων σε μια σκηνή.

Παρενθετικά αναφέρουμε ότι το 1923, ο γάλλος μαθηματικός Hadamard (1865-1963) πρότεινε το χαρακτηρισμό των μαθηματικών προβλημάτων ως **καλώς τεθειμένων (well-posed)** αν (1) υπάρχει λύση, (2) η λύση αυτή είναι μοναδική και (3) η λύση είναι συνεχής συνάρτηση των δεδομένων (στην πρακτική ορολογία, η λύση είναι **σθεναρή απέναντι σε θόρυβο (robust against noise)**.

---

[1] Η ακτινοβόληση (irradiance) ορίζεται ως η επιφανειακή πυκνότητα της ενέργειας που λαμβάνει το αισθητήριο. Μετριέται σε μονάδες $W.m^{-2}$ και αποτελεί την ποσότητα που μετράει το αισθητήριο και κωδικοποιεί σε τιμές γκρίζας απόχρωσης για κάθε εικονοστοιχείο.

Τα προβλήματα της προβολικής οπτικής που αφορούν στην υλοποίηση της απεικόνισης *A* είναι συνήθως καλώς τεθειμένα. Από την άλλη πλευρά, για μοναδικότητα της απεικόνισης $s = A^{-1}(E)$ προϋποτίθεται μια σχέση ένα-προς-ένα κατά την απεικόνιση *A*, πράγμα που σημαίνει ότι μια εικόνα *E* παράγεται από ακριβώς μια σκηνή *s*. Στην πράξη, αυτό δεν ισχύει ποτέ και μικρές αλλαγές σε παρατηρούμενες εικόνες μπορούν να οδηγήσουν σε σημαντικά διαφορετικές ερμηνείες της σκηνής. Θα δούμε στη συνέχεια μερικούς από τους περιορισμούς στην ανακατασκευή του σχήματος αντικειμένων.

6.3.Βασικοί περιορισμοί στην ανακατασκευή σχήματος

Από την πλευρά του σχήματος, ένα αντικείμενο σε μια σκηνή χαρακτηρίζεται ως ένα τοπολογικώς συνεκτικό συμπαγές (κλειστό και φραγμένο)[2] σύνολο σημείων του R^3, το οποίο θεωρείται αναλλοίωτο σε ευκλείδιες γεωμετρικές απεικονίσεις. Η χρονική εξάρτηση του σχήματος είναι αδιάφορη για **στερεά (rigid)** αντικείμενα. Αλλα χαρακτηριστικά του αντικειμένου, όπως η υφή, το χρώμα και η ανακλαστικότητα της επιφάνειας του, δεν συμπεριλαμβάνονται στη γεωμετρική τοπολογική προσέγγιση.

Η ανακατασκευή πλευρών ενός τριδιαστάτου αντικειμένου από διδιάστατες οπτικές απεικονίσεις του είναι δυνατή μόνο για ορατές πλευρές του. Ενα αντικείμενο είναι **εξωτερικώς πλήρως ορατό (entirely visible from outside)** αν υπάρχει ακτίνα για κάθε σημείο **P** της επιφάνειάς του που τέμνει το αντικείμενο μόνο στο σημείο **P**. Επομένως ανακατασκευή επιφάνειας μπορεί να πραγματοποιηθεί μόνο για εξωτερικώς πλήρως ορατά αντικείμενα.[3] Συνήθως όμως τα αντικείμενα προβάλλονται μόνο κατά μια ή ένα μικρό αριθμό κατευθύνσεων.

[2] Ενα τοπολογικώς συμπαγές σύνολο είναι κλειστό ως προς την τοπολογία του ευκλείδιου χώρου και κάθε άπειρη ακολουθία σημείων του περιέχει μια συγκλίνουσα υπακολουθία της. Ενα αντικείμενο στον τριδιάστατο χώρο μπρεί να είναι μονοδιάστατο (δηλαδή μια καμπύλη), διδιάστατο (δηλαδή μια επιφάνεια) ή τριδιάστατο. Για ένα τριδιάστατο αντικείμενο, ένα σημείο είναι **εξωτερικό σημείο του αντικειμένου** αν δεν ανήκει στο αντίστοιχο σύνολο σημείων που ορίζει το αντικείμενο, ενώ είναι ένα **επιφανειακό σημείο** αν σε κάθε περιοχή του ανήκει τουλάχιστον ένα εξωτερικό σημείο του αντικειμένου.

[3] Η ανακατασκευή του εσωτερικού ενός τριδιαστάτου αντικειμένου αποτελεί αντικείμενο της **υπολογιστικής τομογραφίας (computer tomography)**.

7. Μοντέλο γεωμετρικών προβολών

7.1.Γενικά

Ο μηχανισμός δημιουργίας ψηφιακών εικόνων παίζει πρωταρχικό ρόλο στο σχεδιασμό μια αλυσίδας καθηκόντων επεξεργασίας και ανάλυσης εικόνας και καθορίζεται από φυσικούς και γεωμετρικούς παράγοντες της διαδικασίας συλλογής των δεδομένων των εικόνων. Πρώτα-πρώτα, οι συναρτησιακές τιμές της ψηφιακής εικόνας εξαρτώνται από το υλικό και τα ανακλαστικά χαρακτηριστικά της κατεγραμμένης σκηνής. Επειτα, οι συναρτησιακές τιμές της ψηφιακής εικόνας εξαρτώνται κατευθείαν από τα αισθητηριακά χαρακτηριστικά της κάμερας, τα οπτικά χαρακτηριστικά των φακών, τη διαδικασία ψηφιοποίησης της αναλογικής εικόνας και από τους γεωμετρικούς νόμους που χαρακτηρίζουν τη διαδικασία συλλογής των δεδομένων της εικόνας. Θα εξετάσουμε εδώ μερικά από τα χαρακτηριστικά του τελευταίου αυτού παράγοντα.

Η προβολή ενός επιφανειακού σημείου μιας τριδιάστατης σκηνής πάνω στο διδιάστατο επίπεδο εικόνας μπορεί να περιγραφεί είτε ως κεντρική (δηλαδή προοπτική) προβολή είτε ως παράλληλη προβολή. Η προβολική σχέση μεταξύ συστημάτων συντεταγμένων (παγκοσμίων, της κάμερας, της εικόνας κλπ.) έχει σημασία για την ανάπτυξη τεχνικών για την ανακατασκευή της επιφανείας αντικειμένων. Θα περιγράψουμε πρώτα μαθηματικά (ιδανικά) μοντέλα προβολής και στη συνέχεια δυο παραλλαγές του που χρησιμοποιούνται ως μοντέλα κάμερας.

7.2.Κεντρικές και παράλληλες προβολές

Η **κεντρική (προοπτική) προβολή** και η ***(ορθογώνια) παράλληλη προβ***_ολή_ (ονομαζόμενη και *ορθογραφική προβολή*) είναι εξιδανικευμένα μοντέλα προβολής σκηνών *s* σε εικόνες *E*. Υποθέτοντας μια ιδανική κάμερα, η προβολή σημείων **P** = (X, Y, Z) μιας σκηνής του χώρου XYZ σε σημεία **p** = (x, y) του επιπέδου εικόνας xy είναι μια κεντρική (προοπτική) προβολή. Αν το σύστημα συντεταγμένων του χώρου XYZ οριστεί να είναι το **σύστημα συντεταγμένων με κέντρο την κάμερα** (το εστιακό σημείο της κάμερας αποτελεί το προβολικό κέντρο και ο οπτικός άξονας συμπίπτει με τον άξονα Z, τότε οι εξισώσεις προβολής γίνονται

$$x = f\,X / Z \quad \text{και} \quad y = f\,Y / Z.$$

Αυτές οι εξισώσεις συνάγονται κατευθείαν από τη γεωμετρία. Αν η αρχή του συστήματος συντεταγμένων τοποθετηθεί πάνω στο επίπεδο εικόνας, οπότε και έχουμε ένα **σύστημα συντεταγμένων με κέντρο την κάμερα**, τότε οι εξισώσεις προβολής γίνονται αντίστοιχα

$$x = f\,X / (f + Z) \quad \text{και} \quad y = f\,Y / (f + Z).$$

Γενικά, για μια κεντρική προβολή, πρέπει να οριστούν το **προβολικό κέντρο** και το **προβολικό επίπεδο**. Τα σημεία του **εξαφανιζομένου επιπέδου (vanishing plane)**, δηλαδή του επιπέδου που διέρχεται από το προβολικό κέντρο και είναι παράλληλο προς το προβολικό επίπεδο, προβάλλονται στο άπειρο πάνω στο προβολικό επίπεδο. Για μια γενική και απλή παρουσίαση της έννοιας της κεντρικής προβολής είναι κατάλληλη η **χρήση ομογενών συντεταγμένων (homogeneous coordinates)**.

Ο πραγματικός χώρος R^3, αντίστοιχα ο χώρος *XYZ*, περιγράφεται από ομογενείς συντεταγμένες ως ο **προβολικός χώρος** P^3. Οι ομογενείς συντεταγμένες

$$(t, u, v, w)$$

προβολικών σημείων αυτού του προβολικού χώρου πρέπει να ερμηνευθούν ως εξής:

<ins>Περίπτωση $w \neq 0$:</ins> Το προβολικό σημείο (t, u, v, w) ταυτοποιείται από το σημείο

$$(X, Y, Z) = (t/w, u/w, v/w)$$

του χώρου *XYZ*.

<ins>Περίπτωση $w = 0$:</ins> Ενα προβολικό σημείο $(t, u, v, 0)$ ταυτοποιείται από ένα **σημείο στο άπειρο** του χώρου *XYZ*.

Αν ένα σημείο του χώρου *XYZ* πρέπει να προσδιοριστεί σε προβολικές συντεταγμένες, τότε μπορεί να χρησιμοποιηθεί οποιαδήποτε τιμή w /= 0. Για παράδειγμα, ένα επίπεδο

$$Z = p\,X + q\,Y + r,$$

έχει την προβολή

$$v = p\,t + q\,u + r\,w$$

σε ομογενείς συντεταγμένες. Για $w = 0$, η προβολή είναι ένα επίπεδο στο άπειρο του χώρου P^3.

8. Το πρόβλημα ανακατασκευής σχήματος

8.1.Γενικά

Το **γενικό πρόβλημα ανακατασκευής σχήματος (general shape reconstruction problem)** αποτελείται από τρία βασικά βήματα:

- Το πρώτο βήμα του γενικού προβλήματος ανακατασκευής σχήματος συνίσταται στη δημιουργία μερικών ή πλήρων **χαρτών βάθους (depth maps)** ή **χαρτών ύψους (height maps)** (δηλαδή πινάκων βαθμωτών δεδομένων αποστάσεων) ή τουλάχιστον μερικών ή πλήρων **χαρτών βαθμίδας (gradient maps)** για όλα τα ορατά επιφανειακά σημεία και για ένα επιλεγμένο σύνολο προβολών.
- Η **καταγραφή (registration)** των διαφόρων χαρτών που αντιστοιχούν σε διαφορετικές προβολικές κατευθύνσεις αποτελεί το δεύτερο βήμα του γενικού προβλήματος ανακατασκευής σχήματος. Στο βήμα αυτό, οι επιμέρους χάρτες συνδιάζονται σε ένα ομοιόμορφο σύστημα συντεταγμένων για τη δεδομένη σκηνή ή το δεδομένο αντικείμενο.
- Η **ολοκλήρωση (integration)** των καταγραμμένων χαρτών αποτελεί το τρίτο βήμα του γενικού προβλήματος ανακατασκευής σχήματος και μπορεί να οδηγήσει στη δημιουργία ενός πλήρους τριδιάστατου μοντέλου για τη δεδομένη σκηνή ή το δεδομένο αντικείμενο.

8.2 Εικόνα απόστασης, χάρτης βάθους, χάρτης ύψους και χάρτης βαθμίδας

Το προβολικό κέντρο για την κεντρική προβολή είναι η αρχή $\mathbf{O} = (0, 0, 0)$ του συστήματος συντεταγμένων *XYZ*. Από το σημείο αυτό αναδύεται μια **ακτίνα προβολής (ray of projection)**

$$\gamma(t) = \mathbf{O} + t\,(\mathbf{Q} - \mathbf{O}) = t\,\mathbf{Q} = (t\,x, t\,y, t\,f), \quad t \geq 1$$

προς το χώρο της σκηνής και περνάει από κάποιο σημείο $\mathbf{Q} = (x, y, f)$ του επιπέδου εικόνας *xy*.

Στην περίπτωση παράλληλης προβολής, η ακτίνα προβολής στο χώρο σκηνής μπορεί να μοντελοποιηθεί ως

$$\gamma(t) = (x, y, t), \quad \text{με} \quad t \geq f.$$

Αυτή η ακτίνα περνάει από κάποιο σημείο εικόνας $\mathbf{Q} = (x, y, f)$ και είναι παράλληλη στον οπτικό άξονα. Σε αυτή την περίπτωση, το προβολικό κέντρο είναι ένα σημείο στο άπειρο.

Ας υποθέσουμε τώρα ότι έχει επιλεγεί κάποιο προβολικό μοντέλο. Ένα επιφανειακό σημείο $\mathbf{P} = (x, y, z)$ προβάλλεται στο σημείο εικόνας $\mathbf{p} = (x, y)$ του επιπέδου εικόνας αν υπάρχει προβολική ακτίνα που τέμνει την επιφάνεια του αντικειμένου στο σημείο **P** = (x, y, z) «για πρώτη φορά» (δηλαδή για το μικρότερο δυνατό $t \geq 1$, αντίστοιχα $t \geq f$) αφού περάσει από το σημείο εικόνας **Q** = (x, y, f).

Αυτός ο ορισμός της προβολής πάνω στο επίπεδο εικόνας προσδιορίζει το πολύ ένα σημείο για κάθε σημείο εικόνας. Επομένως αποκλείονται καταστάσεις όπως η επόμενη. Στην περίπτωση της παράλληλης προβολής όλα τα επιφανειακά σημεία με ταυτόσημες συντεταγμένες *XY* προβάλλονται γεωμετρικά στο ίδιο σημείο (x,y), όπως για παράδειγμα επιφανειακά σημεία σε μια πλευρά κάθετη στο επίπεδο εικόνας. Ο ορισμός μας προσδιορίζει το «πλησιέστερο σημείο» αν υπάρχουν πολλά «υποψήφια». Γενικά (και στην περίπτωση της κεντρικής προβολής), αν πολλά σημεία βρίσκονται στην ίδια προβολική ακτίνα, τότε επιλέγουμε εκείνο το σημείο με τη μικρότερη συντεταγμένη t. Αν ένα επιφανειακό σημείο προβάλλεται σε ένα σημείο εικόνας, τότε αυτό το σημείο είναι **ορατό (visible)** αναφορικά με το προβολικό μοντέλο.

Ο χάρτης βάθους, ο χάρτης ύψους και ο χάρτης βαθμίδας είναι συναρτήσεις ορισμένες πάνω στο πεδίο ορισμού $\{ (x, y): 1 \leq x \leq M, 1 \leq y \leq N \}$ της εικόνας. Η κατανομή τιμών σε χάρτες βάθους ή χάρτες ύψους μπορεί να οπτικοποιηθεί εύκολα με εικόνες γκρίζας απόχρωσης σύμφωνα με τη σύμβαση: «λαμπερό» αντιστοιχεί σε «τιμή μεγάλου μεγέθους» και «σκοτεινό» αντιστοιχεί σε «τιμή μικρού μεγέθους».

Στην ιδανική περίπτωση, σε κάθε σημείο (x,y), ο **χάρτης βάθους** $Z=Z(x,y)$ υποδηλώνει το βάθος Z του επιφανειακού σημείου **P** = (X, Y, Z) που προβλέπεται στο σημείο εικόνας **p** = (x,y).

Ο **χάρτης ύψους** $H(x,y)$ ορίζεται σχετικά με κάποιο επίπεδο αναφοράς (μηδενικού ύψους), που είναι παράλληλο στο επίπεδο εικόνας. Στην ιδανική περίπτωση, κάθε σημείο (x,y) η τιμή $H(x,y)$ ισούται με το ύψος του επιφανειακού σημείου **P**=(X,Y,Z) που προβάλλεται στο σημείο εικόνας **p** = (x,y). Το ύψος μετριέται σε σχέση με το επιλεγμένο επίπεδο αναφοράς. Οι οπτικοποιήσεις ενός χάρτη βάθους και ενός χάρτη ύψους συμπεριφέρονται ως μια θετική και μια αρνητική φωτογραφία.

Μια τρίτη επιλογή είναι ο υπολογισμός μιας εικόνας απόστασης **(range image)**. Όμως, η ευκλείδια απόσταση

$$d_2(P,O) = \sqrt{X^2 + Y^2 + Z^2}$$

ανάμεσα στο ορατό επιφανειακό σημείο **P** = (X, Y, Z) και την αρχή του συστήματος συντεταγμένων της κάμερας ή η ευκλείδια απόσταση

$$d_2(P,Q) = \sqrt{(X-x)^2 + (Y-y)^2 + (Z-f)^2}$$

ανάμεσα στο ορατό σημείο **P** = (X, Y, Z) και στο προβολικό σημείο **Q** = (x, y, z) στο επίπεδο εικόνας είναι γεωμετρικά ισοδύναμες με το βάθος Z.

Συγκριτικά με τους χάρτες ύψους, βάθους ή απόστασης, η επόμενη αναπαράσταση επιφανειών αντικειμένων είναι κάπως «ασθενέστερη». Στην ιδανική περίπτωση, σε κάθε σημείο εικόνας (x, y), ένας **χάρτης βαθμίδας (gradient map)** δηλώνει τη βαθμίδα (p, q) του επιφανειακού σημείου **P** = ($X, Y,$

Z) που προβάλλεται στο σημείο εικόνας (x, y). Επομένως, ο χάρτης βαθμίδας μπορεί να οπτικοποιηθεί με δύο εικόνες γκρίζας απόχρωσης που περιγράφουν τις συναρτήσεις $p(x,y)$ και $q(x,y)$.

Ο χάρτης βάθους και ο χάρτης ύψους από τη μια και ο χάρτης βαθμίδας από την άλλη δεν αποτελούν γεωμετρικά ισοδύναμες αναπαραστάσεις ορατών επιφανειακών σημείων. Για παράδειγμα, υποθέτουμε ότι μια κάμερα παρακολουθεί μα σειρά από σκαλοπάτια σε τρόπο ώστε να «βλέπει» μόνο τις πλευρές που είναι παράλληλες στο επίπεδο εικόνας της κάμερας. Όλες αυτές οι επιφάνειες έχουν την ίδια βαθμίδα και επομένως ο χάρτης βαθμίδας θα είχε την ίδια διανυσματική τιμή σε κάθε θέση. Όμως τα σκαλοπάτια θα εμφανίζονταν σε διαφορετικές θέσεις στους χάρτες ύψους ή βάθους. Τα δεδομένα βαθμίδας μπορούν να υπολογιστούν προσεγγιστικά από διακριτούς χάρτες ύψους ή βάθους, αλλά η δημιουργία χαρτών ύψους ή βάθους από χάρτες βαθμίδας είναι πολύ δύσκολη.

Παράδειγμα : Θεωρούμε μια σφαίρα στο χώρο XYZ με κέντρο (0, 0, a) στον άξονα Z και ακτίνα r. Τα σημεία **P** = (X, Y, Z) στην επιφάνεια της σφαίρας ικανοποιούν τη συναρτησιακή εξίσωση

$$X^2 + Y^2 + (Z - a)^2 = r^2 ,$$

όπου για να βεβαιωθούμε ότι η σφαίρα και το επίπεδο εικόνας δεν τέμνονται υποθέτουμε

$$a - r > f .$$

Στην περίπτωση παράλληλης προβολής, $x = X$ και $y = Y$ και ένα σημείο τομής (X,Y,Z) της επιφάνειας της σφαίρας με μια ακτίνα προβολής $\gamma(t) = (X, Y, t)$ ορίζεται ως η πραγματική λύση ως προς t της εξίσωσης

$$X^2 + Y^2 + (t - a)^2 = r^2 .$$

Όταν υπάρχουν δύο πραγματικές λύσεις $t \geq f$, η μικρότερη $t_{\min}$ από αυτές λαμβάνεται ως τιμή βάθους.

$$Z(x, y) = t_{\min} = a - \sqrt{r^2 - x^2 - y^2}$$

Ισχύει ακόμα:

$$grad(Z)(x, y) = (p(x, y), q(x, y)) = (\frac{x}{\sqrt{r^2 - x^2 - y^2}}, \frac{y}{\sqrt{r^2 - x^2 - y^2}}).$$

Για την περίπτωση της κεντρικής προβολής όπου

$$x = \frac{fx}{Z} \quad και \quad x = \frac{fy}{Z}$$

οι τομές με τις ακτίνες προβολής *γ*(*t*) προκύπτουν από τις πραγματικές λύσεις ως προς $t \geq 1$ της εξίσωσης

$$t^2x^2 + t^2y^2 + (tf - a)^2 = r^2$$

Η επιλογή της μικρότερης λύσης οδηγεί στο ορατό σημείο τομής με την επιφάνεια της σφαίρας

$$t_{\min} = \frac{af - \sqrt{r^2(x^2 + y^2 + f^2) - a^2(x^2 + y^2)}}{x^2 + y^2 + f^2}$$

και επομένως ο χάρτης βάθους είναι $Z(x,y) = tf$. Αντικαθιστώντας την τιμή αυτή στις προβολικές εξισώσεις, οδηγούμαστε σε τιμές για τις συντεταγμένες *X*, *Y* και σε έκφραση για το χάρτη βαθμίδας.

Το παράδειγμα αυτό δείχνει πως μπορούν να δημιουργηθούν συνθετικά δεδομένα χαρτών ύψους, βάθους ή βαθμίδας για κάποιο γνωστό αντικείμενο. Το πρόβλημα όμως της τεχνητής όρασης είναι ακριβώς το αντίστροφο, δηλαδή επιχειρεί να ανακατασκευάσει πλευρές τριδιάστατων αντικειμένων από προβολικά δεδομένα. Θα δούμε στη συνέχεια κατά πόσο η δημιουργία χαρτών ύψους, βάθους ή βαθμίδας αποτελεί ένα χρήσιμο ενδιάμεσο βήμα.

8.3 Αντίστροφη προβολή (backprojection)

Οι επιφάνειες των αντικειμένων μιας σκηνής προβάλλονται στο επίπεδο εικόνας με βάση κάποιο από τα προβολικά μοντέλα που είδαμε σε προηγούμενα σημεία. Η αντίστροφη απεικόνιση των προβεβλημένων επιφανειών στον τριδιάστατο χώρο ονομάζεται γενικά **αντίστροφη προβολή (backprojection)**. Οι χάρτες βάθους (ύψους) και βαθμίδας παρέχουν σημαντική και διαφορετική πληροφορία για τη διαδικασία της αντίστροφης προβολής.

Ως παράδειγμα ανακατασκευής μιας πλευράς αντικειμένου, θεωρούμε ένα τμήμα του επιπέδου $Z = pX + qY + r$ και υποθέτουμε κεντρική προβολή. Είναι γνωστό ότι οι συντεταγμένες τριών σημείων που δεν βρίσκονται στην ίδια ευθεία ορίζουν μοναδικά ένα επίπεδο. Συνήθως όμως, τα δεδομένα δεν είναι αυτής της μορφής. Μας ενδιαφέρει να προβάλλουμε αντίστροφα μεμονωμένα σημεία εικόνας (x,y) στον τριδιάστατο χώρο. Η «τοπική γνώση γύρω» από αυτό το σημείο θα πρέπει να υποστηρίζει ένα τρόπο για ανακατασκευή του ορατού επιφανειακού σημείου (X,Y,Z) που προβάλλεται στο (x,y) καθώς και κάποιου κομματιού της «επιφάνειας γύρω από το (X,Y,Z)».

Γενικά, οι συντεταγμένες ενός ορατού σημείου (*X,Y,Z*) στην επιφάνεια ενός αντικειμένου μπορούν να υπολογιστούν ως συναρτήσεις των συντεταγμένων εικόνας *xy,* των πλευρικών παραμέτρων *p*(*x,y*), *q*(*x,y*), *r*(*x,y*) και του εστιακού μήκους *f* αν η πλευρά δεν είναι παράλληλη με τις ακτίνες προβολής. Οι συντεταγμένες εικόνας *x* και *y* είναι γνωστές. Το εστιακό μήκος *f* συνήθως δίνεται. Επομένως, αν μία υποθετική τεχνική ανάλυσης εικόνας απέτρεπε τον υπολογισμό της βαθμίδας, δηλαδή των *p*(*x,y*) και *q*(*x,y*), όπως και του *r*(*x,y*) στα σημεία εικόνας (*x,y*), τότε η ορατή πλευρά θα μπορούσε να ανακατασκευαστεί κατά μοναδικό τρόπο στον τριδιάστατο χώρο με βάση το επόμενο θεώρημα:

Θεώρημα αντίστροφης προβολής πλευράς : Για μια κεντρική προβολή του επιπέδου *Z = pX + qY + r* στο επίπεδο εικόνας *xy*, ισχύουν

$$X = \frac{rx}{f - gy - px}, \quad Y = \frac{ry}{f - gy - px}, \quad \kappa\alpha\iota \quad Z = \frac{rf}{f - gy - px}$$

όπου *f* είναι το εστιακό μήκος και $f - qy - px \neq 0$.

Καταρχήν, το θεώρημα και η υποτιθέμενη τεχνική ανάλυσης εικόνας επιτρέπουν την αντίστροφη προβολή όλων των «ορατών πλευρών αντικειμένων», δηλαδή πλευρών που έχουν μοναδική τομή σε ένα ορατό σημείο με τουλάχιστον μια ακτίνα προβολής. Αν διάφορα ορατά επιφανειακά σημεία προβάλλονται σε διαφορετικά σημεία εικόνας, τότε οι ίδιες οι τιμές βαθμίδας *p* και *q* και οι ίδιες τιμές του *r* μπορούν να χρησιμοποιηθούν για την ανακατασκευή αυτών των επιφανειακών σημείων. Η υποτιθέμενη τεχνική ανάλυσης εικόνας δεν είναι γνωστή μέχρι τώρα στην τεχνητή όραση και, επομένως, το παραπάνω θεώρημα ενθαρρύνει το σχεδιασμό τεχνικών ανακατασκευής σχήματος με βάση υπολογισμούς των *p*, *q* και *r*.

Αντί της υποθετικής τεχνικής ανάλυσης εικόνας, μπορεί να διαθέτουμε μόνο δεδομένα βαθμίδας και τη σταθερά (εστιακό μήκος) *f* ή τιμές απόστασης σε λίγα σημεία εικόνας. Αυτές οι καταστάσεις προσδιορίζουν διαφορετικά προβλήματα αντίστροφης προβολής. Γενικά, έχουμε να απαντήσουμε στην ερώτηση: πώς μπορούν να χρησιμοποιηθούν τα υπολογισμένα δεδομένα για να υποστηρίξουν κάποια τεχνική αντίστροφης προβολής; Ακόμα καλύτερα: Να προσδιοριστεί ένα σύνολο παραμέτρων που υποστηρίζει την ανακατασκευή σχήματος (να μπορεί να αποδειχθεί αυτό) και να σχεδιαστεί μια τεχνική τεχνητής όρασης που προσφέρει «καλής ποιότητας» υπολογισμό αυτών των παραμέτρων.

Μια γενική τεχνική αντίστροφης προβολής μιας πλευράς αντικειμένων μπορεί επίσης να συμπεριλάβει και την ανάλυση γειτονικών πλευρών αντικειμένων. Γειτονικές πλευρές αντικειμένων τέμνονται σε ένα ευθύγραμμο τμήμα. Δύο ευθύγραμμα τμήματα ενός πολυγωνικού συνόρου ή τρεις ακμές στο σύνορο μιας πλευράς προσδιορίζουν μια εξίσωσης για το επίπεδο της πλευράς. Τέτοιες διασυνδέσεις μπορεί να χρησιμοποιηθούν για να υποστηρίξουν τη διαδικασία ανακατασκευής.

