

ΥΠΕΡΘΕΣΗ ΕΙΚΟΝΩΝ ΜΕ ΜΙΚΤΕΣ ΚΑΤΑΝΟΜΕΣ

Η ΜΕΤΑΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ ΕΞΕΙΔΙΚΕΥΣΗΣ

υποβάλλεται στην

ορισθείσα από την Γενική Συνέλευση Ειδικής Σύνθεσης

του Τμήματος Πληροφορικής Εξεταστική Επιτροπή

από τον

Δημήτριο Γερογιάννη

ως μέρος των Υποχρεώσεων για τη λήψη του

ΜΕΤΑΠΤΥΧΙΑΚΟΥ ΔΙΠΛΩΜΑΤΟΣ ΣΤΗΝ ΠΛΗΡΟΦΟΡΙΚΗ

ΜΕ ΕΞΕΙΔΙΚΕΥΣΗ

ΣΤΙΣ ΤΕΧΝΟΛΟΓΙΕΣ & ΕΦΑΡΜΟΓΕΣ

Σεπτέμβριος 2007

ΑΦΙΕΡΩΣΗ

Στους αγαπημένους μου γονείς για τις θυσίες που έκαναν για μένα. Ελπίζω να εκπλήρωσα τα όνειρά τους.

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω δημόσια τον επιβλέποντα την εργασία μου κ. Νίκου Χριστόφορο, Λέκτορα του τμ. Πληροφορικής, για τη συνεργασία που είχαμε κατά την εκπόνηση της. Σημαντική ήταν και η συνεισφορά του κ. Λύκα Αριστείδη, Αναπληρωτή Καθηγητή του τμ. Πληροφορικής. Η βοήθειά τους ήταν πολύτιμη και ουσιαστική για την ολοκλήρωση της μεταπτυχιακής μου εργασίας. Τέλος να ευχαριστήσω τους φίλους μου Αλέξανδρο και Γιώργο από το εργαστήριο Μοριακής Βιολογίας της Ιατρικής σχολής του Πανεπιστημίου Ιωαννίνων για τις βιολογικές εικόνες που μου προσέφεραν και τις οποίες χρησιμοποίησα στα πειράματά μου.

ΠΕΡΙΕΧΟΜΕΝΑ

1	Εισαγωγή στην Υπέρθυση Εικόνας	1
1.1	Εισαγωγή	1
1.2	Ορισμός του προβλήματος	4
1.3	Μετασχηματισμοί υπέρθεσης	5
1.4	Μέθοδοι υπέρθεσης εικόνας	6
1.4.1	Μέθοδοι που βασίζονται στην συσχέτιση	6
1.4.2	Μέθοδοι που βασίζονται στο μετασχηματισμό Fourier	7
1.4.3	Μέθοδοι που βασίζονται στην αντιστοίχιση σημείων	7
1.4.4	Μέθοδοι που βασίζονται στην αμοιβαία πληροφορία	8
2	Μικτές Κατανομές	10
2.1	Αρχές και εκπαίδευση Μικτών Κατανομών - Ο αλγόριθμος Expectation Maximization (EM)	10
2.2	Μικτές Κανονικές Κατανομές	13
2.3	Μικτές Student's t - Κατανομές	16
2.4	Αποστάσεις Μεταξύ Κατανομών	18
2.4.1	Απόσταση Chernoff	18
2.4.2	Απόσταση Battacharyya	19
2.4.3	Τετραγωνική L_2 απόσταση	19
3	Υπέρθυση Εικόνας με Μικτές Κατανομές	24
3.1	Περιγραφή της μεθόδου	24
3.2	Πειραματικά αποτελέσματα	25
3.3	Ανάλυση αποτελεσμάτων	28
4	Υπέρθυση νεφών σημείων	40
4.1	Εισαγωγή στην υπέρθεση σημείων - Ο αλγόριθμος Iterative Closest Point - ICP	40
4.2	Περιγραφή της μεθόδου	42
4.3	Πειραματικά αποτελέσματα	43
4.4	Ανάλυση αποτελεσμάτων	44
5	Επίλογος	47

ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ

2.1	Γραφικό μοντέλο αναπαράστασης Μικτών Κατανομών.	11
2.2	Ο τύπος της συνάρτησης πυκνότητας πιθανότητας της κανονικής κατανομής όπως απλοποιείται για τις διάφορες μορφές του πίνακα συμμεταβλητότητας Σ	15
2.3	Γραφική παράσταση συνάρτησης πυκνότητας πιθανότητας Κανονικής Κατανομής. Παρατηρείστε το κωδωνοειδές σχήμα της. Από τη θεωρία ισχύει ότι στο διάστημα $[\mu - 2\sigma, \mu + 2\sigma]$ ανήκει το 98% των δεδομένων. Επιπλέον όσο μεγαλύτερη είναι η διασπορά τόσο πιο πολύ απλώνονται τα άκρα της κατανομής.	15
2.4	Παραδείγματα κατηγοριοποίησης δεδομένων με χρήση Μοντέλου Μικτών Κανονικών Κατανομών. (α) Σφαιρικός πίνακας συμμεταβλητότητας, (β) Διαγώνιος πίνακας συμμεταβλητότητας, (γ) Πλήρης πίνακας συμμεταβλητότητας.	16
2.5	Κατηγοριοποίηση σημείων με Μικτές Κανονικές Κατανομές, (α) σημεία χωρίς θόρυβο, (β) τα σημεία του σχήματος (α) με προσθήκη θορύβου. Παρατηρήστε την έντονη επιρροή των αποκλινόντων σημείων που μεταφράζεται σε μετατόπιση των κέντρων, αλλά και σε διεύρυνση της έλλειψης ώστε να καλύψει και τα σημεία θορύβου.	20
2.6	Κατηγοριοποίηση σημείων με μικτές Student's t -Κατανομές, (α) σημεία χωρίς θόρυβο, (β) τα σημεία του σχήματος (α) με προσθήκη θορύβου. Σε αντίθεση με τις Κανονικές Κατανομές η Student's t -Κατανομή είναι πιο ανθεκτική. Το αποτέλεσμα είναι ένα πιο αξιόπιστο μοντέλο.	21
2.7	Γραφική παράσταση της συνάρτησης πυκνότητας πιθανότητας Κανονικής (κυανό) και Student's t -Κατανομής (κόκκινο).	22
2.8	Γραφική παράσταση της συνάρτησης πυκνότητας πιθανότητας της Student's t -Κατανομής για διάφορες τιμές της παραμέτρου ν . Παρατηρήστε ότι καθώς το ν μεγαλώνει η καμπύλη προσεγγίζει αυτή της Κανονικής Κατανομής.	22
3.1	Διάγραμμα ροής αλγορίθμου υπέρθεσης εικόνων με χρήση μικτών κατανομών.	25
3.2	Διαγράμματα της αντικειμενικής συνάρτησης που ελαχιστοποιείται κατά τη διαδικασία της υπέρθεσης. Το διάγραμμα αναφέρεται στις εικόνες (α),(β) του σχήματος 3.3. (α),(β) Αμοιβαία πληροφορία. (γ),(δ) απόσταση Battacharyya.	28
3.3	(α) Αρχική Εικόνα Υπέρθεσης. Εικόνα 3 περιοχών σταθερής φωτεινότητας κατά τμήματα. (β) Τελική Εικόνα Υπέρθεσης. Εικόνα 3 περιοχών σταθερής φωτεινότητας κατά τμήματα. (γ) Διάγραμμα σφάλματος υπέρθεσης (Γκαουσιανός θόρυβος). (δ) Διάγραμμα σφάλματος υπέρθεσης (Ομοιόμορφος θόρυβος). (ε) Στατιστικά αποτελέσματα πειράματος. Οι τιμές είναι ομαδοποιημένες με βάση τον τύπο του θορύβου (Γκαουσιανός, Ομοιόμορφος).	31

3.4	(α) Αρχική Εικόνα Υπέρθεσης. Τεχνητή εικόνα 8bit χρώματος. (β) Τελική Εικόνα Υπέρθεσης. Τεχνητή εικόνα 8bit χρώματος. (γ) Διάγραμμα σφάλματος υπέρθεσης (Γκαουσιανός θόρυβος). (δ) Διάγραμμα σφάλματος υπέρθεσης (Ομοιόμορφος θόρυβος). (ε) Διάγραμμα σφάλματος υπέρθεσης (Αποκλίνοντα σημεία). (στ) Στατιστικά αποτελέσματα πειράματος. Οι τιμές είναι ομαδοποιημένες με βάση τον τύπο του θορύβου (Γκαουσιανός, Ομοιόμορφος, Αποκλίνοντα σημεία).	33
3.5	(α) Αρχική Εικόνα Υπέρθεσης. Μαγνητική αξονική τομογραφία εγκεφάλου. (β) Τελική Εικόνα Υπέρθεσης. Μαγνητική αξονική τομογραφία εγκεφάλου. (γ) Διάγραμμα σφάλματος υπέρθεσης (Γκαουσιανός θόρυβος). (δ) Διάγραμμα σφάλματος υπέρθεσης (Ομοιόμορφος θόρυβος). (ε) Στατιστικά αποτελέσματα πειράματος. Οι τιμές είναι ομαδοποιημένες με βάση τον τύπο του θορύβου (Γκαουσιανός, Ομοιόμορφος).	35
3.6	(α) Αρχική Εικόνα Υπέρθεσης. Βλαστοκύτταρα ποντικού σε ορατό φώς. (β) Τελική Εικόνα Υπέρθεσης. Βλαστοκύτταρα ποντικού σε υπέρυθρο φώς. (γ) Διάγραμμα σφάλματος υπέρθεσης (Γκαουσιανός θόρυβος). (δ) Διάγραμμα σφάλματος υπέρθεσης (Ομοιόμορφος θόρυβος). (ε) Στατιστικά αποτελέσματα πειράματος. Οι τιμές είναι ομαδοποιημένες με βάση τον τύπο του θορύβου (Γκαουσιανός, Ομοιόμορφος).	37
3.7	(α) Αρχική Εικόνα Υπέρθεσης. Δορυφορική εικόνα της Ευρώπης στις 8/1/2007 (01h00). (β) Τελική Εικόνα Υπέρθεσης. Δορυφορική εικόνα της Ευρώπης στις 9/1/2007 (01h00). Οι εικόνες προέρχονται από το δορυφόρο Meteosat. (γ) Διάγραμμα σφάλματος υπέρθεσης (Γκαουσιανός θόρυβος). (δ) Διάγραμμα σφάλματος υπέρθεσης (Ομοιόμορφος θόρυβος). (ε) Διάγραμμα σφάλματος υπέρθεσης (Αποκλίνοντα σημεία). (στ) Στατιστικά αποτελέσματα πειράματος. Οι τιμές είναι ομαδοποιημένες με βάση τον τύπο του θορύβου (Γκαουσιανός, Ομοιόμορφος, Αποκλίνοντα σημεία).	39
4.1	Δεδομένα που χρησιμοποιήσαμε στα πειράματα ελέγχου της μεθόδου μας σε νέφη σημείων. (α) Αρχικό σύνολο σημείων (μοντέλο - σημεία που παράχθησαν από ομοιόμορφη κατανομή). (β) Ένα στιγμιότυπο του τελικού συνόλου σημείων υπέρθεσης με προσθήκη 2% αποκλινόντων σημείων.	44
4.2	Πειραματικά αποτελέσματα για τα νέφη σημείων που περιγράφονται στο σχήμα 4.1 (α) Διάγραμμα σφάλματος υπέρθεσης (Γκαουσιανή κατανομή + αποκλίνοντα σημεία). Σύγκριση Μικτών κατανομών χωρίς αντιστοιχίες. (β) Διάγραμμα σφάλματος υπέρθεσης (Γκαουσιανή κατανομή + αποκλίνοντα σημεία). Σύγκριση Μικτών κατανομών με αντιστοιχίες (Hungarian αλγόριθμος).	45
4.3	Πειραματικά αποτελέσματα για τα νέφη σημείων που περιγράφονται στο σχήμα 4.1 (α) Διάγραμμα σφάλματος υπέρθεσης (Ομοιόμορφη κατανομή + αποκλίνοντα σημεία). Σύγκριση Μικτών κατανομών χωρίς αντιστοιχίες. (β) Διάγραμμα σφάλματος υπέρθεσης (Ομοιόμορφη κατανομή + αποκλίνοντα σημεία). Σύγκριση Μικτών κατανομών με αντιστοιχίες (Hungarian αλγόριθμος).	46

ΕΥΡΕΤΗΡΙΟ ΑΛΓΟΡΙΘΜΩΝ

1	Σύντομη απόδοση του αλγορίθμου EM	14
2	Αλγόριθμος EM για εκπαίδευση Μικτών Κανονικών Κατανομών.	17
3	Αλγόριθμος EM για εκπαίδευση Μικτών Student's t -Κατανομών.	23
4	Αλγόριθμος υπέρθεσης εικόνων με Μικτές Κατανομές	24
5	Ο αλγόριθμος ICM.	26
6	Αλγόριθμος υπέρθεσης εικόνων βασισμένος σε Μοντέλα Μικτών Κατανομών.	26
7	Ο αλγόριθμος ICP.	42
8	Μέθοδος για υπέρθεση νεφών σημείων.	43
9	Ο Hungarian αλγόριθμος.	52

ΕΠΕΞΗΓΗΣΕΙΣ ΣΥΜΒΟΛΙΣΜΩΝ

Συμβάσεις για τη σημειογραφία της εργασίας:

- Με τα γράμματα N, K θα εννοούμε το πλήθος των προτύπων ενός συνόλου εκπαίδευσης και το πλήθος των ομάδων αντίστοιχα σε ένα πρόβλημα ταξινόμησης.
- $p(z_j = 1) = P(j)$, είναι η πιθανότητα στη δυαδική μεταβλητή Z , το στοιχείο στη θέση j να είναι μονάδα.
- $p(z_j = 1|x) = p(j|x)$.
- Με κεφαλαία γράμματα συμβολίζονται τα διανύσματα και με πεζά οι συνιστώσες τους, π.χ. $X = \{x_1, x_2, \dots, x_n\}$.
- Θ είναι το σύνολο των παραμέτρων ενός Μοντέλου Μεικτών Κατανομών.
- θ_i είναι το σύνολο των παραμέτρων της i -οστής κατανομής σε ένα Μοντέλο Μεικτών Κατανομών.
- MI - Mutual Information, Αμοιβαία Πληροφορία.
- GMM - Gaussian Mixture Model, Μοντέλο Μεικτών Κανονικών Κατανομών.
- tMM - Student's t Mixture Model, Μοντέλο Μεικτών Student's t Κατανομών.

ΠΕΡΙΛΗΨΗ

Δημήτριος Γερογιάννης του Παναγιώτη και της Αλεξάνδρας. MSc, Τμήμα Πληροφορικής, Πανεπιστήμιο Ιωαννίνων, Οκτώβριος, 2007. Υπέρθεση εικόνας με μικτές κατανομές. Επιβλέπων: Νίκου Χριστόφορος.

Το θέμα της παρούσας εργασίας προέρχεται από τον χώρο της Υπολογιστικής Όρασης και ειδικότερα σχετίζεται με την υπέρθεση εικόνας. Το πρόβλημα της υπέρθεσης είναι αρκετά διαδεδομένο σε διάφορους επιστημονικούς τομείς (π.χ. Ιατρική εικόνα, Μετεωρολογία, Τηλεπισκόπηση, Πληροφοριακά Γεωγραφικά Συστήματα, Ρομποτική Όραση κ.α.). Σε πολλές περιπτώσεις η υπέρθεση είναι ένα βήμα προεπεξεργασίας κάποιας μεθόδου και μάλιστα η επιτυχία της εξαρτάται από το πόσο καλή ήταν η υπέρθεση.

Με τον όρο υπέρθεση εικόνων εννοούμε τη χωρική γεωμετρική στοίχιση των συντεταγμένων των εικονοστοιχείων των εικόνων έτσι ώστε περιοχές με νοηματική συγγένεια της μίας να συμπέσουν με τις αντίστοιχες της άλλης. Για παράδειγμα στην υπέρθεση δύο ιατρικών εικόνων του εγκεφάλου που έχουν ληφθεί με διαφορετικό τρόπο και σε διαφορετικές χρονικές στιγμές, στόχος της υπέρθεσης θα ήταν να έχουμε όσο το δυνατόν μεγαλύτερη ταύτιση έτσι ώστε εικονοστοιχεία της μίας που απεικονίζουν για παράδειγμα τη φαιά ουσία στη μία εικόνα, να απεικονίζουν επίσης τη φαιά ουσία στην άλλη. Είναι προφανές ότι ο λόγος που επιθυμούμε την υπέρθεση είναι γιατί λόγω διαφόρων παραγόντων, οι δύο εικόνες που πρόκειται να υπερτεθούν διαφέρουν χωρικά εξαιτίας της παραμόρφωσης κάποιου μετασχηματισμού. Τις παραμέτρους αυτού του μετασχηματισμού αναζητάμε σε ένα πρόβλημα υπέρθεσης.

Στη σχετική βιβλιογραφία υπάρχουν πολλές μέθοδοι επίλυσης του προβλήματος, με κυριότερη και ευρέως χρησιμοποιούμενη αυτή που βασίζεται στη μεγιστοποίηση της αμοιβαίας πληροφορίας. Στην παρούσα μεταπτυχιακή εργασία πρόκειται να παρουσιαστεί και αναλυθεί μία διαφορετική προσέγγιση που βασίζεται στις Μικτές Κατανομές. Η βασική ιδέα πίσω από τον αλγόριθμο που προτείνουμε είναι ότι αν υποθέσουμε κάποιες κατανομές στη φωτεινότητα των δύο εικόνων, τότε στην περίπτωση της τέλει στοίχισης, η απόσταση μεταξύ των δύο αυτών κατανομών θα είναι η ελάχιστη. Με βάση λοιπόν την παραπάνω λογική παρατήρηση, αρχικά εκτιμούμε την κατανομή της φωτεινότητας στη μία εικόνα (εξού και η χρήση των μικτών κατανομών) και εν συνεχεία προβάλλοντας τη διαμέριση που πήραμε από το Μοντέλο Μικτών Κατανομών στην άλλη παίρνουμε μία εκτίμηση για την κατανομή της πιθανότητας στην άλλη εικόνα. Εκτελλόντας αυτή τη διαδικασία για διάφορες τιμές των παραμέτρων μετασχηματισμού, προσπαθούμε να εντοπίσουμε εκείνες που ελαχιστοποιούν την απόσταση μεταξύ των δύο κατανομών που εκτιμήσαμε.

Από τεχνικής άποψης, όσον αφορά στα Μετά Μοντέλα χρησιμοποιήσαμε Μίξη Κανονικών και Student's t -Κατανομών, ενώ στην περίπτωση της ελαχιστοποίησης εφαρμόσαμε τον αλγόριθμο SIMPLEX. Εφαρμόσαμε τη μέθοδό μας σε διάφορους τύπους εικόνων (Ιατρικές, Μετεωρολογικές, Τηλεπισκόπησης κ.α.). Οι μελετούμενες εικόνες παραμορφώθηκαν με χρήση συμπαγών (*Rigid*)

μετασχηματισμών, ενώ επιπλέον προστέθηκαν και διάφορα επίπεδα θορύβου για να μελετηθεί και η ευρωστία της μεθόδου. Το γενικό συμπέρασμα είναι ότι η υπέρθεση με Μιχτές Κατανομές είναι πιο αποδοτική από ότι αυτή με μεγιστοποίηση της αμοιβαίας πληροφορίας, ιδίως παρουσία θορύβου. Όσον αφορά δε στις Μιχτές Κατανομές, οι Student's t - εμφανίζουν πολύ μεγαλύτερη ευρωστία έναντι των Κανονικών ακόμα και σε πολύ υψηλά επίπεδα θορύβου.

Η εργασία κλείνει με την παρουσίαση της επέκτασης της μεθόδου σε υπέρθεση νεφών σημείων. Και στην περίπτωση αυτή ακολουθήσαμε την ίδια τακτική (διαμέριση με Μιχτά Μοντέλα), μόνο που αντί του σταδίου της προβολής, χρησιμοποιήσαμε τον Hungarian αλγόριθμο για τον εντοπισμό των αντιστοιχιών μεταξύ των συνιστωσών κατάτμησης των επιμέρους ομάδων της διαμέρισης. Εν συνεχεία, εφαρμόζοντας τεχνικές ελαχιστοποίησης αναζητήσαμε τους μετασχηματισμούς υπέρθεσης, ελαχιστοποιώντας μία συνάρτηση παρόμοια με αυτή της περίπτωσης των εικόνων. Τα πειράματα που κάναμε έδειξαν ότι η υπέρθεση με Μιχτές Κατανομές, και δει με Μιχτές Student's t -Κατανομές, είναι πιο αξιόπιστα και πιο ακριβή, ιδίως παρουσία θορύβου, σε σχέση με τον ευρεως χρησιμοποιούμενο αλγόριθμο ICP.

EXTENDED ABSTRACT IN ENGLISH

Gerogiannis, Demetrios, Panagiotis. MSc, Computer Science Department, University of Ioannina, Greece. October, 2007. Image registration using mixture models. Thesis Supervisor: Nikou Christophoros.

This research work is focused on the problem of image registration. Image registration is a common problem in many scientific fields, e.g. medical imaging, remote sensing, robot vision, geographical information systems. In all these fields, image registration is considered to be a preprocessing step, upon which the success of a method is based. The problem of registration is to find the transformation parameters that bring into alignment the two images.

Many methods have been proposed to solve the image registration problem. The most common is based on the maximization of mutual information. Herein, we propose a pixel similarity-based algorithm enabling accurate rigid registration between single and multimodal images presenting gross dissimilarities due to noise, missing data or outlying measures. The method relies on the partitioning of a reference image by a mixture model, MM. This partition is then projected onto the image to be registered. The main idea is that a MM component in the reference image corresponds to a MM component in the image to be registered. If the images are correctly registered the weighted sum of distances between the corresponding components is minimized. We used both Gaussian mixture models (GMM) and Student's t-mixture models (tMM). The use of tMM components is justified by the property that they have heavier tails than standard Gaussians, thus providing robustness to outliers. Experimental results indicate that, even in the case of images presenting low SNR or important amount of dissimilarities due to temporal changes, the proposed algorithm compares favorably to the histogram-based mutual information method that is widely used in a variety of applications.

In general the proposed algorithm can be summarized as follows:

- Segment the reference image with a mixture model.
- Projection of the partition map onto the image to be registered by applying a transformation.
- Keep this instance of transformations' parameters that minimizes an energy function.

Finally, we have extended our method to the registration of clouds of points. The same experiments were conducted. The only difference was that the reference method in this circumstance was the ICP algorithm. The method was tested for its robustness in the presence of outliers. Again the results were in favor of the registration based on Mixture Models, while comparing the Gaussian mixture models with the Student's t-mixture model, the latter demonstrated a better behavior in the case of outliers' existence.

ΚΕΦΑΛΑΙΟ 1

ΕΙΣΑΓΩΓΗ ΣΤΗΝ ΥΠΕΡΘΕΣΗ ΕΙΚΟΝΑΣ

-
- 1.1 Εισαγωγή
 - 1.2 Ορισμός του προβλήματος
 - 1.3 Μετασχηματισμοί υπέρθεσης
 - 1.4 Μέθοδοι υπέρθεσης εικόνας
 - 1.4.1 Μέθοδοι που βασίζονται στην συσχέτιση
 - 1.4.2 Μέθοδοι που βασίζονται στο μετασχηματισμό Fourier
 - 1.4.3 Μέθοδοι που βασίζονται στην αντιστοίχιση σημείων
 - 1.4.4 Μέθοδοι που βασίζονται στην αμοιβαία πληροφορία
-

1.1 Εισαγωγή

Πολλές φορές σε διάφορα προβλήματα της Υπολογιστικής Όρασης, αλλά και άλλων επιστημονικών κλάδων, εμφανίζεται η ανάγκη για γεωμετρική στοίχιση δύο ή και περισσότερων εικόνων. Οι εικόνες αυτές απεικονίζουν την ίδια σκηνή, έχουν όμως ληφθεί με διαφορετικές παραμέτρους και ενδεχομένως σε διαφορετικές χρονικές στιγμές. Ένα πολύ κλασικό παράδειγμα, από τον χώρο της ιατρικής επεξεργασίας εικόνας, όπου το πρόβλημα της υπέρθεσης είναι αρκετά κοινό, θα μπορούσε να είναι η γεωμετρική στοίχιση της MRI (Αξονική Τομογραφία, *Magnetic Resonance Imaging*) εικόνας ενός ασθενούς που απεικονίζει τον εγκέφαλο με την αντίστοιχη SPECT (Σπινθηρογράφημα 3Δ, *Single Photon Emission Computed Tomography*). Όπως γίνεται αντιληπτό και οι δύο εικόνες απεικονίζουν την ίδια πληροφορία (εγκέφαλος), ωστόσο κάθε μία ακολουθεί διαφορετικό τρόπο στην οπτικοποίησή της. Ας υποθέσουμε ότι ο θεράπων ιατρός επιθυμεί να προβάλλει την μία εικόνα πάνω στην άλλη έτσι ώστε οι γεωμετρικές δομές της μίας, που αντιστοιχούν σε βιολογικά χαρακτηριστικά του εγκεφάλου, να συμπέσουν στις αντίστοιχές τους στην άλλη. Αυτή λοιπόν η διεργασία καλείται υπέρθεση των δύο εικόνων και για να έχει νόημα θα πρέπει αφενός μεν οι δύο εικόνες να απεικονίζουν κοινή πληροφορία, αφετέρου δε θα πρέπει να έχουν τέλεια γεωμετρική στοίχιση, έτσι ώστε εικονοστοιχεία

της μίας που προβάλλονται στην άλλη να αντιστοιχούν στις ίδιες δομές της αρχικής σκηνής που απεικονίζουν οι εικόνες. Γενικότερα, το πρόβλημα της υπέρθεσης δύο ή και περισσότερων εικόνων συνίσταται στον εντοπισμό εκείνων των γεωμετρικών μετασχηματισμών, που αν εφαρμοστούν στη μία από τις δύο εικόνες θα τη μετασχηματίσουν με τέτοιο τρόπο, έτσι ώστε τα εικονοστοιχεία της μετασχηματισμένης αν προβληθούν επί της αρχικής θα αντιστοιχούν σε περιοχές με συγκεκριμένη 'νοηματική' πληροφορία. Η αναγωγή στο προαναφερθέν παράδειγμα θα σήμαινε εικονοστοιχεία που στη μετασχηματισμένη εικόνα αντιστοιχούν στη φαιά ουσία, όταν προβληθούν στην αρχική θα δείχνουν στα αντίστοιχά τους της φαιάς ουσίας της αρχικής.

Όπως ενδεχομένως γίνεται αντιληπτό, το πρόβλημα της υπέρθεσης μπορεί να εμφανιστεί σε διάφορους επιστημονικούς τομείς, όπως η επεξεργασία ιατρικής εικόνας, η ρομποτική όραση, η παρακολούθηση της εξέλιξης διαφόρων φαινομένων σε δορυφορικές εικόνες - παραδείγματος χάριν της τρύπας του όζοντος. Για να καταλάβει ο αναγνώστης τη σπουδαιότητα του προβλήματος αναφέρουμε απλά ότι τη δεκαετία 1993 - 2003, περισσότερες από 1000 δημοσιεύσεις έγιναν πάνω στο αντικείμενο αυτό [4]. Στο πόνημα που κρατάτε στα χέρια σας θα παρουσιάσουμε μία μέθοδο για υπέρθεση εικόνας και θα τη συγκρίνουμε με αυτή που πρότειναν οι Viola και Wells [26] και Maes *et al.* [15] και βασίζεται στην μεγιστοποίηση της αμοιβαίας πληροφορίας των δύο εικόνων. Ο λόγος που επιλέξαμε αυτή τη σύγκριση είναι γιατί η υπέρθεση που βασίζεται στην αμοιβαία πληροφορία είναι η πιο διαδεδομένη μέθοδος ενώ μπορεί να επιτύχει αρκετά καλά αποτελέσματα.

Σε γενικές γραμμές οι εφαρμογές της υπέρθεσης εικόνας μπορούν να κατηγοριοποιηθούν σε τέσσερις κύριες ομάδες [4], ανάλογα με τον τρόπο λήψης των εικόνων:

- **Οπτική Γωνία:** Η είσοδος στον αλγόριθμο υπέρθεσης περιλαμβάνει εικόνες της ίδιας σκηνής, ληφθείσες από διαφορετική οπτική γωνία. Σε αυτή την ομάδα μπορούν να υπαχθούν προβλήματα ανασύνθεσης (*mosaicing*) μίας περιοχής που παρακολουθείται (ειδικότερο πρόβλημα της κατηγορίας εξ αποστάσεως επισκόπησης).
- **Χρόνος:** Οι ληφθείσες εικόνες αναφέρονται στην ίδια σκηνή, αλλά ο χρόνος λήψης τους διαφέρει, ενώ και οι συνθήκες λήψης μπορούν να διαφέρουν. Στην κατηγορία αυτή ένα κλασσικό παράδειγμα θα μπορούσαν να είναι οι ιατρικές εικόνες, όπου έχουμε μία χρονική ακολουθία εικόνων (π.χ. ακτινογραφιών) που δείχνουν την εξέλιξη της θεραπείας μίας νόσου σε έναν ασθενή.
- **Μηχανισμός Λήψης:** Στην κατηγορία αυτή οι εικόνες που θα υπερτεθούν έχουν ληφθεί με διαφορετικό μηχανισμό (π.χ. η μία αναπαριστά τη σκηνή στο υπέρυθρο και η άλλη στη υπεριώδες φάσμα). Η ανάγκη υπέρθεσης εικόνων αυτής της κατηγορίας είναι το γεγονός ότι με τον τρόπο αυτό είναι δυνατός ο συνδιασμός διαφόρων πληροφοριών από ξεχωριστές πηγές σε ένα μοντέλο. Ένα πρακτικό πρόβλημα αυτής της κατηγορίας είναι η υπέρθεση δύο διαφορετικού τύπου εικόνων (π.χ. μίας ακτινογραφίας και μίας αξονικής τομογραφίας) για τον εντοπισμό των αποτελεσμάτων μίας πάθησης σε κάποιον ασθενή. Είναι πολύ σημαντικό να αναφερθεί ότι οι υπερτιθέμενες εικόνες στην κατηγορία αυτή μπορεί να έχουν πλεονάζουσα πληροφορία, αυτή όμως είναι συμπληρωματική και επομένως μπορεί να οδηγήσει στην εξόρυξη νέας γνώσης, η οποία διαφορετικά δε θα ήταν προσιτή.
- **Μοντέλο περιγραφής των δεδομένων:** Μία αρκετά ευρύα κατηγορία προβλημάτων, όπου από τη μία μεριά υπάρχει η εικόνα της σκηνής και από την άλλη ένα μοντέλο περιγραφής της. Για

παράδειγμα στην ομάδα αυτή των προβλημάτων ανήκει και αυτό της υπέρθεσης μίας αεροφωτογραφίας επί ενός χάρτη της αντίστοιχης περιοχής. Η μεν αεροφωτογραφία είναι η εικόνα της σκηνής μας ο δε χάρτης αποτελεί το μοντέλο περιγραφής της σκηνής.

Είναι άμεσα αντιληπτό ότι υπάρχει μία ποικιλία στα δεδομένα εισόδου ενός αλγορίθμου υπέρθεσης, γεγονός που αντανακλάται σε μία ποικιλομορφία στον τρόπο αλλοίωσης των δεδομένων (ποιότητα θορύβου). Έτσι είναι δύσκολο να επινοήσει κάποιος μία μέθοδο γενικής χρήσης, δεδομένου ότι κάθε φορά θα υπάρχουν διαφορετικά μοντέλα προσέγγισης του μετασχηματισμού υπέρθεσης (δηλαδή διαφορετικές γεωμετρίες), διαφορετικά μοντέλα προσομοίωσης του θορύβου και γενικά μία πληθώρα παραμέτρων που διαφοροποιούνται αναλόγως του προβλήματος [4].

Ωστόσο όλες οι μέθοδοι που έχουν προταθεί μπορούν να συνοψιστούν στα κάτωθι τέσσερα γενικά βήματα [4]:

- Εντοπισμός χαρακτηριστικών. Στο βήμα αυτό κάποιος ειδικός, ή αν αυτό είναι εφικτό με αυτόματο τρόπο, εξάγονται τα βασικά χαρακτηριστικά σημεία της εικόνας (περιγράμματα, περιοχές με σαφώς καθορισμένα όρια, ακμές, κ.ά.). Γενικά τα σημεία αυτά πολλές φορές μπορούν να περιγραφούν από ένα επιμέρους σύνολο σημείων (π.χ. το κεντροειδές, στην περίπτωση ομάδας σημείων, το σημείο αρχής και τέλους, στην περίπτωση ευθγράμμων τμημάτων κ.τ.λ.). Στη βιβλιογραφία τα σημεία αυτά ονομάζονται σημεία ελέγχου (*Control Points - CP*).
- Ταίριασμα χαρακτηριστικών. Στο βήμα αυτό απότερος στόχος είναι να εντοπιστεί η αντιστοίχιση των χαρακτηριστικών που εντοπίστηκαν στο προηγούμενο βήμα στις δύο εικόνες. Για να επιτευχθεί αυτό έχουν προταθεί διάφορες μέθοδοι που αξιοποιούν την ομοιότητα των χαρακτηριστικών, την χωρική τους συγγένεια κ.α.
- Εκτίμηση του μετασχηματισμού υπέρθεσης. Στο βήμα αυτό γίνεται εκτίμηση τόσο του τύπου του μετασχηματισμού - θα αναφερθούμε στη συνέχεια στο θέμα αυτό - όσο και των παραμέτρων του μετασχηματισμού. Είναι ιδιαίτερα σημαντικό να τονιστεί ότι η αυτή η εκτίμηση των παραμέτρων γίνεται με βάση την αντιστοίχιση των χαρακτηριστικών που προηγήθηκαν στο ανωτέρω βήμα.
- Μετασχηματισμός εικόνας. Η μία εκ των δύο εικόνων μετασχηματίζεται, με βάση την εκτίμηση των παραμέτρων που βρέθηκαν στο προηγούμενο βήμα. Να σημειωθεί ότι η διαδικασία αυτή είναι επαναληπτική, δηλαδή τα παρουσιαζόμενα τέσσερα βήματα εκτελούνται επαναληπτικά έως ότου συγκλίνουμε στη λύση του προβλήματος (εύρεση του μετασχηματισμού υπέρθεσης). Επομένως είναι κατανοητό ότι σε κάθε βήμα επανάληψης υπάρχει το ενδεχόμενο, λόγω του μετασχηματισμού, να εμφανιστούν μη ακέραιες συντεταγμένες για τα εικονοστοιχεία της εικόνας. Στην περίπτωση αυτή υπεισέρχεται η έννοια της παρεμβολής, προκειμένου να εξαχθεί η τελική τιμή του χρώματος. Περισσότερα για το θέμα της παρεμβολής θα αναλύσουμε στην ενότητα 3.3, όπου θα περιγράψουμε πώς διέξάγαμε τα πειράματα σύγκρισης της μεθόδου μας.

Κατ' αντιστοιχία με τα παραπάνω βήματα μπορούμε να πούμε ότι ένας αλγόριθμος υπέρθεσης περιλαμβάνει [21]:

- έναν χώρο των χαρακτηριστικών. Είναι το σύνολο των δεδομένων από τις δύο εικόνες που θα χρησιμοποιηθούν από τον αλγόριθμο υπέρθεσης για να εκτιμηθούν οι παράμετροι του μετασχηματισμού.

- έναν χώρο αναζήτησης. Το διάστημα στο οποίο θα αναζητήσουμε τη βέλτιστη τιμή κάθε παραμέτρου του μετασχηματισμού υπέρθεσης.
- μία στρατηγική αναζήτησης. Δεδομένων των δύο ανωτέρων στοιχείων (παραμέτρων και χώρων αναζήτησης) πρέπει να αποφασίσουμε και τον τρόπο με τον οποίο θα γίνει σε κάθε βήμα η επιλογή της βέλτιστης λύσης.
- ένα μέτρο ομοιότητας. Είναι η μετρική που θα χρησιμοποιηθεί σε κάθε επανάληψη έτσι ώστε να εκτιμηθεί η ποιότητα της υπολογισμένης υπέρθεσης.

Οι διάφορες μέθοδοι που έχουν προταθεί στη βιβλιογραφία, ακολουθούν για κάθε βήμα μία προσέγγιση ανάλογα με την υφή του προβλήματος. Για παράδειγμα, το βήμα της αναζήτησης, εκεί όπου δηλαδή γίνεται η εκτίμηση των παραμέτρων του μετασχηματισμού υπέρθεσης για την τρέχουσα επανάληψη, είναι στην ουσία μία μέθοδος ελαχιστοποίησης. Κάθε μέθοδος λοιπόν μπορεί να υιοθετεί διαφορετική στρατηγική για τον υπολογισμό ελαχίστων (ή μεγίστων), ανάλογα με την αντικειμενική συνάρτηση που μελετάται.

Στο κεφάλαιο 3 θα δούμε πώς τα ανωτέρω βήματα εξειδικεύονται στην μέθοδο υπέρθεσης που προτείνουμε.

1.2 Ορισμός του προβλήματος

Η υπέρθεση δύο εικόνων ορίζεται ως μία χωρική αντιστοίχιση χωρικών συντεταγμένων μεταξύ δύο εικόνων [21]. Για παράδειγμα, αν υποθέσουμε ότι οι εικόνες αναπαρίστανται ως 2Δ μήτρες I_1, I_2 , οπότε $I_i(x, y)$ είναι η φωτεινότητα του εικονοστοιχείου (x, y) της εικόνας I_i (ή οποιαδήποτε άλλη μέτρηση που σχετίζεται με τον τρόπο λήψης της εικόνας), τότε η υπέρθεση μπορεί να οριστεί ως η αντιστοίχιση:

$$I_2(x, y, z) = g(I_1(\Omega(x, y, z)))$$

όπου Ω είναι ένας τρισδιάστατος χωρικός μετασχηματισμός, δηλαδή οι χωρικές συντεταγμένες (x, y, z) αντιστοιχίζονται στις (x', y', z') , και g είναι ένας μονοδιάστατος μετασχηματισμός της φωτεινότητας της εικόνας. Περισσότερες λεπτομέρειες για τους μετασχηματισμούς υπέρθεσης θα δώσουμε στην επόμενη ενότητα αυτού του κεφαλαίου. Απλά να σημειώσουμε ότι στην παρούσα εργασία δε θα ασχοληθούμε με μετασχηματισμούς που αλλοιώνουν την φωτεινότητα της εικόνας, αν και στα πειράματα που θα παρουσιάσουμε θα χρησιμοποιήσουμε εικόνες που διαφέρουν μεταξύ τους κατά τον τρόπο λήψης, για παράδειγμα η μία εικόνα θα αναπαριστά πληροφορία στο ορατό φως και η άλλη την ίδια σκηνή στο υπεριώδες. Έτσι μπορούμε να ισχυριστούμε ότι κατά τρόπον τινά μπορούμε να βρούμε και μία αντιστοίχιση στις φωτεινότητες. Στόχος κάθε μεθόδου υπέρθεσης είναι να εντοπίσει τους βέλτιστους μετασχηματισμούς έτσι ώστε οι δύο εικόνες να ταυτιστούν πλήρως, είτε γιατί το ζητούμενο είναι οι τιμές αυτών των παραμέτρων, είτε γιατί επιθυμούμε να ταυτοποιήσουμε περιοχές ενδιαφέροντος στις δύο εικόνες, για παράδειγμα στην ιατρική υπέρθεση εικόνας θα θέλαμε να βρούμε σε μία MRI και σε μία SPECT εικόνα την περιοχή που αναπαριστά έναν δεδομένο καρκινικό όγκο, προκειμένου να συνδιαστεί ιατρική πληροφορία από δύο διαφορετικές πηγές. Ο εντοπισμός των βέλτιστων μετασχηματισμών είναι το κλειδί για μία επιτυχή υπέρθεση και ουσιαστικά πρόκειται για ένα πρόβλημα βελτιστοποίησης. Όσο πιο αποδοτική είναι η βελτιστοποίηση τόσο πιο κοντά στη βέλτιστη λύση θα βρεθούμε.

1.3 Μετασχηματισμοί υπέρθεσης

Ορίσαμε στην προηγούμενη ενότητα την υπέρθεση ως ένα πρόβλημα αναζήτησης βέλτιστων, ως προς ένα κριτήριο, μετασχηματισμών. Στην ενότητα αυτή θα δούμε αναλυτικά ποιοι είναι οι μετασχηματισμοί αυτοί και πώς μπορούν να μοντελοποιηθούν μαθηματικά.

Οι μετασχηματισμοί υπέρθεσης χωρίζονται σε δύο μεγάλες κατηγορίες, αν λάβουμε υπόψη μας τήν περιοχή εφαρμογής, τους καθολικούς (*Global*) και τους τοπικούς (*Local*). Οι καθολικοί μετασχηματισμοί εφαρμόζονται σε όλη την εικόνα, σε αντίθεση με τους τοπικούς, οι οποίοι θεωρούν την εικόνα σαν ένα σύνολο από μικρά χωρία, οπότε για καθένα από αυτά οι μετασχηματισμοί υπέρθεσης έχουν διαφορετικές τιμές. Οι τοπικοί μετασχηματισμοί χρησιμοποιούνται σπάνια και εν γένει, στην κατηγορία αυτή ανήκουν όσοι μετασχηματισμοί δεν μπορούν να χαρακτηριστούν ως καθολικοί. Έτσι για παράδειγμα ένα σύνολο μετασχηματισμών που εφαρμόζεται σε ένα μέρος μίας εικόνας μόνο, δε χαρακτηρίζεται ως τοπικό αλλά ως ταθολικό, αφού μπορούμε να απομονώσουμε εκείνο το τμήμα της εικόνας στο οποίο εφαρμόστηκαν οι μετασχηματισμοί και να το επεξεργαστούμε ως μία νέα εικόνα [19].

Με δεδομένη την φύση των μετασχηματισμών, δηλαδή το είδος της μεταβολής που επιφέρουν με την εφαρμογή τους, μπορούμε να τους διακρίνουμε σε Συμπαγείς (*Rigid*) Συναφείς (*Affine*), Προβολικούς (*Projective*) και Ελαστικούς (*Elastic*). Με βάση την προηγούμενη σειρά αναφοράς κάθε μετασχηματισμός αποτελεί ειδική περίπτωση του επόμενου. Έτσι στην περίπτωση που ο μετασχηματισμός υπέρθεσης είναι συνδυασμός των ανωτέρω, μπορούμε απλά να τον κατατάξουμε στην κατηγορία που ανήκει ο πιο σύνθετος. Δηλαδή αν στην υπέρθεση έχουν λάβει χώρα Συναφείς και Προβολικοί μετασχηματισμοί, μπορούμε χωρίς βλάβη της γενικότητας να πούμε ότι τελικά έχει εφαρμοστεί ένας Προβολικός μετασχηματισμός. Κατά τη διαδικασία της υπέρθεσης θα υπολογιστούν εκείνες οι τιμές των παραμέτρων, έτσι ώστε να συγχωνευτεί ο μετασχηματισμός Συναφείας μέσα στον Προβολικό. Ενδεχομένως η διαδικασία αυτή να είναι υπολογιστικά πιο απαιτητική, ωστόσο στο τέλος θα οδηγήσει στο επιθυμητό αποτέλεσμα.

Οι Συμπαγείς μετασχηματισμοί προκαλούνται είτε από την κίνηση των αισθητήρων λήψης των εικόνων, είτε από την κίνηση αντικειμένων μέσα στη σκηνή. Σε κάθε περίπτωση όμως το σχετικό σχήμα και μέγεθος διατηρείται. Επομένως ένας συμπαγής μετασχηματισμός είναι μία σύνθεση περιστροφής, μετατόπισης και κλιμάκωσης. Το μαθηματικό μοντέλο περιγραφής του μετασχηματισμού στις τρεις διαστάσεις μπορεί να δωθεί με τη χρήση πινάκων:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = (sI)R \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} T_x \\ T_y \\ T_z \end{pmatrix}$$

όπου το διάνυσμα $(T_x, T_y, T_z)^T$ εκφράζει τη μετατόπιση ως προς x, y, z αντίστοιχα, ενώ R είναι ο πίνακας περιστροφής και ορίζεται ως

$$R = \begin{pmatrix} 1 & 0 & 0 \\ \sigma\upsilon\nu\theta_x & 0 & -\eta\mu\theta_x \\ 0 & -\eta\mu\theta_x & \sigma\upsilon\nu\theta_x \end{pmatrix} \begin{pmatrix} \sigma\upsilon\nu\theta_y & 0 & -\eta\mu\theta_y \\ 0 & 1 & 0 \\ \eta\mu\theta_y & 0 & \sigma\upsilon\nu\theta_y \end{pmatrix} \begin{pmatrix} \sigma\upsilon\nu\theta_z & \eta\mu\theta_z & 0 \\ -\eta\mu\theta_z & \sigma\upsilon\nu\theta_z & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

με θ_x να εκφράζει τη γωνία περιστροφής ως προς τον άξονα x , θ_y να εκφράζει τη γωνία περιστροφής ως προς τον άξονα y και θ_z να εκφράζει τη γωνία περιστροφής ως προς τον άξονα z . Τέλος η σταθερά s περιγράφει την κλιμάκωση. Προφανώς τιμές κατ' απόλυτη τιμή μεγαλύτερες της μονάδος οδηγούν σε μεγέθυνση των αντικειμένων, ενώ τιμές κατ' απόλυτη τιμή μικρότερες της μονάδος επιφέρουν

σμίχρυνση των αντικειμένων. Να σημειωθεί ότι η κλιμάκωση είναι ίδια προς όλους τους άξονες του συστήματος συντεταγμένων. Οι μετασχηματισμοί αυτής της κατηγορίας είναι και οι πιο διαδεδομένοι στις τεχνικές υπέρθεσης εικόνας, δεδομένου ότι υπάρχει ένα σαφές μοντέλο περιγραφής τους. Η εργασία που εκπονήσαμε και παρουσιάζουμε εδώ μελετά αυτού του είδους τους μετασχηματισμούς.