Ένα **πολυεδρικό αντικείμενο σκηνής (polyhedral scene object)** είναι προφανώς η φυσική υλοποίηση ενός πολυέδρου. Πολλά αντικείμενα σε σκηνές του πραγματικού κόσμου μπορούν να ερμηνευθούν ως πολύεδρα. Αυτά τα αντικείμενα του πραγματικού κόσμου προσδιορίζουν κάποιο «πολυεδρικό κόσμο». Το θεώρημα αντίστροφης προβολής χρησιμεύει στο σχεδιασμό τεχνικών ανακατασκευής σχήματος για ένα πολυεδρικό κόσμο όπου οι παράμετροι των ανακατασκευασμένων πλευρών βελτιστοποιούνται σε κάθε τμήμα εικόνας και η ανακατασκευή γειτονικών πλευρών (π.χ. γεωμετρικοί περιορισμοί μεταξύ αυτών των τμημάτων) χρησιμοποιούνται για να ελέγξουν την όλη διαδικασία. Οι ευθείες τομής των ανακατασκευασμένων επιπέδων αναδεικνύονται σε ακμές προσανατολισμού των πολυεδρικών αντικειμένων σκηνής.

8.4.Οπτικοποίηση χαρτών βαθμίδας

Για την οπτικοποίηση χαρτών βαθμίδας μπορούν να χρησιμοποιηθούν οι ονομαζόμενοι **χάρτες βελόνων**. Σε αυτούς, γνωστές κάθετοι σε επιφάνειες υπεισέρχονται ως ευθύγραμμα τμήματα με αρχή τα σημεία εικόνας για τα οποία προσδιορίστηκαν. Μια διανυσματική αναπαράσταση των επιφανειακών καθέτων (με βέλη) δεν είναι, γενικά, χρήσιμη για την επιλεγμένη διακριτική ικανότητα επειδή δεν θα είναι οπτικά αναγνωρίσιμη. Σε ορισμένες περιπτώσεις όμως, η τεχνητή όραση επιτρέπει τη δημιουργία πυκνών **χαρτών βελόνων**.

Απαίτηση : Υποθέτοντας ότι οι τιμές των καθέτων ή της βαθμίδας δίνονται ως είσοδος για σημεία εικόνας (x, y), να αναπαρασταθούν στην οθόνη ως χάρτης βελόνων. Οι δεδομένες κάθετοι θεωρούνται είτε ως κάθετοι αυθαίρετου μήκους σε συντεταγμένες *XYZ* είτε ως μοναδιαίοι κάθετοι σε συντεταγμένες μοναδιαίας σφαίρας.
Λύση : Ορίζουμε ότι μια βελόνα αναπαριστάνει μια κάθετο ως ένα ευθύγραμμο τμήμα. Μια δεδομένη κάθετος $\mathbf{n} = (n_x, n_y, n_z)$ στο σύστημα συντεταγμένων *XYZ* με μη κανονικοποιημένο μήκος αναπαριστάνεται με την αντίστοιχη μοναδιαία κάθετο $\mathbf{n}^0 = (n_1, n_2, n_3)$, αφού μια μη κανονικοποιημένη κάθετος θα περιέπλεκε την οπτική ερμηνεία του χάρτη βελόνων. Για τη μοναδιαία κάθετο, χρησιμοποιείται η αναπάρασταση μοναδιαίας σφαίρας ως βάση. Η κατεύθυνση του ευθύγραμμου τμήματος ορίζεται από τη γωνία θ (***tilt***) και το μήκος του τμήματος προσδιορίζεται από το ημίτονο $\sin(\sigma)$ της γωνίας σ (***slant***).
Αλγόριθμος : Δίνεται ο ψευδοκώδικας την επόμενη σελίδα.
Σχόλιο : Για πλευρές με προσανατολισμό παράλληλο προς το επίπεδο εικόνας (δηλαδή με κάθετο ορθογώνια προς το επίπεδο εικόνας και $\sin(\sigma) = 0$), οι βελόνες είναι κουκίδες. Πλευρές σχεδόν ορθογώνιες προς το επίπεδο εικόνας αναπαριστάνονται με βελόνες μέγιστου μήκους. Αν διατίθεται έγχρωμη οθόνη, πληροφορία όπως το μήκος $\sin(\sigma)$ ή η κατεύθυνση της βαθμίδας μπορεί επίσης να αναπαρασταθεί με ψευδοχρώμα.

9. Στατική στερεοσκοπική ανάλυση

9.1 Γενικά

Οι διαδικασίες που αποσκοπούν στην κατανόηση ή ανάλυση επιφανειών τριδιάστατων ορατών αντικειμένων με βάση δεδομένα εικόνας υποδηλώνονται ως **στερεοσκοπική όραση (stereo vision)** σε διάφορα πεδία της επιστήμης. Τα συστήματα όρασης των ζώων και των ανθρώπων αποδεικνύουν ότι η στερεοσκοπική όραση λειτουργεί σε περίπλοκα περιβάλλοντα. Στο ανθρώπινο σύστημα όρασης, δύο μάτια δημιουργούν αντίληψη του βάθους με βάση την **στερεοψία (stereopsis)**. Αυτό είναι γνωστό από τον 19^{o} αιώνα. Παρόλο που μια ζωηρή αντίληψη του περιβάλλοντος μπορεί να αποκτηθεί από ένα μάτι, η ανθρώπινη όραση είναι βασικά μια **διοφθαλμική διαδικασία (binocular process)** που μετασχηματίζει δύο εικόνες από ελαφρά διαφορετικά σημεία σε μια αντίληψη του τριδιάστατου χώρου.

Η τεχνική υλοποίηση της διοφθαλμικής όρασης στην κλασική προσέγγιση της ανακατασκευής σχήματος χρησιμοποιήθηκε στη **φωτογραμμομετρία (photogrammetry)** για πολλές δεκαετίες. Αλληλεπιδραστικές και ακριβές λύσεις επιτρέπουν πολύ ακριβείς ανακατασκευές σχήματος (τηλεπισκόπηση, αντίστροφη μηχανική, αρχιτεκτονική, κ.λ.π.). Η προσέγγιση της τεχνητής όρασης προσανατολίζεται σε αυτοματοποιημένες λύσεις. Η **στατική στερεοσκοπική ανάλυση (static stereo analysis)** υποδηλώνει ένα πολύ ενεργό ερευνητικό πεδίο της τεχνητής όρασης, όπου υποτίθεται ότι δύο τουλάχιστον κάμερες φωτογραφίζουν μία σκηνή ταυτόχρονα ή μέσα σε κάποιο χρονικό διάστημα. Ένα σημαντικό πλεονέκτημα της στατικής στερεοσκοπικής ανάλυσης είναι ότι εκτός από μια δεύτερη (ή τρίτη κ.λ.π.) κάμερα δεν απαιτούνται ούτε επιπλέον εξοπλισμός (π.χ. πηγές φωτός) ούτε ειδικές προβολές της ενέργειας στη σκηνή (π.χ. δομημένος φωτισμός). Βασικά, η στατική στερεοσκοπική ανάλυση είναι η μόνη τεχνική ανακατασκευής σχήματος που μπορεί να ονομαστεί **παθητική (passive)** επειδή δεν απαιτεί αλληλεπίδραση με την τριδιάστατη σκηνή.

Για την στατική στερεοσκοπική ανάλυση, υποτίθεται ότι δεν λαμβάνουν χώρα μετακινήσεις αντικειμένων κατά το χρονικό διάστημα συλλογής των δύο στερεοσκοπικών εικόνων. Αν υπάρχει εγγύηση ότι οι δύο εικόνες λαμβάνονται ακριβώς την ίδια στιγμή, τότε ικανοποιείται αυτή η υπόθεση ακόμα και αν οι κάμερες είναι προσαρμοσμένες πάνω σε κινητά ρομπότ ή τα αντικείμενα της σκηνής είναι κινητά.

Σε σύγκριση με τη **δυναμική στερεοσκοπική ανάλυση (dynamic stereo analysis)**, η διαφορά έγκειται στο ότι πιθανές μετακινήσεις δεν επηρεάζουν τους υπολογισμούς και στο ότι η στατική στερεοσκοπική ανάλυση συνήθως δεν αποσκοπεί στην κατανόηση της κίνησης. Ο σκοπός της στατικής στερεοσκοπικής ανάλυσης είναι μάλλον ο προσδιορισμός πληροφορίας βάθους ή η ανακατασκευή σχήματος. Καταρχήν, κάθε στατική στερεοσκοπική ανάλυση συμπεριλαμβάνει την επόμενη ακολουθία βημάτων επεξεργασίας **(στατική στερεοσκοπική σωλήνωση – static stereo pipeline)**:

1. **Συλλογή εικόνας (image acquisition) :** Αυτή η διαδικασία επηρεάζεται από το περιβάλλον (συμπεριλαμβανομένων των πηγών φωτός, των υπολογιστικών μηχανημάτων κ.λ.π.) και από την εφαρμογή.
2. **Μοντελοποίηση αισθητηρίου (camera modeling)** ή **βαθμονόμηση (calibration) :** Προσδιορισμός των παραμέτρων του αισθητηρίου.
3. **Εξαγωγή χαρακτηριστικών (feature extraction) :** Σε μερικές τεχνικές απαιτείται υπολογισμός των σημαντικών χαρακτηριστικών της εικόνας, όπως ακμών.
4. **Ανάλυση αντιστοιχίας (correspondence analysis) :** Η τεχνητή όραση προσανατολίζεται προς τον αυτοματοποιημένο προσδιορισμό των αντίστοιχων σημείων εικόνων.
5. **Τριγωνοποίηση (triangulation) :** Η τιμή βάθους μπορεί να προσδιοριστεί για ένα σημείο του τριδιάστατου χώρου σκηνής ξεκινώντας από αντίστοιχα σημεία στη δεξιά και αριστερή εικόνα.
6. **Παρεμβολή (interpolation)** ή **προσέγγιση (approximation) :** Συνήθως, τα υπολογισμένα σημεία του τριδιάστατου χώρου σκηνής πρέπει να μετασχηματιστούν σε κάποια αναπαράσταση της επιφάνειας ενός εντικειμένου στον τριδιάστατο χώρο.

Στο κεφάλαιο αυτό, εστιάζουμε κυρίως στην αυτοματοποιημένη διαδικασία ανάλυσης της αντιστοιχίας σε ένα ζευγάρι στερεοσκοπικών εικόνων.

9.2 Γεωμετρία στατικής στερεοσκοπικής ανάλυσης

Συχνά μία κάμερα μπορεί να θεωρηθεί ότι αποτελείται από ένα επίπεδο εικόνας σε απόσταση *f* από ένα προβολικό κέντρο. Το επίπεδο εικόνας βρίσκεται ανάμεσα στο προβολικό κέντρο και την προβαλλόμενη σκηνή. Το προβολικό κέντρο ορίζεται φυσικά από το **εστιακό σημείο (focal point)** ή **οπτικό κέντρο (optical center)** της κάμερας και η απόσταση *f* από το **εστιακό μήκος (focal length)** της κάμερας. Η προβαλλόμενη εικόνα είναι ένας πεπερασμένος πίνακας στο επίπεδο εικόνας. Ο **οπτικός άξονας (optical axis)** της κάμερας τέμνει το επίπεδο εικόνας κατά προσέγγιση στο κέντρο της προβαλλόμενης εικόνας.

Στη συνηθισμένη γεωμετρία στερεοσκοπικής ανάλυσης, θεωρούμε δύο κάμερες του ίδιου εστιακού μήκους *f* έτσι ώστε η απόσταση μεταξύ των εστιακών σημείων τους $\mathbf{O}_L$ και $\mathbf{O}_R$ να ισούται με *b*. Επιτρέπεται οι δύο κάμερες να έχουν μια κλίση προς αλλήλους, πράγμα που επιτρέπει να γίνονται ορατές περισσότερες πλευρές αντικειμένων και στις δύο εικόνες, με αντάλλαγμα κάποιους επιπλέον υπολογισμούς.

Για να απλοποιήσουμε τη συζήτηση, υποθέτουμε ότι οι δύο κάμερες έχουν παράλληλους άξονες *Y*. Χωρίς απώλεια της γενικότητας, υποθέτουμε ότι οι δύο κάμερες έχουν την ίδια κλίση, ίση με γωνία θ, όπως στο σχήμα. Στην απολύτως γενική περίπτωση, οι κάμερες έχουν αυθαίρετη κατεύθυνση και διαφορετικά εστιακά μήκη. Με τις υποθέσεις που κάναμε όμως, γίνεται φανερό πως πρέπει να αντιμετωπιστεί η γενικότερη περίπτωση.

Η γραμμή ανάμεσα στα δύο οπτικά κέντρα **O**$_L$ και **O**$_R$ ονομάζεται **γραμμή βάσεως (base line)** και η τιμή *b* **απόσταση βάσεως (base distance)**. Η γωνία μεταξύ των δύο οπτικών αξόνων είναι 2*θ*. Τα συστήματα συντεταγμένων $X_L Y_L Z_L$ και $X_R Y_R Z_R$ για την αριστερή και τη δεξιά κάμερα, αντίστοιχα, ορίζονται όπως στο σχήμα. Το σύστημα συντεταγμένων *XYZ* ορίζεται έτσι ώστε ο άξονας *Z* να χωρίζει τη γωνία μεταξύ των αξόνων Z_L και Z_R σε δύο ακριβώς ίσες γωνίες *θ*. Το σύστημα συντεταγμένων *XYZ* μπορεί να μετασχηματιστεί στο σύστημα συντεταγμένων $X_L Y_L Z_L$ μέσω μιας περιστροφής κατά γωνία -*θ* γύρω από τον άξονα *Y*

$$\begin{bmatrix} \cos(-\theta) & 0 & -\sin(-\theta) \\ 0 & 1 & 0 \\ \sin(-\theta) & 0 & \cos(-\theta) \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}$$

και μιας μετάθεσης κατά *b*/2 προς τα αριστερά

$$\begin{pmatrix} X - \frac{b}{2} \\ Y \\ Z \end{pmatrix}$$

Άρα, συνολικά ισχύει:

$$\begin{bmatrix} X_L \\ Y_L \\ Z_L \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \begin{pmatrix} X - \frac{b}{2} \\ Y \\ Z \end{pmatrix}$$

Ανάλογα, το σύστημα συντεταγμένων *XYZ* μπορεί να μετασχηματιστεί στο σύστημα συντεταγμένων $X_R Y_R Z_R$ μέσω μιας περιστροφής κατά *θ* γύρω από τον άξονα *Y* και μιας μετάθεσης κατά *b*/2 προς τα δεξιά οπότε και έχουμε:

$$\begin{bmatrix} X_R \\ Y_R \\ Z_R \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix} \begin{pmatrix} X + \frac{b}{2} \\ Y \\ Z \end{pmatrix}$$

Ας υποθέσουμε ότι ένα σημείο **P** = (*X,Y,Z*) στον τριδιάστατο χώρο σκηνής προβάλλεται στα σημεία (X_L,Y_L) και (X_R,Y_R) των επιπέδων εικόνων της κάμερας. Υποθέτοντας κεντρικές προβολές, έχουμε

$$X_L = \frac{fX_L}{Z_L} \; , \; Y_L = \frac{fY_L}{Z_L}$$
$$X_R = \frac{fX_R}{Z_R} \; , \; Y_R = \frac{fY_R}{Z_R}$$

Θεωρούμε την περίπτωση όπου τα αντίστοιχα σημεία εικόνας (X_L,Y_L) και (X_R,Y_R) προσδιορίζονται αντίστοιχα. Τότε, η τριδιάστατη θέση του σημείου **P** = (X,Y,Z) μπορεί να υπολογιστεί από τις παραπάνω εξισώσεις. Έχουμε:

$$X_L = f\,\frac{(x-b/2)\cos\theta + Z\sin\theta}{-(\mathrm{X}-b/2)\sin\theta + Z\cos\theta}$$

$$X_R = f\,\frac{(x+b/2)\cos\theta - Z\sin\theta}{(\mathrm{X}+b/2)\sin\theta + Z\cos\theta}$$

$$Y_L = f\,\frac{Y}{-(\mathrm{X}-b/2)\sin\theta + Z\cos\theta}$$

$$Y_R = f\,\frac{Y}{(\mathrm{X}+b/2)\sin\theta + Z\cos\theta}$$

Οι εξισώσεις αυτές αποτελούν ένα γραμμικό σύστημα που μπορεί να λυθεί ως προς X, Y και Z. Πράγματι ορίζοντας:

$$a_0 = -[\frac{b}{2}X_L\sin\theta + \frac{b}{2}f]$$
$$a_1 = -X_L\sin\theta - f\cos\theta$$
$$a_2 = 0$$
$$a_3 = x_L\cos\theta - f\sin\theta$$
$$b_1 = x_R\sin\theta - f\cos\theta$$
$$b_2 = 0$$
$$b_3 = x_R\cos\theta + f\sin\theta$$
$$b_0 = -[\frac{b}{2}X_R\sin\theta + \frac{b}{2}f]$$
$$c_0 = \frac{b}{2}y_L\sin\theta$$
$$c_1 = -y_L\sin\theta$$
$$c_2 = -f$$
$$c_3 = y_L\cos\theta$$
$$d_0 = -[\frac{b}{2}y_R\sin\theta]$$

$$d_1 = y_R \sin\theta$$
$$d_2 = -f$$
$$d_3 = y_R \cos\theta \,,$$

καταλήγουμε στο σύστημα

$$\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} a_0 \\ b_0 \\ c_0 \\ d_0 \end{bmatrix}$$

το οποίο μπορεί να λυθεί ως προς *X,Y* και *Z*. Η διαδικασία αυτή ονομάζεται **τριγωνοποίηση (triangulation).**

Παράδειγμα : Ας θεωρήσουμε ότι η γωνία *θ* μεταξύ των οπτικών αξόνων ισούται με μηδέν, δηλαδή τα δύο συστήματα συντεταγμένων διαφέρουν μόνο κατά παράλληλη μετάθεση. Ισχύει ότι $Z_L = Z_R = Z$ και η τριγωνοποίηση είναι αρκετά απλή. Η τιμή του βάθους προσδιορίζεται από τη σχέση

$$Z = \frac{fb}{x_L - x_R}$$

για δύο αντίστοιχα σημεία (X_L, Y_L) και (X_R, Y_R) στην αριστερή και δεξιά εικόνα ($Y_L = Y_R$ στην περίπτωση αυτή). Το βάθος είναι αντιστρόφως ανάλογο της βαθμωτής **ανισοτιμίας (disparity)** $\Delta(\mathbf{p}_L) = (x_L - x_R)$ αφού το εστιακό μήκος *f* στις δύο κάμερες και η απόσταση βάσεως *b* ανάμεσα στα προβολικά κέντρα είναι σταθερές στην παραπάνω εξίσωση. Η γεωμετρία αυτού του παραδείγματος ονομάζεται **συνήθης στερεοσκοπική γεωμετρία (standard stereo geometry)** και, συχνά, περιπτώσεις άλλων γενικότερων γεωμετριών μετασχηματίζονται πρώτα σε αυτήν.

Ο επόμενος γεωμετρικός περιορισμός είναι χρήσιμος στο σχεδιασμό χρονικά αποτελεσματικών διαδικασιών ανάλυσης αντιστοιχίας. Όπως προηγουμένως, θα υποθέσουμε ότι ένα σημείο σκηνής **P** προβάλλεται στο σημείο $\mathbf{p}_L$ στο αριστερό επίπεδο εικόνας και στο σημείο $\mathbf{p}_R$ στο δεξί επίπεδο εικόνας. Υποθέτουμε ότι ξεκινάει η διαδικασία ανάλυσης αντιστοιχίας με ένα σημείο στο ένα επίπεδο εικόνας, έστω $\mathbf{p}_L$. Η αναζήτηση για το αντίστοιχο σημείο $\mathbf{p}_R$ στο δεύτερο επίπεδο εικόνας μπορεί να απλοποιηθεί σημαντικά χρησιμοποιώντας την **επιπολική γεωμετρία (epipolar geometry)** που εμφανίζεται κατά τη διοφθαλμική συλλογή εικόνων. Αυτό περιορίζει το χώρο αναζήτησης. Αντί για αναζήτηση σε ολόκληρη την εικόνα της δεύτερης κάμερας, μπορούμε να περιορίσουμε τη διαδικασία αναζήτησης πάνω στην ευθεία της τομής αυτού του δεύτερου επιπέδου εικόνας με το λεγόμενο επιπολικό επίπεδο.

Η γραμμή βάσεως διέρχεται από τα προβολικά κέντρα $\mathbf{O}_L$ και $\mathbf{O}_R$. Γενικά, η γραμμή βάσεως τέμνει κάθε επίπεδο εικόνας σε ένα σημείο (εννοείται, εφόσον τα δύο επίπεδα εικόνας δεν βρίσκονται στο ίδιο επίπεδο). Αυτά τα δύο σημεία τομής είναι οι **επιπόλοι (epipoles)** της δεδομένης διάταξης διοφθαλμικής κάμερας. Αν και τα δύο επίπεδα εικόνας βρίσκονται πάνω στο ίδιο επίπεδο, τότε οι επιπόλοι είναι σημεία στο άπειρο και μπορούν να προσδιοριστούν σε ομογενείς συντεταγμένες. Η γραμμή βάσεως ορίζεται επίσης από το προβολικό κέντρο μιας από τις κάμερες και τον επιπόλο στο επίπεδο εικόνας αυτής της κάμερας.

Ένα *επιπολικό επίπεδο* ορίζεται από ένα σημείο **P** στον τριδιάστατο χώρο και τα δύο οπτικά κέντρα $\mathbf{O}_L$ και $\mathbf{O}_R$. Το ίδιο επίπεδο ορίζεται από τις επόμενες δύο γνωστές ευθείες γραμμές: η μια ευθεία είναι η ακτίνα από το προβολικό κέντρο O_L στο δεδομένο σημείο εικόνας $\mathbf{p}_L$ στο επίπεδο εικόνας της μιας κάμερας. Η γραμμή βάσεως που διέρχεται από τα $\mathbf{O}_L$ και $\mathbf{O}_R$ μπορεί να χρησιμοποιηθεί ως η δεύτερη ευθεία. Επομένως, το επιλεγμένο σημείο $\mathbf{p}_L$ προσδιορίζει το αντίστοιχο επιπολικό επίπεδο για δεδομένη γεωμετρική βαθμολόγηση της κάμερας.

Μια **επιπολική γραμμή (epipolar line)** είναι η τομή ενός επιπολικού επιπέδου με ένα επίπεδο εικόνας. Όλα τα αντικείμενα στη σκηνή που βρίσκονται πάνω σε ένα επιπολικό επίπεδο που ορίζεται από το $\mathbf{p}_L$ προβάλλονται πάνω στο δεύτερο επίπεδο εικόνας σε σημεία αυτής της επιπολικής γραμμής. Σημειώνουμε ότι ο επιπόλος ενός επιπέδου εικόνας βρίσκεται επίσης πάνω στην επιπολική γραμμή.

Σε συμφωνία με τις παραπάνω γεωμετρικές σχέσεις, το πρόβλημα της στατικής στερεοσκοπικής ανάλυσης καταλήγει στο πρόβλημα του **ταιριάσματος αντίστοιχων σημείων εικόνας** στην αριστερή και τη δεξιά εικόνα.

9.3 Υποθέσεις και περιορισμοί

Γενικά λαμβάνουν χώρα απροσδιοριστίες κατά τη διαδικασία προσδιορισμού αντιστοίχων σημείων σε στερεοσκοπικές εικόνες. Για ένα επιλεγμένο εικονοστοιχείο της μιας εικόνας, γενικά υπάρχουν περισσότερα από ένα εικονοστοιχεία στην άλλη εικόνα ως πιθανά υποψήφια για αντιστοιχία. Η μείωση αυτών των απροσδιοριστιών μπορεί να βασιστεί, για παράδειγμα, σε εκ των προτέρων γνώση των παραμέτρων αντικειμένων της σκηνής ή των διαστημάτων των τιμών παραμέτρων αντικειμένων. Αν, για παράδειγμα, γνωρίζουμε ότι τα αντικείμενα της σκηνής είναι πολύεδρα ή ότι έχουν καμπύλες επιφάνειες, τότε αυτή η πληροφορία μπορεί να χρησιμοποιηθεί κατευθείαν για τη σχεδίαση μιας διαδικασίας ανάλυσης αντιστοιχίας.

Όμως, η πληροφορία αυτή είτε δεν διατίθεται η δεν είναι επιθυμητό να περιοριστεί η γεωμετρία των αντικειμένων. Και σε αυτές τις περιπτώσεις όμως, είναι δυνατό να προσδιοριστούν περιορισμοί για να μειωθούν οι απροσδιοριστίες των διαδικασιών ανάλυσης αντιστοιχίας. Μια πρώτη επίδειξη έγινε με την επιπολική γεωμετρία. Τέτοιοι περιορισμοί προκύπτουν από μοντέλα για τη δημιουργία εικόνων, όπως για παράδειγμα η προβολική γεωμετρία ή οι φωτομετρικές ιδιότητες, και επίσης από τη μελέτη των χαρακτηριστικών αντικειμένων στο φυσικό μας περιβάλλον. Η μελέτη μοντέλων δημιουργίας εικόνων επιτρέπει τη δημιουργία γεωμετρικών περιορισμών για

- τη θέση των επιπολικών γραμμών **(επιπολικός περιορισμός – epipolar constraint)**
- τη μοναδικότητα της ανάθεσης **(περιορισμός μοναδικότητας – uniqueness constraint)**
- τη συμβατότητα των τιμών της έντασης στην εικόνα **(περιορισμός φωτομετρικής συμβατότητας – photometric compatibility constraint)**
- τη γεωμετρική ομοιότητα κάποιων χαρακτηριστικών **(περιορισμός γεωμετρικής ομοιότητας – geometric similarity constraint)**

Με βάση ιδιότητες των αντικειμένων σκηνής, επιπλέον περιορισμοί προκύπτουν για

- τη διάταξη των προβαλλομένων σημείων στις εικόνες **(περιορισμός διάταξης – ordering constraint)**
- τη συνέχεια των ανισοτήτων **(περιορισμός συνέχειας – continuity constraint)**
- τη συνέχεια κατά μήκος ακμών προσανατολισμού **(περιορισμός σχηματικής συνέχειας – figural continuity constraint)**
- τη συμβατότητα κάποιων χαρακτηριστικών **(περιορισμός συμβατότητας – compatibility constraint)**
- το όριο ανισοτιμίας **(disparity limit)** ή
- το όριο ανισοτιμίας βαθμίδας **(disparity gradient limit)**

Παρακάτω ορίζονται μερικοί από αυτούς τους περιορισμούς. Σε μια τεχνική ανάλυσης αντιστοιχίας, δεν μπορούν να ληφθούν υπόψη όλοι αυτοί οι περιορισμοί. Η επιβολή όμως έστω και μερικών από αυτούς βοηθάει στη μείωση απροσδιοριστιών, αλλά και επιταχύνει την εκτέλεση των υπολογισμών.

1. **Επιπολικός περιορισμός :** Ένα σημείο $\mathbf{p}_L$ στην αριστερή εικόνα μπορεί να αντιστοιχεί μόνο σε ένα σημείο της δεξιάς εικόνας που βρίσκεται πάνω στην αντίστοιχη επιπολική γραμμή της δεξιάς εικόνας που προσδιορίζει μοναδικά το σημείο $\mathbf{p}_L$. Όταν χρησιμοποιείται η συνήθης στερεοσκοπική γεωμετρία, οι επιπολικές γραμμές στα δύο επίπεδα εικόνας συμπίπτουν με τις οριζόντιες γραμμές σάρωσης των εικόνων.
2. **Υπόθεση μοναδικότητας :** Συνήθως κάθε εικονοστοιχείο της μιας εικόνας σε ένα στερεοσκοπικό ζευγάρι αντιστοιχεί ακριβώς σε ακριβώς ένα εικονοστοιχείο της άλλης εικόνας.

3. **Υπόθεση συμβατότητας των τιμών έντασης εικόνας :** Δύο γειτονικά εικονοστοιχεία (X_{L1},Y_{L1}) και (X_{L2},Y_{L2}) στην E_L και δύο γειτονικά εικονοστοιχεία (X_{R1},Y_{R1}) και (X_{R2},Y_{R2}) στην E_R μπορεί να αντιστοιχούν προς άλληλα μόνο αν οι απόλυτες διαφορές $|E_L(X_{Li},Y_{Li}) - E_R(X_{Ri},Y_{Ri})|$, i = 1,2, μεταξύ των τιμών των εντάσεών τους βρίσκονται κάτω από μια τιμή κατωφλίου και η απόλυτη

διαφορά $|[E_L(X_{L1},Y_{L1}) - E_L(X_{L2},Y_{L2})] - [E_R(X_{R1},Y_{R1}) - E_R(X_{R2},Y_{R2})]|$ στις αλλαγές έντασης είναι μικρότερη από κάποιο κατώφλι.