Οι Συναφείς μετασχηματισμοί είναι πιο γενικοί από τους Συμπαγείς και επομένως πιο πολύπλοκοι. Μπορούν να επιφέρουν πιο σύνθετες παραμορφώσεις στην εικόνα, όπως αλλαγή στο λόγο λήψης (*aspect ratio*) της εικόνας. Ένας τύπος Συναφούς μετασχηματισμού είναι η Αποκοπή (*Shear*). Ο μετασχηματισμός αυτός μπορεί να επιδρά με διαφορετικό τρόπο σε κάθε άξονα, οδηγώντας σε πιο πολύπλοκες παραμορφώσεις. Μαθηματικά ο δισδιάστατος μετασχηματισμός Αποκοπής μπορεί να παρασταθεί με τους πίνακες:

$$Shear_x = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \quad Shear_y = \begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix}$$

Ένας ακόμη μετασχηματισμός αυτής της κατηγορίας είναι η κλιμάκωση, αλλά με διαφορετικές παραμέτρους ως προς κάθε άξονα, π.χ. στις τρεις διαστάσεις

$$s = \begin{pmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & s_z \end{pmatrix}.$$

Αναφορικά με τους προβολικούς μετασχηματισμούς, αυτοί μπορούν να προκαλέσουν παραμορφώσεις λόγω προβολής αντικειμένων επί του επιπέδου της εικόνας.

Τέλος η ελαστική παραμόρφωση, που είναι και ο πιο γενικός τύπος μετασχηματισμού, δεν μπορεί εν γένει να παρασταθεί με έναν γενικό τύπο. Στις πλείστες των περιπτώσεων, είτε για κάθε σημείο εφαρμόζεται ένας μετασχηματισμός μετατόπισης, $y_i = x_i + t_i(x)$ είτε ο μετασχηματισμός προσεγγίζεται στο σύνολό του ως πολυωνυμικός, συναρτήσε των χωρικών συντεταγμένων.

1.4 Μέθοδοι υπέρθεσης εικόνας

Στην ενότητα αυτή θα παρουσιάσουμε περιληπτικά τις γενικές αρχές για διάφορες μεθόδους που έχουν προταθεί. Αφού έχουμε αποφασίσει το είδος του μετασχηματισμού που θα αναζητήσουμε, θα πρέπει να ορίσουμε ένα κριτήριο με βάση το οποίο θα γίνει η αναζήτηση των βέλτιστων μετασχηματισμών. Ανάλογα με το κριτήριο που επιλέγεται μία μέθοδος κατατάσσεται στην αντίστοιχη κατηγορία.

1.4.1 Μέθοδοι που βασίζονται στην συσχέτιση

Η συσχέτιση είναι ένα στατιστικό μέτρο ομοιότητας και μπορεί να εκτιμήσει το βαθμό ταύτισης μίας εικόνας και ενός προτύπου. Στην περίπτωση της υπέρθεσης εικόνων, η συσχέτιση μπορεί να οριστεί ως εξής:

Δοθέντος ενός προτύπου I_1 και μίας εικόνας I_2 , ο συντελεστής συσχέτισης ορίζεται από τον τύπο:

$$C(u, v) = \frac{\sum_{x,y} I_1(x, y) I_2(x - u, y - v)}{\sqrt{\sum_{x,y} I_1^2(x, y) \sum_{x,y} I_2^2(x, y)}}$$

Η μέτρηση της συσχέτισης υπολογίζεται με τη χρήση παραθύρων στην εικόνα αναφοράς (*reference*) και στην εικόνα υπέρθεσης (*sensed*). Το ζητούμενο είναι η αναζήτηση της μέγιστης τιμής. Τα ζεύγη παραθύρων στις δύο εικόνες για τα οποία η συσχέτιση είναι μέγιστη, θεωρούνται ως αντιστοιχισμένα. Οι μέθοδοι συσχέτισης μπορούν να υπολογίσουμε με πολύ καλή ακρίβεια μετασχηματισμούς μετατόπισης, ενώ μπορούν να δώσουν καλές προσεγγίσεις και για μετασχηματισμούς περιστροφής, με μικρές γωνίες περιστροφής, και κλιμάκωσης, με μικρές τιμές στην παράμετρο κλιμάκωσης. Γενικευμένες εκδοχές του παραπάνω αλγορίθμου υπάρχουν και για πιο σύνθετους μετασχηματισμούς, ωστόσο οι υπολογιστικές απαιτήσεις αυξάνονται ραγδαία όσο πιο πολύπλοκος είναι ο μετασχηματισμός.

Στην ίδια λογική με την προαναφερθείσα τεχνική είναι και ο αλγόριθμος Εντοπισμού Συνεχόμενης Ομοιότητας (*Sequential Similarity Detection Algorithm*) [13]. Η μέθοδος αυτή χρησιμοποιεί συνεχόμενη αναζήτηση και μία απλούστερη εκδοχή της απόστασης ομοιότητας σε σχέση με την (1.4.1). Ειδικότερα, υπολογίζεται το άθροισμα της απόλυτης διαφοράς στις φωτεινότητες των εικόνων (l_2 στάθμη) και εφαρμόζεται ένα κριτήριο κατωφλίσωσης. Αν το συνολικό άθροισμα των δύο παραθύρων είναι μικρότερο από την τιμή κατωφλίου, τότε θεωρείται ότι υπάρχει αντιστοίχιση, αλλιώς το τρέχω παράθυρο που εξετάζεται απορρίπτεται και ο έλεγχος προχωρά σε επόμενα ζεύγη παραθύρων. Η μέθοδος αυτή έχει ελάχιστα χειρότερα αποτελέσματα από την κλασική συσχέτιση, εμφανίζει όμως πολύ καλύτερη χρονική απόδοση.

Δύο βασικά μειονεκτήματα αυτής της κατηγορίας είναι το γεγονός ότι το μέτρο που χρησιμοποιείται για την ομοιότητα εμφανίζει πεπλατισμένα μέγιστα, και το ότι η υπολογιστική πολυπλοκότητα είναι πολύ υψηλή.

1.4.2 Μέθοδοι που βασίζονται στο μετασχηματισμό Fourier

Οι τεχνικές που βασίζονται στο μετασχηματισμό Fourier, ενδείκνυνται κυρίως για τις περιπτώσεις που επιθυμούμε να επιταχύνουμε τους υπολογισμούς, όπως επίσης και περιπτώσεις όπου οι εικόνες έχουν ληφθεί υπό διαφορετικές συνθήκες ή έχουν αλλοιωθεί από θόρυβο σχετιζόμενο με τη συχνότητα.

Οι μέθοδοι της κατηγορίας αυτής εκμεταλλεύονται την αναπαράσταση της εικόνας στο πεδίο της συχνότητας. Η μέθοδος συσχέτισης φάσης αρχικά χρησιμοποιήθηκε για προβλήματα παραμόρφωσης λόγω μετατόπισης. Η ένταση του φάσματος των δύο εικόνων υπολογίζεται και αναζητάται η περιοχή όπου εμφανίζεται κορυφή στην αντίστροφη ποσότητα, $\frac{F(f)F(g)^*}{|F(f)F(g)^*|} = e^{2\pi i(ux_0+vy_0)}$. Η μέθοδος αυτή είναι αρκετά εύρωστη σε συσχετισμένο θόρυβο και σε μη ομοιόμορφες αλλοιώσεις στο φωτισμό. Η δε ωφέλεια στον υπολογιστικό χρόνο είναι ανάλογη του μεγέθους των εικόνων. Ορισμένες ενδιαφέρουσες εργασίες που έχουν δημοσιευθεί στο χώρο αυτό είναι των De Castro και Morandi [14], όπου περιγράφεται μία μέθοδος για την επίλυση προβλημάτων υπέρθεσης, όπου ο μετασχηματισμός παραμόρφωσης περιλαμβάνει και περιστροφές, ενώ στην εργασία των Wolberg και Zokai [18], περιγράφεται μία τεχνική που βασίζεται σε πολικές απεικονίσεις για υπέρθεση με μετασχηματισμούς συνάφειας. Παραπέμπουμε τον αναγνώστη στις εργασίες [4] και [21] για περισσότερες πληροφορίες σχετικά με μεθόδους αυτής της κατηγορίας.

1.4.3 Μέθοδοι που βασίζονται στην αντιστοίχιση σημείων

Οι μέθοδοι που αναλύσαμε στις προηγούμενες υποενότητες, βασίζονται στην ταύτιση περιοχών της εικόνας. Οι μέθοδοι που βασίζονται στην αντιστοίχιση σημείων προϋποθέτουν ότι έχει προηγηθεί

ένα βήμα στο οποίο έχουν ανακτηθεί σημεία ελέγχου και στις δύο εικόνες. Στη βιβλιογραφία έχουν προταθεί διάφορες μέθοδοι για την επίλυση αυτού του προβλήματος, όπως για παράδειγμα η ταύτιση γράφων [1], η ομαδοποίηση [17], κ.α. Παραπέμπουμε τον αναγνώστη στις αναφορές [4] και [21] για λεπτομερή αναφορά των διαφόρων μεθόδων. Δε θα αναφερθούμε εκτενέστερα σε αυτή την κατηγορία μεθόδων, δεδομένου ότι απέχει από την προτεινόμενη μέθοδο, όπου η χρήση σημείων ελέγχου δεν είναι απαραίτητη.

1.4.4 Μέθοδοι που βασίζονται στην αμοιβαία πληροφορία

Στην ενότητα αυτή θα περιγράψουμε μία κατηγορία μεθόδων με μεγάλη απήχηση στον κλάδο της υπέρθεσης. Βασικό χαρακτηριστικό είναι η χρήση της αμοιβαίας πληροφορίας ως μετρικού για τη σύγκριση δύο εικόνων. Θεμέλιο λίθο για την προώθηση αυτού του τομέα απετέλεσαν οι εργασίες των Viola και Wells [26] και Maes *et al.* [15]. Ο λόγος που γίνεται ιδιαίτερη μνεία στην κατηγορία αυτή είναι η ευρεία διάδοση των τεχνικών της αλλά και η σχετική ευρωστία που παρουσιάζουν σε δεδομένα που έχουν αλλοιωθεί από θόρυβο. Μάλιστα στην αναφορά [4] οι μέθοδοι της κατηγορίας αυτής χαρακτηρίζονται ως πρωτοπόρες σε προβλήματα υπέρθεσης πολυτροπικών (*multimodal*) εικόνων, εικόνων δηλαδή που έχουν ληφθεί με διαφορετικό τρόπο. Για το λόγο αυτό άλλωστε αποφασίσαμε να χρησιμοποιήσουμε την κατηγορία αυτή ως μέτρο σύγκρισης της απόδοσης της μεθόδου που προτείνουμε.

Η αμοιβαία πληροφορία, έννοια της Θεωρίας Πληροφορίας, είναι ένα μέτρο σύγκρισης της στατιστικής ομοιότητας δύο μεγεθών. Ο ακριβής ορισμός δίνεται στη συνέχεια:

Θεώρημα 1.1 (Ορισμός της αμοιβαίας πληροφορίας). *Δοθέντων δύο τυχαίων μεταβλητών X, Y με περιθώριες κατανομές $p(X), p(Y)$ ως αμοιβαία πληροφορία των δύο μεταβλητών ορίζεται η ποσότητα*

$$I(X, Y) = H(X) - H(X|Y) = H(X) + H(Y) - H(X, Y)$$

όπου

- $H(X)$ είναι η εντροπία της μεταβλητής X , $H(X) = E_x(\log(P(X)))$.
- $H(X, Y)$ είναι η από κοινού εντροπία των μεταβλητών X, Y , $H(X, Y) = E_{x,y}(\log(P(X, Y)))$.
- $P(X, Y)$ είναι από κοινού κατανομή των X, Y .

Οι μέθοδοι που βασίζονται στην αμοιβαία πληροφορία συνίστανται στην αναζήτηση εκείνων των παραμέτρων των μετασχηματισμών που μεγιστοποιούν την αμοιβαία πληροφορία. Όπως γίνεται άμεσα αντιληπτό, η χρήση των μεθόδων αυτής της κατηγορίας υπονοεί τη δυνατότητα υπολογισμού της από κοινού κατανομής των δύο εικόνων. Μία λύση στο πρόβλημα αυτό μπορεί να δοθεί με τη χρήση ιστογραμμάτων. Αυτό όμως προαπαιτεί τον υπολογισμό του ιστογράμματος για κάθε σύνολο παραμέτρων υπέρθεσης, γεγονός που επιφέρει υψηλό υπολογιστικό φόρτο. Όπως έχει προταθεί στη βιβλιογραφία, μπορεί να επιτευχθεί επιτάχυνση των υπολογισμών αν εφαρμοστεί μία πυραμιδοειδής διάταξη για μείωση της διάστασης των εικόνων [4]. Όπως προαναφέραμε, από τις πρώτες εργασίες στον τομέα αυτό ήταν των Viola και Wells [26] και Maes *et al.* [15]. Έκτοτε πλήθος εργασιών δημοσιεύθηκαν. Ενδεικτικά αναφέρουμε τους Thévenaz και Unser [25], Ritter *et al.* [22]). Για

μία αναλυτική σύγκριση της αμοιβαίας πληροφορίας με άλλα μέτρα ομοιότητας, παραπέμπουμε τον αναγνώστη στην αναφορά [16]. Τέλος αποφασίσαμε να κάνουμε ιδιαίτερη αναφορά στην εργασία του Studholme *et al.* [10], όπου εμφανίζεται ο ισχυρισμός ότι η κανονικοποιημένη αμοιβαία πληροφορία υπολογισμένη σε επικαλυπτόμενες περιοχές, μπορεί να αποτελέσει καλύτερο μέτρο σύγκρισης. Έτσι λοιπόν στο εξής όποτε αναφερόμαστε στην αμοιβαία πληροφορία, θα εννοούμε την κανονικοποιημένη της μορφή, υπολογισμένη σε περιοχές επικάλυψης (όσο εφικτό μπορεί να είναι αυτό).

ΚΕΦΑΛΑΙΟ 2

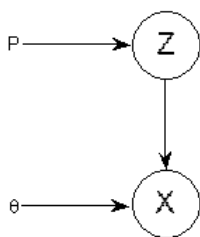
ΜΙΚΤΕΣ ΚΑΤΑΝΟΜΕΣ

-
- 2.1 Αρχές και εκπαίδευση Μικτών Κατανομών - Ο αλγόριθμος Expectation Maximization (EM)
 - 2.2 Μικτές Κανονικές Κατανομές
 - 2.3 Μικτές Student's t -Κατανομές
 - 2.4 Αποστάσεις Μεταξύ Κατανομών
 - 2.4.1 Απόσταση Chernof
 - 2.4.2 Απόσταση Battacharyya
 - 2.4.3 Τετραγωνική απόσταση - L_2 στάθμη
-

2.1 Αρχές και εκπαίδευση Μικτών Κατανομών - Ο αλγόριθμος Expectation Maximization (EM)

Η έννοια της ταξινόμησης είναι ευρέως διαδεδομένη στο χώρο της Μηχανικής Μάθησης. Στο κεφάλαιο αυτό θα παρουσιάσουμε μία τεχνική για την επίλυση του εν λόγω προβλήματος, η οποία βασίζεται σε αρχές της θεωρίας πιθανοτήτων. Υποθέτουμε ένα στατιστικό παραμετρικό μοντέλο που παράγει τα δεδομένα και στόχος μας είναι να εκτιμήσουμε τις παραμέτρους του μοντέλου αυτού. Στα πλαίσια των μικτών κατανομών, το προαναφερθέν μοντέλο προσεγγίζεται από ένα άθροισμα κατανομών και επομένως περιγράφεται μέσω του συνόλου των παραμέτρων των κατανομών [7], [8]. Έτσι η αρχική υπόθεση περί στατιστικής διαδικασίας παραγωγής των δεδομένων συνοψίζεται στην ύπαρξη K ανεξάρτητων κατανομών, που αποτελούν τους γεννήτορες κάθε ομάδας του συνόλου των δεδομένων. Είναι προφανές ότι κάθε ομάδα πλέον περιγράφεται μέσω των παραμέτρων της κατανομής που παράγει τα δεδομένα της. Ο γράφος που περιγράφει ένα μοντέλο μικτών κατανομών του συνόλου δεδομένων X , δίνεται στο σχήμα 2.1, όπου Z είναι το πλήρες σύνολο των δεδομένων, $Z = \{(z_i, x_i)\}$, $i = 1..N$, $x_i \in X$.

Η μεταβλητή z , η οποία ονομάζεται κρυφή μεταβλητή, είναι δυαδική και ακολουθεί αναπαράσταση 1-από- K , δηλαδή αν έχουμε K ομάδες στο σύνολο δεδομένων και το πρότυπο x ανήκει στην



Σχήμα 2.1: Γραφικό μοντέλο αναπαράστασης Μικτών Κατανομών.

κατηγορία j , τότε το $z \in \{0, 1\}^K$ έχει στη θέση j μονάδα και στις υπόλοιπες θέσεις μηδενικά, δηλαδή

$$z = [0, 0, \dots, 0, \underbrace{1}_{\text{θέση } j}, 0, \dots, 0]^T.$$

Με άλλα λόγια η μεταβλητή z καθορίζει την κατηγορία του προτύπου x .

Με βάση τα παραπάνω λοιπόν μπορούμε να πούμε ότι, στις μίχτες κατανομές, η συνάρτηση πυκνότητας πιθανότητας για ένα πρότυπο x δίνεται από τον παρακάτω τύπο:

$$p(x) = \sum_{j=1}^K P(j)p(x|\theta_j) \quad (2.1)$$

όπου:

- $P(j)$, είναι η πιθανότητα κάθε κατηγορίας, ανάλογη με το πλήθος των δεδομένων που παράγει. Προφανώς $\sum_{j=1}^K P(j) = 1$ ενώ $P(j) \in [0, 1] \quad j = 1, 2, \dots, K$. Οι παράμετροι $P(j)$ καλούνται στη βιβλιογραφία συντελεστές μίξης (mixing coefficients) και στην ουσία είναι τα βάρη συμμετοχής της κάθε επιμέρους κατανομής στο μοντέλο εκτίμησης της συνολικής πιθανότητας.
- $p(x|\theta_j)$, είναι η πιθανότητα δοθήσης της κατανομής θ_j το πρότυπο x να έχει παραχθεί από αυτή. x .

Από την εξίσωση (2.1) γίνεται αντιληπτό αμέσως ότι για να μπορέσουμε να κατασκευάσουμε ένα μικτό μοντέλο κατανομών, θα πρέπει εκ των προτέρων να ορίσουμε τον τύπο της κατανομής $p(x|\theta_j)$. Ανάλογα με τον τύπο κατανομής που υιοθετούμε το μοντέλο παίρνει και το αντίστοιχο όνομα. Στα επόμενα μέρη του κεφαλαίου θα δούμε δύο αρκετα συχνούς τύπους κατανομών.

Όμως πώς είναι εφικτή η κατηγοριοποίηση ενός προτύπου μέσω των μίχτων κατανομών; Η απάντηση στο ερώτημα είναι απλή και βασίζεται στη λογική έννοια της πιθανότητας, ένα πρότυπο ανήκει στην ομάδα με την μεγαλύτερη πιθανότητα. Και έτσι εγείρεται ένα νέο ερώτημα. Πώς μπορώ να καθορίσω με ποια πιθανότητα ένα πρότυπο x ανήκει σε μία ομάδα; Τη λύση στο πρόβλημα αυτό δίνει ένα από τα θεμελιώδη θεωρήματα της θεωρίας Πιθανοτήτων, το θεώρημα του Bayes, του οποίου η γενική διατύπωση είναι η ακόλουθη.

Θεώρημα 2.1 (θεώρημα του Bayes). Δοθέντων δύο τυχαίων μεταβλητών A, B ισχύει η παρακάτω σχέση μεταξύ της εκ των προτέρων και της εκ των υστέρων πιθανότητας του A .

$$p(A|B) = \frac{p(B|A)p(A)}{p(B)} \quad (2.2)$$

Επιστρέφοντας στο πρόβλημα της ταξινόμησης μπορούμε να αξιοποιήσουμε τον κανόνα του Bayes και να πάρουμε τον παρακάτω τύπο για την πιθανότητα ενός προτύπου x να ανήκει στην κατηγορία θ_j (εκ των υστέρων πιθανότητα):

$$p(\theta_j|x) = \frac{p(x|\theta_j)P(j)}{p(x)} = \frac{p(x|\theta_j)P(j)}{\sum_{j=1}^K \{P(j)p(x|\theta_j)\}} \quad (2.3)$$

Για τις εκ των υστέρων πιθανότητες ισχύει

$$\sum_{j=1}^K p(\theta_j|x) = 1 \quad (2.4)$$

Ανακεφαλαιώνοντας μπορούμε να πούμε ότι ένας ταξινομητής που βασίζεται σε μίξη προκαθορισμένων κατανομών λειτουργεί με βάση την ακόλουθη αρχή:

$$C(x) = \arg \max_{j=1\dots K} \{p(\theta_j|x)\} = \arg \max_{j=1\dots K} \left\{ \frac{p(x|\theta_j)P(j)}{p(x)} \right\} = \arg \max_{j=1\dots K} \left\{ \frac{p(x|\theta_j)P(j)}{\sum_{j=1}^K \{P(j)p(x|\theta_j)\}} \right\} \quad (2.5)$$

όπου ως $C(x)$ ορίζεται η κατηγορία του προτύπου x ενώ η εκπαίδευσή του συνίσταται στην εκτίμηση των παραμέτρων $\theta_j, P(j) | j = 1 \dots K$. Αφού καθορίσαμε το πώς θα λειτουργήσει το μοντέλο μας θα περιγράψουμε τώρα έναν γενικό τρόπο για το πώς θα καθορίσουμε τις παραμέτρους του. Ο τρόπος αυτός θα εξειδικευθεί στις επόμενες ενότητες, αναλόγως του τύπου της κατανομής που επιλέγουμε. Η μέθοδος που θ' αναλύσουμε στη συνέχεια βασίζεται στη μεγιστοποίηση της πιθανοφάνειας (Maximum Likelihood) του μοντέλου. Η πιθανοφάνεια ενός μοντέλου Θ δοθέντων των παρατηρήσεων (δεδομένα) $X = \{x_1, x_2, \dots, x_N\}$ ορίζεται ως:

$$L = p(X; \Theta) = p(X).$$

Ποιοτικά, είναι ένα μέτρο του πόσο καλή εκτίμηση των παραμέτρων έχουμε, ή με άλλα λόγια εκφράζει την πιθανότητα το εκτιμηθέν μοντέλο να παράγει τις παρατηρήσεις X . Θυμίζουμε ότι $\Theta = \{P(j), \theta_j | j = 1, 2, \dots, K\}$. Τελικά, λόγω των σχέσεων 2.1 και 2.1 ισχύει:

$$L = p(X; \Theta) = p(X) = \prod_{i=1}^N p(x_i) = \prod_{i=1}^N \sum_{j=1}^K P(j)p(x_i|\theta_j) \quad (2.6)$$

Ας δούμε τώρα πώς μπορεί να μεγιστοποιηθεί η ποσότητα αυτή. Νωρίτερα στην ενότητα αυτή ορίσαμε την κρυμμένη μεταβλητή Z . Στην περίπτωση που γνωρίζαμε εκ των προτέρων σε ποια κατηγορία ανήκει το κάθε πρότυπο, τότε η μεγιστοποίηση της πιθανοφάνειας θα κατέληγε στις παρακάτω παραμέτρους, που είναι και οι βέλτιστες:

- $P(j) = \frac{|J|}{N}$, όπου $|J|$ είναι το πλήθος των προτύπων της κατηγορίας j .
- θ_j θα είναι οι παράμετροι της κάθε κατανομής.

Η πιθανοφάνεια που προκύπτει με τη χρήση της μεταβλητής Z λέγεται πλήρης πιθανοφάνεια γιατί για τόν υπολογισμό της χρησιμοποιήσαμε τόσο τις παρατηρήσεις όσο και την κρυφή γνώση που υπάρχει στα δεδομένα Z .

Όμως σε ένα πρόβλημα ομαδοποίησης η τιμή των δεδομένων Z είναι το ζητούμενο και επομένως δεν είναι διαθέσιμη κατά την διαδικασία της εκπαίδευσης. Άρα θα πρέπει να αναζητήσουμε έναν

αλγόριθμο που θα υπολογίζει ένα μέγιστο της πιθανοφάνειας χωρίς την κρυμμένη γνώση. Αφού η αντικειμενική συνάρτηση είναι γνωστή πλέον το πρόβλημα έχει αναχθεί σε πρόβλημα βελτιστοποίησης. Δεδομένου ότι ισχύει $\max\{f(x)\} = \min\{-f(x)\}$, μπορούμε αντί για μεγιστοποίηση της πιθανοφάνειας, να ελαχιστοποιήσουμε την ποσότητα $E = -\ln(L)$, που ονομάζεται αρνητική λογαριθμική πιθανοφάνεια, μιάς και η λογαριθμική συνάρτηση είναι αύξουσα. Ο λόγος της χρήσης του λογαρίθμου είναι γιατί διευκολύνει τις πράξεις, αφού πλέον τα γινόμενα στην πιθανοφάνεια γίνονται αθροίσματα.

$$E = -\ln(L) = -\sum_{i=1}^N \ln(p(x_n)) = -\sum_{i=1}^N \ln \left\{ \sum_{j=1}^K P(j)p(x^n|j) \right\} \quad (2.7)$$

Η μέθοδος που θα εφαρμόσουμε για την εκτέλεση της ελαχιστοποίησης είναι ο αλγόριθμος Expectation-Maximization - EM, ένας αρκετά διαδεδομένος αλγόριθμος στο πεδίο της Μηχανικής Μάθησης. Πριν προχωρήσουμε στη μελέτη του αλγορίθμου να σημειώσουμε ότι η συνάρτηση πιθανοφάνειας, είναι μία πλήρως παραγωγίσιμη συνάρτηση, και επομένως μπορεί να εφαρμοσθεί οποιοσδήποτε αναλυτικός ή αριθμητικός αλγόριθμος ελαχιστοποίησης. Ωστόσο η χρήση του EM στο συγκεκριμένο πρόβλημα ενδείκνυται γιατί είναι αρκετά απλός και εύκολα υλοποιήσιμος, δεν απαιτεί υπολογισμό παραγώγων και συνήθως συγκλίνει πιο γρήγορα στη λύση [7].

Ο EM εκτελείται σε δύο φάσεις, εξού και το όνομά του. Στην πρώτη φάση υπολογίζεται η υπό συνθήκη πιθανότητα $p(z_j = 1|x)$. Η ποσότητα αυτή καλείται και υπευθυνότητα (*responsibility*) [8] και είναι μία εκτίμηση της πιθανότητας το πρότυπο x να ανήκει στην ομάδα j . Στο εξής θα ισχύει $p(z_j = 1|x) = p(j|x)$. Στη δεύτερη φάση γίνεται ενημέρωση των παραμέτρων με τέτοιο τρόπο έτσι ώστε να μεγιστοποιείται η πιθανοφάνεια (ή κατ'αντιστοιχία, να ελαχιστοποιείται η αρνητική λογαριθμική πιθανοφάνεια). Η διαδικασία επαναλαμβάνεται μέχρις ότου ένα προκαθορισμένο κριτήριο τερματισμού εκπληρωθεί (για παράδειγμα αν εκτελεστεί ένα μέγιστο πλήθος επαναλήψεων ή αν ο ρυθμός μείωσης της αρνητικής λογαριθμικής πιθανοφάνειας γίνει μικρότερος από μία τιμή κατωφλίου). Με απλά λόγια η λογική του EM αλγορίθμου είναι να εκτιμηθεί μία κατηγοριοποίηση των παρατηρήσεων και με βάση αυτή να υπολογιστούν εκείνες οι παράμετροι των κατανομών που είναι πιο πιθανές, δηλαδή προσεγγίζουν το πραγματικό μοντέλο καλύτερα. Μία σύντομη περιγραφή του EM αποτελεί ο αλγόριθμος (1).

2.2 Μικτές Κανονικές Κατανομές

Στην προηγούμενη ενότητα αναφερθήκαμε γενικά στο Μοντέλο των Μικτών Κατανομών, όπου και θέσαμε ως βασική προϋπόθεση τον καθορισμό του τύπου της κατανομής που θα χρησιμοποιήσουμε, προκειμένου να είναι εφικτή η χρήση του μοντέλου. Στην ενότητα αυτή θα αναφερθούμε στο Μοντέλο Μικτών Κανονικών Κατανομών, όπου όπως φαίνεται και από τον τίτλο του μοντέλου, οι επιμέρους κατανομές που θα χρησιμοποιήσουμε είναι Κανονικές.

Πριν προχωρήσουμε όμως στην μελέτη του συγκεκριμένου μοντέλου, θυμίζουμε στον αναγνώστη τον τύπο της Κανονικής Κατανομής. Έτσι λοιπόν μία τυχαία μεταβλητή $X = x, x \in R^d$ που ακολουθεί Κανονική Κατανομή με μέση τιμή μ και πίνακα συμμεταβλητότητας Σ , δηλαδή $X \sim N(\mu, \Sigma)$ έχει συνάρτηση πυκνότητας πιθανότητας που δίνεται από τον τύπο:

$$p(x) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp \left\{ -\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu) \right\} \quad (2.8)$$

EM ($X, K, \Theta^0, \epsilon, \max$)

X : Το σύνολο του οποίου την κατανομή θέλουμε να εκτιμήσουμε

K : Το πλήθος των κατανομών που χρησιμοποιούμε στο Μοντέλο Μικτών Κατανομών.

ϵ : ελάχιστη μεταβολή στην πιθανοφάνεια μεταξύ επαναλήψεων.

Θ^0 : τυχαία αρχικοποίηση των παραμέτρων. $\Theta = \{P(j), \theta_j\} j = 1, 2, \dots, K$

\max : μέγιστος αριθμός επαναλήψεων.

• **while** ($k \leq \max$ και $\Delta L \geq \epsilon$)

– Expectation

* Υπολογισμός της αναμενόμενης τιμής της υπό συνθήκης πιθανότητας $p(j|x)$, ανάλογα με τον τύπο κατανομής που έχουμε επιλέξει, και με βάση τον κανόνα του Bayes, σχέση 2.2.

– Maximization

* Υπολογισμός των νέων τιμών των παραμέτρων, Θ_{new} , έτσι ώστε να μεγιστοποιείται η πιθανοφάνεια (ή αντίστοιχα να ελαχιστοποιείται η αρνητική λογαριθμική πιθανοφάνεια).

– $\Theta_{old} \leftarrow \Theta_{new}$

– $k := k + 1$

• **end while**

Αλγόριθμος 1: Σύντομη απόδοση του αλγορίθμου EM

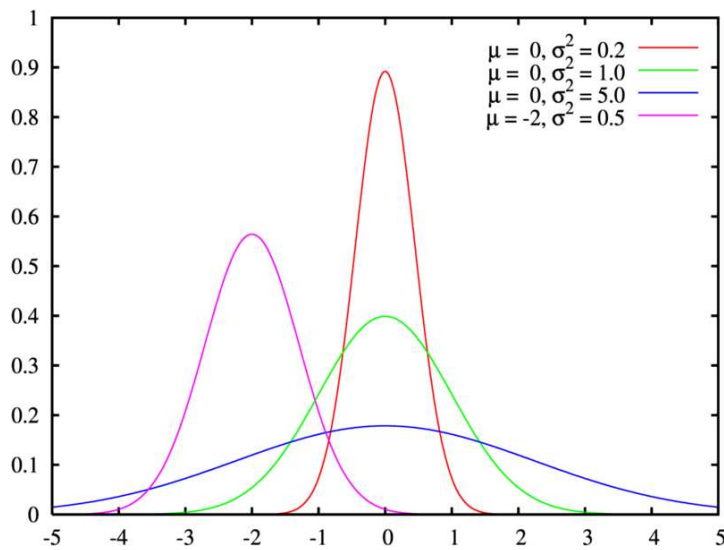
Η Κανονική Κατανομή πολλές φορές καλείται και Γκαουσιανή (Gaussian), από το όνομα του μεγάλου Γερμανού μαθηματικού Carl Friedrich Gauss, ο οποίος και τη χρησιμοποίησε στη μελέτη αστρονομικών δεδομένων. Από τον ορισμό της Κανονικής Κατανομής, φαίνεται ότι οι παράμετροί της είναι η μέση τιμή μ και ο πίνακας συμμεταβλητότητας Σ .

Η ιδανική περίπτωση, οπότε και το προτεινόμενο Μικτό Μοντέλο είναι το βέλτιστο, προκύπτει όταν ο πίνακας συμμεταβλητότητας Σ είναι πλήρης. Δεδομένης όμως της υπολογιστικής δυσκολίας (υπολογισμός ορίζουσας, αντιστροφή, η οποία ενδεχομένως να μην είναι και εφικτή) που έχει ή εκπαίδευση ενός τέτοιου μοντέλου, ειδικά σε μεγάλες διαστάσεις και μεγάλο όγκο δεδομένων, ακολουθούμε απλούστερες μορφές για τον ορισμό του Σ , οπότε γίνεται απλούστερη και η συνάρτηση πυκνότητας πιθανότητας (2.8). Στον πίνακα (2.2) περιγράφονται οι διάφορες μορφές του πίνακα Σ και ο αντίστοιχος τύπος που παίρνει η συνάρτηση πυκνότητας πιθανότητας της Κανονικής κατανομής.

Στο σχήμα (2.4) φαίνεται το αποτέλεσμα της κατηγοριοποίησης ενός συνόλου δισδιάστατων δεδομένων, για τις διάφορες μορφές του πίνακα συμμεταβλητότητας. Παρατηρήστε το γεωμετρικό ανάλογο που προκύπτει για την κάθε επιλογή. Ο σφαιρικός πίνακας συμμεταβλητότητας ομαδοποιεί τα δεδομένα σχεδιάζοντας σφαίρες με κέντρο τη μέση τιμή μ της κάθε κατανομής και ακτίνα τη διακύμανση σ . Αντίθετα στις περιπτώσεις του διαγώνιου και του πλήρους πίνακα Σ η ομαδοποίηση συνίσταται στο σχεδιασμό ελλείψεων με κέντρο τη μέση τιμή μ της κάθε κατανομής και άξονες ανάλογους προς τα ιδιοδιανύσματα και τις ιδιοτιμές του πίνακα Σ . Η μόνη διαφορά έγκυται στον προσανατολισμό της έλλειψης. Έτσι στην περίπτωση του διαγώνιου πίνακα συμμεταβλητότητας, οι άξονες της έλλειψης είναι παράλληλοι προς τους κύριους άξονες συντεταγμένων, ενώ στην περίπτωση

μορφή πίνακα Σ	τύπος Κανονικής κατανομής
σφαιρικός $\Sigma = \sigma^2 I$	$p(x) = \frac{1}{(2\pi\sigma^2)^{d/2}} \exp \left\{ -\frac{\ x-\mu\ ^2}{2\sigma^2} \right\}$
διαγώνιος $\Sigma = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_d^2)$	$p(x) = \frac{1}{(2\pi)^{\frac{d}{2}} (\prod_{i=1}^d \sigma_i)^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu) \right\}$
πλήρης $\Sigma \in R^{d \times d}$ συμμετρικός και θετικά ορισμένος	$p(x) = \frac{1}{(2\pi)^{\frac{d}{2}} \Sigma ^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu) \right\}$

Σχήμα 2.2: Ο τύπος της συνάρτησης πυκνότητας πιθανότητας της κανονικής κατανομής όπως απλοποιείται για τις διάφορες μορφές του πίνακα συμμεταβλητότητας Σ



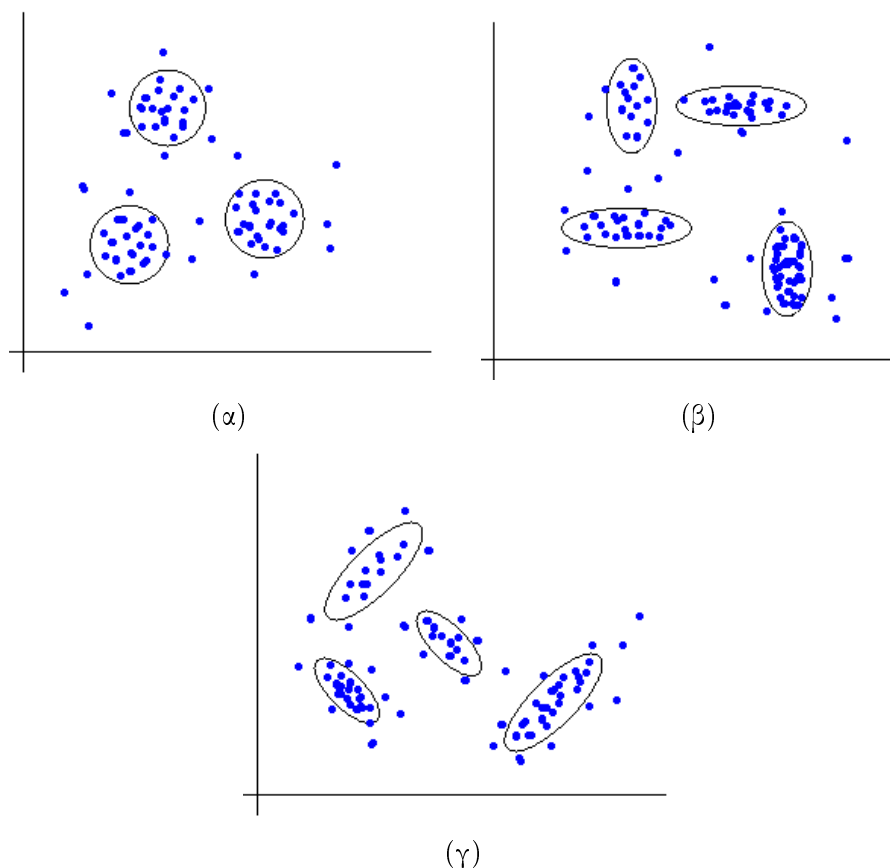
Σχήμα 2.3: Γραφική παράσταση συνάρτησης πυκνότητας πιθανότητας Κανονικής Κατανομής. Παρατηρείστε το κωδωνοειδές σχήμα της. Από τη θεωρία ισχύει ότι στο διάστημα $[\mu - 2\sigma, \mu + 2\sigma]$ ανήκει το 98% των δεδομένων. Επιπλέον όσο μεγαλύτερη είναι η διασπορά τόσο πιο πολύ απλώνονται τα άκρα της κατανομής.

του πλήρους πίνακα συμμεταβλητότητας, η έλλειψη μπορεί να έχει οποιοδήποτε προσανατολισμό στο χώρο.

Καθένα από τα παραπάνω μοντέλα μπορούν να προσεγγίσουν σε οποιοδήποτε επιθυμητό βαθμό οποιαδήποτε συνάρτηση πυκνότητας πιθανότητας, δεδομένου ότι υπάρχει ικανοποιητικός αριθμός από κατανομές (McLachlan and Basford, 1988). Είναι αυτονόητο ότι ανάλογα με το βαθμό πολυπλοκότητας του μοντέλου, θα χρειάζονται και λιγότερες κατανομές για να επιτύχουμε την επιθυμητή προσέγγιση. Έτσι το μοντέλο με πλήρη Σ θα απαιτήσει λιγότερες κατανομές, θα έχει όμως περισσότερες παραμέτρους που θα πρέπει να καθοριστούν.

Αφού ορίσαμε τον τύπο των κατανομών, μπορούμε να προχωρήσουμε στην εκπαίδευσή τους, δηλαδή στον προσδιορισμό των παραμέτρων τους. Στην ουσία θα προσαρμόσουμε την γενική εκδοχή του EM αλγορίθμου (αλγόριθμος 1) στην περίπτωση των Μικτών Κανονικών Κατανομών. Το αποτέλεσμα είναι ο αλγόριθμος 2 για εκπαίδευση ενός Μοντέλου Μικτών Κανονικών Κατανομών.

Υπάρχουν δύο τεχνικές δυσκολίες που προκύπτουν κατά την ελαχιστοποίηση της αρνητικής



Σχήμα 2.4: Παραδείγματα κατηγοριοποίησης δεδομένων με χρήση Μοντέλου Μικτών Κανονικών Κατανομών. (α) Σφαιρικός πίνακας συμμεταβλητότητας, (β) Διαγώνιος πίνακας συμμεταβλητότητας, (γ) Πλήρης πίνακας συμμεταβλητότητας.

λογαριθμικής πιθανοφάνειας. Πρώτον η συνάρτηση E έχει ολικό ελάχιστο στο $-\infty$, η περίπτωση που μία κατανομή ταυτίζεται με ένα σημείο x_i του συνόλου εκπαίδευσης, οπότε θα ισχύει $\mu = x_i, \Sigma = O$. Το πρόβλημα λύνεται με το να ελέγχουμε κάθε φορά τον πίνακα Σ για πολύ μικρές τιμές και αν τις εντοπίσουμε να τις αντικαθιστούμε με μεγάλες. Το δεύτερο πρόβλημα είναι ότι η αντικειμενική συνάρτηση έχει πολλά τοπικά ελάχιστα, που οδηγούν σε κακή προσέγγιση της συνάρτησης πυκνότητας πιθανότητας. Η συνηθισμένη πρακτική για αποφυγή του προβλήματος είναι να εκκινούμε τον αλγόριθμο κάθε φορά από διαφορετικό σημείο και να κρατούμε την καλύτερη λύση στο τέλος. Ως ένα επιπλέον βήμα για εύρεση καλύτερων ελαχίστων, είναι ν' αρχικοποιούμε τις παραμέτρους θ^0 χρησιμοποιώντας τον αλγόριθμο K-Κέντρων, σαν ένα στάδιο προεπεξεργασίας.

2.3 Μικτές Student's t - Κατανομές

Είδαμε στην προηγούμενη ενότητα πώς λειτουργεί ένα Μικτό Μοντέλο Κανονικών Κατανομών. Το μοντέλο αυτό είναι αρκετά διαδεδομένο, έχει όμως ένα βασικό μειονέκτημα. Είναι αρκετό ευαίσθητο σε δεδομένα με θόρυβο, και μάλιστα ακόμα και η παραμικρή εμφάνιση αποκλινόντων σημείων (outliers), έχει ως αποτέλεσμα την μετατόπιση της έλλειψης που χαρακτηρίζει την ομάδα, προς τα αποκλινόντα σημεία, μειώνοντας έτσι την αξιοπιστία του μοντέλου. Χαρακτηριστικά παραδείγματα

GMM-EM ($X, K, \Theta^0, \epsilon, max$)

X : Το σύνολο του οποίου την κατανομή θέλουμε να εκτιμήσουμε

K : Το πλήθος των κατανομών που χρησιμοποιούμε στο Μοντέλο Μικτών Κατανομών.

Θ^0 : τυχαία αρχικοποίηση των παραμέτρων.

ϵ : ελάχιστη μεταβολή στην πιθανοφάνεια μεταξύ επαναλήψεων.

max : μέγιστος αριθμός επαναλήψεων.

- $E_{old} = 0, E_{new} = +\infty$

- **while** ($k \leq max$ και $|E_{new} - E_{old}| \geq \epsilon$)

- $p(i|x_j)^{k+1} = z_{ij}^{k+1} = \frac{P(j)N(x_j|\mu_j, \Sigma_j)}{\sum_{l=1}^K P(l)N(x_j|\mu_l, \Sigma_l)}$ όπου $N(x|\mu, \Sigma)$ είναι η συνάρτηση πυκνότητας πιθανότητας της Κανονικής κατανομής (βλ. πίνακα 2.2).

- Ενημέρωση των παραμέτρων $\Theta = \{(P(j), \mu_j, \Sigma_j)\}, j = 1 \dots K$:

- * $P_j^{(k+1)} = \frac{1}{N} \sum_{n=1}^N P^{(k)}(\theta_j|x^n)$

- * $\mu_j^{(k+1)} = \frac{\sum_{n=1}^N P^{(k)}(\theta_j|x^n)x^n}{\sum_{n=1}^N P^{(k)}(\theta_j|x^n)}$

- * $\Sigma_j^{(k+1)} = \frac{1}{d} \frac{\sum_{n=1}^N P^{(k)}(\theta_j|x^n)(x^n - \mu_j^{(k+1)})(x^n - \mu_j^{(k+1)})^T}{\sum_{n=1}^N P^{(k)}(\theta_j|x^n)}$

- $E_{old} = E_{new}$

- $E_{new} = -\ln \left(\sum_{i=1}^N \sum_{j=1}^K P^{k+1}(j)N(x_i|\mu_j^{k+1}, \Sigma_j^{k+1}) \right)$

- $k := k + 1$

- **end while**

Αλγόριθμος 2: Αλγόριθμος EM για εκπαίδευση Μικτών Κανονικών Κατανομών.