4. **Υπόθεση γεωμετρικής ομοιότητας :**

4.1. **Κριτήριο γωνίας :** Ένα ευθύγραμμο τμήμα S_L με προσανατολισμό W_L στην αριστερή εικόνα αντιστοιχεί σε ένα ευθύγραμμο τμήμα S_R με προσανατολισμό W_R στη δεξιά εικόνα μόνο αν η διαφορά $|W_L\text{-}W_R|$ είναι μικρότερη από κάποιο κατώφλι.

4.2. **Κριτήριο μήκους :** Ένα ευθύγραμμο τμήμα S_L με μήκος L_L στην αριστερή εικόνα αντιστοιχεί σε ένα ευθύγραμμο τμήμα S_R με μήκος L_R στη δεξιά εικόνα αν η διαφορά $|L_L\text{-}L_R|$ είναι μικρότερη από κάποιο κατώφλι.

5. **Υπόθεση συνέχειας για ανισότητες :** Αν δύο σημεία (X_{L1},Y_{L1}) και (X_{R1},Y_{R1}), στην αριστερή και δεξιά εικόνα αντιστοιχούν μεταξύ τους, τότε ένα σημείο (X_{L2},Y_{L2}) στην τοπική γειτονιά του σημείου (X_{L1},Y_{L1}) στην αριστερή εικόνα αντιστοιχεί συνήθως μόνο σε σημείο (X_{R2},Y_{R2}) στη δεξιά εικόνα για το οποίο το μήκος του διανύσματος της διαφοράς των δύο διανυσμάτων ανισοτιμίας είναι μικρότερο από κάποιο κατώφλι. Στην περίπτωση της συνήθους στερεοσκοπικής γεωμετρίας πρέπει η διαφορά $||x_{L1}\text{-}x_{R1}| - |x_{L2}\text{-}x_{R2}||$ να είναι μικρότερη από κάποιο κατώφλι.

6. **Περιορισμός σχηματικής συνέχειας :** Αν δύο σημεία ακμών (X_{L1},Y_{L1}) και (X_{R1},Y_{R2}) στην αριστερή και τη δεξιά εικόνα αντιστοιχιών μεταξύ τους, τότε ένα σημείο ακμής (X_{L1},Y_{L1}) στην τοπική περιοχή του σημείου (X_{L1},Y_{L1}) στην αριστερή εικόνα μπορεί να αντιστοιχεί μόνο σε ένα σημείο ακμής (X_{R2},Y_{R2}) στη δεξιά εικόνα για το οποίο το μήκος

$$\left|\sqrt{(X_{L1}-X_{R1})^2}-\sqrt{(X_{L2}-X_{R2})^2}\right|$$

του διανύσματος της διαφοράς των δύο διανυσμάτων ανισότητας έχει τιμή μικρότερη από κάποιο κατώφλι.

7. **Υπόθεση συμβατότητας των χαρακτηριστικών :** Τα χαρακτηριστικά (ακμές, σκιές, επικαλύψεις κ.λ.π.) στις εικόνες ενός στερεοσκοπικού ζεύγους μπορούν να αντιστοιχούν μεταξύ τους μόνο αν έχουν την ίδια φυσική αρχή στη σκηνή.

8. **Όριο ανισοτιμίας :** Υπάρχει μια μέγιστη τιμή d_{max} για το μήκος των διανυσμάτων ανισοτιμίας για όλα τα αντίστοιχα σημεία $\mathbf{p}_L = (x_L,y_L)$ και $\mathbf{p}_L = (x_R,y_R)$ στην αριστερή και δεξιά εικόνα, δηλαδή

$$\left\|\Delta(X_L,Y_L)\right\| = \sqrt{(X_L-X_R)^2+(Y_L-Y_R)^2} \le d_{\max}$$

9. **Όριο βαθμίδας ανισοτιμίας :** Για δύο διαδοχικά σημεία a_L και b_L στην αριστερή εικόνα, υποθέτουμε ότι a_R και b_R είναι τα αντίστοιχα σημεία τους. Τότε, b_R είναι υποψήφιο σημείο του b_L αν

$$\frac{\left\|\Delta(a_L)-\Delta(b_{L)}\right\|}{d_{cs}(a_L,b_L)} \leq \gamma_{\max}$$

όπου γ_{max} είναι το όριο βαθμίδας ανισοτιμίας και

$$d_{cs}(a_L,b_L)=d_2\left(\frac{a_L+a_R}{2},\frac{b_L+b_R}{2}\right)$$

10. **Περιορισμός διάταξης :** Σημεία που βρίσκονται σε μια επιπολική γραμμή σε μια εικόνα ενός στερεοσκοπικού ζεύγους προβάλλονται στην αντίστοιχη επιπολική γραμμή της άλλης εικόνας με ακριβώς την ίδια διάταξη.

9.4.Ο αλγόριθμος Shirai

Θεωρούμε πάλι δύο κάμερες σε διάταξη συνήθους στερεοσκοπικής γεωμετρίας. Αν θεωρήσουμε ότι οι παράμετροι f και b είναι γνωστές, ο επόμενος αλγόριθμος (Shirai) επιτρέπει την ανακατασκευή ορατών επιφανειακών σημείων αντικειμένων στον τριδιάστατο χώρο σκηνής.

Απαίτηση : Να υπολογιστούν αντίστοιχα σημεία εικόνας και ο χάρτης ανισοτήτων για ζεύγη στερεοσκοπικών εικόνων γκρίζας απόχρωσης. Υποθέτουμε ότι χρησιμοποιείται η συνήθως στερεοσκοπική γεωμετρία για τις δεδομένες εικόνες, δηλαδή ότι επαρκεί ο περιορισμός της αναζήτησης για αντίστοιχο σημείο στην ίδια γραμμή της δεύτερης εικόνας. Επιπλέον, υποθέτουμε ότι η αναζήτηση αρχίζει πάντα από ένα εικονοστοιχείο ακμής.
Λύση : Χωρίς καμιά απώλεια γενικότητας, υπολογίζουμε την εικόνα των ακμών της αριστερής εικόνας, δηλαδή η αναζήτηση αρχίζει πάντα με αρχικό σημείο στην αριστερή εικόνα. Για κάθε εικονοστοιχείο ακμής

$$(\mathbf{p}, E_{left}(\mathbf{p})) = ((x_{left},y), E_{left}(x_{left},y))$$

της αριστερής εικόνας ξεκινάει μια διαδικασία αναζήτησης για αντίστοιχο εικονοστοιχείο

$$(\mathbf{q}, E_{right}(\mathbf{q})) = ((x_{right},y), E_{right}(x_{right},y))$$

στη δεξιά εικόνα. Για το σκοπό αυτό, μπορούμε γενικά να υποθέσουμε ότι

$$x_{right} \leq x_{left}$$

Επομένως, αρκεί να περιοριστεί η αναζήτηση αντίστοιχου εικονοστοιχείου με τιμή x_{right} στο διάστημα [1, x_{left}].

Κατά τη διάρκεια της αναζήτησης, υπολογίζεται ένα **μέτρο ομοιότητας (similarity measure)** *SIMILARITY*(**p**,**q**) για κάθε εικονοστοιχείο (**q**, E_{right} (**q**)) στη δεξιά εικόνα που αποτελεί υποψήφιο με βάση το διάστημα αναζήτησης. Τα πρόσωπα τιμών γκρίζας απόχρωσης σε παράθυρα $F(E_{left}$, **p**) της αριστερής και $F(E_{right}$, **q**) της δεξιάς εικόνας συγκρίνονται μεταξύ τους για να προσδιοριστεί το μέτρο ομοιότητας. Ο Y. Shirai (1989) πρότεινε το επόμενο μέτρο ομοιότητας *SIMILARITY*(**p**,**q**) (στην πραγματικότητα πρόκειται για μέτρο ανομοιότητας, αλλά ιστορικά χρησιμοποιήθηκε ο όρος «μέτρο ομοιότητας»).

Στην αριστερή εικόνα, ένα σταθερό παράθυρο $F(E_{left}$, **p**) συμπεριφέρεται ως ένας πίνακας σύγκρισης. Κατά την αναζήτηση, το παράθυρο εικόνας $F(E_{right}$, **q**) συγκρίνεται σε όλες τις δυνατές θέσεις στη γραμμή αναζήτησης y με το σταθερό παράθυρο εικόνας χρησιμοποιώντας το **τετραγωνικό σφάλμα (squared error)**

$$SE(p,q) = \sum_{i=-k}^{k}\sum_{j=-k}^{k}\left(E_{left}\left(X_{left}+i, y+j\right) - E_{rightt}\left(X_{right}+i, y+j\right)\right)$$

Αυτή η σύγκριση κανονικοποιείται σε σχέση με την ποσότητα

$$VARIANCE(p) = \frac{1}{(2k+1)^2}\sum_{i=-k}^{k}\sum_{j=-k}^{k}[E_{left}\left(X_{left}+i, y+j\right) - AVERAGE(p)]^2 =$$

$$\frac{1}{(2k+1)^2}\sum_{i=-k}^{k}\sum_{j=-k}^{k}E_{left}{}^2\left(X_{left}+i, y+j\right) - AVERAGE^2(p)_j$$

που αποτελεί την εκτίμηση της διακύμανσης του σταθερού παράθυρου εικόνας στην αριστερή εικόνα. Η τιμή *AVERAGE*(**p**) υποδηλώνει τον αριθμητικό μέσο του παράθυρου εικόνας $F(E_{left}$, **p**) και, κατά συνέπεια, μια εκτίμηση της αναμενόμενης τιμής της γκρίζας απόχρωσης σε αυτό το παράθυρο εικόνας. Συνοπτικά, το μέτρο ομοιότητας του Shirai ορίζεται ως

$$SIMIARITY(p,q) = \frac{SE(p,q)}{VARIANCE(p)+1}.$$

Αν τα δύο παράθυρα είναι ταυτόσημα, τότε το τετραγωνικό σφάλμα είναι μηδέν, δηλαδή το μέτρο ομοιότητας μηδενίζεται. Αλλιώς, οι τιμές του μέτρου ομοιότητας είναι πάντα θετικές.

Αλγόριθμος : Δίνεται ο ψευδοκώδικας στην επόμενη σελίδα.
Σχόλιο : Η απευθείας υλοποίηση του αλγορίθμου μπορεί να είναι μη αποτελεσματική. Είναι σημαντικό να γίνει σωστή επιλογή του τελεστή ακμών, της χρήσης της εκ των προτέρων γνώσης, του περιορισμού των διαστημάτων αναζήτησης κ.λ.π.

Ο αλγόριθμος Shirai χρησιμοποιεί τον επιπολικό περιορισμό, τον περιορισμό μοναδικότητας και τον περιορισμό φωτομετρικής συμβατότητας. Επειδή χρησιμοποιεί ως υποδιαδικασία αλγόριθμο ανίχνευσης ακμών, μπορεί να τροποποιηθεί ώστε να χρησιμοποιεί και τον περιορισμό γεωμετρικής ομοιότητας. Επίσης μπορεί να τροποποιηθεί ώστε να περιλαμβάνει τον περιορισμό συνέχειας.

9.5.Στερεοσκοπική ανάλυση με τρεις κάμερες

Μια απλή επέκταση της διοφθαλμικής στερεοσκοπικής γεωμετρίας επιτυγχάνεται με την προσθήκη μιας τρίτης κάμερας. Τρεις διαφορετικές διατάξεις είναι συνηθισμένες για τέτοιες **τριοφθαλμικές (trinocular)** καταστάσεις. Για την συνεπίπεδη ορθής γωνίας (a) και τη συγγραμμική (b) διάταξη, οι οπτικοί άξονες είναι παράλληλοι και τα επίπεδα εικόνας συνεπίπεδα. Επομένως, οι επιπολικές γραμμές και οι οριζόντιες γραμμές σάρωσης και στις τρεις κάμερες ταυτίζονται (διάταξη (b)) ή οι επιπολικές γραμμές και οι κατακόρυφες γραμμές σάρωσης σε ένα ζεύγος καμερών ταυτίζονται (διάταξη (a)). Το μειονέκτημα της συγγραμμικής διάταξης (b) είναι ότι οι οριζόντιες γραμμές στη σκηνή προβάλλονται στις ίδιες επιπολικές γραμμές και στα τρία επίπεδα. Αυτό δεν προσφέρει επιπρόσθετη πληροφορία για να διευκρινιστούν απροσδιοριστίες στις διαδικασίες ανάλυσης αντιστοιχίας.

Από την άλλη πλευρά, στη συνεπίπεδη διάταξη ορθής γωνίας, οι οριζόντιες όπως και οι κατακόρυφες γραμμές σάρωσης μπορούν να χρησιμοποιηθούν για ανάλυση αντιστοιχίας. Η ανάθεση μιας ακμής μπορεί να γίνει ανεξάρτητα από τον προσανατολισμό της στα επίπεδα εικόνας. Η ακριβής τοποθέτηση των καμερών πρέπει να ικανοποιεί επιπροσθέτως περιορισμούς ακρίβειας σε σχέση με τη διοφθαλμική περίπτωση.

Με τη χρήση μιας διάταξης χωρίς περιορισμούς, η στερεοσκοπική ανάθεση μπορεί να διεξαχθεί με βάση μόνο γεωμετρικούς νόμους. Αυτό επιτρέπει να αποφευχθούν διάφορες περιοριστικές υποθέσεις. Το μειονέκτημα αυτής της διάταξης είναι ότι οι επιπολικές γραμμές δεν αντιστοιχούν με γραμμές σάρωσης στην εικόνα και πρέπει να επαναπροσδιοριστούν για κάθε διάταξη. Ένας τρόπος είναι να μετασχηματιστεί η γεωμετρία σε κάποια συνήθη.

Μια τριοφθαλμική διάταξη καμερών μπορεί να επεκταθεί περαιτέρω με την προσθήκη και άλλων καμερών. Είναι γνωστές διατάξεις με μέχρι οκτώ κάμερες. Σε γενικές γραμμές, οι περιορισμοί που προκύπτουν είναι συνήθως ανάλογοι με αυτούς της τριοφθαλμικής διάταξης.

Οι στρατηγικές ανάλυσης αντιστοιχίας σε τριοφθαλμικές διατάξεις διακρίνονται σε δύο κατηγορίες:

(A) Στρατηγικές που αναθέτουν μόνο εκείνα τα εικονοστοιχεία για τα οποία έχει προσδιοριστεί αντιστοιχία για όλες τις εικόνες.

(B) Στρατηγικές που αναθέτουν όλα τα εικονοστοιχεία για τα οποία έχει προσδιοριστεί αντιστοιχία σε δύο τουλάχιστον εικόνες.

Η πρώτη προσέγγιση (A) είναι να επαληθεύονται οι αναθέσεις. Ελαττώνει τον αριθμό των λαθών σημαντικά, αλλά ταυτόχρονα ελαττώνει και τον αριθμό των αναθέσεων, ειδικά σε περιπτώσεις ακμών που μερικά καλύπτονται σε κάποια εικόνα. Από την άλλη πλευρά, το βασικό κίνητρο για την επιλογή της δεύτερης προσέγγισης (B) είναι συνήθως η επιθυμία να είναι δυνατή η ανάθεση ακμών που μερικά επικαλύπτονται σε μία από τις εικόνες. Αυτό αυξάνει τον αριθμό των αναθέσεων και επομένως και τον αριθμό των ανακατασκευασμένων τριδιάστατων επιφανειακών σημείων.

10. Δυναμική στερεοσκοπική ανάλυση

10.1 Γενικά

Αυτό το κεφάλαιο ασχολείται σχεδόν αποκλειστικά με κινήσεις αντικειμένων σκηνών που προκαλούν αλλαγές στις σκηνές. Εναλλακτικές δυνατότητες αποτελούν κινούμενες κάμερες ή τροποποιήσεις του φωτισμού. Σύμφωνα με αποτελέσματα της φυσιολογίας, η αντίληψη της κίνησης είναι ένα σημαντικό γνωσιακό στοιχείο της οπτικής ερμηνείας του τριδιάστατου κόσμου από ζώα ή ανθρώπους. Η γεωμετρική ανάλυση δυναμικών χώρων σκηνών αποδεικνύει ότι προβαλλόμενα διανύσματα κίνησης μπορούν να υποστηρίξουν τον υπολογισμό τιμών βάθους ή διανυσμάτων βαθμίδας. Στην ιδανική περίπτωση, η κίνηση του αντικειμένου σκηνής στον τριδιάστατο χώρο (τριδιάστατη κίνηση) αντιστοιχεί σε μια διδιάστατη κίνηση σε μια ακολουθία εικόνων. Αυτές οι προβαλλόμενες κινήσεις μπορούν να αναπαρασταθούν στο επίπεδο εικόνας ως ένα **πεδίο διανυσμάτων τοπικής μετατόπισης (field of local displacement vectors).** Αν μια τεχνική επιτρέπει την ακριβή μέτρηση τέτοιων πεδίων τοπικής μετατόπισης, τότε μπορεί να επιτευχθεί ή να υποστηριχτεί μια ανακατασκευή σχήματος στη βάση αυτών των διανυσμάτων τοπικής μετατόπισης. Για παράδειγμα, η κατανόηση των κινήσεων και των σχημάτων του ανθρώπινου σώματος μπορεί να βασιστεί σε επιφανειακά σημάδια που επιτρέπουν την αναμφίβολη παρακολούθηση μεμονωμένων επιφανειακών σημείων όπως είναι γνωστό από τη δεκαετία του 1960. Αυτή η τεχνική της παρακολούθησης επιφανειακών σημαδιών χρησιμοποιείται σε πρόσφατες εφαρμογές της τεχνητής όρασης στις αθλητικές επιστήμες.

Παρ' όλα αυτά, είναι δυνατό μόνο κατά ένα μέρος να υπολογιστούν σχεδόν χωρίς λάθος σύνολα διανυσμάτων τοπικής μετατόπισης. Το κύριο πρόβλημα στην υλοποίηση μιας προσέγγισης δυναμικής στερεοσκοπικής όρασης είναι η σωστή μέτρηση των προβαλλόμενων κινήσεων και η ανάλυση των μερικώς παραμορφωμένων πεδίων διανυσμάτων τοπικής μετατόπισης. Πεδία διανυσμάτων τοπικής μετατόπισης μπορούν να υπολογιστούν προσεγγιστικά ως πεδία οπτικής ροής. Τα πεδία οπτικής ροής επιτρέπουν γενικές εκτιμήσεις του βάθους ή της απόστασης, για παράδειγμα για την πλοήγηση ρομπότ ή για την παρακολούθηση κινουμένων αντικειμένων σε μια σκηνή.

10.2 Διανύσματα μετατόπισης και ανακατασκευή

Σε αυτό το κομμάτι, ασχολούμαστε με επιλεγμένες θεωρητικές αρχές δυναμικής στερεοσκοπικής ανάλυσης. Η τριδιάστατη κίνηση στερεών σωμάτων υποτίθεται ότι δεν υπόκειται σε περιορισμούς, δηλαδή είναι δυνατές αυθαίρετες μετατοπίσεις και/ή περιστροφές. Επίσης, δεν υποθέτουμε καταστάσεις βαθμονομημένης συλλογής εικόνας. Κάτω από αυτές τις γενικές υποθέσεις, οι κάθετοι σε αντικείμενα μπορούν (θεωρητικά) να υπολογίσουν ή τουλάχιστον να οριστεί η περιοχή των τιμών τους.

Διανύσματα τοπικής μετατόπισης

Ας θεωρήσουμε την **ακολουθία εικόνων (image sequence)** E_1, E_2, E_3, …που παράγεται από εικόνες E_i που έχουν συλληφθεί σε κάποιο σταθερό χρονικό διάστημα δt_{const} μεταξύ τους. Για στατικές σκηνές, **σχετικές κινήσεις (relative motions)** μπορούν να παρατηρηθούν σε μια τέτοια ακολουθία εικόνων ως αποτέλεσμα κίνησης της κάμερας ή αλλαγής του φωτισμού μεταξύ των εικόνων E_i και E_{i+1}. Από την άλλη πλευρά, για δυναμικές σκηνές **απόλυτες κινήσεις (absolute motions)**, που είναι αποτέλεσμα μετακινήσεων αντικειμένων προβάλλονται στην ακολουθία εικόνων. Μια σωστή προβολή απόλυτων κινήσεων υποθέτει σταθερή θέση και προσανατολισμό κάμερας και σταθερό φωτισμό. Ταυτόχρονες κινήσεις της κάμερας, αλλαγές στο φωτισμό και μετακινήσεις αντικειμένων δημιουργούν τη δυσκολία να διαχωριστούν οι σχετικές από τις απόλυτες κινήσεις. Δεν θα ασχοληθούμε με τέτοιες περίπλοκες καταστάσεις.

Ας υποθέσουμε στερεά αντικείμενα σκηνής και καθόλου αλλαγές στο φωτισμό ή στις παραμέτρους της κάμερας. Η τριδιάστατη κίνηση ενός επιφανειακού σημείου **P** που είναι ορατό σε δύο διαδοχικές εικόνες E_i και E_{i+1} απεικονίζεται σε ένα διδιάστατο διάνυσμα κίνησης στο επίπεδο εικόνας. Η απεικόνιση εξαρτάται από το επιλεγμένο προβολικό μοντέλο. Ένα επιφανειακό σημείο **P** προβάλλεται στο σημείο εικόνας $\mathbf{p}_{old}=(x,y)$ στην εικόνα E_i και στο $\mathbf{p}_{new}$ στην E_{i+1} . Το διάνυσμα

$$d_{xy}(t_i) = \mathbf{p}_{new} - \mathbf{p}_{old} = \delta t_{const} \cdot (\xi_{xy}(ti), \psi_{xy}(ti)),$$

ορίζεται για μια διακριτή χρονική στιγμή $t_i = i \cdot \delta t_{const}$ και για σημείο εικόνας (x,y). Υποθέτουμε ότι η κίνηση του σημείου μπορεί να διαφοριστεί ως προς το χρόνο. Τότε

$$v_{xy}(t) = \frac{dp}{dt}(t) = \dot{p}(t) = (\xi_{xy}(t), \psi_{xy}(t))$$

και

$$d_{xy}(t) = \delta t_{const} \cdot v_{xy}(t)$$

είναι τα **διανύσματα τοπικής ταχύτητας (local velocity vectors)** και τα **διανύσματα τοπικής μετατόπισης (local displacement vectors)**, αντίστοιχα για την προβολή του τριδιάστατου επιφανειακού σημείου **P** στο επίπεδο εικόνας. Το διάνυσμα μετατόπισης αναπαριστάνει την μετατόπιση (κανονικοποιημένη στη μονάδα χρόνου δt_{const}) του προβαλλόμενου επιφανειακού σημείου στο επίπεδο εικόνας από την εικόνα E_i στην εικόνα E_{i+1}. Τα ξ και ψ υποδηλώνουν ταχύτητες κατά τις διευθύνσεις x και y, αντίστοιχα. Στη διακριτή περίπτωση, ισχύει

$$\mathbf{p}_{new} = \mathbf{p}_{old} + \delta t_{const} \cdot (\xi_{xy}(t_i), \psi_{xy}(t_i)) =$$
$$= (x + \delta t_{const}\, \xi_{xy}(t_i), y + \delta t_{const}\, \psi_{xy}(t_i))$$

και στη γενική περίπτωση

$$\mathbf{p}_{new} = \mathbf{p}_{old} + \delta t_{const} \cdot \dot{p}\,(t_i).$$

Τα διανύσματα τοπικής μετατόπισης $d_{xy}(t)$ αναπαριστάνουν την τριδιάστατη κίνηση στο επίπεδο εικόνας σύμφωνα με τη δεδομένη ή υποτιθέμενη προβολή. Για στερεά αντικείμενα σκηνής, ένα σύνολο από τέτοια διανύσματα $d_{xy}(t)$ ορίζει ένα **πεδίο τοπικής μετατόπισης (local displacement field)** στο επίπεδο εικόνας. Στην περίπτωση διακριτών χρονικών στιγμών t_i , το πεδίο τοπικής μετατόπισης ανατίθεται σε ένα ζευγάρι εικόνων E_i και E_{i+1}. Όμως, θα μπορούσε επίσης να αναπαριστάνει τη μέση κίνηση για πολλές εικόνες της ακολουθίας εικόνων για παράδειγμα, πεδία τοπικής μετατόπισης αποτελούν συχνά τα αρχικά δεδομένα για διερευνήσεις της γεωμετρίας της κίνησης στις περιπτώσεις των κινήσεων αντικειμένων ή κάμερας.

Οι κινήσεις που πρέπει να αναπαρασταθούν στο επίπεδο εικόνας δεν είναι μόνο οι μετατοπίσεις επιφανειακών σημείων αντικειμένων σκηνής αλλά επίσης και περιστροφές. Παρ' όλα αυτά, η κίνηση ενός προβαλλόμενου επιφανειακού σημείου μεταξύ δύο διαδοχικών εικόνων μπορεί προσεγγιστικά να περιγραφεί τοπικά στο επίπεδο εικόνας με ένα μόνο διάνυσμα μετατόπισης. Το **πρόβλημα του ανοίγματος (aperture problem)** δημιουργείται από το γεγονός ότι μόνο τοπικά περιορισμένα ανοίγματα (για παράδειγμα, μια ολόκληρη εικόνα διαστάσεων *M*x*N* είναι επίσης μια χωρικά περιορισμένη προβολή μιας σκηνής) συχνά δεν παρέχουν ικανή πληροφορία για κίνηση αντικειμένων που συμβαίνει ολικά και απεικονίζεται σε μια ακολουθία εικόνων. Αυτό σημαίνει ότι οι αλγόριθμοί μας είναι πάντα περιορισμένων δυνατοτήτων σχετικά με την κλίμακα των ανακατασκευάσιμων κινήσεων αντικειμένων. Πεδία τοπικής μετατόπισης μπορούν να υπολογιστούν από ακολουθίες εικόνων αν μεμονωμένα επιφανειακά σημεία μπορούν να παρακολουθηθούν χωρίς απροσδιοριστία στην ακολουθία εικόνων. Αυτό είναι το **πρόβλημα αντιστοιχίας δυναμικής στερεοσκοπικής ανάλυσης (correspondence problem of dynamic stereo analysis)**. Κατ' αρχή, οι συγκρίσεις τοπικής αντιστοιχίας που συζητήθηκαν για στατική στερεοσκοπική ανάλυση μπορούν επίσης να εφαρμοστούν. Όμως, μια μείωση του διαστήματος αναζήτησης για τα αντίστοιχα σημεία εικόνας (όπως ήταν δυνατόν για στατική ανάλυση θεωρώντας επιπολική γεωμετρία) μπορεί να βασιστεί μόνο σε πληροφορία για περιορισμούς στην τριδιάστατη κίνηση. Για παράδειγμα, θα μπορούσε να υποτεθεί ότι η δυναμική των αντικειμένων σκηνής περιορίζεται σε περιστροφή πάνω σε ένα τραπέζι. Η λύση του προβλήματος αντιστοιχίας για δυναμική στερεοσκοπική ανάλυση μπορεί, για παράδειγμα να υποστηριχτεί επίσης από τη χρήση επιφανειακών σημαδιών, όπως σε συστήματα τεχνητής όρασης που χρησιμοποιούνται σε αθλητικές επιστήμες και στα οποία προστίθενται σημάδια στο δέρμα των αθλητών.