αυτής της συμπεριφοράς δίνονται στο σχήμα 2.5. Λύση στο πρόβλημα αυτό μπορεί ν' αποτελέσει η Student's t -κατανομή, η οποία παρουσιάζει ανθεκτικότητα στην παρουσία σημείων που αποκλίνουν από το μοντέλο (*outliers*). Επιπλέον η μίξη Student's t -κατανομών, μπορεί να χρησιμοποιηθεί και στην περίπτωση που υπάρχουν ελλιπή δεδομένα στο σύνολο εκπαίδευσης (*missing data problem*). Στη συνέχεια δίνεται η κατηγοριοποίηση που γίνεται με χρήση πλέον μικτών Student's t -κατανομών, στα ίδια σύνολα που χρησιμοποιήσαμε και προηγουμένως (σχήμα 2.5) τις Κανονικές Κατανομές. Μία εξήγηση αυτής της ευρωστίας της Student's t -κατανομής μπορεί να γίνει κατανοητή από τη γραφική παράσταση του σχήματος 2.7 των συναρτήσεων πυκνότητας πιθανότητας για μία Κανονική και μία Student's t -Κατανομή. Παρατηρείστε πώς προσαρμόζονται τα άκρα της Student's t -κατανομής για να μοντελοποιήσει τα αποκλίνοντα σημεία. Επιπλέον, όπως φαίνεται και στο σχήμα 2.8 ανάλογα με την τιμή της παραμέτρου ν η καμπύλη της γραφικής παράστασης απλώνει τα άκρα της για να μοντελοποιήσει τα αποκλίνοντα σημεία, αντί να μετακινήσει τη μέση τιμή της. Όσο μικρότερη είναι η τιμή του ν τόσο πιο θορυβώδη είναι τα δεδομένα, άρα υπάρχουν πολλές αποκλίσεις από το Γκαουσιανό μοντέλο περιγραφής τους.

Όπως πράξαμε και στις μικτές κατανομές, προτού προχωρήσουμε στη μελέτη του μικτού μοντέλου Student's t -κατανομών, θα αναφερθούμε στην ίδια την Student's t -κατανομή. Έτσι λοιπόν όταν λέμε ότι μία p -διάστατη τυχαία μεταβλητή Y ακολουθεί την Student's t -κατανομή δηλαδή $Y \sim T(\mu, \Sigma, \nu)$, εννοούμε ότι, πρώτον δοθέντος ενός διανύσματος βαρών τ , η Y ακολουθεί μία

πολυδιάστατη κανονική κατανομή με μέση τιμή μ και πίνακα συμμεταβλητότητας Σ και δεύτερον η τυχαία μεταβλητή νT ακολουθεί την κατανομή χ^2_ν με ν βαθμούς ελευθερίας [12]. Πιο συγκεκριμένα:

- $Y|\mu, \Sigma, \nu \sim N(\mu, \Sigma/\tau)$
- $\tau|\mu, \Sigma, \nu \sim \text{Gamma}(\nu/2, \nu/2)$, όπου η κατανομή $\text{Gamma}(\alpha, \beta)$ έχει συνάρτηση πυκνότητας πιθανότητας

$$p(x|\alpha, \beta) = \beta^\alpha x^{\alpha-1} \frac{\exp\{-\beta x\}}{\Gamma(\alpha)}, \quad x > 0, \alpha > 0, \beta > 0$$

ενώ $\Gamma(\cdot)$ είναι η συνάρτηση Γάμμα, $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$.

Ολοκληρώνοντας ως προς τ την από κοινού συνάρτηση πυκνότητας πιθανότητας, παίρνουμε την περιθώρια κατανομή. Η συνάρτηση πυκνότητας πιθανότητας της Student's t -κατανομής δίνεται από τη σχέση:

$$p(Y = (y_1, y_2, \dots, y_p)^T | \mu, \Sigma, \nu) = \frac{\Gamma(\frac{\nu+p}{2}) |\Sigma|^{-1/2}}{(\pi\nu)^{p/2} \Gamma(\frac{\nu}{2}) [1 + (Y - \mu)^T \Sigma^{-1} (Y - \mu)]^{(\nu+p)/2}} \quad (2.9)$$

όπου η ποσότητα $(Y - \mu)^T \Sigma^{-1} (Y - \mu)$ είναι η τετραγωνική απόσταση Mahalanobis.

Οι παράμετροι της κατανομής είναι η μέση τιμή μ , ο πίνακας συμμεταβλητότητας Σ και οι βαθμοί ελευθερίας ν . Για την παράμετρο ν μπορεί να δειχθεί ότι καθώς $\nu \rightarrow \infty$, $\tau \rightarrow 1$ με πιθανότητα 1, οπότε η Y γίνεται οριακά κανονική, $Y \sim N(\mu, \Sigma)$.

Η έννοια της παραμέτρου ν σχετίζεται με την ποσότητα θορύβου που υπάρχει στα δεδομένα. Είναι δηλαδή ένα μέτρο της απόκλισης των τελευταίων από το μοντέλο που τα περιγράφει. Όσο πιο μικρή τιμή έχει η παράμετρος ν τόσο πιο πολύ θόρυβος υπάρχει στα δεδομένα. Εδώ με την έννοια του θορύβου εννοούμε την απόκλιση από το μοντέλο περιγραφής των δεδομένων.

Ας δούμε τώρα πώς εκπαιδεύεται ένα μικτό μοντέλο Student's t -κατανομών, αλγόριθμος 3. Όπως και στις μικτές κανονικές κατανομές θα παρουσιάσουμε την αντίστοιχη εκδοχή του EM [12].

2.4 Αποστάσεις Μεταξύ Κατανομών

Όπως θα έχει γίνει κατανοητό από την ανάγνωση της παρούσας εργασίας, βασικό ρόλο στην επιτυχία της προτεινόμενης μεθόδου παίζει ο τρόπος με τον οποίο θα μπορέσουμε να αξιολογήσουμε τις διάφορες κατανομές των δύο επεξεργαζόμενων εικόνων. Με άλλα λόγια θα πρέπει να οριστεί ένα μέτρο που θα εκτιμά πόσο κοντά ή μακριά είναι μία κατανομή από μία άλλη. Δηλαδή απαιτείται ο ορισμός μίας απόστασης μεταξύ κατανομών. Στην ενότητα αυτή θα γίνει λόγος για το μέτρο αυτό που χρησιμοποιήσαμε, και ειδικότερα θα παρουσιάσουμε ορισμένες ευρέως διαδεδομένες συναρτήσεις απόστασης κατανομών.

2.4.1 Απόσταση Chernoff

Η απόσταση Chernoff μεταξύ δύο Κανονικών Κατανομών $N_1(\mu_1, \Sigma_1)$ και $N_2(\mu_2, \Sigma_2)$ ορίζεται ως:

$$\mu(s) = \frac{s(1-s)}{2} (\mu_2 - \mu_1)^T [s\Sigma_1 + (1-s)\Sigma_2]^{-1} (\mu_2 - \mu_1) + \frac{1}{2} \ln \frac{|s\Sigma_1 + (1-s)\Sigma_2|}{|\Sigma_1|^s |\Sigma_2|^{1-s}} \quad (2.10)$$

όπου $s \in (0, 1]$.

2.4.2 Απόσταση Battacharyya

Η απόσταση Battacharyya προκύπτει από την Chernoff αν θέσουμε $s = 0,5$, δηλαδή

$$\mu(0,5) = \frac{1}{8}(\mu_2 - \mu_1)^T \left[\frac{\Sigma_1 + \Sigma_2}{2} \right]^{-1} (\mu_2 - \mu_1) + \frac{1}{2} \ln \frac{|\frac{\Sigma_1 + \Sigma_2}{2}|}{\sqrt{|\Sigma_1| |\Sigma_2|}} \quad (2.11)$$

Να σημειωθεί ότι δεδομένης της προέλευσης της Battacharyya από την απόσταση Chernoff, είναι λογικό να ισχύει η πρώτη ως μέτρο σύγκρισης Κανονικών Κατανομών. Ωστόσο μπορεί να χρησιμοποιηθεί και για μη Κανονικές Κατανομές, όμως όσο πιο πολύ αποκλίνουν οι κατανομές από μία Κανονική τόσο λιγότερη πληροφορία υπάρχει στο επιστρεφόμενο αποτέλεσμα της Battacharyya.

Μπορούμε να επεκτείνουμε τη χρήση της Battacharyya απόστασης ως μέτρο σύγκρισης όχι μόνο δύο κατανομών αλλά και δύο Μικτών Κατανομών. Με άλλα λόγια αν έχουμε δύο Μικτές Κανονικές Κατανομές $M_1 = \{a_i, \mu_i, \Sigma_i\}$, $i = 1 \dots N$ και $M_2 = \{b_j, \nu_j, \Gamma_j\}$, $j = 1 \dots M$, τότε η απόσταση D των δύο μοντέλων μπορεί να οριστεί ως:

$$D(M_1, M_2) = \sum_{i=1}^N \sum_{j=1}^M a_i b_j \text{Batt}(\{\mu_i, \Sigma_i\}, \{\nu_j, \Gamma_j\})$$

όπου $\text{Batt}(\{\mu, \Sigma\}, \{\nu, \Gamma\})$ είναι η Battacharyya απόσταση μεταξύ των κατανομών με μέσες τιμές μ, ν και πίνακες συμμεταβλητότητας Σ, Γ αντίστοιχα.

2.4.3 Τετραγωνική L_2 απόσταση

Η ποσότητα

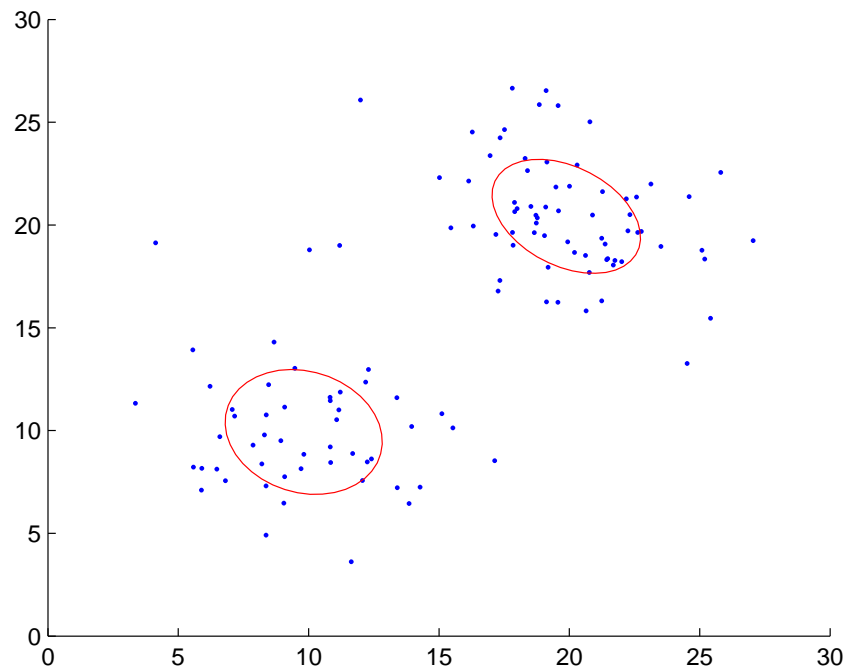
$$D(f_1, f_2) = \int (f_1(x) - f_2(x))^2 dx = \int (f_1^2(x) - 2f_1(x)f_2(x) + f_2^2(x)) dx$$

καλείται τετραγωνική απόσταση των κατανομών f_1, f_2 . Στην περίπτωση όπου οι κατανομές f_1, f_2 είναι κανονικές ισχύει, $f_1(x)f_2(x) = \phi(0)$ όπου η ϕ είναι επίσης μία κανονική κατανομή με μέση τιμή $\mu_\phi = \mu_{f_1} - \mu_{f_2}$ και πίνακα συμμεταβλητότητας $\Sigma_\phi = \Sigma_{f_1} + \Sigma_{f_2}$.

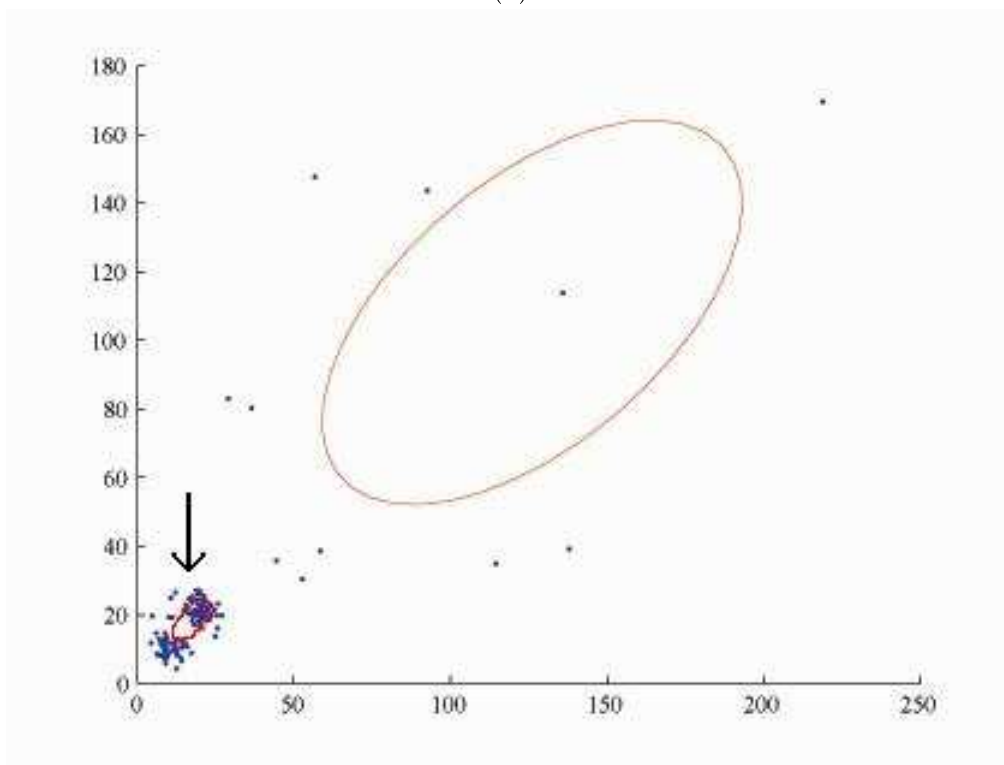
Μπορούμε επίσης να χρησιμοποιήσουμε την L_2 απόσταση για να υπολογίσουμε αποστάσεις μεταξύ Μικτών Κανονικών Κατανομών. Ακολουθώντας τη σημειογραφία της προηγούμενης ενότητας, η απόσταση D ορίζεται ως:

$$D(M_1, M_2) = \sum_{i=1}^N \sum_{j=1}^M a_i b_j L_2(\{\mu_i, \Sigma_i\}, \{\nu_j, \Gamma_j\})$$

όπου $L_2(\{\mu, \Sigma\}, \{\nu, \Gamma\})$ είναι η L_2 απόσταση μεταξύ των κατανομών με μέσες τιμές μ, ν και πίνακες συμμεταβλητότητας Σ, Γ αντίστοιχα.

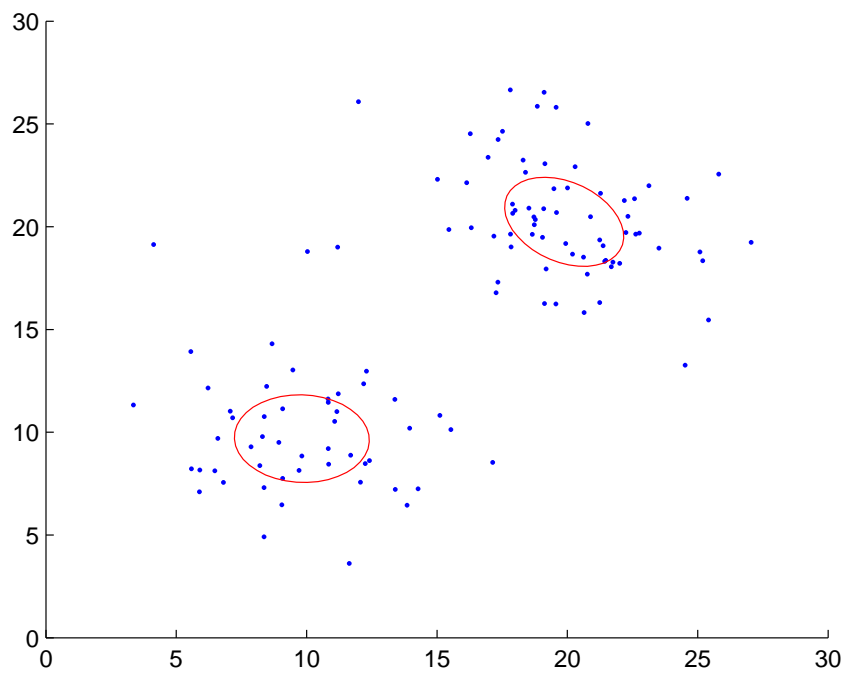


(α)

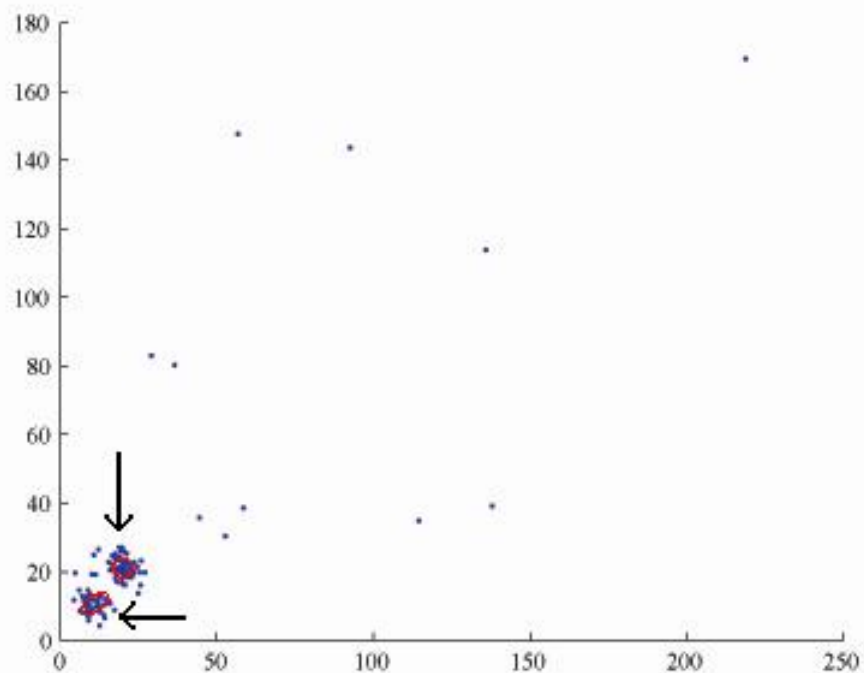


(β)

Σχήμα 2.5: Κατηγοριοποίηση σημείων με Μιχτές Κανονικές Κατανομές, (α) σημεία χωρίς θόρυβο, (β) τα σημεία του σχήματος (α) με προσθήκη θορύβου. Παρατηρήστε την έντονη επιρροή των αποκλιόντων σημείων που μεταφράζεται σε μετατόπιση των κέντρων, αλλά και σε διεύρυνση της έλλειψης ώστε να καλύψει και τα σημεία θορύβου.

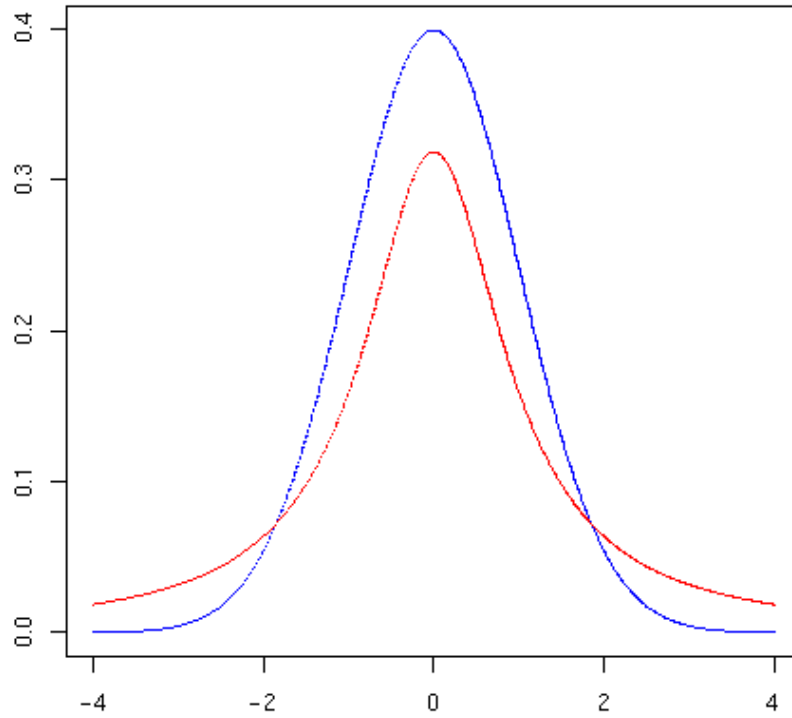


(α)

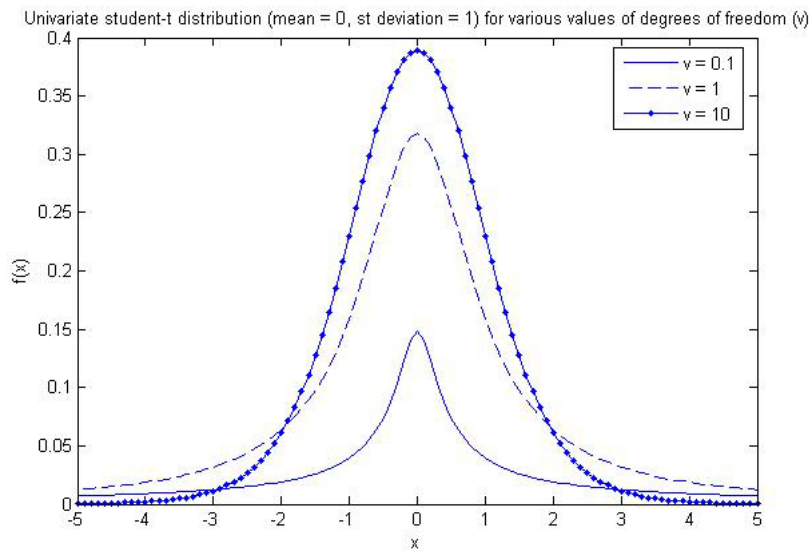


(β)

Σχήμα 2.6: Κατηγοριοποίηση σημείων με μικτές Student's t -Κατανομές, (α) σημεία χωρίς θόρυβο, (β) τα σημεία του σχήματος (α) με προσθήκη θορύβου. Σε αντίθεση με τις Κανονικές Κατανομές η Student's t -Κατανομή είναι πιο ανθεκτική. Το αποτέλεσμα είναι ένα πιο αξιόπιστο μοντέλο.



Σχήμα 2.7: Γραφική παράσταση της συνάρτησης πυκνότητας πιθανότητας Κανονικής (κυανό) και Student's t -Κατανομής (κόκκινο).



Σχήμα 2.8: Γραφική παράσταση της συνάρτησης πυκνότητας πιθανότητας της Student's t -Κατανομής για διάφορες τιμές της παραμέτρου ν . Παρατηρήστε ότι καθώς το ν μεγαλώνει η καμπύλη προσεγγίζει αυτή της Κανονικής Κατανομής.

tMM-EM ($X, K, \Theta^0, \epsilon, max$)

X : Το σύνολο του οποίου την κατανομή θέλουμε να εκτιμήσουμε

K : Το πλήθος των κατανομών που χρησιμοποιούμε στο Μοντέλο Μικτών Κατανομών.