Κίνηση αντικειμένων και τοπική μετατόπιση

Ο υπολογισμός πεδίων τοπικής μετατόπισης για γεωμετρικά (συνθετικά) αντικείμενα μπορεί να χρησιμοποιηθεί, για παράδειγμα, για τον έλεγχο αλγορίθμων ανακατασκευής με αυτά τα συνθετικά δεδομένα εισόδου. Πρώτα πρέπει να προσδιοριστεί ένα προβολικό μοντέλο για την προβολή μετακινήσεων αντικειμένων στο επίπεδο εικόνας. Τότε, προβάλλονται στο επίπεδο εικόνας τα τριδιάστατα διανύσματα κίνησης για ορατά επιφανειακά σημεία. Αυτά τα τριδιάστατα διανύσματα κίνησης δεν είναι σταθερά για ένα στερεό σώμα (π.χ. μετάθεση και κεντρική προβολή). Έστω

$$v_P(t) = v_{XYZ}(t)$$

η **ταχύτητα (velocity)** του επιφανειακού σημείου **P** =(X,Y,Z) τη στιγμή t. Τα τριδιάστατα διανύσματα κίνησης εξαρτώνται από τη θέση (X,Y,Z) του αντίστοιχου επιφανειακού σημείου και ισούνται με το γινόμενο του διανύσματος ταχύτητας $v_P(t)$ και το χρονικό διάστημα δt_{const} ανάμεσα σε δύο εικόνες της ακολουθίας εικόνων. Το διάνυσμα τοπικών μετατοπίσεων του προβαλλομένου σημείου **p** = (x,y) στο επίπεδο εικόνας τη στιγμή t υποδηλώνεται ως

$$d_P(t) = d_{xy}(t) = \delta t_{const} \cdot (\xi_{xy}(t), \psi_{xy}(t)) = \delta t_{const} \cdot v_P(t)$$

Το επίπεδο εικόνας εντοπίζεται στο χώρο XYZ σύμφωνα με το προβολικό μοντέλο με κέντρο την κάμερα, δηλαδή ως το επίπεδο $Z = f$ σε απόσταση f από το κέντρο προβολής. Στο σύστημα συντεταγμένων XYZ τα σημεία **P** και **p** έχουν συντεταγμένες **P** = (X,Y,Z) και p = (X,Y,f), επίσης ταυτοποιούνται με τα διανύσματα θέσης **P** = (X,Y,Z) και **p** = (X,Y,f), αντίστοιχα από την αρχή **O** = (0,0,0) στο αντίστοιχο σημείο. Ισχύει

$$v_{XYZ}(t) = \frac{dp}{dt}(t) = \dot{p}(t) \text{ και } d_{xy}(t) = \frac{dp}{dt}(t) = \dot{p}(t)$$

Για το βαθμωτό γινόμενο του **P**(t) = $(X(t), Y(t), Z(t))$ και του μοναδιαίου διανύσματος **k** = (0,0,1) τον άξονα Z, ισχύει ότι **P**$(t) \cdot k = Z(t)$. Επιπλέον, υποθέτουμε κεντρική προβολή και σύστημα συντεταγμένων με κέντρο την κάμερα. Επομένως, θα έχουμε:

$$\mathbf{p}(t)=\frac{f}{Z(t)}P(t) = \frac{f}{P(t)k}P(t)$$

Παραγωγίζοντας αυτή την εξίσωση ως προς την μεταβλητή του χρόνου, βρίσκουμε:

$$d_{xy}(t)=\frac{dp}{dt}(t) = \frac{f}{Z^2(t)}(P(t)xV_{XYZ}(t))xk$$

όπου έχουμε χρησιμοποιήσει τον κανόνα παραγώγισης πηλίκου.

Αυτό αποτέλεσμα για το μοντέλο κεντρικής προβολής μπορεί να χρησιμοποιηθεί για να δημιουργηθούν συνθετικά αντικείμενα και ένα πεδίο $d_{xy}(t)$

διανυσμάτων τοπικών μετατοπίσεων. Για το σκοπό αυτό, πρέπει να παρέχονται επιλεγμένα επιφανειακά σημεία **P** = (*X,Y,Z*) και τα διανύσματα ταχύτητάς τους

$$v_P(t) = (\dot{X}, \dot{Y}, \dot{Z})$$

ως είσοδος. Από την άλλη, με αυτή την αναπαράσταση του $d_{xy}(t)$ μπορεί να διεξαχθεί μια συζήτηση της συμπεριφοράς των διανυσμάτων τοπικής μετατόπισης σχετικά με την γεωμετρία των αντικειμένων της σκηνής. Επιφανειακά σημεία γειτονικών αντικειμένων έχουν γενικά σχεδόν ταυτόσημα διανύσματα μετατόπισης. Ασυνέχειες στο πεδίο των διανυσμάτων τοπικής μετατόπισης εμφανίζονται σε σύνορα επικάλυψης.

Κίνηση αντικειμένων και βαθμίδες

Για το χαρακτηρισμό της ταχύτητας $v_P(t)$ ενός επιφανειακού σημείου **P**, υποθέτουμε ότι το σημείο αυτό βρίσκεται σε μια πλευρά που αποτελεί μέρος του επιπέδου $Z = pX+qY+r$. Επομένως, οι συντεταγμένες του **P** είναι της μορφής **P** = (A, B, $pA+qB+r$) και η επιφάνεια του αντικειμένου έχει βαθμίδα (p,q) στο σημείο **P**. Αντίστοιχα, η κάθετος στο επιφανειακό σημείο **P** είναι το διάνυσμα (p,q,-1). Υποθέτουμε, επιπλέον, ένα συνδυασμό μιας περιστροφής και στη συνέχεια μετάθεσης ως ομοιόμορφη κίνηση της πλευράς. Για στερεά αντικείμενα, ολικές παραμέτρους κίνησης μιας περιστροφής και μιας μετάθεσης (για παράδειγμα, προσδιορισμένες ως προς το κέντρο βάρους και τους κύριους άξονες του αντικειμένου) και από τη θέση του σημείου **P** ή της πλευράς.

Έστω ότι η περιστροφή της πλευράς γύρω από το κέντρο περιστροφής (o,o,r) χαρακτηρίζεται από ένα διάνυσμα ($\omega_1,\omega_2,\omega_3$) η κατεύθυνση του οποίου περιγράφει τον άξονα περιστροφής και το μήκος του την ταχύτητα περιστροφής. Η μετάθεση ορίζεται από ένα διάνυσμα (a,b,c) του οποίου το μήκος αναπαριστάνει την ταχύτητα μετάθεσης.

Η ταχύτητα ενός σημείου **P**= ($X,Y,pX+qY+r$) αυτής της πλευράς είναι ίση με

$$v_P(t) = (\dot{X}, \dot{Y}, \dot{Z}) = (a,b,c) + (\omega_1,\omega_2,\omega_3) \times (X,Y,Z-r)$$

$$= \begin{pmatrix} a + p\omega_2 X + q\,\omega_2 Y - \omega_3 Y \\ b + \omega_3 X - p\omega_1 X - q\omega_1 Y \\ c - \omega_2 X + \omega_1 Y \end{pmatrix}$$

αφού $Z - r = pX + qY$. Αν, αντί για το (o,o,r), είχε υποτεθεί άλλο σημείο ως κέντρο περιστροφής, τότε η σταθερά r θα υπεισέρχετο στην παραμετρικοποίηση της μετάθεσης και θα έπρεπε να περιληφθεί και στην συνιστώσα της περιστροφής κατά τον άξονα Z.

10.3 Οπτική ροή και ο αλγόριθμος Horn - Schunk

Προσεγγιστικοί υπολογισμοί των πεδίων τοπικής μετατόπισης μπορούν να βασιστούν στην ανάλυση της **οπτικής ροής (optical flow)** που ορίζεται από τις αλλαγές στις ακτινοβολήσεις από την εικόνα E_i στην εικόνα E_{i+1} μιας ακολουθίας εικόνων. Ας υποθέσουμε ότι αυτές οι αλλαγές παράγονται από απόλυτες κινήσεις αντικειμένων. Με αυτή την υπόθεση η οπτική ροή επιτρέπει προσεγγίσεις των πεδίων τοπικής μετατόπισης. Η «κλασσική» μέθοδος Horn–Schunk θα αναπτυχθεί με λεπτομέρεια ως ένα παράδειγμα τέτοιας τεχνικής.

Η στρατηγική της λύσης

Η οπτική ροή μπορεί να αναλυθεί για αυθαίρετες ακολουθίες εικόνων, π.χ. για παραμορφώσιμα αντικείμενα σκηνής όπως σύννεφα σε μετεωρολογικές φωτογραφίες ή για καλλιέργειες βακτηρίων σε φωτογραφίες βιολογικών μικροσκοπίων. Η οπτική ροή έχει, επίσης, ενδιαφέρον για τον υπολογισμό εκτιμήσεων της κίνησης στη περιοχή της **ενεργούς όρασης (active vision).**

Η οπτική ροή δεν μπορεί να ταυτιστεί με τοπικές μετατοπίσεις. Για παράδειγμα, θεωρούμε μια σφαίρα χωρίς επιφανειακή υφή που περιστρέφεται μπροστά στην κάμερα. Καθόλου οπτική ροή δεν μπορεί να παρατηρηθεί παρ' όλο που λαμβάνει χώρα κίνηση σημείων των αντικειμένων και, επομένως, ορίζεται το πεδίο των τοπικών μετατοπίσεων. Παρόμοιες δυσκολίες ανακύπτουν γενικά για κινήσεις επιφανειών χωρίς επιφανειακή υφή. Από την άλλη πλευρά, για μια στατική σφαίρα και μεταβλητές συνθήκες φωτισμού μπορεί να παρατηρηθεί οπτική ροή για την επιφάνεια της σφαίρας παρ' όλο που δεν λαμβάνει χώρα καθόλου κίνηση επιφανειακών σημείων. Η οπτική ροή μπορεί επίσης να εξαρτάται από αστάθειες του αισθητηρίου της κάμερας, αλλαγές φωτισμού ή διαφορετική εμφάνιση της επιφάνειας από διαφορετικές κατευθύνσεις επόπτευσης. Επομένως, ξεκαθαρίζεται από την αρχή ότι η οπτική ροή επιτρέπει μόνο να προσεγγιστούν πεδία τοπικών μετατοπίσεων.

Έστω ότι η ακτινοβόληση της εικόνας E_i στο σημείο $\mathbf{p}=(x,y)$ είναι $E(x,y,t_i)$, $i=0,1,2,\ldots$, όπου $t_i= i \cdot \delta t_{const}$ και δt_{const} είναι το σταθερό χρονικό διάστημα ανάμεσα στη σύλληψη δύο διαδοχικών εικόνων της ακολουθίας. Ας υποθέσουμε ότι η οπτική ροή ορίζεται μοναδικά (στην πραγματικότητα διαφορετικά μοντέλα οπτικής ροής οδηγούν σε διαφορετικά διανυσματικά πεδία) και έστω

$$\mathbf{u}_i\,(x,y)= (u_i(x,y),\, v_i(x,y))$$

η οπτική ροή από την εικόνα E_i στην εικόνα E_{i+1} που χαρακτηρίζει την αλλαγή της ακτινοβόλησης από την εικόνα E_i στην εικόνα E_{i+1}. Στο μοντέλο οπτικής ροής υποθέτουμε ότι ικανοποιείται η εξίσωση

$$E(x+ u_i(x,y),\, y+v_i(x,y),\, t_{i+1}) = E(x,y,\, t_i), \qquad t_{i+1}= t_i + \delta t_{const}.$$

Αυτή η υποτιθέμενη πιστότητα των τιμών εικόνας της οπτικής ροής δεν μπορεί από μόνη της να υποστηρίξει το σχεδιασμό αλγορίθμων για τον υπολογισμό των διανυσμάτων (u,v) αφού οι μετρούμενες τιμές ακτινοβόλησης επηρεάζονται πάντα από θόρυβο. Πρέπει να διαμορφωθούν επιπλέον υποθέσεις ή περιορισμοί για να υποστηριχτεί μια στρατηγική επίλυσης που θα επιτρέπει τον υπολογισμό μιας μοναδικής λύσης.

Γενικά, περιμένουμε ότι η υπολογιζόμενη οπτική ροή έχει μεγάλο βαθμό συσχέτισης με την τοπική μετατόπιση. Ιδανική συσχέτιση υπάρχει όταν η οπτική ροή από εικόνα σε εικόνα της ακολουθίας ταυτίζεται με τις τοπικές μετατοπίσεις

$$d_{xy}\,(t_i)= \mathbf{p}(t_i)= (\xi_{xy}(t_i),\ \psi_{xy}(t_i)\ )$$

των σημείων προβολής $\mathbf{p}=(x,y)$ στην εικόνα, δηλαδή αν οι εξισώσεις

$$\xi_{xy}(t_i) = u_i(x,y) \quad \text{και} \quad \psi_{xy}(t_i) = v_i(x,y)$$

ισχύουν για τη χρονική στιγμή t_i στην ακολουθία εικόνων. Μια τέτοια υποτιθέμενη ταυτότητα, που ονομάζεται **πιστότητα κίνησης (motion fidelity)** της οπτικής ροής, μπορεί να χρησιμοποιηθεί για τον ορισμό ενός μέτρου λάθους για αξιολόγηση της υπολογιζόμενης οπτικής ροής όπως επίσης και ως μια επιπρόσθετη υπόθεση για την εύρεση μιας στρατηγικής επίλυσης της οπτικής ροής.

Η μέθοδος Horn-Schunck

Όλα τα σημεία εικόνας θα έχουν την ίδια αντιμετώπιση, ανεξάρτητα από την (άγνωστη) απόσταση των προβαλλομένων επιφανειακών σημείων αντικειμένων από το επίπεδο εικόνας. Αυτή η υπόθεση αντιστοιχεί στο προβολικό μοντέλο της παράλληλης προβολής.

Υποθέτουμε ότι η συνάρτηση εικόνας $E(x,y,t_i)$ μπορεί να αναπαρασταθεί τοπικά από ένα ανάπτυγμα Taylor

$$E(x+\delta x,\ y+\ \delta x,\ t_i+\delta t)=$$

$$=E(x,y,t_i)+\ \delta x\frac{\partial E}{\partial x}(x,y,t_i)+\delta y\frac{\partial E}{\partial y}(x,y,t_i)+\delta t\frac{\partial E}{\partial t}(x,y,t_i)+e$$

για ένα μικρό βήμα ($\delta x,\ \delta y,\ \delta t$). Θεωρούμε το ανάπτυγμα Taylor της $E(x,y,t_i)$ για ένα συγκεκριμένο βήμα ($u_i(x,y)$, $v_i(x,y)$, δt_{const}). Από μια υποτιθέμενη πιστότητα τιμών εικόνας της οπτικής ροής, προκύπτει ότι

$$o= u_i(x,y)\frac{\partial E}{\partial x}(x,y,t_i)+v_i(x,y)\frac{\partial E}{\partial y}(x,y,t_i)+\frac{\partial E}{\partial t}(x,y,t_i)+e$$

ή, σε πιο απλή μορφή,

$$o= u_i(x,y)\ \frac{\partial E}{\partial x}(x,y,t_i)+v_i(x,y)\frac{\partial E}{\partial y}(x,y,t_i)+\frac{\partial E}{\partial t}(x,y,t_i)$$

για $e = 0$ και $\delta t_{const}= 1$. Αυτή η εξίσωση ονομάζεται **περιορισμός των Horn-Schunck** και η κανονικοποίηση $\delta t_{const}= 1$ δεν θυσιάζει τη γενικότητα.

Ο περιορισμός των Horn-Schunck είναι απόρροια της πιστότητας τιμών εικόνας της οπτικής ροής και της υπόθεσης ότι αυτά τα διανύσματα οπτικής ροής περιγράφουν μόνο μικρά βήματα για τα οποία ικανοποιείται η υπόθεση γραμμικότητας για την $E(x,y,t_i)$, δηλαδή για τα οποία μπορεί να υποτεθεί $e = 0$.

Εναλλακτικά, ο περιορισμός Horn-Schunck μπορεί επίσης να εξαχθεί υποθέτοντας τον πιο κατάλληλο περιορισμό πιστότητας της κίνησης της οπτικής ροής. Για το ανάπτυγμα Taylor της συνάρτησης εικόνας, υποθέτουμε πρώτα ότι ο όρος λάθους e είναι περίπου της τάξεως της διαφοράς

$$E(x+\delta x,\ y+\delta y,\ t_i+\delta t) - E(x,y,t_i).$$

Τότε, οδηγούμαστε στην απλουστευμένη εξίσωση

$$0 = \delta\chi \frac{\partial E}{\partial x}(x,y,t_i) + \delta y \frac{\partial E}{\partial y}(x,y,t_i) + \delta t \frac{\partial E}{\partial t}(x,y,t_i)$$

Διαιρώντας με δt, έχουμε

$$0 = \frac{\delta x}{\delta t}\frac{\partial E}{\partial X}(x,y,t_i) + \frac{\delta y}{\delta t}\frac{\partial E}{\partial y}(x,y,t_i) + \frac{\partial E}{\partial t}(x,y,t_i)$$

και, καθώς $\delta t \to 0$, οδηγούμαστε στη διαφορική εξίσωση

$$0 = \dot{x}(i)\frac{\partial E}{\partial x}(x,y,t_i) + \dot{y}(i)\frac{\partial E}{\partial y}(x,y,t_i) + \frac{\partial E}{\partial t}(x,y,t_i).$$

Στη συνέχεια, θα θέσουμε $\delta t_{const} = 1$ για απλοποίηση. Επομένως, οι τιμές

$$T_i = i, \text{ για } I = 0, 1, 2, \ldots.$$

προσδιορίζουν τις διακριτές χρονικές στιγμές στις οποίες λαμβάνονται οι εικόνες της ακολουθίας εικόνων.

Για κάποια συγκεκριμένη χρονική στιγμή στην ακολουθία εικόνων, ο περιορισμός Horn-Schunk έχει τη μορφή

$$u \cdot E_x + v \cdot E_y = - E_{ti}$$

όπου τα u και v εξαρτώνται από τα (x,y) και η παράγωγοι E_x, E_y και E_t της συνάρτησης ακτινοβόλησης εικόνας εξαρτώνται από τα (x,y,t). Οι παράγωγοι της συνάρτησης ακτινοβόλησης εικόνας δίνονται από προσεγγιστικές τιμές, δηλαδή τιμές τοπικών τελεστών εικόνας που υπολογίζουν πρώτες παραγώγους. Επομένως, για ένα συγκεκριμένο σημείο εικόνας (x,y), οι τιμές των u και v περιορίζονται από αυτή τη γραμμική εξίσωση αν $E_x \neq 0$ ή $E_y \neq 0$ σ' αυτό το σημείο. Μέχρι εδώ, οποιοδήποτε σημείο σ' αυτήν την ευθεία γραμμή στο χώρο *uv* **(χώρο ταχύτητας (velocity space))** μπορεί να επιλεχθεί ως λύση (u,v), όπως, για παράδειγμα, το σημείο τομής με τη γραμμή $E_y \cdot u - E_x \cdot v = 0$. Όμως, μια καλύτερη λύση είναι να επιλεγεί το επιθυμητό σημείο σ' αυτήν την ευθεία με βάση μια άλλη (λογική) τοπική ή ολική υπόθεση για την οπτική ροή. Η απροσδιοριστία της λύσης (u,v) στη γραμμή $u \cdot E_x + v \cdot E_y = -E_t$ αντιστοιχεί στο πρόβλημα του ανοίγματος της κάμερας που αναφέρθηκε προηγουμένως.

Ξεκινώντας με την αντίληψη ότι γειτονικά επιφανειακά σημεία ενός δυναμικού στερεού αντικειμένου έχουν κατά προσέγγιση τα ίδια διανύσματα τοπικής μετατόπισης, η **ομαλότητα (smoothness) του πεδίου οπτικής ροής** μπορεί να τεθεί ως μια ολική υπόθεση. Αυτή η ομαλότητα επιτυγχάνεται αν οι πρώτες παράγωγοι των πεδίων των διανυσμάτων οπτικής ροής έχουν τιμές «κοντά στο μηδέν», δηλαδή αν η έκφραση

$$u_x^2(x,y) + u_y^2(x,y) + v_x^2(x,y) + v_y^2(x,y)$$

παίρνει ελάχιστη τιμή για τα σημεία εικόνας $\mathbf{p} = (x,y)$ για μια δεδομένη ακολουθία εικόνων σε κάποια συγκεκριμένη χρονική στιγμή t. Έστω Ω το σύνολο όλων των σημείων εικόνων για τα οποία πρέπει να υπολογιστεί η οπτική ροή. Η ολική βελτιστοποίηση σημαίνει ότι συναρτησιακό

$$F_s(u,v) = \iint_\Omega (u_x^2(x,y) + u_y^2(x,y) + v_x^2(x,y) + v_y^2(x,y))\,dxdy$$

πρέπει να ελαχιστοποιηθεί. Αφού για όλα τα σημεία $\mathbf{p} = (x,y)$ της δεδομένης ακολουθίας εικόνων τη στιγμή t πρέπει να ικανοποιηθεί ο περιορισμός Horn-Schunck, η λύση (u,v) για τη στιγμή t πρέπει επίσης να ελαχιστοποιηθεί το συναρτησιακό

$$F_h(u,v) = \iint_\Omega (u(x,y)E_x(x,y,t) + (v(x,y)E_y(x,y,t) + E_t(x,y,t))^2\,dxdy$$

Αυτό το συναρτησιακό υποδηλώνει το λάθος σχετικά με την εγκυρότητα του περιορισμού Horn-Schunck για ένα ζευγάρι (u,v). στην επιρροή των δύο συναρτησιακών F_s και F_h μπορεί να δοθεί διαφορετικό βάρος με τη χρήση μιας παραμέτρου $\lambda \geq 0$. Συνολικά, μια λύση μπορεί να βασιστεί στην ελαχιστοποίηση ενός συνδυασμένου συναρτησιακού λάθους

$$F_c(u,v) = F_s(u,v) + \lambda F_h(u,v).$$

Για την ελαχιστοποίηση του συναρτησιακού F_c χρησιμοποιείται ο επαναληπτικός αλγόριθμος Horn-Schunck που περιγράφεται στη συνέχεια.

Απαίτηση : Υποθέτοντας μια δεδομένη ακολουθία ψηφιακών εικόνων, όπως για παράδειγμα δύο διαδοχικές εικόνες $E(i,j,t)$ και $E(i,j,t+1)$ διαστάσεων MxN, να υπολογιστεί η οπτική ροή χρησιμοποιώντας τη μέθοδο Horn-Schunck. Ο παράγοντας βάρους λ και ο αριθμός των επαναλήψεων είναι ελεύθερες παράμετροι της τεχνικής, ενώ θα χρησιμοποιηθούν απλοί τυπικοί προσεγγιστικοί τύποι για τον υπολογισμό των παραγώγων E_x και E_y.

Λύση : Αρχή προσδιορίζονται αρχικές τιμές, π.χ.

$$u_{ij}^{o} = 0 \text{ και } v_{ij}^{o} = 0,$$

ή συγκεκριμένες τιμές στην ευθεία γραμμή που αντιστοιχεί στον περιορισμό Horn-Schunck. Ο υπολογισμός των νέων τιμών

$$u_{ij}^{n+1} \text{ και } v_{ij}^{n+1}$$

του επαναληπτικού βήματος n+1 βασίζεται στους αριθμητικούς μέσους επί των τεσσάρων γειτονικών εικονοστοιχείων που υπολογίστηκαν στο προηγούμενο επαναληπτικό βήμα n. Επομένως, τα βήματα επαναληπτικής λύσης δίνονται συνολικά από τους επαναληπτικούς τύπους:

$$u_{ij}^{n+1} = \bar{u}_{ij}^{n} - \frac{E_x(i,j,t)\bar{u}_{ij}^{n} + E_y(i,j,t)\bar{\bar{r}}_{ij}^{n} + E_t(i,j,t)}{1+\lambda(E_x^2(i,j,t)+E_y^2 9i,j,t)}\lambda E_x(i,j,t)$$

$$v_{ij}^{n+1} = \bar{v}_{ij}^{n} - \frac{E_x(i,j,t)\bar{u}_{ij}^{n} + E_y(i,j,t)\bar{\bar{r}}_{ij}^{n} + E_t(i,j,t)}{1+\lambda(E_x^2(i,j,t)+E_y^2 9i,j,t)}\lambda E_x(i,j,t)$$

όπου n = 0, 1, 2, … υποδηλώνει επαναληπτικό βήμα. Σε κάθε επαναληπτικό βήμα, οι τιμές της οπτικής ροής υπολογίζονται για όλα τα σημεία εικόνας (i,j).

Για αυτούς τους τύπους, πρέπει να προσδιορίσουμε διαδικασίες για την προσέγγιση των παραγώγων $E_x(i,j,t)$, $E_y(i,j,t)$ και $E_t(i,j,t)$ των δεδομένων συναρτήσεων ακτινοβόλησης εικόνας. Ο προφανής τρόπος είναι η χρήση τελεστών τοπικών διαφορών. Για παράδειγμα:

$$E_x(i, j, t)= \frac{1}{4}\,(E(i+1, j, t) + E(i+1, j, t+1) + E(i+1, j+1, t) +$$
$$+ E(i+1, j+1, t+1) - \frac{1}{4}\,(E(i, j, t) + E(i, j, t+1) +$$
$$+ E(i, j+1, t) + E(i, j+1, t+1)$$

με ανάλογες εκφράσεις για $E_y(i, j, t)$ και $E_t(i, j, t)$ είναι επιτρεπτές προσεγγιστικές συναρτήσεις.

Αλγόριθμος : Δίνεται ο ψευδοκώδικας στην επόμενη σελίδα.

Σχόλιο : Συνολικά, περιγράφτηκε μια τεχνική για τον υπολογισμό της οπτικής ροής σε μια ακολουθία εικόνων. Διαφορετικές επιλογές μπορούν να γίνουν για τον προσδιορισμό των αρχικών τιμών, του παράγοντα βάρους λ και των προσεγγίσεων των συναρτήσεων E_x, E_y και E_t. Το βάθος επανάληψης μπορεί να ελεγχθεί από ένα μέτρο λάθους, π.χ., την ολική απόκλιση μεταξύ των λύσεων των επαναληπτικών βημάτων n και n+1. Εναλλακτικά, μπορεί να οριστεί ένα επίπεδο τερματισμού.

12. Σύντηξη αισθητηρίων

12.1.Γενικά

Ένας μεγάλος αριθμός σημαντικών εφαρμογών εξαρτάται από την ικανότητα υπολογιστών να διατηρούν **διεπαφή (interface)** με τον πραγματικό κόσμο. Αυτές οι εφαρμογές περιλαμβάνουν στρατιωτικά, ιατρικά και κατασκευαστικά συστήματα, όπως και συστήματα μεταφορών, ασφάλειας και περιβαλλοντικού σχεδιασμού. Πολλές από αυτές τις εφαρμογές υλοποιήθηκαν με δυσκολία εξαιτίας προβλημάτων στην είσοδο δεδομένων από αισθητήρια κατευθείαν σε αυτοματοποιημένα συστήματα.

Η **σύντηξη αισθητηρίων (sensor fusion)** εμφανίστηκε ως η μέθοδος που επιλέχθηκε για να επιλύσει αυτά τα προβλήματα. Η **ευφυής επεξεργασία δεδομένων από αισθητήρια (intelligent sensor processing)** αποτελεί μία ταχέως αναπτυσσόμενη περιοχή της επιστημονικής έρευνας και τεχνολογίας και βοηθάει τους ερευνητές να αναπτύξουν συστήματα που «βλέπουν» και «αντιλαμβάνονται» το περιβάλλον τους χρησιμοποιώντας νέες υπολογιστικές μεθόδους και εργαλεία. Παραδείγματα των μεθόδων που χρησιμοποιούνται για την επεξεργασία δεδομένων από αισθητήρια περιλαμβάνουν τη συγχώνευση δεδομένων πολλαπλών διαστάσεων, την ανάπτυξη δυναμικών μεθόδων για την επεξεργασία των δεδομένων και την ανάπτυξη μοντέλων οπτικής αντίληψης που απλοποιούν τη λήψη αποφάσεων μετά από ερμηνεία των δεδομένων. Εφαρμογές της επεξεργασίας δεδομένων αισθητηρίων στην επεξεργασία εικόνας και σήματος ασχολούνται με αναπαραστάσεις δυναμικών δεδομένων αισθητηρίων με πολλαπλές διαστάσεις και αλληλεπιδράσεις μεταξύ μεμονωμένων αισθητηρίων. Τα αισθητήρια μπορεί να μετρούν βαθμωτές τιμές (π.χ. πίεση) ή διανυσματικές τιμές (π.χ. τη θέση στον τριδιάστατο χώρο). Τις περισσότερες φορές, οι μετρήσεις αυτές είναι συναρτήσεις και του χρόνου και του χώρου.