ϵ : ελάχιστη μεταβολή στην πιθανοφάνεια μεταξύ επαναλήψεων.

Θ^0 : τυχαία αρχικοποίηση των παραμέτρων.

max : μέγιστος αριθμός επαναλήψεων.

• $E_{new} = +\infty, E_{old} = 0$

• **while** ($k \leq max$ και $|E_{new} - E_{old}| \geq \epsilon$)

$$- p(i|x_j) = z_{ij}^{k+1} = \frac{\pi_i^k p(x_j; \mu_i^k, \Sigma_i^k, \nu_i^k)}{\sum_{m=1}^K p(x_j; \mu_m^k, \Sigma_m^k, \nu_m^k)}$$

$$- \tau_{ij}^{k+1} = \frac{\nu_i^k + d}{\nu_i^k + (x_j - \mu_i^k)^T \Sigma_i^{-1} (x_j - \mu_i^k)}$$

- Ενημέρωση των παραμέτρων:

$$* P_i^{(k+1)} = \frac{1}{N} \sum_{n=1}^N P^{(k)}(i|x^n)$$

$$* \mu_i^{k+1} = \frac{\sum_{j=1}^N z_{ij}^k \tau_{ij}^k x_j}{\sum_{j=1}^N z_{ij}^k \tau_{ij}^k}$$

$$* \Sigma_i^{k+1} = \frac{\sum_{j=1}^N z_{ij}^k \tau_{ij}^k (x_j - \mu_i^{k+1})(x_j - \mu_i^{k+1})^T}{\sum_{j=1}^N z_{ij}^{k+1}}$$

* Οι βαθμοί ελευθερίας στην επανάληψη ($k + 1$) προκύπτουν από την επίλυση της παρακάτω εξίσωσης:

$$\log\left(\frac{\nu_i^{k+1}}{2}\right) - \psi\left(\frac{\nu_i^{k+1}}{2}\right) + 1 - \log\left(\frac{\nu_i^k + d}{2}\right) + \frac{\sum_{j=1}^N z_{ij}^k (\log \tau_{ij}^k - \tau_{ij}^k)}{\sum_{j=1}^N z_{ij}^k} + \psi\left(\frac{\nu_i^k + d}{2}\right) = 0$$

- $E_{old} = E_{new}$

- $E_{new} = -\ln\left(\sum_{i=1}^N \sum_{j=1}^K P^{k+1}(j) t(x_i | \mu_j^{k+1}, \Sigma_j^{k+1}, \nu_j^{k+1})\right)$

- $k := k + 1$

• **end while**

Αλγόριθμος 3: Αλγόριθμος EM για εκπαίδευση Μικτών Student's t -Κατανομών.

ΚΕΦΑΛΑΙΟ 3

ΥΠΕΡΘΕΣΗ ΕΙΚΟΝΑΣ ΜΕ ΜΙΚΤΕΣ ΚΑΤΑΝΟΜΕΣ

-
- 3.1 Περιγραφή της μεθόδου
 - 3.2 Πειραματικά αποτελέσματα
 - 3.3 Ανάλυση αποτελεσμάτων
-

3.1 Περιγραφή της μεθόδου

Έχοντας καλύψει την θεωρία πάνω στην οποία βασίζεται η μέθοδός μας, στην ενότητα αυτή θα προχωρήσουμε στην περιγραφή της.

Η βασική ιδέα είναι το γεγονός ότι αν οι εικόνες περιγράφονται από κάποιες κατανομές, τότε στην περίπτωση της τέλει υπέρθεσης θα έχουμε την ελάχιστη δυνατή απόσταση μεταξύ των κατανομών αυτών. Με αυτή τη διαπίστωση λοιπόν δεν έχουμε παρά να βρούμε τις κατανομές τις κάθε εικόνας και εν συνεχεία χρησιμοποιώντας κάποια απόσταση ως μέτρο σύγκρισης δύο κατανομών, μπορούμε να εκτιμήσουμε τις παραμέτρους του μετασχηματισμού υπέρθεσης ελαχιστοποιώντας την απόσταση αυτή. Σε γενικές γραμμές ο αλγόριθμος που προτείνουμε είναι ο αλγόριθμος 4.

Υπέρθεση Με Μικτές Κατανομές (I_{ref}, I_{reg})

- I_{ref} : Η αρχική εικόνα, επί της οποίας θα γίνει η υπέρθεση.
- I_{reg} : Η εικόνα την οποία θέλουμε να υπερέσουμε.
- R, T : Οι παράμετροι του μετασχηματισμού υπέρθεσης τις οποίες θα εκτιμήσουμε (R : Διάνυσμα γωνιών περιστροφής, T : Διάνυσμα μετατόπισης).

- Εκτίμηση της κατανομής των δύο εικόνων (Μοντέλα Μικτών Κατανομών - Αλγόριθμος EM).
- Ελαχιστοποίηση της συνολικής απόστασης των δύο κατανομών ως προς τις παραμέτρους R, T του μετασχηματισμού υπέρθεσης.

Αλγόριθμος 4: Αλγόριθμος υπέρθεσης εικόνων με Μικτές Κατανομές

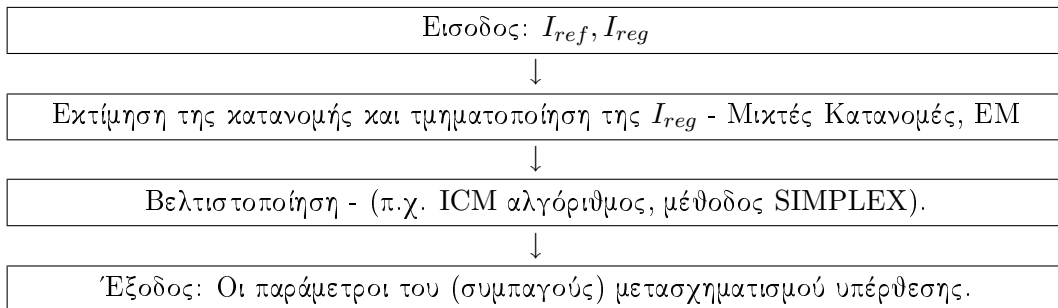
Όμως πώς θα εκτιμήσουμε τις κατανομές στις εικόνες και πώς θα μετρήσουμε τις αποστάσεις μεταξύ των κατανομών; Με βάση λοιπόν τη θεωρία που παραθέσαμε στα προηγούμενα κεφάλαια, αντιλαμβάνεται κανείς ότι η απάντηση στο πρώτο ερώτημα είναι η χρήση των Μοντέλων Μικτών Κατανομών, ενώ για το δεύτερο πρόβλημα η λύση είναι μία από τις αποστάσεις που περιγράφονται στο κεφάλαιο 2. Στα πλαίσια της εργασίας μας χρησιμοποιήσαμε δύο τύπους Μοντέλων Μικτών Κατανομών, (Μίξη Κανονικών και Student's T Κατανομών), ενώ για την εκτίμηση της απόστασης μεταξύ κατανομών χρησιμοποιήσαμε τις αποστάσεις Battacharyya και L2. Όπως θα δείξουμε στην επόμενη ενότητα, τα αποτελέσματα για την κάθε επιλογή έχουν αξιοσημείωτη διαφορά, ειδικά στην περίπτωση θορύβου. Ειδικότερα η συνολική απόσταση μεταξύ ομοίων κατανομών που χρησιμοποιείται στη μέθοδό μας, δίνεται από τον ακόλουθο τύπο:

$$E(I_{ref}(\cdot), I_{reg}(\cdot)) = \sum_{k=1}^K \pi_k B(\theta_k^{ref}, \theta_k^{reg}) \quad (3.1)$$

όπου $B(\theta_1, \theta_2)$ είναι η Battacharyya απόσταση μεταξύ των κατανομών με παραμέτρους θ_1, θ_2 , ενώ $\theta_k^{ref}, \theta_k^{reg}$ είναι αντίστοιχα οι παράμετροι των Μικτών Κατανομών της ομάδας k για τις δύο εικόνες που επιθυμούμε να υπερθέσουμε.

Ένα επιπλέον πρόβλημα που πρέπει να επιλυθεί είναι και αυτό της ελαχιστοποίησης της συνολικής απόστασης μεταξύ των κατανομών των δύο εικόνων. Μπορούν να χρησιμοποιηθούν διάφορες τεχνικές βελτιστοποίησης. Στα πειράματά μας χρησιμοποιήσαμε τη μέθοδο SIMPLEX, που είναι ενσωματωμένη στο πακέτο βελτιστοποίησης του Matlab (fminsearch). Μία αρκετά διαδεδομένη είναι ο ICM αλγόριθμος [9], ο οποίος στη γενική του μορφή είναι ο αλγόριθμος 5.

Έτσι λοιπόν καταλήγουμε στην προτεινόμενη μέθοδο υπέρθεσης εικόνας βασισμένη σε μίξη κατανομών, αλγόριθμος 6.



Σχήμα 3.1: Διάγραμμα ροής αλγορίθμου υπέρθεσης εικόνων με χρήση μικτών κατανομών.

3.2 Πειραματικά αποτελέσματα

Στην ενότητα αυτή θα παρουσιάσουμε ορισμένα πειραματικά αποτελέσματα προκειμένου να δείξουμε τις δυνατότητες της μεθόδου μας. Τα πειράματα που εκτελέσαμε έχουν ως στόχο να συγκρίνουν τη μέθοδό μας με εκείνες που βασίζονται στη μεγιστοποίηση της αμοιβαίας πληροφορίας (βλ. κεφάλαιο 1), ιδιαίτερα παρουσία θορύβου. Να σημειώσουμε στο σημείο αυτό ότι προκειμένου να μην υπάρχει κάποια πόλωση στα αποτελέσματα ως προς το είδος του θορύβου, αποφασίσαμε να διεξάγουμε τα πειράματά μας για διάφορους τύπους θορύβου. Πιο αναλυτικά χρησιμοποιήσαμε:

- Θόρυβο κανονικής κατανομής με μέση τιμή μηδέν και μεταβλητή διακύμανση.

ICM (S_0, E, ϵ, \max)

S_0 : Τυχαία αρχικοποίηση των παραμέτρων.

E : Η συνάρτηση υπολογισμού της ενέργειας.

ϵ : Ελάχιστη επιθυμητή μεταβολή της ενέργειας μεταξύ δύο επαναλήψεων.

\max : Μέγιστο πλήθος επαναλήψεων.

- **while** ($k \geq \max$ ή $\Delta E > \epsilon$)

- Καθορισμός τυχαίας σειράς βελτιστοποίησης των παραμέτρων

- Για τη σειρά που ορίσαμε στο προηγούμενο βήμα

- * Κρατάμε σταθερές τις μη βελτιστοποιούμενες παραμέτρους

- * Για την προς βελτιστοποίηση παράμετρο, δημιουργούμε ένα σύνολο από διάφορες τιμές της

- * Για κάθε τιμή του συνόλου που δημιουργήσαμε στο προηγούμενο βήμα, υπολογίζουμε την ενέργεια E , θεωρώντας τις υπόλοιπες παραμέτρους σταθερές.

- * Για την παράμετρο που βελτιστοποιούμε επιλέγουμε εκείνη την τιμή του προαναφερθέντος συνόλου, για την οποία η τιμή της E που υπολογίσαμε στο προηγούμενο βήμα είναι βέλτιστη.

- * Η προς βελτιστοποίηση παράμετρος καθιλώνεται στην τιμή που υπολογίσαμε στο προηγούμενο βήμα.

- **end while**

Αλγόριθμος 5: Ο αλγόριθμος ICM.

Υπέρθεση Με Μικτές Κατανομές (I_{ref}, I_{reg})

I_{ref} : Η αρχική εικόνα, επί της οποίας θα γίνει η υπέρθεση.

I_{reg} : Η εικόνα την οποία θέλουμε να υπερθέσουμε.

- Τμηματοποίηση της I_{reg} εικόνας σε K ομάδες με χρήση Μοντέλων Μικτών Κατανομών (Gaussian ή Student's T).

- \forall ομάδα $k = 1 \dots K$ της εικόνας I_{reg} :

Εφαρμογή ενός αλγορίθμου ελαχιστοποίησης, όπου σε κάθε επανάληψή του υπολογίζω μία ενέργεια (Battacharyya απόσταση) ως εξής:

- Προβολή της μάσκας κατάτμησης επί της I_{ref} .

- Σε κάθε επιμέρους τμήμα της προβολής, υπολογίζω μέση τιμή μ και πίνακα συμμεταβλητότητας Σ . Έτσι έχω μία συνολική εκτίμηση της κατανομής της προβολής.

- Υπολογισμός της συνολικής απόστασης Battacharyya, με βάση τον τύπο (3.1), μεταξύ της εκτίμησης της κατανομής στην I_{reg} , όπως προέκυψε από τα Μικτά Μοντέλα Κατανομών, και της εκτίμησης της κατανομής της προβολής.

Αλγόριθμος 6: Αλγόριθμος υπέρθεσης εικόνων βασισμένος σε Μοντέλα Μικτών Κατανομών.

- Θόρυβο ομοιόμορφης κατανομής.
- Αποκλίνοντα εικονοστοιχεία σε προκαθορισμένο ποσοστό επί του συνόλου της εικόνας (outliers).

Η δομή με την οποία θα παρουσιάσουμε τα αποτελέσματα είναι η ακόλουθη: Δίνεται η αρχική εικόνα, η τελική εικόνα, ένα διάγραμμα με το μέσο σφάλμα υπέρθεσης σε εικονοστοιχεία (*pixels*) και ένας πίνακας με στατιστικά στοιχεία για το σφάλμα αυτό (μέση τιμή, απόκλιση, κ.τ.λ.). Το σφάλμα υπέρθεσης υπολογίζεται ως η μέση απόκλιση - απόσταση των τεσσάρων κορυφών της εικόνας, δηλαδή η απόσταση του κάθε σημείου από την πραγματική του θέση και την υπολογισμένη. Είναι προφανές ότι όσο μεγαλύτερη είναι η εικόνα τόσο μεγαλύτερη επίδραση έχει το σφάλμα στον υπολογισμό των παραμέτρων των μετασχηματισμών. Θεωρήσαμε ότι η παράθεση διαγραμμάτων για το σφάλμα που περιγράψαμε είναι πιο ενδεικτική σε σχέση με την απλή αναφορά στην απόκλιση των τιμών των παραμέτρων των μετασχηματισμών, π.χ. σφάλμα στον υπολογισμό της γωνίας. Ο λόγος είναι ότι δεν μπορεί να γίνει συνολική εκτίμηση της απόδοσης της κάθε μεθόδου, μιας και τα σφάλματα αυτά ερμηνεύονται διαφορετικά, ανάλογα με το μέγεθος της εικόνας, ενώ δεν είναι εύκολα αντηληπτό από τον αναγνώστη η ποιοτική και σημασιολογική σημασία της κάθε τιμής.

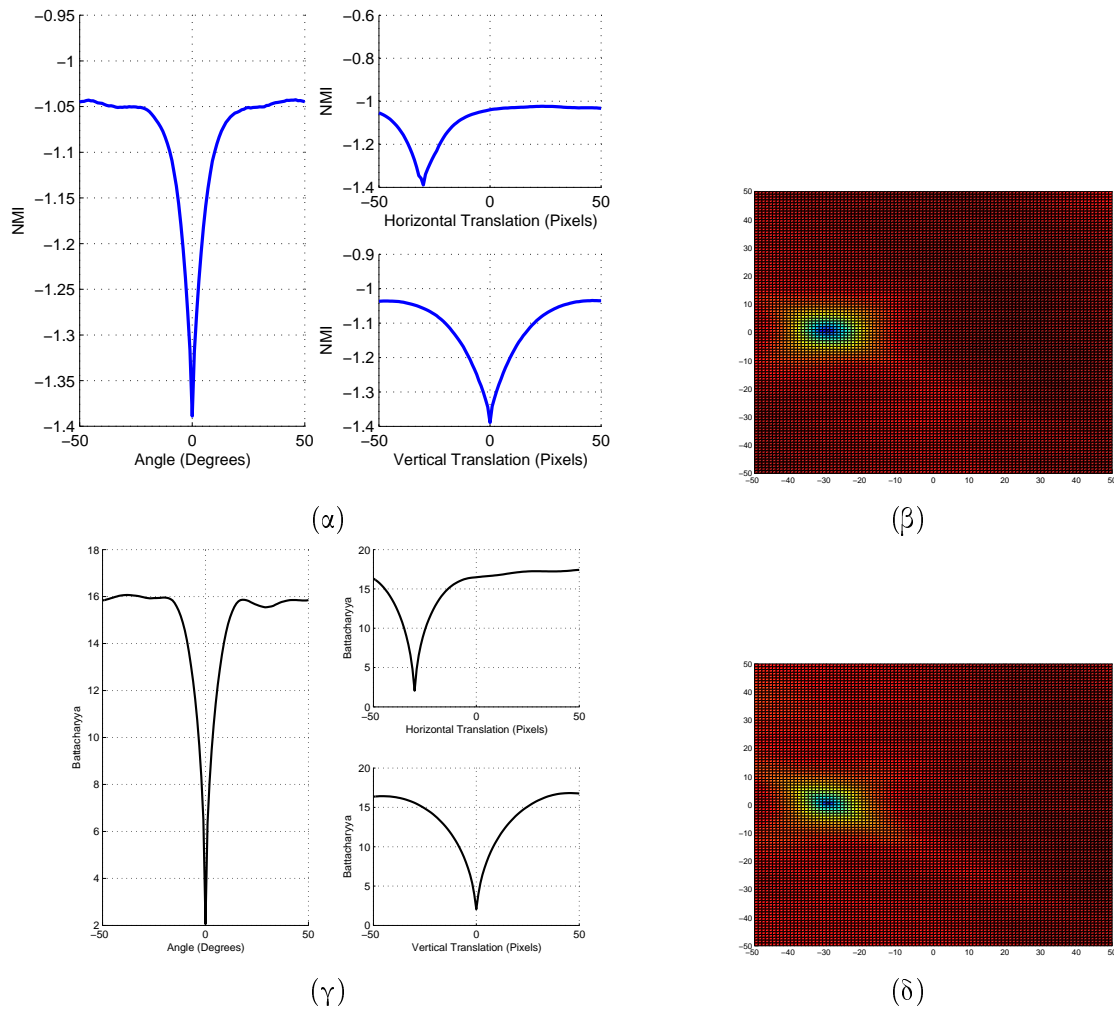
Τέλος να αναφέρουμε ότι για την εξαγωγή των παρακάτω αποτελεσμάτων, θεωρήσαμε συμπαγείς μετασχηματισμούς (περιστροφή και μετατόπιση), ενώ δε συμπεριλάβαμε την κλιμάκωση. Αρχικά δοκιμάσαμε τα όρια στα οποία η μέθοδός μας δίνει ορθά αποτελέσματα. Στην περίπτωση αυτή χρησιμοποίησαμε αυτούσιες τις εικόνες εισόδου, χωρίς να προσθέσουμε παραμόρφωση λόγω θορύβου, αλλά μεταβάλλοντας τις παραμέτρους του μετασχηματισμού. Με βάση τα πειράματα που κάναμε παρατηρήσαμε ότι για διαστήματα $\theta \in [-45, +45]$ μοίρες και διάνυσμα μετατόπισης $T \in [-30, +30] \times [-30, +30]$ εικονοστοιχεία η μέθοδός μας δίνει ικανοποιητικά αποτελέσματα. Όσον αφορά στον έλεγχο της ανοχής στο θόρυβο, κρατήσαμε σταθερή την παραμετροποίηση του μετασχηματισμού υπέρθεσης (γωνία 0 μοίρες, μετατόπιση $[0, 0]^T$) και προσθέτοντας σε κάθε στιγμιότυπο κατάλληλο θόρυβο. Για την εκτέλεση της ελαχιστοποίησης, σε κάθε περίπτωση χρησιμοποιήσαμε τον αλγόριθμο SIMPLEX.

Όσον αφορά στις εικόνες με τις οποίες πειραματιστήκαμε, προσπαθήσαμε να χρησιμοποιήσουμε από διάφορους επιστημονικούς τομείς (Ιατρική, Βιολογία, Μετεωρολογία, Τηλεπισκόπηση, κ.ά.) προκειμένου να δοκιμαστεί η μέθοδός μας σε μια ποικιλία δεδομένων εισόδου, με διαφορετικά χαρακτηριστικά και ιδιομορφίες. Πιο αναλυτικά οι εικόνες που χρησιμοποιήσαμε στα πειράματά μας είναι οι ακόλουθες:

- Πείραμα 1, σχήμα 3.3. Εικόνα 3 περιοχών σταθερής φωτεινότητας κατά τμήματα.
- Πείραμα 2, σχήμα 3.4. Τεχνητή εικόνα 8bit χρώματος απεικόνισης ενός δωματίου.
- Πείραμα 3, σχήμα 3.5. Μαγνητική αξονική τομογραφία εγκεφάλου (MRI).
- Πείραμα 4, σχήμα 3.6. Βλαστοκύτταρα ποντικού σε ορατό και υπέρυθρο φως.
- Πείραμα 5, σχήμα 3.7. Δορυφορικές μετεωρολογικές εικόνες.

Τέλος σχετικά με την παρεμβολή, χρησιμοποιήσαμε κυβική, έτσι ώστε να αποφύγουμε την είσοδο σφαλμάτων λόγω ατελειών της παρεμβολής. Για περισσότερες πληροφορίες επί του θέματος ο αναγνώστης παραπέμπεται στην αναφορά [20]. Ενδεικτικά στο σχήμα 3.2 δίνουμε μία αναπαράσταση της αντικειμενικής συνάρτησης (αμοιβαία πληροφορία και απόσταση Battacharyya που αντιστοιχεί στις εικόνες (α), (β) του σχήματος 3.3. Οι δύο εικόνες διαφέρουν μόνο στον άξονα x κατά 30

εικονοστοιχεία, οπότε αναμένουμε τα ακρότατα να βρίσκονται στις 0 μοίρες, -30 εικονοστοιχεία στον οριζόντιο άξονα και 0 εικονοστοιχεία στον κατακόρυφο άξονα.



Σχήμα 3.2: Διαγράμματα της αντικειμενικής συνάρτησης που ελαχιστοποιείται κατά τη διαδικασία της υπέρθεσης. Το διάγραμμα αναφέρεται στις εικόνες (α),(β) του σχήματος 3.3. (α),(β) Αμοιβαία πληροφορία. (γ),(δ) απόσταση Battacharyya.