Η σύντηξη πολλαπλών αισθητηρίων έχει αποτελέσει το στόχο μεγάλου όγκου έρευνας. Άρθρα με νέες μεθόδους και θεωρίες δημοσιεύονται τακτικά σε περιοδικά που ειδικεύονται σε ένα μεγάλο φάσμα θεμάτων, όπως αεροναυπηγική, ρομποτική, στατιστική και τεχνητή νοημοσύνη. Σε πρώτη ματιά, το θέμα μπορεί να εμφανιστεί ως παρωχημένο. Όμως, τα αισθητήρια της σημερινής τεχνολογίας δεν αποτελούν αξιόπιστες διεπαφές για αρκετούς λόγους με συνέπεια εφαρμογές όπως για παράδειγμα σε συστήματα αυτόματης πλοήγησης να υλοποιούνται με δυσκολία.

12.2 Ιεραρχία επεξεργασιών δεδομένων αισθητηρίων

Ένα **αισθητήριο (sensor)** είναι μια ηλεκτρική ή μηχανική συσκευή που απεικονίζει την τιμή κάποιας περιβαλλοντικής ιδιότητας σε μια ποσοτική μέτρηση. Κάθε αισθητήριο πρέπει να ανιχνεύει κάποια άποψη της κατάστασης του περιβάλλοντός του. Αυτή η ανίχνευση περιλαμβάνει άμεση αλληλεπίδραση με το περιβάλλον. Η αλληλεπίδραση μπορεί να περιλαμβάνει ανίχνευση ηλεκτρομαγνητικής ακτινοβολίας που προέρχεται από το περιβάλλον, καταγραφή ανακλάσεων ακτινοβολίας που έχει εκπέμψει το αισθητήριο (π.χ. σόναρ ή ραντάρ) ή άμεση φυσική επαφή με εμπόδια στο περιβάλλον. Όταν το αισθητήριο καταγράφει πληροφορία που είναι ήδη παρούσα γύρω του, τότε το αισθητήριο είναι **παθητικό (passive)**. Αν το αισθητήριο καταγράφει πληροφορία που είναι αποτέλεσμα κάποιας δράσης που ανέλαβε το ίδιο, τότε το αισθητήριο είναι **ενεργό (active).** Γενικά, οι αλληλεπιδράσεις προκαλούν φθορά στις συσκευές αίσθησης που οδηγεί τελικά σε βλάβη εξαρτημάτων του. Αυτή η βλάβη μπορεί να είναι μηχανικής, χημικής ή ηλεκτρικής φύσεως. Παράλληλα, οποιαδήποτε μέθοδος ανίχνευσης έχει περιορισμένη ακρίβεια.

Μετά την **ανίχνευση (detection)** απαιτείται **προεπεξεργασία (preprocessing).** Καθήκοντα που περιλαμβάνονται στην προεπεξεργασία γενικά περιλαμβάνουν μείωση του θορύβου, επαναβαθμονόμηση του αισθητηρίου, αφαίρεση της μεροληψίας σε μετρήσεις και δραστηριότητες που αναφέρονται στο συγκεκριμένο τύπο αισθητηρίου. Επίσης, μπορούν να περιλαμβάνουν ανίχνευση ακμών και χαρακτηριστικών ή άλλες πιο ευφυείς δραστηριότητες. Αυτές οι δραστηριότητες κατατάσσονται και πάλι στο στάδιο της προεπεξεργασίας, σύμφωνα με τον ορισμό που χρησιμοποιούμε εδώ, εφόσον είναι δραστηρίοτητες που ταιριάζουν σε μετρήσεις ενός μεμονωμένου αισθητηρίου.

Όταν μια μέτρηση έχει διέλθει από το στάδιο της προεπεξεργασίας, έχει τεθεί σε μια μορφή υπολογιστικά κατάλληλη και μετρήσεις από διάφορα ανεξάρτητα αισθητήρια μπορούν να συντηχθούν σε μια μόνο μέτρηση. Αυτή η συντηγμένη μέτρηση θα έχει την ίδια μορφή με τα προεπεξεργασμένα δεδομένα. Οι τιμές που βρίσκονται με τη διαδικασία της εισόδου και τα συντηγμένα δεδομένα πρέπει να αναπαριστάνουν καθαρά τις αβεβαιότητες που είναι έμφυτες σε μερικώς αντιφατικά δεδομένα. Αυτή η διαδικασία σύντηξης πρέπει να βασίζεται σε μαθηματικά αυστηρές μεθόδους που αποφεύγουν την αθώα διάδοση λαθών στο σύστημα. Για να επιτευχθεί αυτό, πρέπει να σχεδιαστεί ένα δίκτυο αισθητηρίων όσο είναι δυνατό πιο αξιόπιστο.

Υπάρχουν τόσο στατιστικές όσο και νομοτελειακές μέθοδοι σύντηξης. Εδώ χρησιμοποιούμε τον ορισμό της σύντηξης των Luo και Kay που διακρίνει ανάμεσα σε ολοκλήρωση και σύντηξη πολλαπλών αισθητηρίων. Η **σύντηξη (fusion) πολλαπλών αισθητηρίων** αναφέρεται στο συνδυασμό των μετρήσεων διαφόρων αισθητηρίων στο συνδυασμό μετρήσεων διαφόρων αισθητηρίων σε μια ομοιόμορφη δομή δεδομένων. Η **ολοκλήρωση (integration) πολλαπλών αισθητηρίων** αναφέρεται στη χρήση πληροφορίας από πολλά αισθητήρια με σκοπό να επιτευχθεί μια συγκεκριμένη εφαρμογή. Παρ' όλο που οι δύο διαδικασίες συνδέονται στενά, απαιτείται διάκρισή τους.

Όταν έχουν συγχωνευθεί δεδομένα από διάφορα αισθητήρια θα πρέπει να ερμηνευθούν. Η διαδικασία της ερμηνείας εξαρτάται από τον απώτερο σκοπό και συνίσταται στην εύρεση του καλύτερου δυνατού ταιριάσματος των δεδομένων στα πλαίσια των πληροφοριακών απαιτήσεων του συστήματος. Αυτό είναι ένα προκλητικό υπολογιστικό καθήκον και κατά πάσα πιθανότητα δεν υπάρχει ένας βέλτιστος τρόπος αυτοματοποίησης της κατανόησης του περιβάλλοντος για όλες τις εφαρμογές.

Για συστήματα με περισσότερα από ένα αισθητήρια, η διαδικασία της αίσθησης ακολουθεί το παραπάνω πρότυπο: ανίχνευση, προεπεξεργασία, σύντηξη και ερμηνεία δεδομένων. Κάθε βήμα εξαρτάται από τα αποτελέσματα του αμέσως προηγούμενου βήματος. Τα δύο μεσαία βήματα, προεπεξεργασία και σύντηξη, μπορεί να μην εμφανίζονται σε απλοποιημένα συστήματα αίσθησης. Παρ' όλο που υπάρχει μία φυσική ιεραρχία στη διαδικασία της αίσθησης, η ιεραρχία αυτή δεν είναι άκαμπτη. Η διαδικασία πρέπει να ξεκινάει με ανίχνευση και να καταλήγει με ερμηνεία δεδομένων, αλλά είναι δυνατό να περιέχει διάφορα στρώματα προεπεξεργασίας, σύντηξης και ερμηνείας ενδιάμεσα.

12.3 Μοντέλα αισθητηρίων

Προκειμένου να διαμορφώσουμε μεθόδους ικανές να απεικονίζουν αισθητηριακά δεδομένα του πραγματικού κόσμου σε μορφές κατάλληλες για επεξεργασία από υπολογιστή, πρέπει να διαχωρίσουμε τις φυσικές συσκευές από τα καθήκοντα που πρέπει να επιτελέσουν. Ο Marzullo ονομάζει αυτή τη διαδικασία ενασχόληση με **αφηρημένα αισθητήρια (abstract sensors)** σε αντίθεση με **συμπαγή αισθητήρια (concrete sensors).** Άλλοι ερευνητές αναφέρονται σε **λογικά αισθητήρια (logical sensors).** Αυτή η αφαίρεση μας επιτρέπει να βρούμε τα θεωρητικά όρια των αισθητηρίων και να δουλέψουμε χωρίς τον περιορισμό των μη αναγκαίων λεπτομερειών υλοποίησης οποιουδήποτε συγκεκριμένου είδους αισθητηρίου.

Στη συνέχεια, ο όρος **αισθητήριο (sensor)** θα αναφέρεται σε μια αφαίρεση. Τα φυσικά αισθητήρια θεωρούνται ως ένα στιγμιότυπο μιας λειτουργίας αφηρημένου αισθητηρίου. Αφού αποκτήσουμε ένα σαφή ορισμό του τι είναι αισθητήριο και τι μπορούμε να περιμένουμε από αυτό, τότε ο ορισμός αυτός μπορεί να χρησιμοποιηθεί για να αναπτυχθούν συστήματα που είναι αξιόπιστα και αποτελεσματικά.

Σημειώνουμε ότι τα αφηρημένα (λογικά) αισθητήρια δεν αντιστοιχούν κατ' ανάγκη σε ένα μόνο συμπαγές αισθητήριο. Μετρήσεις από πολλά αισθητήρια μπορούν να συντηχθούν για να δημιουργήσουν ένα αφηρημένο αισθητήριο. Τα αισθητήρια που αντιστοιχούν σε ένα μόνο συμπαγές αισθητήριο θα αναφέρονται ως **απλά αισθητήρια (simple sensors)** και εκείνα που αποτελούνται από διάφορα συμπαγή αισθητήρια θα αναφέρονται ως **δίκτυα αισθητηρίων (sensors networks).**

Μαθηματική περιγραφή αισθητηρίων

Ένα αισθητήριο μπορεί να θεωρηθεί ως μια μαθηματική συνάρτηση δύο ορισμάτων, του περιβάλλοντος E και του χρόνου t. Αυτή η συνάρτηση S απεικονίζει το περιβάλλον σε αριθμητικές τιμές που αναπαριστάνονται ως μια μεταβλητή V. Η μεταβλητή V είναι ένα σημείο σε ένα χώρο D διαστάσεων και περιβάλλεται από μια αβεβαιότητα ε. Έτσι έχουμε την εξίσωση

$$S(E,t) = \{V(t), \varepsilon(t)\}$$

Για παράδειγμα ένα θερμόμετρο αντιστοιχεί σε μια συνάρτηση S η οποία, όταν θα αναγνωστεί το θερμόμετρο τη χρονική στιγμή t, επιστρέφει μια μέτρηση θερμοκρασίας V με δυνατότητα σφάλματος ε. Ανάλογα για ένα τριδιάστατο πλέγμα από θερμόμετρα.

Κάνουμε ορισμένες υποθέσεις για το περιβάλλον E. Το περιβάλλον E είναι μια αφηρημένη οντότητα που περιέχει διάφορες φυσικές ιδιότητες που μπορούν αν μετρηθούν. Αυτές οι ιδιότητες μπορεί να έχουν τιμές σε ένα συνεχές σύνολο (θερμοκρασία, ταχύτητα, απόσταση κ.λ.π.) ή σε ένα διακριτό σύνολο (αριθμός αεροσκαφών, αριθμός πελατών σε μια ουρά κ.λ.π.) ή να είναι δίτιμες (παρουσία ή όχι ενός οχήματος, θερμοκρασία ανώτερη κάποιου ορίου κ.λ.π.).

Ο χρόνος t θα μπορούσε να θεωρηθεί ως μια άλλη ιδιότητα του περιβάλλοντος E. Θεωρείται εκπεφρασμένος για να δοθεί μια περισσότερο διαισθητική παρουσίαση της λειτουργίας του αισθητηρίου. Οι μετρούμενες τιμές αλλάζουν με τον χρόνο επειδή και το περιβάλλον του αισθητηρίου αλλάζει με το χρόνο. Επειδή οι περισσότερες διαδικασίες αισθητηρίων έχουν περιορισμούς πραγματικού χρόνου, αξίζει να ασχοληθεί κανείς εκπεφρασμένα με ερωτήματα συγχρονισμού.

Ως σημείο σε ένα D-διάστατο χώρο, η μεταβλητή V μπορεί να αναπαρασταθεί επίσης με ένα πίνακα. Τα στοιχεία του πίνακα μπορεί να είναι δίτιμη, διακριτά ή συνεχή στις τιμές τους ανάλογα με τις φυσικές μεταβλητές που μετρούν.

Αβεβαιότητα και ακρίβεια

Τα αισθητήρια έχουν μεταβλητή ακρίβεια. Σπάνια, ίσως ποτέ, μπορούν αισθητήρια να παράσχουν μια απολύτως ακριβή τιμή για μια φυσική μεταβλητή. Ακόμα και αν το αισθητήριο μπορούσε να κάνει τέτοια μέτρηση, οι φυσικές τιμές που μετριούνται γενικά μεταβάλλονται με το χρόνο. Μια μέτρηση αισθητηρίου που καλύπτει μια πεπερασμένη χρονική περίοδο θα έπρεπε να επιστρέψει μια περιοχή τιμών προκειμένου να αναπαραστήσει με ακρίβεια το περιβάλλον.

Η έννοια της περιορισμένης ακρίβειας είναι ένας παράγοντας σε κάθε επεξεργασία δεδομένων. Όλες οι κοινές αναπαραστάσεις δεδομένων έχουν άνω και κάτω όρια για τις τιμές που μπορούν να αναπαραστήσουν και ένα μέγιστο αριθμό ψηφίων ακρίβειας. Αν δεν αντιμετωπιστούν σωστά τα όρια στην ακρίβεια, μπορεί να χαθεί σημαντική πληροφορία ή τα αποτελέσματα των υπολογισμών να περιέχουν καταστροφικό λάθος.

Γενικά αναπαριστούμε τα ε και V ως ένα ανώτερο και κατώτερο όριο για κάθε διάσταση. Αυτή η αναπαράσταση υπονοείται στο συμβολισμό που χρησιμοποιούμε: ε είναι μια περιοχή με άνω και κάτω όριο που περιέχει το V. Το V δεν είναι αναγκαστικά το μεσαίο σημείο του διαστήματος ε και, αν το αισθητήριο λειτουργεί σωστά, η φυσική τιμή που μετριέται πρέπει να βρίσκεται μέσα στην περιοχή ε γύρω από το V. Η μέτρηση και η ακρίβεια του αισθητηρίου μπορούν να συσχετίζονται, αλλά όχι αναγκαστικά.

Πιθανότητα και αξιοπιστία

Εκτός από την περιορισμένη ακρίβεια κάθε αισθητηρίου, παραμένει η δυνατότητα βλάβης ενός αισθητηρίου κατά τη διάρκεια της ζωής ενός συστήματος ή προσωρινής παροχής ανακριβών μετρήσεων για οποιοδήποτε λόγο. Όσο καλά και να σχεδιαστεί το σύστημα, εξαρτήματά του θα πάψουν τελικά να λειτουργούν σωστά. Οι μηχανικοί πρέπει να προσπαθούν να ελαχιστοποιούν το αποτέλεσμα της βλάβης εξαρτημάτων.

Πέρα από την κανονική φθορά, οι ικανότητες οποιουδήποτε εξαρτήματος είναι περιορισμένες ειδικά όταν το σύστημα τοποθετείται σε ένα εχθρικό και / ή θορυβώδες περιβάλλον. Τέτοια περιβάλλοντα αποτελούν πυρηνικοί αντιδραστήρες, πεδία μαχών, το διάστημα και εργοστάσια, ακριβώς δηλαδή περιβάλλοντα για τα οποία είναι πολύ σημαντικά ρομποτικά συστήματα.

Για τους λόγους αυτούς, κάθε αισθητήριο S χαρακτηρίζεται από μια πιθανότητα P η μέτρηση V για να βρίσκεται στην περιοχή ε της πραγματικής φυσικής τιμής που μετριέται. Η πιθανότητα P εξαρτάται από το αισθητήριο S, το περιβάλλον E και το χρόνο t, δηλαδή είναι μια συνάρτηση P(S,E,t). Γενικά, καθώς αυξάνεται το διάστημα ε επιτρεπτών τιμών του αισθητηρίου, μειώνεται η πιθανότητα P η πραγματική τιμή να εμπίπτει στην επιτρεπτή περιοχή.

Ακόμα και για απλά αισθητήρια, η πιθανότητα αυτή δεν είναι ταυτόσημη με παραδοσιακές μετρήσεις αξιοπιστίας, όπως το μέσο χρόνο μέχρι βλάβη, την αξιοπιστία συστήματος και τη διαθεσιμότητα συστήματος. Βλάβες και λάθη σε ένα θορυβώδες περιβάλλον λειτουργίας προσθέτουν τη δυνατότητα μεταβατικών λαθών που δεν περιλαμβάνονται σε πολλά μοντέλα αξιοπιστίας. Ο περιβαλλοντικός θόρυβος δεν είναι ομοιόμορφα παρών στο χρόνο και έτσι παραβιάζει πολλές από τις υποθέσεις που κάνει η πλειονότητα των μεθόδων ανάλυσης της αξιοπιστίας συστημάτων. Από την άλλη πλευρά, αυτή η πιθανότητα σχετίζεται στενά με παραδοσιακές μετρήσεις της αξιοπιστίας συστημάτων.

Η πιθανότητα P(S,E,t) γενικά θα μειώνεται με το χρόνο για συστήματα που δεν επισκευάζονται ή δεν επαβαθμονομούνται. Όμως και η επισκευή και η επαναβαθμονόμηση είναι ακριβείς διαδικασίες και απαιτούν να απενεργοποιηθεί το σύστημα για ένα χρονικό διάστημα. Ένα πλεονέκτημα της σύντηξης αισθητηρίων είναι ότι μπορεί να χρησιμοποιηθεί για να δημιουργήσει συστήματα αισθητηρίων είναι ότι μπορεί αν χρησιμοποιηθεί για να δημιουργήσει συστήματα αισθητηρίων που αυτοβαθμολογούνται και προσαρμόζουν περιοδικά τις παραμέτρους τους σε τρόπο ώστε να αυξάνεται η ακρίβειά τους.

Κόστος

Κάθε αισθητήριο έχει κάποιο συναφές κόστος C(S) έχει μεγάλη σημασία στο σχεδιασμό δικτύων αισθητηρίων αφού οι σχεδιαστές πρέπει να κατασκευάσουν δίκτυα αισθητηρίων που εμπίπτουν μέσα σε ορισμένα όρια κόστους και ακρίβειας. Για να έχει ένα μοντέλο πρακτική αξία, ο παράγοντας του κόστους δεν μπορεί να αγνοηθεί.

Αντιμετωπίζουμε το κόστος ως μια δεδομένη τιμή για κάθε απλό αισθητήριο και το κόστος ενός δικτύου αισθητηρίων ως το άθροισμα των κοστών των απλών αισθητηρίων και του κόστους ολοκλήρωσής τους. Αυτό το μοντέλο είναι το πιο γενικό μοντέλο κόστους που μπορεί να παρουσιαστεί και δείχνει πως οι περιορισμοί κόστους μπορούν να ολοκληρωθούν σε αποφάσεις σχεδιασμού ενός δικτύου αισθητηρίων.

Γενικά, το κόστος μετριέται σε μονάδες κάποιου νομίσματος, αλλά σε ορισμένες εφαρμογές μπορεί να είναι χρησιμότερες άλλες μονάδες. Για παράδειγμα, πολλές αεροναυπηγικές εφαρμογές θέτουν σοβαρούς περιορισμούς βάρους σε συστήματα.

12.4 Χρήση πολλαπλών αισθητηρίων

Ένα δίκτυο αισθητηρίων δημιουργείται από απλά αισθητήρια. Μετά από σύντηξη, τα δεδομένα ενός δικτύου αισθητηρίων μπορούν να αντιμετωπιστούν σαν να είχαν προέλθει από ένα μόνο αισθητήριο. Αυτά τα δίκτυα μπορεί να έχουν διάφορα επίπεδα πολυπλοκότητας και έχουν προταθεί αρκετές διαφορετικές αρχιτεκτονικές. Διαφορετικές αρχιτεκτονικές δικτύου έχουν σημαντικά διαφορετικές απαιτήσεις επικοινωνίας δεδομένων και υπολογιστικής πολυπλοκότητας.

Υπάρχουν αρκετά διαφορετικά κίνητρα για την κατασκευή δικτύων αισθητηρίων. Γενικά, διάφορα ανεξάρτητα αισθητήρια συνδυάζονται σε ένα δίκτυο με σκοπό να παράσχουν ικανότητες που ένα απλό αισθητήριο δεν μπορεί από μόνο του να παράσχει. Η αλληλεπίδραση μεταξύ των αισθητηρίων καθορίζει πώς ακριβώς θα λάβει χώρα σύντηξη των μετρήσεών τους.

Τα δίκτυα πολλαπλών αισθητηρίων έχουν διαιρεθεί σε **συμπληρωματικά (complementary), ανταγωνιστικά (competitive)** και **συνεργαζόμενα (cooperative).** Προκειμένου η σύντηξη να είναι προσαρμόσιμη σε όλες τις δυνατές διατάξεις, προσθέτουμε μια νέα κατηγορία στις παραπάνω τρεις που θα ονομάζουμε **ανεξάρτητα (independent)** δίκτυα.

Αυτός ο διαχωρισμός ορίζεται από τη λειτουργικότητα του δικτύου αισθητηρίων και διατάξεις αισθητηρίων μπορεί να ταιριάζουν ταυτόχρονα σε πολλές από αυτές τις κατηγορίες.

Συμπληρωματικά αισθητήρια
Τα αισθητήρια είναι συμπληρωματικά όταν δεν εξαρτώνται μεταξύ τους άμεσα, αλλά μπορούν να συνδυαστούν για να δώσουν μια πιο πλήρη εικόνα των φαινομένων που μελετούμε. Ένα παράδειγμα συμπληρωματικής διάταξης αποτελούν οι μετρητές αποστάσεων και ταχύτητας σε ένα αυτοκίνητο. Μετρούν διαφορετικά φαινόμενα, αλλά χρησιμοποιούνται από κοινού για να πλοηγηθεί το όχημα.

Τα συμπληρωματικά δίκτυα μπορεί να παρέχουν μια αναπαράσταση των δεδομένων σε μια ευρεία περιοχή ή να παρέχουν διάφορες απόψεις του ίδιου φαινομένου. Γενικά, είναι εύκολη η σύντηξη συμπληρωματικών δεδομένων. Τα δεδομένα από ανεξάρτητα αισθητήρια μπορεί να προσαρτηθούν μεταξύ τους παρέχοντας μια πιο πλήρη απεικόνιση των φυσικών ιδιοτήτων που μελετώνται.

Ανταγωνιστικά αισθητήρια
Τα ανταγωνιστικά αισθητήρια παρέχουν ανεξάρτητες μετρήσεις της ίδιας πληροφορίας για κάποιο φυσικό φαινόμενο. Αφού παρέχουν δεδομένα που θα έπρεπε να είναι ταυτόσημα, τα αισθητήρια βρίσκονται σε ανταγωνισμό για το

ποια μέτρηση θα γίνει πιστευτή από το σύστημα όταν αυτές δεν συμφωνούν. Τα ανταγωνιστικά αισθητήρια μπορεί να είναι ταυτόσημα ή να χρησιμοποιούν διαφορετικές μεθόδους μέτρησης της φυσικής ιδιότητας. Αισθητήρια τοποθετούνται σε ανταγωνιστική διάταξη για να παρέχουν μεγαλύτερη αξιοπιστία ή ανθεκτικότητα σε βλάβες σε ένα σύστημα. Παραδείγματα αυτού του τύπου διάταξης αισθητηρίων αποτελούν σχήματα πλεονασμού N μονάδων τα οποία ανέχονται έναν αριθμό εξαρτημάτων με βλάβη.

Συνεργαζόμενα αισθητήρια

Τα δίκτυα συνεργαζομένων αισθητηρίων συνδυάζουν δεδομένα από ανεξάρτητα αισθητήρια για να εξάγουν πληροφορία που δεν θα ήταν διαθέσιμη από μεμονωμένα αισθητήρια. Αυτός ο τύπος αλληλεπίδρασης είναι ιδιαίτερα δύσκολος στο σχεδιασμό αφού η πληροφορία που αναζητείται είναι ευαίσθητη σε ανακρίβειες σε όλα τα εξαρτήματα των χρησιμοποιούμενων απλών αισθητηρίων. Η συνεργαζόμενη σύντηξη προκαλεί, γενικά, μείωση με τα συνεργαζόμενα αισθητήρια όπου η ακρίβεια και η αξιοπιστία γενικά αυξάνονται.

Ένα καλό παράδειγμα συνεργαζόμενων αισθητηρίων αποτελεί η χρήση δύο καμερών για την κατασκευή μιας τριδιάστατης άποψης μιας περιοχής. Καμία κάμερα δεν μπορεί να παράσχει πληροφορία βάθους μόνη της, όμως αν αντιστοιχιστούν τα δεδομένα από τις δύο κάμερες μπορεί να συμπεραστεί η τρίτη διάσταση από διδιάστατες εικόνες.

Γενικά, η συνεργασία μεταξύ αισθητηρίων εξαρτάται από τη συγκεκριμένη εφαρμογή και δεν εμπίπτει εύκολα στον ορισμό που υιοθετήσαμε για τη σύντηξη αισθητηρίων. Η συνεργασία μεταξύ αισθητηρίων απαιτεί μια αντιμετώπιση χαμηλού επιπέδου που θα ενέπιπτε μάλλον στη φάση της προεπεξεργασίας.

Ανεξάρτητα αισθητήρια

Δίκτυα ανεξάρτητων ασθητηρίων είναι εκείνα που δεν εμπίπτουν σε καμία από τις τρεις προηγούμενες κατηγορίες. Για παράδειγμα, μια βάση ραντάρ θα μπορούσε να καταχωρεί και θερμικά δεδομένα. Η αναζήτηση μιας σχέσης ανάμεσα στην πληροφορία από την παρακολούθηση αεροσκαφών με τα ραντάρ και τις μετρούμενες θερμοκρασίες δεν θα ήταν ρεαλιστική. Αφού αυτά τα αισθητήρια δεν επιστρέφουν δεδομένα που επηρεάζουν τα δεδομένα από άλλα αισθητήρια, η σύντηξη των δεδομένων τους δεν θα γίνει με την αυστηρότερη έννοια του όρου.

Από την άλλη πλευρά, οι σχεδιαστές συστημάτων πρέπει να υλοποιήσουν μια δομή δεδομένων που αποθηκεύει αποτελεσματικά όλα τα δεδομένα που συλλέγει το σύστημα και δημιουργεί κάποιου είδους χάρτη του περιβάλλοντος του συστήματος. Επομένως, είναι δυνατό να εξοικονομήσουμε πολύ αποθηκευτικό χώρο αποθηκεύοντας δεδομένα από δύο ή περισσότερα αισθητήρια σε ένα χάρτη παρά δημιουργώντας ένα ξεχωριστό χάρτη για κάθε αισθητήριο. Επομένως, δεδομένα αισθητηρίων χωρίς προφανή συσχέτιση μπορούν να αποθηκευτούν μαζί σε μια δομή δεδομένων.

12.5.Κατασκευή αξιόπιστων αφηρημένων αισθητηρίων από απλά αφηρημένα αισθητήρια.

Ένα από τα επιτεύγματα της έρευνας σε σύντηξη αισθητηρίων υπήρξε αυτό που ο Marzullo ονομάζει «αξιόπιστα αφηρημένα αισθητήρια». Παρόλο που ένα απλό αισθητήριο θα αποτύχει αν ένα εξάρτημά του πάθει βλάβη, ένα δίκτυο αισθητηρίων μπορεί να κατασκευαστεί από μερικά ανεξάρτητα απλά αισθητήρια που παρέχουν πλεονασματικά δεδομένα. Αυτό το δίκτυο είναι τότε ικανό να παίρνει τα δεδομένα από κάθε μεμονωμένο αισθητήριό του και να βρίσκει την επαληθεύσιμα σωστή μέτρηση εφόσον έχουν βλάβη λιγότερα αισθητήρια από κάποιο αριθμό.

Η ικανότητα να γίνονται λάθη ανεκτά μπορεί αν προσεγγιστεί διαισθητικά. Έχοντας ένα ρολόι, μπορούμε να είμαστε αρκετά σίγουροι για το τι ώρα είναι. Έχοντας δύο ρολόγια από τα οποία το ένα είναι ελαττωματικό, δεν μπορούμε να γνωρίζουμε την ώρα. Με τρία ρολόγια όμως, μπορούμε να γνωρίζουμε την ώρα προσεγγιστικά ακόμα και αν το ένα ρολόι είναι χαλασμένο. Η αύξηση του αριθμού των ρολογιών αυξάνει και τον αριθμό των ρολογιών που επιτρέπεται να χαλάσουν χωρίς να μειωθεί η ικανότητά μας να γνωρίζουμε την ώρα.

Η ίδια συλλογιστική μπορεί να εφαρμοστεί σε προβλήματα αξιοπιστίας μηχανών. Είναι γνωστή ως **πλεονασμός Ν μονάδων (N – modular redundancy)** και έχει γενικευτεί σε μια γενική λύση που βασίζεται στη θεωρία πληροφοριών. Ο πλεονασμός Ν μονάδων είναι επίσης γνωστός και ως το **πρόβλημα των βυζαντινών στρατηγών.**

Μια γεωμετρική εξαγωγή τεχνικών αναγνώρισης βλαβών στο ποσό που αφορούν τη σύντηξη αισθητηρίων έχει γίνει από το Marzullo. Ο Marzullo ορίζει τα αισθητήρια ως κατά τμήματα συνεχείς συναρτήσεις που χαρακτηρίζονται από το σχήμα και το διάστημα τιμών των δεδομένων που επιστρέφουν. Το **σχήμα** αναφέρεται στη μορφή της αβεβαιότητας γύρω από την τιμή επιστρέφει το αισθητήριο και μπορεί να είναι γραμμικό, κυκλικό κ.λ.π. ανάλογα με τον αριθμό των διαστάσεων που επιστρέφει το αισθητήριο και άλλους παράγοντες. Το **διάστημα τιμών** της ακρίβειας που περιβάλλει τη μέτρηση πρέπει να είναι πεπερασμένο, αφού πληροφορία που περιέχεται σε ένα άπειρο διάστημα τιμών ουσιαστικά δεν αποτελεί πληροφορία.

Απλά αισθητήρια μπορούν να συνδυαστούν και να σχηματίζουν μια συνάρτηση με το ίδιο σχήμα με κάθε μεμονωμένο αισθητήριο του οποίου το διάστημα τιμών δεν είναι μεγαλύτερο του διαστήματος τιμών οποιουδήποτε από τα απλά αισθητήρια. Ο Marzullo ονομάζει αυτή τη νέα συνάρτηση ένα **αξιόπιστο αφηρημένο αισθητήριο.** Στην μονοδιάστατη περίπτωση, η σωστή τιμή της φυσικής μεταβλητής που μετριέται θα είναι πάντα μέσα στα όρια ακρίβειας του αξιόπιστου αφηρημένου αισθητηρίου εφόσον περισσότερα από τα μισά αισθητήρια είναι σωστά. Σε πολυδιάστατα δεδομένα, τα αποτελέσματα του Marzullo ισχύουν εφόσον λιγότερα από $\frac{N}{2D}$ αισθητήρια έχουν βλάβη, όπου Ν είναι ο συνολικός αριθμός αισθητηρίων και D είναι ο αριθμός των διαστάσεων των δεδομένων.

Η κατασκευή αξιόπιστων αφηρημένων αισθητηρίων από μεμονωμένα αισθητήρια είναι ισοδύναμη με την κατασκευή μιας συμφωνίας μεταξύ διαφόρων ανεξαρτήτων μηχανών. Το γενικό πρόβλημα έχει απασχολήσει την έρευνα σε συστήματα κατανεμημένου υπολογισμού και στη γενικότερη μορφή του δηλώνεται ως το «πρόβλημα των βυζαντινών στρατηγών». Το πρόβλημα παρουσιάζεται ως μια μεταφορά που μπορεί να εφαρμοστεί κατευθείαν σε κατανεμημένα υπολογιστικά συστήματα. Ο βυζαντινός στρατός είχε περικυκλώσει μια εχθρική πόλη. Ο επικεφαλής των βυζαντινών δυνάμεων πρέπει να αποφασίσει αν θα επιτεθεί στην πόλη ή θα υποχωρήσει. Αν όλα τα πιστά στρατεύματα επιτεθούν ενιαία ή υποχωρήσουν ενιαία , τότε είναι εγγυημένη η επιτυχία. Δυστυχώς, μερικοί βυζαντινοί στρατηγοί είναι άπιστοι και συνεργάζονται με τον εχθρό. Οι άπιστοι στρατηγοί διαδίδουν ψευδείς πληροφορίες με σκοπό να διχάσουν τις πιστές δυνάμεις. Το πρόβλημα είναι να βρεθεί ένας αλγόριθμος που θα χρησιμοποιήσουν όλοι οι πιστοί στρατηγοί για να συμφωνήσουν στην απόφαση να επιτεθούν ή να αποχωρήσουν όπως τους διατάξει ο επικεφαλής.

Πολλοί αλγόριθμοι έχουν παρουσιαστεί για την επίλυση παραλλαγών του προβλήματος των βυζαντινών στρατηγών. Σημειώνουμε ότι τα αποτελέσματα της σύντηξης αισθητηρίων του Marzullo διαφέρουν από το γενικότερο πρόβλημα των βυζαντινών στρατηγών. Εκείνο προσπαθεί να προωθήσει τη **συμφωνία** μεταξύ των επεξεργαστών και οι σχετικοί αλγόριθμοι ανέχονται βλάβες σε μέχρι το ένα τρίτο των συνιστωσών του συστήματος. Η προσέγγιση του Marzullo ενδιαφέρεται κυρίως για την **ακρίβεια** της τελικής απάντησης και ανέχεται βλάβες σε μέχρι το ένα τρίτο των αισθητηρίων.

12.6.Στατικά και δυναμικά δίκτυα

Όλες οι αβεβαιότητες σε δεδομένα αισθητηρίων πρέπει να αντιμετωπιστούν με ένα ομοιόμορφο τρόπο. Οι αβεβαιότητες θα υπάρχουν πάντα εξαιτίας της περιορισμένης ακρίβειας των αισθητηρίων, αλλά λάθη μπορούν επίσης να προκύψουν από αβεβαιότητας που οφείλονται σε κινήσεις των αισθητηρίων. Τα **στατικά δίκτυα (static networks)** είναι δίκτυα αισθητηρίων στα οποία τα αισθητήρια δεν κινούνται σε σχέση με το περιβάλλον τους ούτε και μεταξύ τους. Δίκτυα αισθητηρίων που δεν είναι στατικά ονομάζονται **δυναμικά δίκτυα (dynamic network)**. Η διαδικασία σύντηξης αισθητηρίων είναι απλούστερη για στατικά παρά για δυναμικά δίκτυα. Αν εκ των προτέρων, τότε οι μετρήσεις μπορούν αξιόπιστα να τεθούν στο ίδιο σύστημα συντεταγμένων με ένα μετασχηματισμό συντεταγμένων. Ένα παράδειγμα στατικού δικτύου θα αποτελούσε ένα σύνολο μη κινητών επιγείων σταθμών ραντάρ.

Αν το δίκτυο αισθητηρίων είναι στατικό, τότε οι σχετικές θέσεις των αισθητηρίων είναι αβέβαιες, αλλά αυτή η σχετικότητα πρέπει να αναπαρασταθεί στη διαδικασία σύντηξης δεδομένων. Όπως εξελίσσεται ο χρόνος, αν τα πλεονάζοντα δεδομένα είναι παρόντα στο σύστημα, θα πρέπει να υπάρχει τρόπος να μειωθεί ή και να εξαλειφθεί αυτή η αβεβαιότητα. Για αυτόν τον τύπο προβλήματος χρησιμοποιούνται στατιστικές μέθοδοι.

Αβεβαιότητα θέσης

Υπάρχουν δύο είδη δυναμικών δικτύων: εκείνα που κινούνται μόνο σε σχέση με το περιβάλλον και εκείνα στα οποία αλλάζουν και οι σχετικές θέσεις των αισθητηρίων του. Και στις δύο περιπτώσεις πρέπει να αντιμετωπιστεί η αβεβαιότητα στη θέση των αισθητηρίων.

Αν τα αισθητήρια κινούνται μόνο σε σχέση με το περιβάλλον, όπως σε ένα κινητό ρομπότ στο οποίο η διάταξη των αισθητηρίων είναι σταθερή, η ελαχιστοποίηση της αβεβαιότητας στη θέση των αισθητηρίων μπορεί να επιτευχθεί με συνεχή σύγκριση των μετρήσεων με προηγούμενες μετρήσεις. Σημάδια στο περιβάλλον μπορούν να χρησιμοποιηθούν για να βελτιωθεί η ακρίβεια των εκτιμήσεων της θέσης των αισθητηρίων. Για αυτό τον τύπο προβλήματος, μπορούν να χρησιμοποιηθούν **φίλτρα Kalman** αν υποτεθεί ότι τα σφάλματα ακολουθούν κανονική κατανομή και αποτελούν, επομένως, γκαουσσιανό θόρυβο.

Το πρόβλημα γίνεται πολύ πιο δύσκολο όταν τα αισθητήρια κινούνται μεταξύ τους. Η σύντηξη αισθητηρίων για τέτοιου τύπου δεδομένα πρέπει να προσθέσει αυτές τις αβεβαιότητες στην αβεβαιότητα που είναι ήδη παρούσα στις μετρήσεις των αισθητηρίων. Για παράδειγμα, τα ρομπότ γραμμών συναρμολόγησης απαιτούν σύντηξη αισθητηρίων όρασης και αφής. Τα αισθητήρια όρασης προσφέρουν πληροφορία θέσης και προσανατολισμού διαφόρων εξαρτημάτων, αλλά απαιτείται επιπλέον πληροφορία αφής, για να μπορούν τα ρομπότ να σηκώνουν και να μεταχειρίζονται εξαρτήματα χωρίς να τα καταστρέφουν. Σε τέτοια διάταξη, το σύστημα αφής τοποθετείται στο βραχίονα των ρομπότ και είναι κινητό, ενώ η θέση της κάμερας παραμένει σταθερή.

Τέλος, δίκτυα αισθητηρίων μπορεί να αλλάζουν με το χρόνο εξαιτίας πρόσθεσης ή αφαίρεσης αισθητηρίων από το δίκτυο. Αυτή η δυνατότητα μπορεί να ελέγχεται από το ίδιο το δίκτυο ή από ένα εξωτερικό παράγοντα. Τέτοιου τύπου αλλαγές μπορεί να αποτελούν μέρος της συντήρησης του συστήματος. Ένα αισθητήριο μπορεί να αφαιρεθεί από το δίκτυο επειδή έχει αναγνωριστεί ως αναξιόπιστο, ενώ μεμονωμένα αισθητήρια μπορούν να αντικαθιστώνται περιοδικά ή όταν διαπιστώνονται λάθη.

Κλιμάκωση

Συχνά, χρησιμοποιούνται δύο διαφορετικοί τύποι αισθητηρίων για αντικείμενα που είναι μακρινά σε σχέση με το σύστημα και για αντικείμενα που είναι σχετικά κοντινά. Σε ένα μεταβαλλόμενο περιβάλλον, οι σχετικές θέσεις των αντικειμένων αλλάζουν και συχνά αντικείμενα που ήταν μακρινά προσεγγίζουν το σύστημα. Εφόσον οι ακρίβειες των αισθητηρίων μεγάλης και μικρής απόστασης είναι συνήθως διαφορετικής τάξης μεγέθους, δεν μπορούν να συντηχθούν με τον ίδιο τρόπο όπως αισθητήρια της ίδιας τάξης ακρίβειας. Από την άλλη πλευρά, είναι ιδιαίτερα χρήσιμο να υπάρχει η δυνατότητα παρακολούθησης αντικειμένων με ακρίβεια καθώς πλησιάζουν. Γενικά, πρέπει να δίνεται προσοχή στην επιλογή της σωστής μεθόδου για το συνδυασμό αυτών των διαφορετικών ειδών δεδομένων κατά το σύντομο χρονικό διάστημα στο οποίο επικαλύπτονται. Ο κίνδυνος είναι το ένα είδος δεδομένων να καταπιέσει το άλλο και να χαθεί χρήσιμη πληροφορία κατά τη διαδικασία αυτή.

12.7.Σύντηξη αισθητηρίων και κατά προσέγγιση συμφωνία

Η ανθεκτικότητα σε αβεβαιότητες και η αξιοπιστία είναι βασικές ιδιότητες που πρέπει να ληφθούν υπόψη στη διαδικασία της ολοκλήρωσης αισθητηρίων για δίκτυα κατανεμημένων αισθητηρίων. Εδώ θα ασχοληθούμε με μεθόδους λήψης της καλύτερης δυνατής απόφασης ολοκλήρωσης αισθητηρίων μέσα από το πλαίσιο του κατανεμημένου υπολογισμού. Αυτό επιτυγχάνεται με το συνδυασμό σύγχρονης έρευνας σε σύντηξη αισθητηρίων και καθιερωμένων μεθόδων κατάληξης σε συμφωνία για κατανεμημένα υπολογιστικά συστήματα.

Το πρόβλημα των βυζαντινών στρατηγών.
Οι βασικές ερωτήσεις που θα μας απασχολήσουν είναι πόσες βλάβες σε εξαρτήματα μπορεί να ανεχτεί ένα δίκτυο και να παραμείνει αξιόπιστο και πώς μπορεί το δίκτυο να ξεχωρίσει την έξοδο των εξαρτημάτων που λειτουργούν σωστά από την έξοδο των ελαττωματικών μηχανών. Μια απαισιόδοξη παρουσίαση αυτής της ερώτησης είναι το πρόβλημα των *βυζαντινών στρατηγών,* όπως το θέσαμε προηγουμένως. Το πρόβλημα προϋποθέτει μια κατανεμημένη διαδικασία λήψης αποφάσεων όπου μερικοί από τους συμμετέχοντες δεν είναι μόνο ελαττωματικοί αλλά επιπλέον προσπαθούν σκόπιμα να επιβάλλουν διαφωνία στην ομάδα. Ένας αλγόριθμος που λύνει το πρόβλημα είναι επαληθεύσιμα αξιόπιστος για χρήση σε κατανεμημένο υπολογισμό, αφού καμία δυνατή βλάβη ενός συγκεκριμένου περιορισμένου αριθμού μηχανών στο δίκτυο δεν μπορεί να προκαλέσει δυσλειτουργία. Ένας τέτοιος αλγόριθμος λέγεται ότι έχει καταλήξει σε **βυζαντινή συμφωνία (Byzantine agreement).**

Το πρόβλημα των βυζαντινών στρατηγών μπορεί να δηλωθεί ως εξής: διατίθενται Ν ανεξάρτητα επεξεργαστικά συστήματα από τα οποία μέχρι t είναι ελαττωματικά. Χρειάζεται να αναπτυχθεί ένα πρωτόκολλο που εγγυάται ότι για μια μετάδοση από έναν επεξεργαστή Χ:

- Οι μη ελαττωματικοί επεξεργαστές συμφωνούν μεταξύ τους για το περιεχόμενο των δεδομένων που λαμβάνουν από τον Χ.

- Αν ο Χ δεν είναι ελαττωματικός, η συμφωνία ισούται με το μήνυμα που μετέδωσε ο Χ.

Αυτή η εγγύηση ονομάζεται επίσης **«αλληλεπιδραστική συμβατότητα στρατηγών (general interactive consistency)»**.

Έχουν αποδειχτεί μερικά ενδιαφέροντα χαρακτηριστικά αυτού του προβλήματος:

- Το πρόβλημα μπορεί να λυθεί μόνο αν ο αριθμός t των «προδοτών» είναι μικρότερος του ενός τρίτου του συνολικού αριθμού N των επεξεργαστικών συστημάτων.
- Ο αριθμός t πρέπει να είναι μικρότερος του μισού της συνεκτικότητας του γράφου.

Οι δύο παραπάνω παρατηρήσεις δηλώνουν ότι για να γίνονται ανεκτές t βλάβες, το σύστημα πρέπει να αποτελείται από τουλάχιστον 3t+1 επεξεργαστικά συστήματα και κάθε επεξεργαστικό σύστημα να συνδέεται άμεσα με τουλάχιστον 2t+1 άλλα επεξεργαστικά συστήματα. Επίσης έχει αποδειχθεί ότι ένας αλγόριθμος που λύνει το πρόβλημα των βυζαντινών στρατηγών πρέπει να εκτελεί τουλάχιστον t+1 γύρους μεταδόσεων μεταξύ των κόμβων.

Πολλοί αλγόριθμοι έχουν βρεθεί για τη λύση του προβλήματος των βυζαντινών στρατηγών, όπως και αρκετές παραλλαγές του προβλήματος έχουν προταθεί. Ένας τυπικός αλγόριθμος προτάθηκε από τον Dolev και αποτελείται από τρία βήματα:

I. Το επεξεργαστικό σύστημα – πομπός N_1 στέλνει την τιμή του σε όλα τα άλλα επεξεργαστικά συστήματα μέσα από 2t+1 διαχωρισμένα μονοπάτια. Ένας μετρητής m τίθεται σε κάποια αρχική τιμή.

II. Κάθε επεξεργαστικό σύστημα συγκρίνει την τιμή που λαμβάνει από τα διάφορα μονοπάτια. Αν διαπιστώσει συμφωνία μεταξύ όλων των τιμών που έλαβε οι οποίες δεν πέρασαν από ένα σύνολο U από t ή λιγότερα επεξεργαστικά συστήματα, τότε αυτή η τιμή είναι η σωστή τιμή και τα επεξεργαστικά συστήματα του συνόλου U είναι «ύποπτα». Αλλιώς ο μετρητής m τίθεται σε κάποια κατά συνθήκη τιμή.

III. Αν m>0 :
   A. Κάθε επεξεργαστικό σύστημα εκτελεί αναδρομικά τον αλγόριθμο αυτό με m=m-1, αποκλείοντας τον εαυτό του από το σύνολο των επεξεργαστικών συστημάτων προς τα οποία μεταδίδει και στέλνοντας την τιμή που προσδιορίστηκε στο βήμα II.
   B. Όλα τα επεξεργαστικά συστήματα καταγράφουν τις τιμές που δέχτηκαν από άλλα επεξεργαστικά συστήματα. Αν από κάποιο επεξεργαστικό σύστημα δεν ληφθεί καμμιά τιμή, τότε ως τιμή αυτού του επεξεργαστικού συστήματος τίθεται κάποια κατά συνθήκη τιμή. Η σωστή τιμή είναι αυτή που έχει μεταδοθεί από την πλειοψηφία των επεξεργαστικών συστημάτων.

Προσεγγιστική βυζαντινή συμφωνία

Πολλές παραλλαγές έχουν προταθεί για το πρόβλημα των βυζαντινών στρατηγών. Θα συζητήσουμε εδώ δύο παραλλαγές που προσπαθούν να βρουν συμφωνία μεταξύ των επεξεργαστικών συστημτων σε μια πραγματική τιμή που εμπίπτει μέσα σε ένα δεδομένο διάστημα. Αυτό το πρόβλημα συναντάται συχνά στο σχεδιασμό συστημάτων. Πιθανές εφαρμογές του προβλήματος περιλαμβάνουν το συγχρονισμό ρολογιών σε κατανεμημένα συστήματα. Το πρόβλημα έχει τεθεί και λυθεί με δύο διαφορετικούς τρόπους. Πρώτα θα δούμε το πρόβλημα όπως τέθηκε από τον Dolev και άλλους και στη συνέχεια θα δούμε την παραλλαγή που διερεύνησαν οι Mahaney και Schneider.

Το πρόβλημα προσεγγιστικής συμφωνίας που έθεσαν ο Dolev και άλλοι υποθέτει N ανεξάρτητα επεξεργαστικά συστήματα, όπου κάθε επεξεργαστικό σύστημα ξεκινάει με τη δική του πραγματική τιμή και όλα τα επεξεργαστικά συστήματα αναζητούν μια τιμή σε απόσταση μικρότερη από ε από τις τιμές των άλλων επεξεργαστικών συστημάτων. Ενας αλγόριθμος προσεγγιστικής συμφωνίας πρέπει να ικανοποιεί τις ακόλουθες δύο απαιτήσεις:

1. **Συμφωνία (agreement)** – Ολα τα επεξεργαστικά συστήματα χωρίς βλάβη πρέπει να τερματίσουν με τιμές εξόδου σε μετα;ξύ τους απόσταση μικρότερη από ε.

2. **Εγκυρότητα (validity)** – Αυτές οι τιμές πρέπει όλες να εμπίπτουν στο διάστημα αρχικών τιμών των επεξεργαστικών συστημάτων χωρίς βλάβη.

Στο σημείο αυτό ορίζουμε τις έννοιες της **ορθότητας (accuracy)** και της **ακρίβειας (precision)** ενός συστήματος. Η ορθότητα αναφέρεται στην απόσταση των αποτελεσμάτων του συστήματος από τα επιθυμητά αποτελέσματα. Για παράδειγμα, όταν κάποιος βάλλει κατά ενός στόχου, η ορθότητα αναφέρεται στην απόσταση της πιο απομακρυσμένης βολής από το κέντρο του στόχου. Η ακρίβεια αναφέρεται στο μέγεθος του διαστήματος τιμών που επιστρέφει το σύστημα. Στο παράδειγμα των βολών κατά στόχου, η ακρίβεια θα μπορούσε να προσδιοριστεί ως η ακτίνα του μικρότερου κύκλου που περικλείει όλες τις βολές.

Η **συμφωνία (agreement)** στον ορισμό του προβλήματος της κατά προσέγγιση συμφωνίας δίνει την τελική ακρίβεια του συστήματος. Ο αλγόριθμος θα τερματίσει με κάθε επεξεργαστικό σύστημα να έχει τιμή που απέχει λίγο από κάποιο δεδομένο (αυθαίρετα μικρό) ε από τις τιμές όλων των άλλων επεξεργαστικών συστημάτων. Η **εγκυρότητα (validity)** αντιστοιχεί στην ορθότητα του συστήματος. Εστω δ το διάστημα τιμών που επιστρέφουν επεξεργαστικά συστήματα που λειτουργούν χωρίς βλάβη. Εφόσον η απάντηση που επιστρέφεται είναι σε απόσταση μικρότερη από δ, είναι τουλάχιστον τόσο ορθή όσο οι απαντήσεις των επεξεργαστικών συστημάτων που λειτουργούν χωρίς βλάβη.

Επομένως, η έρευνα των Dolev και άλλων επιχειρεί να αυξήσει την ακρίβεια του συστήματος. Αυτό το κέρδος σε ακρίβεια δεν επιτρέπεται να μεταβάλλει το σύστημα σε λιγότερο ορθό από την αρχική τιμή του ελάχιστα ορθού από τα επεξεργαστικά συστήματα που λειτουργούν χωρίς βλάβη.

Ο Dolev και άλλοι χρησιμοποιούν μια συνολοθεωρητική προσέγγιση στο πρόβλημα. Η αυστηρή μεθοδολογία τους μπορεί να παρουσιαστεί και διαισθητικά. Από τον ορισμό της ορθότητας ενός συστήματος, η τιμή ενός επεξεργαστικού συστήματος είναι άκυρη μόνο όταν βρίσκεται εκτός της επιτρεπτής ορθότητας που αναπαριστάνουμε με δ.

Εφόσον το πολύ t επεξεργαστικά συστήματα έχουν βλάβη, το πολύ t επεξεργαστικά συστήματα έχουν τιμές μικρότερες από το κάτω όριο δ. Ανάλογα, το πολύ t επεξεργαστικά συστήματα έχουν τιμές μεγαλύτερες από το άνω όριο δ. Ο αριθμός των επεξεργαστικών συστημάτων είναι τουλάχιστον 3 t + 1, επομένως αν αγνοήσουμε τις t μικρότερες και τις t μεγαλύτερες τιμές, μας μένουν τουλάχιστον t + 1 τιμές στο διάστημα δ. Υπολογίζοντας το μέσο όρο ενός υποσυνόλου αυτών των τιμών, καταλήγουμε σε μια τιμή που επίσης εμπίπτει στο διάστημα δ.

Επομένως, ο αλγόριθμος για το σύγχρονο πρόβλημα είναι:

I. Κάθε επεξεργαστικό σύστημα εκπέμπει καθολικά (προς όλα τα επεξεργαστικά συστήματα) την τιμή του.

II. Κάθε επεξεργαστικό σύστημα λαμβάνει τις τιμές από όλα τα άλλα επεξεργαστικά συστήματα και τις ταξινομεί κατά μέγεθος σε ένα διάνυσμα v.

III. Οι t χαμηλότερες τιμές και οι t υψηλότερες τιμές απορρίπτονται από το διάνυσμα V σε κάθε επεξεργαστικό σύστημα.

IV. Κάθε επεξεργαστικό σύστημα σχηματίζει ένα νέο διάνυσμα v΄ που περιέχει τις τιμές v(i*t), για i = 0,1,2,…, δηλαδή τη μικρότερη τιμή και κάθε μια από αυτές που απέμειναν.

V. Η νέα τιμή είναι ο μέσος όρος των τιμών στο v΄.

Ενας αλγόριθμος για το ισοδύναμο ασύγχρονο πρόβλημα έχει επίσης παρουσιαστεί από τον Dolev και άλλους. Είναι σχεδόν ταυτόσημος με τον αλγόριθμο για το σύγχρονο πρόβλημα με τις επόμενες εξαιρέσεις:

- το n πρέπει να είναι μεγαλύτερο από 5 t
- στο βήμα III, απορρίπτονται οι 2 t χαμηλότερες και οι 2 t υψηλότερες τιμές
- στο βήμα IV, το v΄ σχηματίζεται από τις τιμές v(2*i*t), i = 0,1,2,….

Το άρθρο των Dolev και άλλων δείχνει ότι και οι δύο αλγόριθμοι συγκλίνουν. Επίσης, οι αλγόριθμοι μπορούν να εξαναγκαστούν σε σύγκλιση μέσα σε ένα αυθαίρετα μικρό διάστημα ε με την εκτέλεση πολλών επαναλήψεων.

Ανακριβής βυζαντινή συμφωνία

Η παραλλαγή ανακριβούς συμφωνίας του προβλήματος των βυζαντινών στρατηγών έχει μελετηθεί από τους Mahaney και Schneider. Θέτοντας το πρόβλημά τους, οι Mahaney και Schneider λαμβάνουν εκπεφρασμένα υπόψη και την ορθότητα και την ακρίβεια του συστήματος. Οριζουν την ακρίβεια ως τη μέγιστη απόσταση ανάμεσα στις τιμές δύο επεξεργαστικών συστημάτων που λειτουργούν χωρίς βλάβη. Η ορθότητα ορίζεται ως η μέγιστη απόσταση που μπορεί να απέχει ανάμεσα η τιμή ενός επεξεργαστικού συστήματος από την πραγματική τιμή που πρέπει να εκτιμηθεί και να θεωρείται ως ορθή.

Ενας αλγόριθμος ανακριβούς συμφωνίας λαμβάνει τιμές από N επεξεργαστικά συστήματα, μέχρι t από τα οποία μπορεί να έχουν αυθαίρετη βλάβη ($N > 3\ t$) και επιστρέφει μια τιμή για κάθε επεξεργαστικό σύστημα που απέχει

- μέχρι δ από την πραγματική τιμή (ορθότητα)
- μέχρι ε από τις τιμές όλων των άλλων επεξεργαστικών συστημάτων.

Ανάλογα προς την προσεγγιστική συμφωνία, αυτοί οι αλγόριθμοι συγκλίνουν και μπορούν να χρησιμοποιηθούν επαναληπτικά για να παράσχουν αυθαίρετο βαθμό ακρίβειας. Αυτό το κέρδος σε ακρίβεια μπορεί όμως να οδηγήσει σε απώλεια ορθότητας.

Σημειώνουμε ότι ο ορισμός της ορθότητας που χρησιμοποιείται από τους Mahaney και Schneider είναι λιγότερο περιοριστικός απο τον ορισμό των Dolev και άλλων, αλλά και πλησιέστερος προς τη φυσική διαδικασία που προσπαθεί να μοντελοποιήσει.