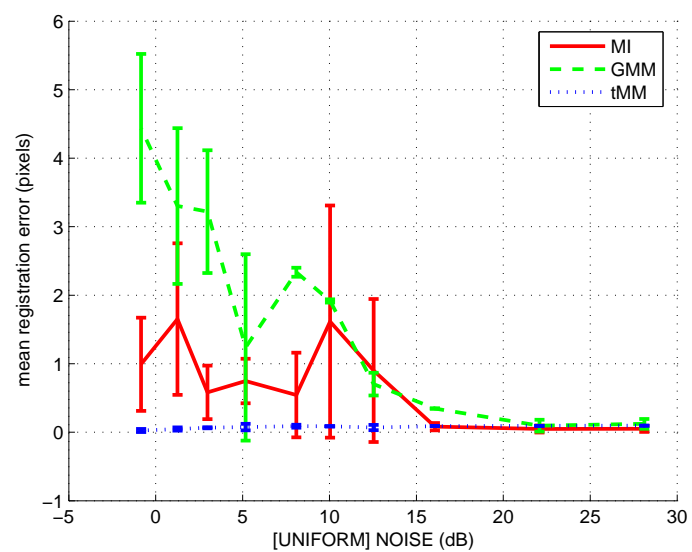
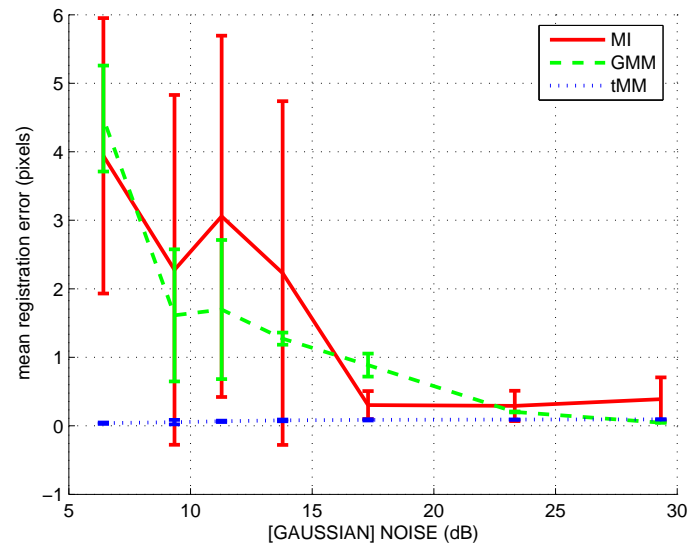
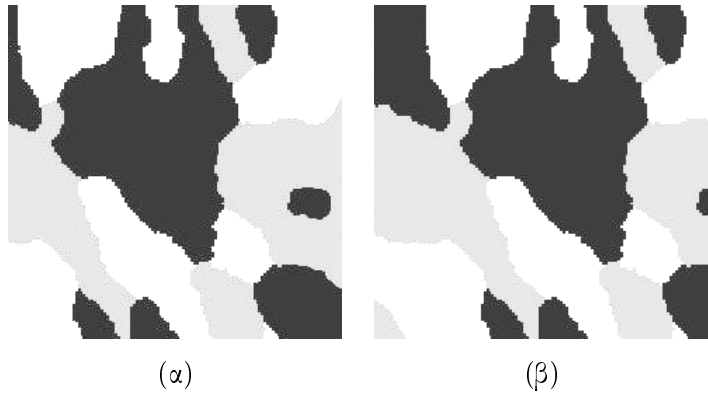
3.3 Ανάλυση αποτελεσμάτων

Κλείνοντας το κεφάλαιο θα θέλαμε να σχολιάσουμε τα αποτελέσματα που παρουσιάσαμε στην προηγούμενη ενότητα. Με μία πρώτη ματιά γίνεται αμέσως αντιληπτό ότι η μέθοδος που προτείνουμε υπερτερεί έναντι της αμοιβαίας πληροφορίας, τόσο στη στοίχιση μονοτροπικών (unimodal) αλλά και πολυτροπικών (multimodal). Η μέθοδος μας επιτυγχάνει στις πλείστες των περιπτώσεων να δώσει ακρίβεια σε υποδιαίρεση του εικονοστοιχείου της εικόνας, σε αντίθεση με την αμοιβαία πληροφορία, γεγονός πολύ σημαντικό στην περίπτωση που η ακρίβεια είναι το ζητούμενο. Η διαφορά γίνεται ακόμα μεγαλύτερη όταν αυξάνει το ποσοστό του θορύβου στο σήμα μας (εικόνα).

Τέλος όσον αφορά στη σύγκριση της υπέρθεσης με βάση τις Κανονικές και τις Student's t -

Κατανομές, είναι φανερό ότι οι τελευταίες είναι πολύ πιο αποδοτικές σχετικά με την ακρίβεια του αποτελέσματος, ειδικά όσο πιο υψηλά επίπεδα θορύβου υπάρχουν στην εικόνα. Επιπλέον η υπέρθεση με μίξη Student's t -Κατανομών παρουσιάζει μία σταθερότητα στη συμπεριφορά, ενώ η απόδοσή της δεν επηρεάζεται από τον τύπο θορύβου που υπάρχει στην εικόνα. Χαρακτηριστικό παράδειγμα ευρωστίας αποτελούν τα πειράματα, των οποίων τα αποτελέσματα δίνονται στα σχήματα 3.4 και 3.7, όπου εκτός του καθιερωμένου τύπου θορύβου, εισαγάγαμε τεχνητά αποκλίνοντα σημεία για να δοκιμάσουμε την απόδοση των δύο τεχνικών. Τα αποτελέσματα, όπως φαίνεται, ευνοούν τη μίξη Student's t -Κατανομών, γεγονός αναμενόμενο, αφού από τη θεωρία γνωρίζουμε ότι η τελευταία είναι εύρωστη στην παρουσία θορύβου.

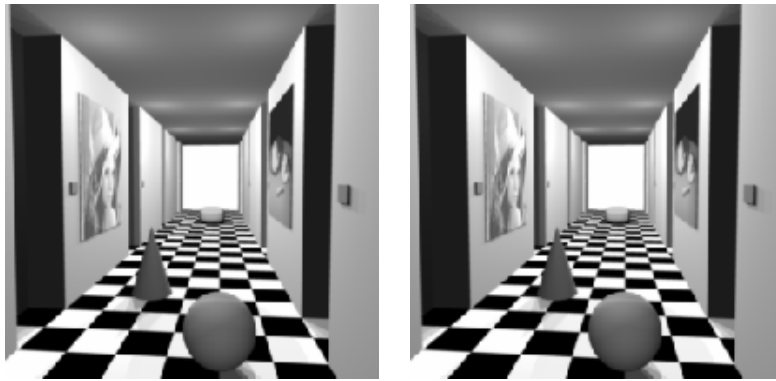
Ολοκληρώνοντας, σχετικά με τις χρονικές αποδόσεις της μεθόδου, μπορούμε να πούμε ότι η υπέρθεση με μιστές κατανομές είναι σαφώς πιο αποδοτική έναντι της αμοιβαίας πληροφορίας. Η διαφορά αυτή μπορεί να γίνει ακόμα πιο σημαντική στην περίπτωση που ενσωματώσουμε μία καθολική μέθοδο ελαχιστοποίησης, που απαιτεί πλήθος κλήσεων στην αντικειμενική συνάρτηση. Στην περίπτωση της αμοιβαίας πληροφορίας κάτι τέτοιο θα ήταν πρακτικά αδύνατο, αφού θα απαιτούσε για κάθε στιγμιότυπο των παραμέτρων των μετασχηματισμών υπέρθεσης τον υπολογισμό του από κοινού ιστογράμματος των δύο εικόνων, υπολογισμός αρκετά χρονοβόρος. Αντίθετα στις Μιστές Κατανομές υπάρχει έτοιμος αναλυτικός τύπος για την απόσταση - ενέργεια. Η μόνη απαιτητική εργασία είναι αυτή της διαμέρισης, η οποία όμως γίνεται άπαξ.



Τύπος Θορύβου	Μέθοδος	mean	std	median	max	min
Γκαουσιανος	MI	0,56	0,44	0,53	1,22	0,09
	GMM	2,23	2,55	1,16	6,98	0,04
	<i>tMM</i>	0,07	0,01	0,07	0,09	0,05
Ομοιόμορφος	MI	0,72	0,59	0,66	1,65	0,04
	GMM	1,77	1,52	1,57	4,43	0,09
	<i>tMM</i>	0,07	0,02	0,08	0,09	0,02

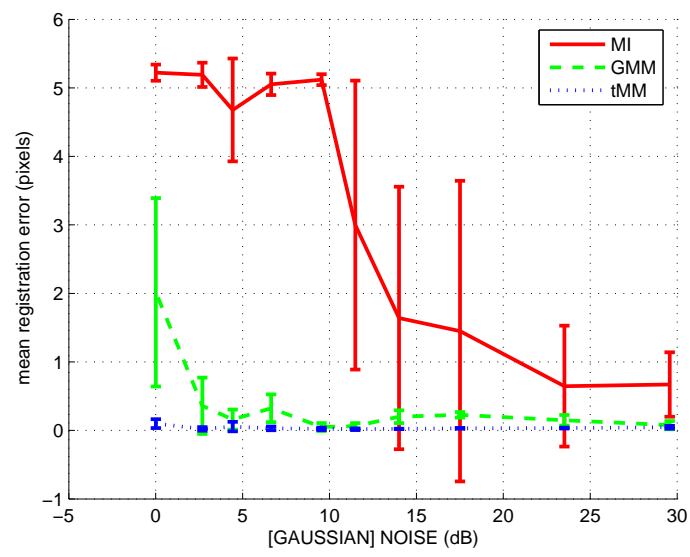
(ε)

Σχήμα 3.3: (α) Αρχική Εικόνα Υπέρθεσης. Εικόνα 3 περιοχών σταθερής φωτεινότητας κατά τμήματα. (β) Τελική Εικόνα Υπέρθεσης. Εικόνα 3 περιοχών σταθερής φωτεινότητας κατά τμήματα. (γ) Διάγραμμα σφάλματος υπέρθεσης (Γκαουσιανός θόρυβος). (δ) Διάγραμμα σφάλματος υπέρθεσης (Ομοιόμορφος θόρυβος). (ε) Στατιστικά αποτελέσματα πειράματος. Οι τιμές είναι ομαδοποιημένες με βάση τον τύπο του θορύβου (Γκαουσιανός, Ομοιόμορφος).

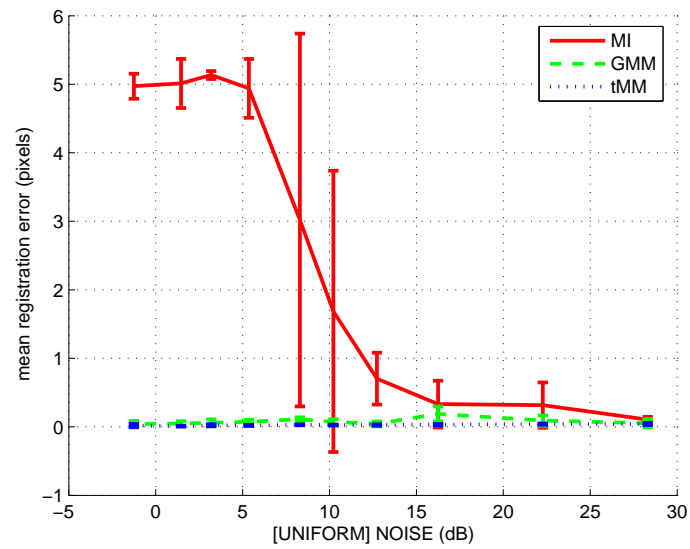


(α)

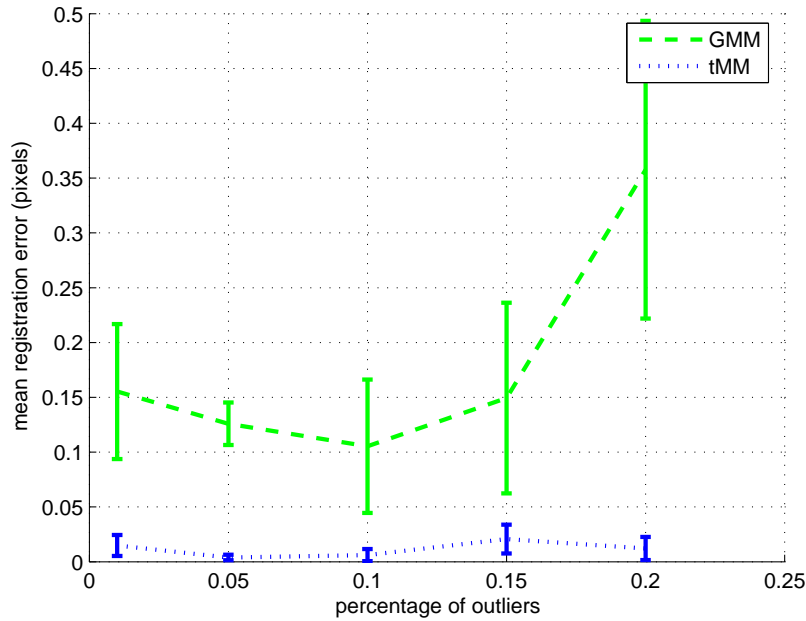
(β)



(γ)



(δ)

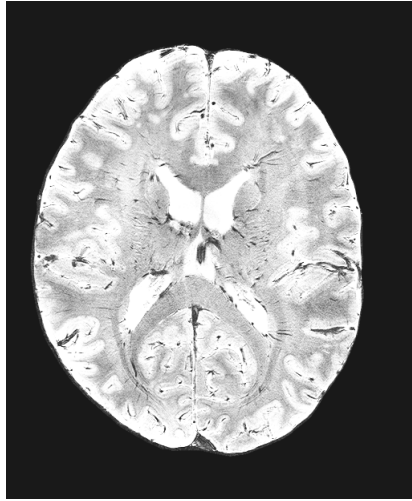


(ε)

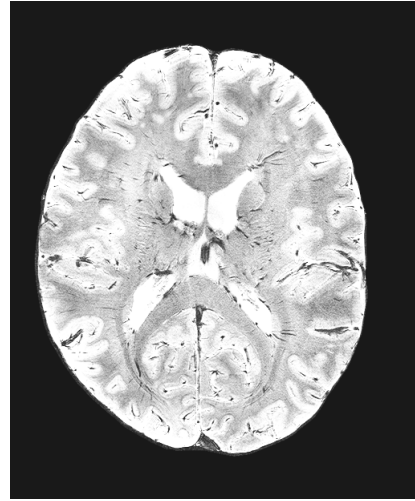
Τύπος Θορύβου	Μέθοδος	mean	std	median	max	min
Γκαουσιανός	MI	3,26	1,99	3,83	5,22	0,64
	GMM	0,36	0,59	0,17	2,01	0,04
	tMM	0,03	0,02	0,03	0,10	0,01
Ομοιόμορφος	MI	2,62	2,22	2,35	5,13	0,10
	GMM	0,07	0,04	0,06	0,18	0,03
	tMM	0,02	0,00	0,02	0,03	0,00
Αποκλίνοντα Σημεία	GMM	0,17	0,10	0,14	0,35	0,14
	tMM	0,01	0,00	0,01	0,02	0,00

(στ)

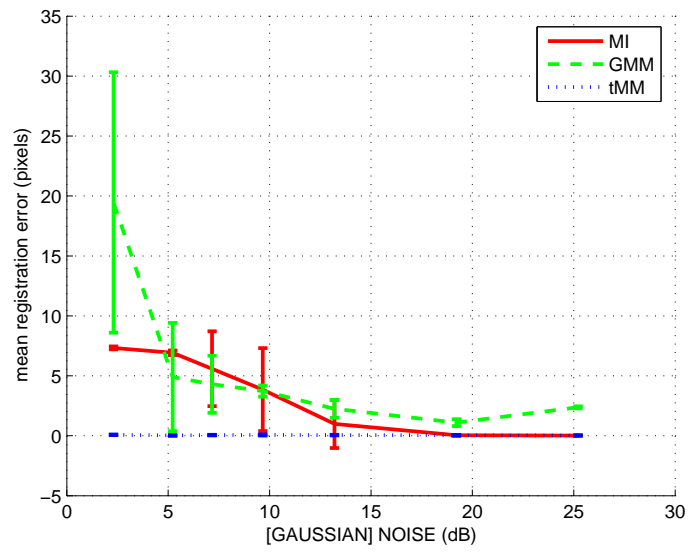
Σχήμα 3.4: (α) Αρχική Εικόνα Υπέρθεσης. Τεχνητή εικόνα 8bit χρώματος. (β) Τελική Εικόνα Υπέρθεσης. Τεχνητή εικόνα 8bit χρώματος. (γ) Διάγραμμα σφάλματος υπέρθεσης (Γκαουσιανός θόρυβος). (δ) Διάγραμμα σφάλματος υπέρθεσης (Ομοιόμορφος θόρυβος). (ε) Διάγραμμα σφάλματος υπέρθεσης (Αποκλίνοντα σημεία). (στ) Στατιστικά αποτελέσματα πειράματος. Οι τιμές είναι ομαδοποιημένες με βάση τον τύπο του θορύβου (Γκαουσιανός, Ομοιόμορφος, Αποκλίνοντα σημεία).



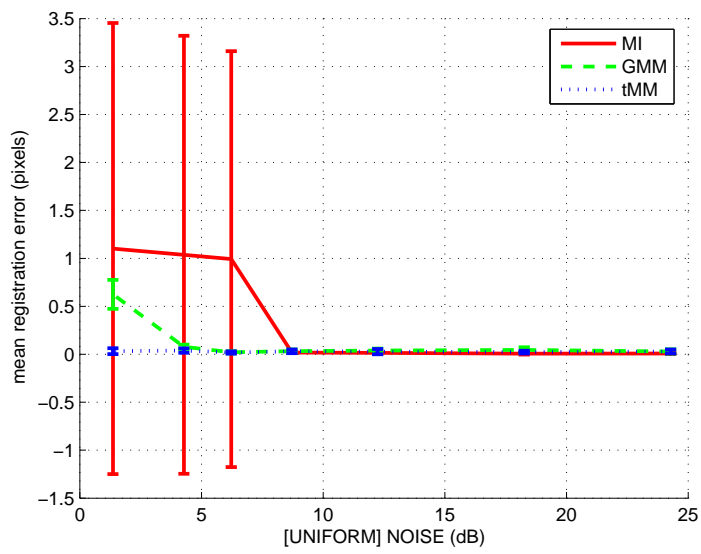
(α)



(β)



(γ)

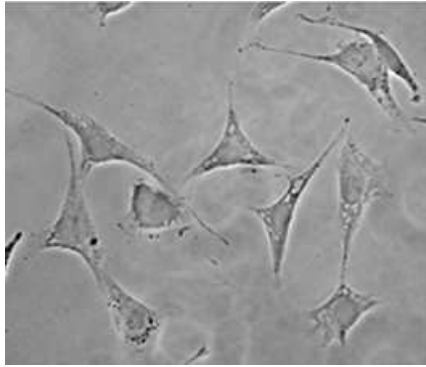


(δ)

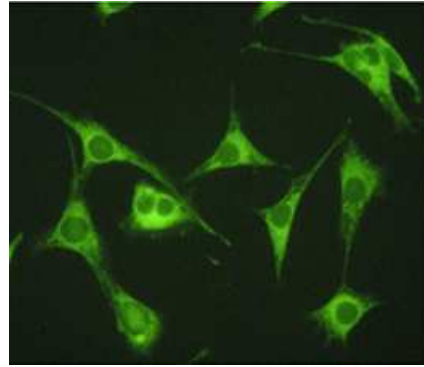
Τύπος Θορύβου	Μέθοδος	mean	std	median	max	min
Γκαουσιανός	MI	0,56	0,44	0,53	1,22	0,09
	GMM	2,23	2,55	1,16	6,98	0,04
	<i>tMM</i>	0,07	0,01	0,07	0,09	0,05
Ομοιόμορφος	MI	0,72	0,59	0,66	1,65	0,04
	GMM	1,77	1,52	1,57	4,43	0,09
	<i>tMM</i>	0,07	0,02	0,08	0,09	0,02

(ε)

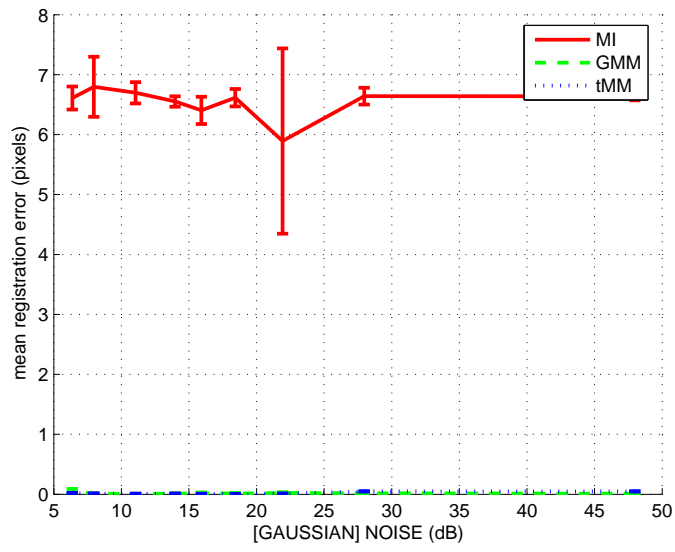
Σχήμα 3.5: (α) Αρχική Εικόνα Υπέρθωσης. Μαγνητική αξονική τομογραφία εγκεφάλου. (β) Τελική Εικόνα Υπέρθωσης. Μαγνητική αξονική τομογραφία εγκεφάλου. (γ) Διάγραμμα σφάλματος υπέρ-θωσης (Γκαουσιανός θόρυβος). (δ) Διάγραμμα σφάλματος υπέρθωσης (Ομοιόμορφος θόρυβος). (ε) Στατιστικά αποτελέσματα πειράματος. Οι τιμές είναι ομαδοποιημένες με βάση τον τύπο του θορύβου (Γκαουσιανός, Ομοιόμορφος).