Οι Mahaney και Schneider παρουσιάζουν δύο αλγόριθμους που επιτελούν ανακριβή συμφωνία. Και οι δύο αλγόριθμοι χρησιμοποιούν σύνολα **αποδεκτών** τιμών. Μια τιμή είναι **αποδεκτή (acceptable)** αν απέχει απόσταση μικρότερη από δ από N-t άλλες τιμές. Αφού ο ορισμός του προβλήματος δηλώνει ότι το πολύ t επεξεργαστικά συστήματα έχουν βλάβη, κάθε τιμή που δεν είναι αποδεκτή δεν μπορεί να είναι ορθή.

Ο **αλγόριθμος ταχείας συγκλίσεως (fast convergence algorithm)** που παρουσιάζουν εκτελείται σε κάθε επεξεργαστικό σύστημα:

I. Κάθε επεξεργαστικό σύστημα λαμβάνει τις τιμές όλων των άλλων επεξεργαστικών συστημάτων και σχηματίζει ένα σύνολο v.

II. Εκείνες οι τιμές που είναι *αποδεκτές* τοποθετούνται σε ένα σύνολο A.

III. Υπολογίζεται (όπως παρακάτω) μια ποσότητα e(A).

IV. Οποιεσδήποτε μη αποδεκτές τιμές αντικαθίστανται στο v με e(A).

V. Η νέα τιμή για το επεξεργαστικό σύστημα είναι ο μέσος όρος των τιμών στο v.

Η τιμή e(A) μπορεί να είναι οποιαδήποτε από έναν αριθμό συναρτήσεων των τιμών που είναι αποθηκευμένες στο A. Οι Mahaney και Schneider προτείνουν το μέσο όρο και ανάλογες συναρτήσεις του διαστήματος τιμών ως πιθανές επιλογές για την τιμή e(A).

Ο δεύτερος αλγόριθμος που προτείνεται είναι **ο αλγόριθμος σύγκλισης σταυροφόρων (crusaders convergence algorithm)**. Ο αλγόριθμος σύγκλισης σταυροφόρων είναι σχεδόν ταυτόσημος με τον αλγόριθμο ταχείας σύγκλισης με την εξαίρεση ότι ο αλγόριθμος σύγκλισης σταυροφόρων που προτάθηκε από τον Dolev εκτελείται στο σύνολο ν μεταξύ των βημάτων 1 και 2.

Και οι δύο αλγόριθμοι συγκλίνουν εγγυημένα όταν $t < N/3$. Η απόδοση των αλγορίθμων μειώνεται ``διακριτικά´´ εφόσον $t < 2N/3$, δηλαδή είτε δεν σχηματίζεται ένα αποδεκτό σύνολο και τερματίζει ο αλγόριθμος είτε οι απαντήσεις που επιστρέφει ο αλγόριθμος δεν είναι τόσο καλές όσο όταν $t < N/3$, αλλά εμπίπτουν σε λογικά όρια ακρίβειας και ορθότητας.

12.8 Σύντηξη αντιφατικής πληροφορίας αισθητηρίων

Ο σύγχρονος κόσμος περιέχει πολλά αυτοματοποιημένα συστήματα που πρέπει να αλληλεπιδρούν με περιβάλλοντα που μεταβάλλονται με το χρόνο. Αφού αυτά τα περιβάλλοντα δεν μπορούν να προκαθοριστούν, τα συστήματα βασίζονται σε αισθητήρια που τους παρέχουν την πληροφορία που χρειάζονται για να επιτελέσουν τα καθήκοντά τους. Δυστυχώς, τα αισθητήρια υπόκεινται σε λάθη, αβεβαιότητες και μηχανικές βλάβες. Προκειμένου να αποφευχθεί η ευαλωσιμότητα ενός συστήματος σε βλάβη ενός εξαρτήματός του, είναι λογική η πλεονασματική χρήση αισθητηρίων. Για παράδειγμα, ένα σύστημα παρακαολούθησης θα μπορούσε να χρησιμοποιεί διαφορετικά είδη αισθητηρίων (ραντάρ, υπέρυθρης ακτινοβολίας, μικροκυματκικής ακτινοβολίας) που δεν είναι εθάλωτα στο ίδιο είδος παρεμβολής. Η πλεονασματική πληροφορία επιτρέπει στο σύστημα να συνεχίσει να λειτουργεί παρά ένα μικρό αριθμό λαθών σε εξαρτήματά του.

Με αυτή τη μέθοδο, ο σχεδιαστής αντιμετωπίζει ένα καινούργιο πρόβλημα. Το σύστημα λαμβάνει μερικές μετρήσεις, που ολικά ή μερικά περιέχουν λάθος, και πρέπει να αποφασίσει ποιά εξαρτήματα έχουν πάθει βλάβη, όπως και πως να ερμηνεύει μετρήσεις που είναι μερικώς τουλάχιστον αντιφατικές.

Ο Marzullo πρότεινε ένα μοντέλο με ``αφηρημένα´´ αισθητήρια, που αποτελούνται από ένα δάστημα τιμών που επιστρέφει ένα αισθητήριο, και ``συμπαγή´´ αισθητήρια, που είναι οι φυσικές συσκευές που επιστρέφουν την τιμή. Εφόσον όλα τα αισθητήρια έχουν περιορισμένη ακρίβεια, η τιμή που επιστρέφει ένα αφηρημένο αισθητήριο αποτελείται από ένα κατώτερο και ένα ανώτερο φράγμα. Με τον τρόπο αυτό, ανακρίβειες αισθητηρίων μπορούν να αντιμετωπιστούν εκπεφρασμένα.

Το πρόβλημα σύντηξης αισθητηρίων που αντιμετωπίζουμε είναι: δεδομένου ενός συνόλου N αισθητηρίων, όλων με περιορισμένη ακρίβεια και το πολύ t με βλάβη, ποιό είναι το μικρότερο διάστημα τιμών όπου βρίσκεται με σιγουριά η ορθή τιμή;

Οπως και με τους αλγόριθμους προσεγγιστικού ταιριάσματος, έχουμε ένα σύνολο N πραγματικών τιμών, η ορθότητα των οποίων εξαρτάται από το να βρίσκονται μέσα σε κάποια απόσταση από τη μοναδική ορθή τιμή, όπως και μέχρι t τιμές που είναι αυθαίρετα λανθασμένες. Σε αντίθεση με αλγόριθμους βασισμένους στο πρόβλημα των βυζαντινών στρατηγών, αυτή η απόσταση μπορεί να είναι διαφορετική για κάθε αισθητήριο (επεξεργαστικό σύστημα).

Σημειώνουμε επίσης ότι δεν τίθενται συγκεκριμένες απαιτήσεις ακρίβειας για αυτό το πρόβλημα.

Κάθε αισθητήριο (επεξεργαστικό σύστημα) αναπαριστάνεται από το ανώτερο και το κατώτερο όριο για την τιμή της μέτρησης. Το μέγεθος αυτού του διαστήματος είναι η ορθότητα δ αυτού του αισθητηρίου. Η χρήση αισθητηρίων διαφορετικού τύπου έχει συχνα πλεονεκτήματα και, επομένως, δεν υποθέτουμε ότι το δ είναι ομοιόμορφο για τα αισθητήρια και δεν θέτουμε περιορισμούς ακρίβειας. Ο Marzullo τονίζει ότι μετρήσεις είναι χρήσιμες μόνο όταν είναι ορθές και το διάστημα τιμών τους αρκετά μικρό. Επομένως, ένα αισθητήριο θεωρείται ορθό όταν το διάστημα τιμών που επιστρέφει είναι περιορισμένο και περιχει την αναζητούμενη φυσική τιμή.

Με βάση αυτούς τους ορισμούς και περιορισμούς, ο Marzullo, έχει αποδείξει ότι μπορεί να βρεθεί η μικρότερη περιοχή που περιέχει τη φυσική μεταβλητή. Το μέγεθος της περιοχής αυτής θα είναι μικρότερο ή ίσο με το μεγαλύτερο διάστημα τιμών που επιστρέφεται από αισθητήριο (επεξεργαστικό σύστημα). Το πρόβλημα λύνεται βρίσκοντας όλες τις περιοχές όπου τέμνονται N-t μετρήσεις αισθητηρίων. Είναι φανερό ότι η ορθή τιμή πρέπει να βρίσκεται σε ένα από αυτά τα διαστήματα. Το σωστό διστημα ορίζεται από την τιμή του μικρότερου κάτω φράγματος και του μεγαλύτερου άνω φράγματος αυτών των τομών.

Στη συνέχεια παρουσιάζουμε τον αλγόριθμο σύντηξης αισθητηρίων του Jayasimha που έχει την ίδια πολυπλοκότητα με τον αλγόριθμο που πρότεινε αρχικά ο Marzullo. Ο αλγόριθμος γράφτηκε για χρήση σε ένα κεντρικό επεξεργαστή που δέχεται δεδομένα κατευθείαν από τα αισθητήρια. Μπορεί όμως να υλοποιηθεί ως ένα κατανεμημένο δίκτυο αισθητηρίων, όπου όλα τα αισθητήρια εκτελούν σύντηξη πληροφορίας παράλληλα χωρίς τροποποιήσεις. Σε ένα κατανεμημένο περιβάλλον, η ακρίβεια του αλγορίθμου έχει το ίδιο όριο με την ορθότητά του.

13. Το προγραμματιστικό περιβάλλον MATLAB

13.1. Τι είναι το MATLAB;

Το MATLAB είναι ένα υπολογιστικό περιβάλλον για αριθμητικούς υπολογισμούς και οπτικοποίηση. Το MATLAB παρέχει αριθμητική ανάλυση, υπολογισμούς με πίνακες, επεξεργασία σήματος και γραφικά σε ένα εύχρηστο περιβάλλον στο οποίο προβλήματα και λύσεις εκφράζονται όπως ακριβώς διατυπώνονται με μαθηματικό τρόπο, χωρίς δηλαδή παραδοσιακό προγραμματισμό.

MATLAB σημαίνει MATrix LABoratory (Εργαστήριο Πινάκων) και αρχικά αναπτύχθηκε για να παρέχει εύκολη πρόσβαση σε λογισμικό πινάκων που είχε αναπτυχθεί στα πλαίσια των προγραμμάτων LINPACK και EISPACK.

Το MATLAB είναι ένα αλληλεπιδραστικό σύστημα στο οποίο το βασικό στοιχείο δεδομένων είναι ένας πίνακας που δεν απαιτεί προσδιορισμό των διαστάσεών του. Αυτό επιτρέπει την επίλυση αριθμητικών προβλημάτων σε ένα κλάσμα του χρόνου που θα απαιτούσε ο προγραμματισμός σε μια γλώσσα όπως η FORTRAN, η BASIC, η C κλπ.

Το MATLAB συμπληρώνεται από ένα σύνολο πακέτων εφαρμογών που ονομάζονται **εργαλειοθήκες (toolboxes)**. Οι εργαλειοθήκες είναι πλήρεις συλλογές από συναρτήσεις MATLAB που επεκτείνουν το περιβάλλον του MATLAB έτσι ώστε να επιλύονται συγκεκριμμένες κλάσεις προβλημάτων. Περιοχές για τις οποίες διατίθενται εργαλειοθήκες περιλαμβάνουν την επεξεργασία σήματος, το σχεδιασμό συστημάτων ελέγχου, την προσομοίωση δυναμικών συστημάτων, την αναγνώριση συστημάτων, τα νευρωνικά δίκτυα, κλπ.

Το πιο σημαντικό χαρακτηριστικό του MATLAB είναι ίσως η επεκτασιμότητά του. Αυτό επιτρέπει στον κάθε χρήστη να συνεισφέρει τις δικές του εφαρμογές. Ως αποτέλεσμα, επιστήμονες, μαθηματικοί και μηχανικοί έχουν συνεισφέρει νέες και ενδιαφέρουσες εφαρμογές χωρίς να γράψουν ούτε μια γραμμή κώδικα σε γλώσσα χαμηλού επιπέδου.

13.2. Το σύστημα του MATLAB

Το σύστημα του MATLAB αποτελείται από πέντε κύρια μέρη:

1. Τη γλώσσα MATLAB : Αυτή είναι μία υψηλού επιπέδου γλώσσα μήτρας/πίνακα με έλεγχο ροής δηλώσεων, συναρτήσεων, δομών δεδομένων, εισόδου/εξόδου, και των **αντικειμενοστρεφών (object-oriented)** χαρακτηριστικών προγράμματος. Επιτρέπει τόσο τη δημιουργία γρήγορων και “βρώμικων” προγραμμάτων **(“programming in the small”)**, όσο και τη δημιουργία ολοκληρωμένων μεγάλων και πολύπλοκων προγραμματιστικών εφαρμογών **(“programming in the large”).**

2. Το λειτουργικό περιβάλλον του MATLAB : Αυτό είναι το σύνολο των εργαλείων και των διευκολύνσεων με το οποίο δουλεύει κανείς με το MATLAB ως χρήστης ή ως προγραμματιστής. Περιλαμβάνει διευκολύνσεις για τον έλεγχο των μεταβλητών στο χώρο εργασίας και την είσοδο και έξοδο δεδομένων. Επίσης, περιλαμβάνει εργαλεία για την εφαρμογή, έλεγχο, διόρθωση λαθών, και κατανομή των Μ-αρχείων.

3. Το σύστημα γραφικών του MATLAB : Αυτό είναι το γραφικό σύστημα του MATLAB. Περιλαμβάνει υψηλού επιπέδου εντολές για διδιάστατη και τριδιάστατη απεικόνιση δεδομένων, επεξεργασία εικόνας, δημιουργία κινουμένων σχεδίων (animation), και παρουσίαση γραφικών. Επίσης, περιλαμβάνει χαμηλού επιπέδου εντολές που επιτρέπουν την κατασκευή και εμφάνιση γραφικών, όπως την κατασκευή πλήρους διεπαφής προς το χρήστη (user interface).

4. Τη βιβλιοθήκη μαθηματικών συναρτήσεων του MATLAB : Αυτή είναι μια απέραντη συλλογή από υπολογιστικούς αλγόριθμους που ξεκινά από στοιχειώδεις συναρτήσεις όπως άθροισμα, ημίτονο, συνημίτονο, και πολύπλοκη αριθμητική, και φτάνει μέχρι πιο πολύπλοκες συναρτήσεις όπως μήτρες, αντίστροφες μητρών, ιδιοτιμές μήτρας, συναρτήσεις Bessel, και μετασχηματισμούς Fourier.

5. Τη διεπαφή προς προγράμματα εφαρμογών (Application Program Interface, API) του MATLAB : Αυτή είναι μία βιβλιοθήκη που σου επιτρέπει να γράψεις προγράμματα σε C, Fortran κλπ. τα οποία αλληλεπιδρούν με το MATLAB. Επίσης, περιέχει διευκολύνσεις για το κάλεσμα ρουτινών της MATLAB (δυναμική συνάρμοση), κάλεσμα του MATLAB σαν μία υπολογιστική μηχανή, και για το διάβασμα και γράψιμο MAT-αρχείων.

Το Simulink: Το Simulink, ένα συνοδευτικό πρόγραμμα του MATLAB, είναι ένα σύστημα που αλληλεπιδρά με προσομοιωμένα μη γραμμικά δυναμικά συστήματα. Είναι ένα γραφικό πρόγραμμα – χρησιμοποιεί ποντίκι – που επιτρέπει τη μοντελοποίηση ενός συστήματος σχεδιάζοντας ένα ιστόγραμμα στην οθόνη και μεταβάλοντας το δυναμικά. Μπορεί να δουλέψει με γραμμικά, μη γραμμικά, συνεχούς χρόνου, διακριτού χρόνου, πολυμεταβλητά και πολυρυθμικά (multirate) συστήματα.

13.3.Ξεκινώντας με το MATLAB

Για να εκκινήσετε το MATLAB σε ένα PC, κάνετε διπλό κλικ στο εικονίδιο του MATLAB. Για να εκκινήσετε το MATLAB σε ένα σύστημα με UNIX, γράψτε matlab στην προτροπή του λειτουργικού συστήματος. Για να κλείσετε το MATLAB οποιαδήποτε στιγμή, γράψτε quit στην προτροπή του MATLAB.

Εάν νιώθετε ότι χρειάζεστε περισσότερη βοήθεια, γράψτε help στην προτροπή του MATLAB ή ανοίξτε το μενού HELP σε ένα PC. Θα μιλήσουμε περισσότερο για την βοήθεια και των κειμένων εντός γραμμής (on line) αργότερα.

Μήτρες και μαγικά τετράγωνα

Ο καλύτερος τρόπος για να ξεκινήσετε με το MATLAB είναι να μάθετε πώς να χειρίζεστε πίνακες. Αυτή η ενότητα θα σας δείξει πώς να το κάνετε αυτό. Στο MATLAB, ένας πίνακας είναι ένα τετραγωνικό διάνυσμα αριθμών. Μερικές φορές δίνουμε ειδική σημασία στους 1 επί 1 πίνακες, που είναι βαθμωτά μεγέθη, και στους πίνακες γραμμής ή στήλης, που είναι διανύσματα. Το MATLAB έχει και άλλους τρόπους να αποθηκεύει τόσο αριθμητικά όσο και μη αριθμητικά δεδομένα, αλλά αρχικά είναι καλύτερα να σκεπτόμαστε τα πάντα ως πίνακες. Οι λειτουργίες του είναι σχεδιασμένες να είναι όσο τον δυνατόν πιο φυσικές. Εκεί που οι άλλες γλώσσες προγραμματισμού δουλεύουν με ένα αριθμό τη φορά, το MATLAB σας επιτρέπει να δουλεύετε με ολόκληρους πίνακες γρήγορα και εύκολα.

Μπορείτε να εισάγετε πίνακες στο MATLAB με διάφορους τρόπους.

1. Εισάγοντας μία λίστα στοιχείων.
2. Φορτώνοντας πίνακες από εξωτερικά αρχεία δεδομένων.
3. Παράγοντας πίνακες χρησιμοποιώντας συναρτήσεις του MATLAB.
4. Δημιουργώντας πίνακες με δικές σας συναρτήσεις σε M-αρχεία.

Αρχίστε εισάγοντας τον πίνακα του Dürer ως λίστα των στοιχείων του. Αρκεί απλά να ακολουθήσετε μερικούς βασικούς κανόνες:

1. Χωρίστε τα στοιχεία κάθε γραμμής με κενά ή κόμμα.
2. Χρησιμοποιήστε το ελληνικό ερωτηματικό, ; , για να δηλώσετε το τέλος κάθε γραμμής.
3. Τοποθετήστε όλη την λίστα με αγκύλες, [ ].

Για να εισάγετε τον πίνακα του Dürer, απλά γράψτε:

```
A = [16 3 2 13; 5 10 11 8; 9 6 7 12; 4 15 14 1]
```

Το MATLAB προβάλει τη μήτρα που μόλις εισάγατε,

```
A =
   16    3    2   13
    5   10   11    8
    9    6    7   12
    4   15   14    1
```

sum, transpose, and diag

Πιθανόν να γνωρίζετε ήδη ότι οι ξεχωριστές ιδιότητες ενός μαγικού τετραγώνου έχουν σχέση με τους διάφορους τρόπους με τους οποίους αθροίζουμε τα στοιχεία του. Το άθροισμα οποιαδήποτε γραμμής ή στήλης, καθώς και των δύο κύριων διαγωνίων του, είναι πάντα ο ίδιος αριθμός. Ας το εξακριβώσουμε αυτό χρησιμοποιώντας το MATLAB. Η πρώτη συνάρτηση που θα δοκιμάσουμε είναι

```
sum(A)
```

Το MATLAB απαντάει με το

```
ans =
   34   34   34   34
```

Όταν δεν καθορίζετε μια μεταβλητή εξόδου, το MATLAB χρησιμοποιεί την μεταβλητή ans, συντομογραφία του answer, για να αποθηκεύσει το αποτέλεσμα ενός υπολογισμού. Έχετε υπολογίσει ένα διάνυσμα που περιέχει τα αθροίσματα των στηλών του A. Φυσικά, κάθε στήλη έχει το ίδιο άθροισμα, τον μαγικό άθροισμα, 34.

Και τα αθροίσματα των γραμμών; Το MATLAB προτιμάει να δουλεύει με τις στήλες ενός πίνακα, έτσι ο ευκολότερος τρόπος για να πάρετε τα αθροίσματα των γραμμών είναι να αντιμεταθέσετε τις γραμμές του πίνακα με τις στήλες, να υπολογίσετε τα αθροίσματα των νέων στηλών και μετά να αντιμεταθέσετε το αποτέλεσμα. Η πράξη της αντιμετάθεσης δηλώνεται με μία απόστροφο, ΄. Περιστρέφει ένα πίνακα γύρω από την κύρια διαγώνιο του και μετατρέπει ένα πίνακα γραμμής σε ένα πίνακα στήλης. Έτσι το

```
A΄
```

παράγει

```
ans =
   16    5    9    4
    3   10    6   15
    2   11    7   14
   13    8   12    1
```

Και το

```
sum(A΄)΄
```

παράγει ένα πίνακα στήλη ο οποίος περιέχει τα αθροίσματα των γραμμών

```
ans =
   34
   34
   34
   34
```

Εύκολα βρίσκουμε το άθροισμα των στοιχείων της κύριας διαγωνίου με τη βοήθεια της συνάρτησης diag η οποία τα ξεχωρίζει από τον πίνακα. Έτσι το

```
diag (A)
```

παράγει

```
ans =
   16
   10
    7
    1
```

και

```
sum(diag(A))
```

παράγει

```
ans =
    34.
```

Η άλλη διαγώνιος, η αντιδιαγώνιος, δεν είναι πολύ σημαντική μαθηματικά και έτσι το MATLAB δεν έχει μία έτοιμη συνάρτηση για αυτή. Υπάρχει όμως μία συνάρτηση, αρχικά προοριζόμενη για χρήση στα γραφικά, η fliplr, η οποία «περιστρέφει» ένα πίνακα γύρω από ένα κατακόρυφο άξονα.

```
sum(diag(fliplr(A)))
ans =
    34
```

Έχετε δείξει ότι ο πίνακας στη ξυλογραφία του Dürer είναι πράγματι ένα μαγικό τετράγωνο και, στην πορεία, δοκιμάσατε μερικές από τις συναρτήσεις πινάκων που σας προσφέρει το MATLAB. Το επόμενο τμήμα συνεχίζει να χρησιμοποιεί τους πίνακες για να δείξει μερικές ακόμα δυνατότητες του MATLAB.

Δείκτες (Subscripts)

Μπορούμε να αναφερθούμε στο στοιχείο γραμμής j και στήλης j με το A(i,j). Για παράδειγμα, το A(4,2) είναι ο αριθμός στην τέταρτη γραμμή και στην δεύτερη στήλη. Στο μαγικό μας τετράγωνο ο A(4,2) είναι το 15. Έτσι είναι δυνατό να υπολογίσουμε το άθροισμα της τέταρτης στήλης γράφοντας

```
A(1,4) + A(2,4) + A(3,4) + A(4,4).
```

Αυτό μας δίνει

```
ans =
    34
```

αλλά δεν είναι ο πιο κομψός τρόπος για να βρούμε το άθροισμα μίας στήλης.

Είναι επίσης δυνατό να αναφερθούμε στα στοιχεία ενός πίνακα με ένα μονό δείκτη, το A(k). Αυτός είναι ο συνήθης τρόπος με τον οποίο αναφερόμαστε σε πίνακες γραμμής και στήλης. Μπορούμε όμως να τον εφαρμόσουμε και σε διδιάστατους πίνακες, όπου θεωρούμε ένα διάνυσμα σαν ένα πίνακα στήλης με στοιχεία τις γραμμές του αρχικού πίνακα. Έτσι, στο μαγικό μας τετράγωνο, το A(8) είναι άλλος ένας τρόπος με τον οποίο μπορούμε να αναφερθούμε στην τιμή 15 που είναι αποθηκευμένη στο A(4,2).

Εάν προσπαθήσετε να χρησιμοποιήσετε την τιμή ενός στοιχείου εκτός της μήτρας, αυτό είναι λάθος. Πράγματι το

```
t = A(4,5)
```

υπερβαίνει τις διαστάσεις του πίνακα.

Από την άλλη πλευρά, εάν αποθηκεύσετε μία τιμή σε ένα στοιχείο εκτός των διαστάσεων του πίνακα αυτός θα διευρυνθεί για να συμπεριλάβει το νέο στοιχείο.

```
X = A;
X(4,5) = 17

X =
   16    3    2   13    0
    5   10   11    8    0
    9    6    7   12    0
    4   15   14    1   17
```

Ο τελεστής άνω και κάτω τελεία

Η άνω και κάτω τελεία, : , είναι ένας από τους πιο σημαντικούς τελεστές στο MATLAB. Την συναντάμε σε πολλές διαφορετικές περιπτώσεις. Η έκφραση

```
1:10
```

είναι ένας πίνακας γραμμής που περιέχει τους ακέραιους από 1 ως 10

```
1    2    3    4    5    6    7    8    9    10
```

Μπορείτε επίσης να καθορίσετε το διάστημα στο οποίο θα διαφέρουν οι αριθμοί. Για παράδειγμα το

```
100:-7:50
```

δίνει

100 93 86 79 72 65 58 51

και το

0:pi/4:pi

μας δίνει

0 0.7854 1.5708 2.3562 3.1416

Οι εκφράσεις δεικτών που χρησιμοποιούν άνω και κάτω τελεία αναφέρονται σε τμήματα ενός πίνακα. Το

A(1 : k, j)

Είναι τα πρώτα k στοιχεία της στήλης j του πίνακα A. Έτσι το

sum(A(1:4,4))

υπολογίζει το άθροισμα της τέταρτης στήλης. Υπάρχει όμως και ένας καλύτερος τρόπος. Η άνω και κάτω τελεία μόνη της αναφέρεται στα στοιχεία μίας γραμμής ή στήλης ενός πίνακα και η λέξη-κλειδί end αναφέρεται στην τελευταία γραμμή ή στήλη. Έτσι το

sum(A(:,end))

υπολογίζει το άθροισμα των στοιχείων της τελευταίας στήλης του A.

```
ans =
    34
```

Γιατί το μαγικό άθροισμα ενός 4 επί 4 πίνακα είναι 34; Εάν χωρίσουμε τους ακέραιους από το 1 ως το 16 σε τέσσερις ομάδες με ίσο άθροισμα αυτό το άθροισμα θα είναι

sum(1:16)/4

το οποίο, φυσικά, είναι

```
ans =
    34
```

Η Συνάρτηση magic
Το MATLAB έχει μία ενσωματωμένη συνάρτηση η οποία δημιουργεί μαγικά τετράγωνα σχεδόν κάθε μεγέθους. Αυτή η συνάρτηση ονομάζεται – πώς αλλιώς – magic.

```
B = magic(4)
B =
   16    2    3   13
    5   11   10    8
    9    7    6   12
    4   14   15    1
```

Ο πίνακας αυτός είναι σχεδόν ο ίδιος με εκείνον στη ξυλογραφία του Dürer και έχει όλες τις «μαγικές» ιδιότητες· η μόνη τους διαφορά τους είναι ότι οι δύο μεσαίες στήλες είναι η μία στην θέση της άλλης. Για να μετατρέψουμε τον B στον πίνακα του Dürer A, ανταλλάσσουμε τις δύο μεσαίες στήλες.

```
A = B(:,[1 3 2 4])
```

Αυτό σημαίνει «για κάθε μία από τις γραμμές του πίνακα B , αναδιάταξε τα στοιχεία με την σειρά 1,3,2,4». Έτσι παίρνουμε

```
A =
   16    3    2   13
    5   10   11    8
    9    6    7   12
    4   15   14    1
```

Γιατί ο Dürer να έμπαινε στον κόπο να αναδιατάξει τις στήλες όταν θα μπορούσε να χρησιμοποιήσει τις υπηρεσίες του MATLAB; Χωρίς αμφιβολία ήθελε να συμπεριλάβει την ημερομηνία της ξυλογραφίας, 1514, στον πάτο του μαγικού τετραγώνου.

13.4.Εκφράσεις

Όπως και οι περισσότερες προγραμματιστικές γλώσσες, το MATLAB παρέχει μαθηματικές εκφράσεις, αλλά αντίθετα από τις γλώσσες αυτές, οι εκφράσεις περιλαμβάνουν ολόκληρους πίνακες. Τα δομικά στοιχεία των εκφράσεων είναι :

Μεταβλητές
Αριθμοί
Τελεστές
Συναρτήσεις

Μεταβλητές
Το MATLAB δεν χρησιμοποιεί δηλώσεις κανενός είδους ή δηλώσεις διαστάσεων. Όταν το MATLAB συναντάει ένα καινούργιο όνομα μεταβλητής, αυτόματα δημιουργεί την μεταβλητή και δεσμεύει τον απαραίτητο χώρο στην μνήμη. Εάν η μεταβλητή ήδη υπάρχει τότε αλλάζει τα περιεχόμενα της και, αν χρειαστεί, δεσμεύει νέο χώρο στην μνήμη. Για παράδειγμα το

num_students = 25

δημιουργεί μία 1-επί-1 μήτρα με το όνομα num_students και τοποθετεί την τιμή 25 στο μοναδικό στοιχείο της.

Τα ονόματα των μεταβλητών αποτελούνται από ένα γράμμα, ακολουθούμενο από οποιοδήποτε αριθμό γραμμάτων, ψηφίων και κάτω παυλών. Το MATLAB χρησιμοποιεί μόνο τους 31 πρώτους χαρακτήρες του ονόματος μίας μεταβλητής. Το MATLAB ξεχωρίζει τα μικρά γράμματα από τα κεφαλαία. Το A και το a δεν είναι το ίδιο. Για να δείτε τον πίνακα που έχει ανατεθεί σε μία μεταβλητή, απλά τυπώστε το όνομα της μεταβλητής.

Αριθμοί
Το MATLAB χρησιμοποιεί τον συμβατικό τρόπο δήλωσης δεκαδικών αριθμών, με προαιρετική μια τελεία για το δεκαδικό μέρος και το συν ή το πλην στην αρχή. Η επιστημονική δήλωση αριθμών γίνεται με το e να ακολουθείται από τον παράγοντα μίας δύναμης του 10. Οι φανταστικοί αριθμοί χρησιμοποιούν είτε το i είτε το j σαν πρόθεμα. Παραδείγματα σωστών αριθμών είναι

```
3           -99          0.0001
9.6397238    1.60210e-20  6.02252e23
1i          -3.14159j    3e5i
```

Όλοι οι αριθμοί αποθηκεύονται εσωτερικά χρησιμοποιώντας την μακριά μορφοποίηση όπως αυτή ορίζεται από το πρότυπο αριθμών μεταβλητής υποδιαστολής IEEE. Οι αριθμοί κινητής υποδιαστολής έχουν μία πεπερασμένη ακρίβεια περίπου 16 σημαντικών δεκαδικών ψηφίων και μία πεπερασμένη ακτίνα περίπου από το 10^{-308} ως το 10^{+308}. (Ο υπολογιστής VAX χρησιμοποιεί διαφορετική μορφοποίηση για του αριθμούς κινητής υποδιαστολής, αλλά η ακρίβεια και η ακτίνα είναι πάνω κάτω η ίδια.)

Load
Η εντολή load διαβάζει δυαδικά αρχεία που περιέχουν πίνακες δημιουργημένους προηγουμένως από το MATLAB ή διαβάζει αρχεία κειμένου τα οποία περιέχουν αριθμητικά δεδομένα. Το αρχείο κειμένου θα πρέπει να είναι οργανωμένο σαν ένας πίνακας αριθμών χωρισμένων με κενά, με μία γραμμή της μήτρας σε κάθε γραμμή κειμένου και ίσο αριθμό στοιχείων σε κάθε γραμμή. Για παράδειγμα,

δημιουργείστε, εκτός MATLAB, ένα αρχείο κειμένου που να περιέχει τις εξής τέσσερις σειρές :

```
16.0   3.0   2.0   13.0
 5.0  10.0  11.0    8.0
 9.0   6.0   7.0   12.0
 4.0  15.0  14.0    1.0
```

Αποθηκεύστε το αρχείο με το όνομα magik.dat. Τότε η εντολή

```
load magik.dat
```

διαβάζει το αρχείο και δημιουργεί μία μεταβλητή, την magik, η οποία περιέχει την παραπάνω μήτρα.

Μ-Αρχεία
Μπορείτε να δημιουργείτε δικούς σας πίνακες χρησιμοποιώντας Μ-αρχεία, τα οποία είναι αρχεία κειμένου που περιέχουν κώδικα σε MATLAB. Απλά δημιουργείστε ένα αρχείο κειμένου το οποίο θα περιέχει τις ίδιες εντολές που θα χρησιμοποιούσατε στην γραμμή εντολής του MATLAB. Αποθηκεύστε το αρχείο με όνομα που τελειώνει σε .m.

Για παράδειγμα, δημιουργείστε ένα αρχείο που να περιέχει τις εξής πέντε γραμμές

```
A = [ ...
16.0   3.0   2.0   13.0
 5.0  10.0  11.0    8.0
 9.0   6.0   7.0   12.0
 4.0  15.0  14.0    1.0 ];
```

αποθηκεύστε το αρχείο με το όνομα magik.m. Τότε η εντολή

```
magik
```

διαβάζει το αρχείο και δημιουργεί μια μεταβλητή, A, η οποία περιέχει το παράδειγμα μας.

Συνένωση

Συνένωση είναι η διαδικασία με την οποία ενώνουμε μικρούς πίνακες για να δημιουργήσουμε μεγαλύτερους. Στην πραγματικότητα δημιουργήσατε την πρώτη σας μήτρα συνενώνοντας ξεχωριστά της στοιχεία. Το ζευγάρι αγκυλών, [ ], είναι ο τελεστής της συνένωσης. Για παράδειγμα ξεκινήστε με το 4-επί-4 μαγικό τετράγωνο και πληκτρολογήστε

```
B = [A  A+32; A+48  A+16]
```

Αυτό θα δώσει ένα 8-επί-8 πίνακα, ο οποίος προήλθε από τους τέσσερεις πίνακες

```
B =

   16    3    2   13   48   35   34   45
    5   10   11    8   37   42   43   40
    9    6    7   12   41   38   39   44
    4   15   14    1   36   47   46   33
   64   51   50   61   32   19   18   29
   53   58   59   56   21   26   27   24
   57   54   55   60   25   22   23   28
   52   63   62   49   20   31   30   17
```

Ο πίνακας αυτός είναι σχεδόν ένα μαγικό τετράγωνο. Τα στοιχεία είναι μία ανακατάταξη των ακεραίων 1:64. Το άθροισμα κάθε στήλης έχει την σωστή τιμή για ένα 8-επί-8 μαγικό τετράγωνο.

Εχουμε:

```
sum(B)

ans =
      260   260   260   260   260   260   260   260
```

αλλά τα αθροίσματα των σειρών του, sum(B΄)΄, δεν είναι όλα τα ίδια. Χρειάζεται περισσότερη επεξεργασία για να κάνουμε τον πίνακα ένα έγκυρο 8-επί-8 μαγικό τετράγωνο.

Διαγραφή Γραμμών και Στηλών

Μπορείτε διαγράψετε γραμμές και στήλες από ένα πίνακα χρησιμοποιώντας απλώς ένα ζευγάρι αγκύλες. Ξεκινήστε με το

```
X = A;
```

Στην συνέχεια, για να διαγράψετε την δεύτερη στήλη του X, χρησιμοποίησε

```
X(:,2) = [ ]
```

Αυτό μετατρέπει το X σε

```
X =
    16     2    13
     5    11     8
     9     7    12
     4    14     1
```

Εάν διαγράψετε ένα μόνο στοιχείο από μία μήτρα, το αποτέλεσμα δεν είναι μήτρα πια. Έτσι εκφράσεις όπως

```
X(1,2) = []
```

θα οδηγήσουν σε ένα λάθος. Πάντως, χρησιμοποιώντας ένα δείκτη, διαγράφουμε ένα στοιχείο ή μια ακολουθία στοιχείων και διαμορφώνουμε τα υπόλοιπα στοιχεία σε ένα διάνυσμα γραμμής. Έτσι το

```
X(2:2:10) = []
```

έχει ως αποτέλεσμα

```
X =
  16    9    2    7   13   12    1
```

Αριθμητικοί Τελεστές + - * / \ ^ `

\+ *Πρόσθεση:* A + B προσθέτει τους πίνακες A και B. Οι πίνακες A και B πρέπει να είναι των ιδίων διαστάσεων, εκτός αν ο ένας είναι βαθμωτό μέγεθος. Ενα βαθμωτό μέγεθος μπορεί να προστεθεί σε πίνακα οποιωνδήποτε διαστάσεων.

\- *Αφαίρεση:* A – B αφαιρεί τον πίνακα B από τον πίνακα A. Οι πίνακες A και B πρέπει να είναι των ιδίων διαστάσεων, εκτός αν ο ένας είναι βαθμωτό μέγεθος. Ενα βαθμωτό μέγεθος μπορεί να αφαιρεθεί από πίνακα οποιωνδήποτε διαστάσεων.

* *Πολλαπλασιασμός:* A * B είναι το γραμμικό αλγεβρικό γινόμενο των πινάκων A και B. Ο αριθμός των στηλών του A πρέπει να ισούται με τον αριθμό των γραμμών του B, εκτός αν ο ένας είναι βαθμωτό μέγεθος. Ενα βαθμωτό μέγεθος μπορεί να πολλαπλασιάσει πίνακα οποιωνδήποτε διαστάσεων.

.* *Διατεταγμένος πολλαπλασιασμός:* A .* B είναι το στοιχείο προς στοιχείο γινόμενο των πινάκων A και B. Οι πίνακες A και B πρέπει να είναι των ιδίων διαστάσεων, εκτός αν ο ένας είναι βαθμωτό μέγεθος.

\ *Αριστερή διαίρεση πινάκων:* Αν A είναι τετραγωνικός πίνακας, A\B είναι σε γενικές γραμμές το ίδιο με inv(A)*B. Αν A είναι ένας m x n πίνακας και B μια στήλη, τότε A\B είναι η λύση ελαχίστων τετραγώνων του συστήματος εξισώσεων AX = B.

.\ *Διατεταγμένη αριστερή διαίρεση πινάκων:* A .\ B είναι ο πίνακας με στοιχεία B(i,j)/A(i,j). Οι πίνακες A και B πρέπει να είναι των ιδίων διαστάσεων, εκτός αν ο ένας είναι βαθμωτό μέγεθος.

/ *Δεξιά διαίρεση πινάκων:* B/A = (A`\B`).

./ *Διατεταγμένη δεξιά διαίρεση πινάκων:* A ./ B είναι ο πίνακας με στοιχεία A(i,j)/B(i,j). Οι πίνακες A και B πρέπει να είναι των ιδίων διαστάσεων, εκτός αν ο ένας είναι βαθμωτό μέγεθος.

^ *Δύναμη:* X^p είναι ο πίνακας X στην p-οστή δύναμη, αν p είναι βαθμωτό μέγεθος. Στη γενική περίπτωση, χρησιμοποιούνται ιδιοτιμές και ιδιοδιανύσματα για τον υπολογισμό. Αν x είναι βαθμωτό μέγεθος και P ένας πίνακας, το x^P υπολογίζεται με χρήση ιδιοτιμών και ιδιοδιανυσμάτων. Η έκφραση X^P, όπου και X και P είναι πίνακες, είναι λάθος.

.^ *Διατεταγμένη δύναμη:* A.^B είναι ο πίνακας με στοιχεία A(i,j) στη δύναμη B(i,j). Οι πίνακες A και B πρέπει να είναι των ιδίων διαστάσεων, εκτός αν ο ένας είναι βαθμωτό μέγεθος.

` *Ανάστροφος πίνακα:* A΄ είναι ο αλγεβρικός ανάστρ οφος του A. Για πίνακες με μιγαδικά στοιχεία υπολογίζεται ο μιγαδικός συζυγής.

.` *Διατεταγμένος ανάστροφος:* A.΄ είναι ο αλγεβρικός ανάστροφος του A. Για πίνακες με μιγαδικά στοιχεία δεν υπολογίζεται ο μιγαδικός συζυγής.

Τελεστές Συσχέτισης

<   >   <=   >=   = =   ~=

Λογικοί Τελεστές

& | ~

Ειδικοί Χαρακτήρες

[   ]   (   )   =   `   .   ;   %   !

Το Παράθυρο Εντολών

Ως τώρα, έχετε χρησιμοποιήσει την γραμμή εντολών του MATLAB, γράφοντας εντολές και εκφράσεις, και βλέποντας τα αποτελέσματα να εμφανίζονται στο παράθυρο εντολών. Αυτός ο τομέας περιγράφει μερικούς τρόπους να αλλάξουμε την εμφάνιση του παράθυρου εντολών. Εάν το σύστημα σας σας επιτρέπει να επιλέξετε την γραμματοσειρά του παράθυρου εντολών σας συμβουλεύουμε να χρησιμοποιήσετε μία σταθερού πλάτους, όπως η Fixedsys ή Courier, ώστε να υπάρχουν κατάλληλα διαστήματα.

Η Εντολή format

Η εντολή format καθορίζει τον τρόπο με το οποίο εμφανίζονται οι αριθμητικές τιμές στο MATLAB. Η εντολή επηρεάζει μόνο τον τρόπο με τον οποίο εμφανίζονται οι αριθμοί και όχι τον τρόπο με τον οποίο το MATLAB κάνει τους υπολογισμούς ή τους αποθηκεύει. Εδώ παρατίθενται οι διάφορες μορφοποιήσεις καθώς και το αποτέλεσμα που παράγεται από ένα πίνακα x.

x = [4/3 1.2345e-6]

format short

1.3333 0.0000

format short e

1.3333e+000 1.2345e-006

format short g

1.3333 1.2345e-006

format long

1.33333333333333 0.00000123450000

format long e

1.333333333333333e+000 1.234500000000000e-006

format long g

1.33333333333333 1.2345e-006

format bank

1.33 0.00

format rat

4/3   1/810045

format hex

3ff5555555555555 3eb4b6231abfd271

Εάν το μεγαλύτερο στοιχείο μίας μήτρας είναι μεγαλύτερο από 10^3 ή μικρότερο από 10^{-3}, το MATLAB χρησιμοποιεί κοινή κλίμακα για τις μορφοποιήσεις short και long.

Μαζί με τις εντολές format που είδατε παραπάνω η

format compact

δεν εμφανίζει πολλές από τις κενές γραμμές που εμφανίζονται. Αυτό σας επιτρέπει να βλέπεται περισσότερες πληροφορίες σε μία οθόνη ή ένα παράθυρο. Εάν θέλετε πλήρη έλεγχο στην μορφή της εξόδου, χρησιμοποιήστε τις συναρτήσεις sprintf και fprintf.

Καταπίεση της Εξόδου
Εάν απλώς τυπώσετε μία εντολή ή έκφραση και πατήσετε το Enter, το MATLAB αυτόματα παρουσιάζει το αποτελέσματα στην οθόνη. Εάν όμως τελειώσετε την γραμμή με ένα ελληνικό ερωτηματικό, το MATLAB κάνει τους υπολογισμούς αλλά δεν παρουσιάζει το αποτέλεσμα. Αυτό είναι ιδιαίτερα χρήσιμο όταν δημιουργείτε μεγάλες πίνακες, π.χ.

A = magic(100)

Μακροσκελείς Εκφράσεις
Εάν μία έκφραση δεν χωράει σε μία γραμμή, χρησιμοποιήστε τρεις τελείες, …, ακολουθούμενες από το Return ή το Enter για να δείξετε ότι η έκφραση συνεχίζεται στην επόμενη γραμμή. Για παράδειγμα

```
s = 1 -1/2 + 1/3 -1/4 + 1/5 - 1/6 + 1/7 ...
     - 1/8 + 1/9 - 1/10 + 1/11 - 1/12;
```

Τα κενά γύρω από τα =, +, - είναι προαιρετικά, αλλά κάνουν πιο ευανάγνωστο το κείμενο.

Δουλεύοντας στην Γραμμή Εντολών
Τα διάφορα βέλη και πλήκτρα ελέγχου στο πληκτρολόγιο σας επιτρέπουν να ανακαλέσετε, να επεξεργαστείτε και να ξαναχρησιμοποιήσετε εντολές που τυπώσατε προηγουμένως. Π.χ. ας πούμε ότι γράψατε κατά λάθος την εντολή

rho = (1 + sqt(5))/2.

Έχετε γράψει λάθος την εντολή sqrt. Το MATLAB απαντάει με:

Undefined function or variable ‘sqt’, δηλαδή, μη καθορισμένη συνάρτηση ή μεταβλητή ‘sqt’

Αντί να ξαναγράψετε ολόκληρη την εντολή, απλώς πατήστε το πάνω βέλος. Η λανθασμένη εντολή θα εμφανιστεί. Χρησιμοποιήστε το αριστερό βέλος για να μετακινήσετε τον κέρσορα και συμπληρώστε το r. Το πάνω βέλος καλεί προηγούμενες εκφράσεις. Αν τυπώσετε μερικούς χαρακτήρες τότε το πάνω βέλος βρίσκει τις εντολές που ξεκινούν με αυτούς τους χαρακτήρες.

Η λίστα των διαθέσιμων πλήκτρων διαφέρει σε διαφορετικούς υπολογιστές. Πειραματιστείτε για να δείτε ποια από τα παρακάτω πλήκτρα είναι διαθέσιμα στον υπολογιστή σας.

Πάνω βέλος	Ctrl + p	Ανάκληση προηγούμενης εντολής
Κάτω βέλος	Ctrl + n	Ανάκληση επόμενης εντολής
Αριστερό βέλος	Ctrl + b	Μπροστά ένας χαρακτήρας
Δεξιό βέλος	Ctrl + f	Πίσω ένας χαραχτήρας
Ctrl + αριστερό βέλος	Ctrl + r	Μπροστά μία λέξη
Ctrl + δεξιό βέλος	Ctrl + l	Πίσω μία λέξη
Home	Ctrl + a	Κίνηση στην αρχή της γραμμής
End	Ctrl + e	Κίνηση στο τέλος της γραμμής
Esc	Ctrl + u	Καθαρίζει την γραμμή
Del	Ctrl + d	Διαγραφή του χαρακτήρα στον κέρσορα
Backspace	Ctrl + h	Διαγραφή του χαρακτήρα πριν από τον κέρσορα
	Ctrl + k	Διαγραφή ως το τέλος της γραμμής

13.5.Γραφικά

Η MATLAB έχει εκτενείς δυνατότητες για την αναπαράσταση διανυσμάτων και πινάκων σε μορφή γραφικών παραστάσεων, όπως επίσης και για την προσθήκη υποσημειώσεων σε αυτά καθώς και για την εκτύπωση τους. Αυτό το κεφαλαίο περιγράφει μερικές από τις πιο σημαντικές συναρτήσεις γραφικών και παραθέτει παραδείγματα κάποιων τυπικών εφαρμογών.

Κατασκευή ενός Σχεδιαγράμματος

Η συνάρτηση plot εμφανίζεται με διαφορετικές μορφές ανάλογα με τις εισαγόμενες παραμέτρους. Εάν υ είναι ένα διάνυσμα τότε η plot(y) παράγει ένα τμηματικά γραμμικό γράφημα των στοιχείων του y ως προς τον δείκτη του y. Εάν δώσετε δυο διανύσματα ως παραμέτρους, η plot(x,y) παράγει ένα γράφημα του y σε συνάρτηση με το x.

Για παράδειγμα, για να σχεδιάσετε τις τιμές της συνάρτησης sin από το 0 έως το 2π χρησιμοποιείστε το ακόλουθο:

```
t = 0:pi/100:2*pi;
y = sin(t);
plot(t,y)
y2 = sin(t-.25);
y3 = sin(t-.5);
plot(t,y,t,y2,t,y3)
```

Είναι δυνατό να προσδιορίσουμε το χρώμα, το στυλ της γραμμής και σύμβολα σήμανσης όπως κύκλους με την εντολή

```
plot(x , y , ’color_style_marker’)
```

Το color_style_marker είναι μία συμβολοσειρά από 1,2 η τρεις χαρακτήρες (το οποίο παρατίθεται με απλά εισαγωγικά) που κατασκευάζετε από ένα χρώμα, ένα στυλ γραμμής και ένα τύπο σημείου σήμανσης

Οι συμβολοσειρές χρώματος είναι ‘c’ , ’m’ ,’ y’ , ’r’, ’ g’ , ’b’ , ’w’ καί ’k’. Αυτές αντιστοιχούν στα χρώματα κυανό, κίτρινο, κόκκινο, πράσινο, άσπρο και μαύρο.

Οι συμβολοσειρές του στυλ γραμμής είναι ’-’ για μια συμπαγή γραμμή , ’--’ για διακεκομμένη γραμμή , ’:’ για γραμμή από τέλειες , ‘-.’ για γραμμή από παύλα-τέλεια και ‘none’ για την μη ύπαρξη γραμμής

Οι πιο συχνοί τύποι σημείων σήμανσης περιλαμβάνουν τα ‘+’ , ’o’ , ’*’ , ’x’
Για παράδειγμα η δήλωση

```
plot( x, y, ’y:+’ )
```

σχεδιάζει μια κιτρίνη γραμμή από τέλειες και τοποθετεί ένα σημείο συν σε κάθε σημείο δεδομένων. Εάν δώσετε το σημείο σήμανσης αλλά όχι το στυλ γραμμής το MATLAB σχεδιάζει μόνο το σημείο σήμανσης.

Παράθυρα Σχεδίασης

Η συνάρτηση plot αυτόματα ανοίγει ένα νέο παράθυρο σχεδίασης εάν δεν υπάρχουν άλλα παράθυρα σχεδίασης ήδη ανοιχτά στην οθόνη. Εάν υπάρχει ένα παράθυρο σχεδίασης, η plot χρησιμοποιεί το παράθυρο αυτό εάν δεν της δοθεί εντολή να χρησιμοποιήσει κάποιο συγκεκριμένο παράθυρο. Για να ανοίξετε ένα νέο παράθυρο σχεδίασης και να το κάνετε το ενεργό σχεδιάγραμμα, πληκτρολογήστε

```
figure
```

Για να κάνετε ένα ήδη υπάρχον παράθυρο το ενεργό παράθυρο πληκτρολογείστε

```
figure(n)
```

όπου (n) είναι ο αριθμός στην ραβδο τίτλου του σχεδιαγράμματος. Τα αποτελέσματα μετέπειτα εντολών γραφικών εμφανίζονται σε αυτό το παράθυρο.

Προσθήκη Σχεδίων σε Υπαρκτό Γράφημα

Η εντολή hold επιτρέπει την προσθήκη σχεδίων σε ένα υπαρκτό γράφημα. Όταν πληκτρολογήσετε

```
hold on
```

δεν αφαιρεί το ήδη υπάρχον γράφημα, αλλά προσθέτει τα νέα δεδομένα στο παρόν γράφημα ανακαλώντας το εάν χρειάζεται. Παραδείγματος χάριν οι ακόλουθες δηλώσεις πρώτα δημιουργούν ένα σχεδιάγραμμα περιγράμματος της συνάρτησης peaks και ύστερα τοποθετεί ένα σχεδιάγραμμα της ίδιας συνάρτησης σε ‘ ψευδοχρώμα ’

```
[x,y,z] = peaks;
contour(x,y,z,20,'k')
hold on
pcolor(x,y,z)
shading interp
```

Η εντολή hold on συνδυάζει το σχεδιάγραμμα pcolor με το σχεδιάγραμμα contour σε μια παράσταση.

Τμηματικά Σχεδιαγράμματα

Η συνάρτηση subplot επιτρέπει την εμφάνιση πολλαπλών σχεδίων στο ίδιο παράθυρο ή την εκτύπωση τους στο ίδιο φύλλο χαρτιού. Πληκτρολογώντας

```
subplot( m,n,p )
```

χωρίζει το παράθυρο σχεδίασης σε έναν πίνακα m επί n ο οποίος αποτελείται από μικρά τμηματικά σχεδιαγράμματα και επιλέγει το p-οστό ως το ενεργό σχεδιάγραμμα. Τα σχεδιαγράμματα είναι αριθμημένα αρχικά κατά μήκος της πιο πάνω σειράς του παραθύρου σχεδίασης, κατόπιν κατά μήκος της δεύτερης , κ.ο.κ. Παραδείγματος χάριν για να σχεδιάσετε δεδομένα σε τέσσερις επιμέρους περιοχές του παραθύρου σχεδίασης πληκτρολογήστε:

```
t = 0:pi/10:2*pi;
[X,Y,Z] = cylinder(4*cos(t));
subplot(2,2,1)
mesh(X)
subplot(2,2,2); mesh(Y)
subplot(2,2,3); mesh(Z)
subplot(2,2,4); mesh(X,Y,Z)
```

Φανταστικά και Μιγαδικά Δεδομένα

Όταν οι τιμές των μεταβλητών της plot είναι μιγαδικές, το φανταστικό μέρος αγνοείτε εκτός εάν δοθεί στην plot μία μόνο μιγαδική τιμή. Σε αυτήν την ειδική περίπτωση , η εντολή είναι μια συντόμευση για ένα σχεδιάγραμμα του πραγματικού ως προς το φανταστικό μέρος.

Συνεπώς η εντολή :

```
plot(z),
```

όπου z είναι ένα μιγαδικό διάνυσμα η μήτρα, είναι ισοδύναμο με την :

```
plot (real(z),imag(z) )
```

Παραδείγματος χάριν το

```
t = 0:pi/10:2*pi;
plot(exp(I*t),'-o')
```

σχεδιάζει ένα εικοσάπλευρο πολύγονο με μικρούς κύκλους στις κορυφές.

Έλεγχος των Αξόνων

Η συνάρτηση axis έχει έναν αριθμό από επιλογές για την προσαρμογή της διαβάθμισης, της κατεύθυνσης και του λόγου μεταξύ των κλιμάκων των αξόνων των σχεδιαγραμμάτων.

Κανονικά, το MATLAB βρίσκει το μέγιστο και ελάχιστο από τα δεδομένα και επιλέγει το κατάλληλο κουτί σχεδίασης και της κατάλληλες επιγραφές αξόνων. Η συνάρτηση axis υπερπηδά την εξ ορισμού ισχύουσα κατάσταση με το να θέτει επιλεγμένα από τον χρήστη όρια αξόνων:

axis([xmin xmax ymin ymax])

Η axis επίσης δέχεται έναν αριθμό από λέξεις κλειδιά για τον έλεγχο των αξόνων. Για παράδειγμα, η εντολή

axis square

κάνει ολόκληρους τους άξονες x και y ισομήκεις και η εντολή

axis equal

κάνει τις αποστάσεις των διαβαθμίσεων των x και y επί των αξόνων ίσες. Έτσι η

plot(exp(i*t)),

ακολουθούμενη είτε από την εντολή axis square είτε από την axis equal, μετατρέπει το οβάλ σε κανονικό κύκλο. Η εντολή

axis auto

επαναφέρει τις διαβαθμίσεις στις εξ ορισμού ισχύουσες τιμές τους. Η εντολή

axis on

ενεργοποιεί της επιγραφές των αξόνων και τις γραμμές διαβαθμίσεων, ενώ η

axis off

απενεργοποιεί τις επιγραφές των αξόνων και τις γραμμές διαβαθμίσεων.

Η εντολή

grid off

απενεργοποιεί τις γραμμές πλέγματος και η

grid on

τις επανενεργοποιεί .

Επιγραφές και Τίτλοι Αξόνων
Οι συναρτήσεις xlabel, ylabel και zlabel προσθέτουν επιγραφές στους άξονες x, y και z. Η συνάρτηση title προσθέτει έναν τίτλο στην κορυφή του σχεδίου και η

συνάρτηση text εισάγει κείμενο οπουδήποτε στο σχεδιάγραμμα. Ένα υποσύνολο του συμβολισμού του TeX παράγει ελληνικά γράμματα, μαθηματικά σύμβολα και εναλλακτικές γραμματοσειρές. Το ακόλουθο παράδειγμα χρησιμοποιεί \leq για ≤ , \pi για π και \it για πλάγιου τύπου γραμματοσειρά:

```
t = -pi:pi/100:pi;
y = sin(t);
plot( t , y );
axis([-pi pi –1 1 ])
xlabel('-\pi \leq {\itt} \leq \pi')
ylabel('sin(t)')
title('Graph of the sine function');
text(1,-1/3,'\it{Note the odd symmetry}')
```