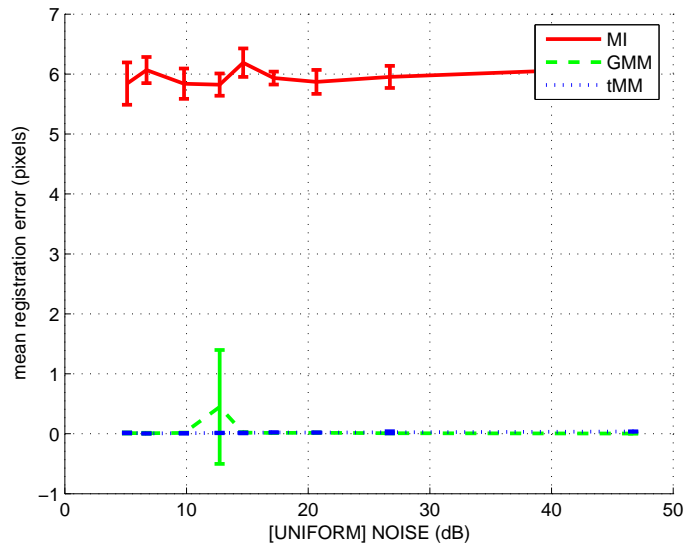
(α)



(β)



(γ)



(δ)

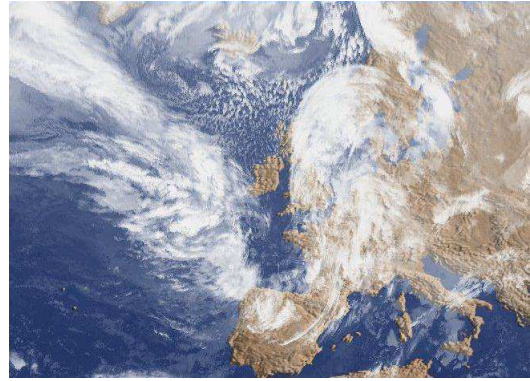
Τύπος Θορύβου	Μέθοδος	mean	std	median	max	min
Γκαουσιανός	MI	6,53	0,26	6,61	6,79	5,89
	GMM	0,01	0,01	0,01	0,05	0,00
	<i>tMM</i>	0,01	0,01	0,00	0,04	0,00
Ομοιόμορφος	MI	5,95	0,13	5,93	6,18	5,82
	GMM	0,06	0,14	0,01	0,44	0,00
	<i>tMM</i>	0,01	0,00	0,01	0,03	0,00

(ε)

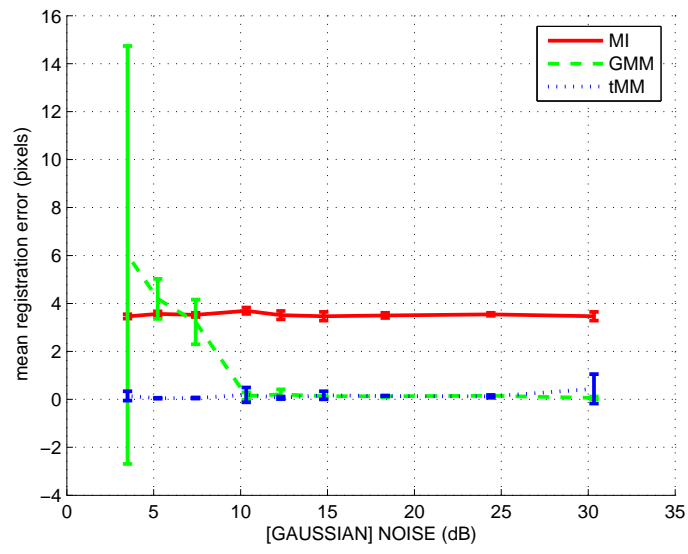
Σχήμα 3.6: (α) Αρχική Εικόνα Υπέρθωσης. Βλαστοκύτταρα ποντικού σε ορατό φώς. (β) Τελική Εικόνα Υπέρθωσης. Βλαστοκύτταρα ποντικού σε υπέρυθρο φώς. (γ) Διάγραμμα σφάλματος υπέρ-θωσης (Γκαουσιανός θόρυβος). (δ) Διάγραμμα σφάλματος υπέρθωσης (Ομοιόμορφος θόρυβος). (ε) Στατιστικά αποτελέσματα πειράματος. Οι τιμές είναι ομαδοποιημένες με βάση τον τύπο του θορύβου (Γκαουσιανός, Ομοιόμορφος).



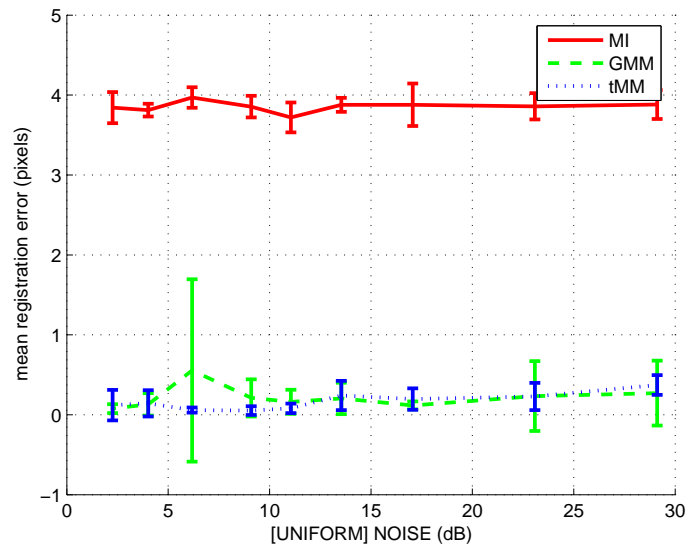
(α)



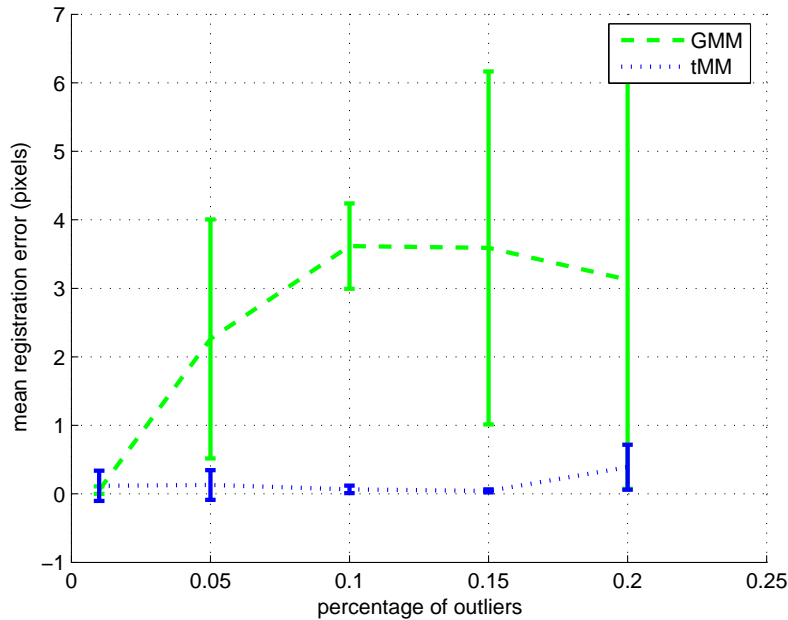
(β)



(γ)



(δ)



(ε)

Τύπος Θορύβου	Μέθοδος	mean	std	median	max	min
Γκαουσιανός	MI	3,52	0,07	3,50	3,69	3,46
	GMM	1,58	2,28	0,14	6,01	0,05
	tMM	0,15	0,11	0,13	0,48	0,04
Ομοιόμορφος	MI	3,85	0,06	3,85	3,96	3,72
	GMM	0,21	0,14	0,20	0,55	0,07
	tMM	0,16	0,10	0,14	0,37	0,05
Αποκλίνοντα Σημεία	GMM	2,52	1,48	3,12	3,61	0,05
	tMM	0,14	0,13	0,11	0,38	0,04

(στ)

Σχήμα 3.7: (α) Αρχική Εικόνα Υπέρθωσης. Δορυφορική εικόνα της Ευρώπης στις 8/1/2007 (01h00). (β) Τελική Εικόνα Υπέρθωσης. Δορυφορική εικόνα της Ευρώπης στις 9/1/2007 (01h00). Οι εικόνες προέρχονται από το δορυφόρο Meteosat. (γ) Διάγραμμα σφάλματος υπέρθεσης (Γκαουσιανός θόρυβος). (δ) Διάγραμμα σφάλματος υπέρθεσης (Ομοιόμορφος θόρυβος). (ε) Διάγραμμα σφάλματος υπέρθεσης (Αποκλίνοντα σημεία). (στ) Στατιστικά αποτελέσματα πειράματος. Οι τιμές είναι ομαδοποιημένες με βάση τον τύπο του θορύβου (Γκαουσιανός, Ομοιόμορφος, Αποκλίνοντα σημεία).

ΚΕΦΑΛΑΙΟ 4

ΥΠΕΡΘΕΣΗ ΝΕΦΩΝ ΣΗΜΕΙΩΝ

-
- 4.1 Εισαγωγή στην υπέρθεση σημείων - Ο αλγόριθμος Iterative Closest Point - ICP
 - 4.2 Περιγραφή της μεθόδου
 - 4.4 Πειραματικά αποτελέσματα
-

4.1 Εισαγωγή στην υπέρθεση σημείων - Ο αλγόριθμος Iterative Closest Point - ICP

Στο κεφάλαιο αυτό θα επεκτείνουμε την έννοια της υπέρθεσης σε νέφη σημείων. Η ουσία είναι και πάλι ίδια: εύρεση εκείνων των μετασχηματισμών και των παραμέτρων τους έτσι ώστε να επιτύχουμε γεωμετρική στοίχιση. Η διαφορά έγκυται στη σημασιολογία των χαρακτηριστικών. Στην περίπτωση των εικόνων είχαμε υπέρθεση εικονοστοιχείων με φωτεινότητες και αχέραιες διατεταγμένες συντεταγμένες, ενώ στα νέφη σημείων η υπέρθεση γίνεται επί ενός συνόλου σημείων με συντεταγμένες σε ένα συνεχή μη διατάξιμο χώρο. Η βιβλιογραφία του συγκεκριμένου τομέα διαθέτει αρκετές αναφορές για τρόπους επίλυσης του προβλήματος. Δε θα αναφερθούμε διεξοδικά στο θέμα αυτό, καθώς κάτι τέτοιο θα ξέφευγε από τους σκοπούς της παρούσης εργασίας. Η μόνη αναφορά που θα γίνει σχετίζεται με τον πιο διαδεδομένο ίσως αλγόριθμο υπέρθεσης σημείων, τον Iterative Closest Point - ICP [24], [11]. Η λογική του αλγοριθμίου βασίζεται στην ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος ως προς τις παραμέτρους του μετασχηματισμού υπέρθεσης T . Ειδικότερα, η αντικειμενική συνάρτηση είναι η ακόλουθη:

$$E(p; x) = \frac{1}{N} \sum_{i=1}^N (y_i - T(x_i \Theta))^2$$

όπου το σημείο y_i έχει προέρθει από την εφαρμογή του μετασχηματισμού υπέρθεσης T με παραμέτρους Θ επί του αρχικού x_i . Μία ενδιαφέρουσα εργασία για το αποτέλεσμα της ελαχιστοποίησης της παραπάνω αντικειμενικής συνάρτησης στην περίπτωση συμπαγούς μετασχηματισμού (Κλιμάκωση - Περιστροφή - Μετατόπιση) δίνεται στην εργασία [28]. Βέβαια απαραίτητη προϋπόθεση είναι να γνωρίζουμε την αντιστοιχία μεταξύ των σημείων του συνόλου Y και X , δηλαδή το σημείο $y_i \in Y$ από ποιο

σημείο $x_i \in X$ έχει προέρθει. Γίνεται άμεσα αντιληπτό ότι ένας καθοριστικός παράγοντας επιτυχίας της μεθόδου είναι το στάδιο αναζήτησης των αντιστοιχιών των σημείων. Εν γέννη η διαδικασία αυτή δεν είναι τετριμμένη και επιπλέον γίνεται δυσκολότερη στην περίπτωση που τα δεδομένα έχουν αλλοιωθεί από θόρυβο.

Κλείνοντας θα θέλαμε να αναφέρουμε την εργασία [3], καθώς κινείται στο πνεύμα της μεθόδου που προτείνουμε. Σε γενικές γραμμές η εν λόγω εργασία προτείνει διαμέριση με Μίξη Κανονικών Κατανομών και ελαχιστοποίηση μίας ενέργειας για εύρεση των παραμέτρων των κατανομών. Η ενέργεια στην περίπτωση αυτή είναι η L_2 τετραγωνική απόσταση. Θυμίζουμε ότι η L_2 απόσταση δύο κατανομών με συναρτήσεις πυκνότητας πιθανότητας f_1, f_2 δίνεται από τον τύπο:

$$D(f_1, f_2) = \int (f_1(x) - f_2(x))^2 dx = \int (f_1^2(x) - 2f_1(x)f_2(x) + f_2^2(x)) dx. \quad (4.1)$$

Εξειδικεύοντας στο πρόβλημά της υπέρθεσης νεφών σημείων, αν υποθέσουμε ότι έχουμε δύο σύνολα S_1 και S_2 , τότε καθένα από αυτά θα περιγράφεται από μία συνάρτηση πυκνότητας πιθανότητας (f, g αντίστοιχα), οι οποίες στην εργασία [3] εκτιμώνται με τη χρήση Μικτών Κανονικών Κατανομών. Δηλαδή

$$f(x) = \sum_{i=1}^{K_1} \alpha_i \phi(x|\mu_i, \Sigma_i)$$

και

$$g(x) = \sum_{i=1}^{K_2} \beta_i \phi(x|\nu_i, \Gamma_i)$$

, όπου $\phi = N(\mu, \Sigma)$ είναι μία Κανονική κατανομή με μέση τιμή μ και πίνακα συμμεταβλητότητας Σ . Αν υποθέσουμε μετασχηματισμό υπέρθεσης λόγω περιστροφής ως προς το σημείο X_g και μετατόπιση (R, T αντίστοιχα) τότε

$$f_{R,T}(x) = \sum_{i=1}^{K_1} \alpha_i \phi(x|R(\mu_i - X_g) + T + X_g, R\Sigma_i R^T) \quad (4.2)$$

Με βάση τις σχέσεις 4.1, 4.2 προκύπτει ότι

$$D(f_{R,T}, g) = \int (f_{R,T}^2 - 2f_{R,T}g + g^2) dx. \quad (4.3)$$

Ο στόχος μας είναι να ελαχιστοποιήσουμε την 4.3 ως προς τις παραμέτρους R, T . Δεδομένου ότι η L_2 στάθμη κάθε συνάρτησης πυκνότητας πιθανότητας είναι αμετάβλητη υπό την εφαρμογή συμπαγούς μετασχηματισμού, δηλαδή $\int f_{R,T}^2 dx = \int f^2 dx$, τελικά στη σχέση 4.3, μας ενδιαφέρει μόνο ο όρος $\int f_{R,T}g dx$. Από τη θεωρία γνωρίζουμε ότι το γινόμενο κανονικών κατανομών είναι επίσης κανονική κατανομή, άρα

$$\int f_{R,T}g dx = \sum_{i=1}^{K_1} \sum_{j=1}^{K_2} \alpha_i \beta_j \phi(x=0|R(\mu_i - X_g) + T + X_g + \nu_j, R\Sigma_i R^T + \Gamma_j).$$

Τελικά η αντικειμενική συνάρτηση που πρέπει να ελαχιστοποιηθεί ως προς τις παραμέτρους R, T είναι η:

$$D(f_{R,T}, g) = - \sum_{i=1}^{K_1} \sum_{j=1}^{K_2} \alpha_i \beta_j \phi(x=0|R(\mu_i - X_g) + T + X_g + \nu_j, R\Sigma_i R^T + \Gamma_j).$$

Αλγόριθμος ICP ($\Sigma_1, \Sigma_2, \epsilon$)

Σ_1 : Το μοντέλο το οποίο θα υπερθέσουμε επί ενός συνόλου σημείων.
 Σ_2 : Το σύνολο το σημείων επί των οποίων θα υπερτεθεί ένα μοντέλο.
 ϵ : Ακρίβεια της μεθόδου.

- (Τυχαία) Αρχικοποίηση των παραμέτρων του μετασχηματισμού υπέρθεσης, Π .
- $E_{new} := +\infty$
- $E_{old} := 0$
- **while** ($|E_{new} - E_{old}| > \epsilon$)
 - $E_{old} = E_{new}$
 - $S := \Upsilon\acute{\pi}\epsilon\rho\theta\epsilon\sigma\eta(\Sigma_2, \Pi)$
 - $CP := \text{Εύρεση-Αντιστοιχιών}(S, \Sigma_1)$
 - $\{\Pi', E_{new}\} := \text{Ενημέρωση-Παραμέτρων}(\Sigma_2, \Sigma_1, CP, \Pi)$
 - $\Pi \leftarrow \Pi'$
- **end while**

Η συνάρτηση Υπέρθωση (Σ_2, Π) εφαρμόζει το μετασχηματισμό με παραμέτρους Π στο σύνολο Σ_2 και επιστρέφει το αποτέλεσμα. Η ρουτίνα Εύρεση-Αντιστοιχιών (S, Σ_1) εντοπίζει τις αντιστοιχίες σημείων μεταξύ S και Σ_1 . Τέλος στη διαδικασία Ενημέρωση-Παραμέτρων ($\Sigma_2, \Sigma_1, CP, \Pi$) με βάση την αντιστοίχιση σημείων CP , αναζητούνται εκείνες οι παράμετροι, οι οποίες ελαχιστοποιούν ένα προκαθορισμένο κριτήριο - ενέργεια. Οι νέες τιμές καθώς και η τιμή του κριτηρίου επιστρέφονται (Π', E_{new} αντίστοιχα).

Αλγόριθμος 7: Ο αλγόριθμος ICP.

4.2 Περιγραφή της μεθόδου

Στόχος μας στο κεφάλαιο αυτό είναι να επεκτείνουμε τη μέθοδο που περιγράψαμε στο κεφάλαιο 3 στην υπέρθεση νεφών σημείων. Η κεντρική ιδέα της διαμέρισης θα χρησιμοποιηθεί και εδώ. Η διαφορά θα προκύψει στο στάδιο της προβολής. Θυμίζουμε στον αναγνώστη ότι στην περίπτωση των εικόνων, το επόμενο βήμα μετά τη διαμέριση ήταν η προβολή των επιμέρους τμημάτων για τον υπολογισμό μίας ενέργειας (απόσταση Battacharyya) και την ελαχιστοποίησή της. Στην περίπτωση των σημείων όμως η έννοια της προβολής δεν είναι αυτονόητη δεδομένης της αλλαγής της σημασιολογίας του χώρου που μελετάμε. Πλέον δεν έχουμε ένα συγκεκριμένο διακριτό διατεταγμένο χώρο (παράβαλε το χώρο των συνεταγμένων των εικονοστοιχείων) αλλά έναν συνεχή χωρίς έννοια διάταξης. Επομένως οι αντιστοιχίες μεταξύ των σημείων των δύο συνόλων δεν είναι αυτονόητες, και επομένως θα πρέπει με κάποιο τρόπο να υπολογισθούν. Σε πρώτη φάση θα ακολουθήσουμε μία άπληστη τεχνική (*greedy*), οπότε θα θεωρήσουμε ότι κάθε συνιστώσα της μίας διαμέρισης, ταιριάζει σε όλες τις άλλες, και έτσι θα υπολογίσουμε την απόστασή της από όλες αυτές. Ωστόσο μπορούμε να πάμε ένα βήμα παραπέρα και να με κάποιο τρόπο να προσομοιώσουμε τη λειτουργία της προβολής στην περίπτωση των εικόνων. Με άλλα λόγια θα μπορούσαμε με έναν αλγόριθμο να εντοπίσουμε για κάθε συνιστώσα της διαμέρισης του ενός συνόλου την αντίστοιχή της - με βάση κάποιο κριτήριο

- στο άλλο.

Η λύση στο πρόβλημα αυτό είναι ο Hungarian αλγόριθμος [6]. Η γενική φιλοσοφία είναι, αφού εκτιμηθεί ένας πίνακας με τα κόστος αντιστοίχισης κάθε συνιστώσας του ενός συνόλου με το άλλο, να γίνει εκείνη η 1-με-1 επιλογή που θα δώσει το ελάχιστο κόστος. Ο αλγόριθμος περιγράφεται στο παράρτημα 5. Η ενσωμάτωσή του στην υπέρθεση σημείων θα μας επιτρέψει να εκτιμήσουμε καλύτερα την ενέργεια και άρα θα δώσει καλύτερα τελικά αποτελέσματα. Τελικά η τροποποιημένη μέθοδος υπέρθεσης που προτείνουμε για νέφη σημείων είναι ο αλγόριθμος 8. Εδώ δίνεται στη γενική του μορφή, δηλαδή δεν εισερχόμαστε σε λεπτομέρειες ως προς το είδος της κατανομής (Κανονική ή Student's t). Για περισσότερες πληροφορίες ο αναγνώστης παραπέμπεται στο κεφάλαιο 2, αλγόριθμοι 2 και 3.

Υπέρθεση Νεφών Σημείων Με Μικτές Κατανομές (Σ_1, Σ_2)

Σ_1 : Το μοντέλο το οποίο θα υπερθέσουμε επί ενός συνόλου σημείων.
 Σ_2 : Το σύνολο το σημείων επί των οποίων θα υπερτεθεί ένα μοντέλο.

- Εκτίμηση της κατανομής των δύο συνόλων.
- Εύρεση της αντιστοίχισης σημείων μεταξύ των δύο συνόλων. Hungarian αλγόριθμος.
- Ελαχιστοποίηση της συνολικής απόστασης των δύο κατανομών ως προς τις παραμέτρους R, T του μετασχηματισμού υπέρθεσης, με βάση την αντιστοίχιση του προηγούμενου βήματος.

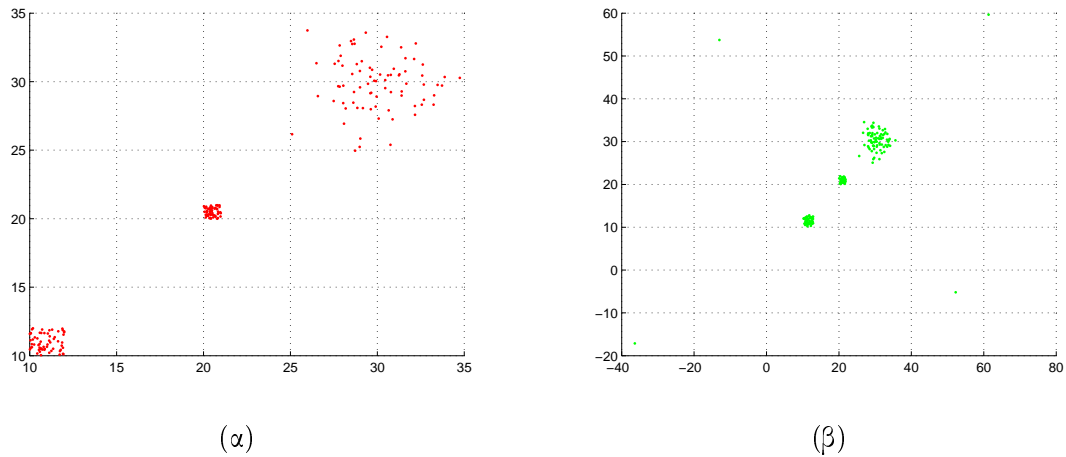
Αλγόριθμος 8: Μέθοδος για υπέρθεση νεφών σημείων.

4.3 Πειραματικά αποτελέσματα

Κλείνοντας το κεφάλαιο θα παρουσιάσουμε και κάποια πειραματικά αποτελέσματα για τη συμπεριφορά της μεθόδου μας, σχήματα 4.1, 4.3. Η σύγκριση γίνεται ως προς τον αλγόριθμο ICP [29], [24], [11] αλλά και ως προς τη χρήση Κανονικών Μικτών Κατανομών, όπως προτείνεται στη δημοσίευση [3] και με βάση τα όσα περιγράψαμε παραπάνω για την απόσταση L_2 .

Για την παρουσίαση των αποτελεσμάτων θα ακολουθήσουμε την ίδια δομή με την οποία παρουσιάσαμε τη μέθοδο για τις εικόνες. Δηλαδή αρχικά θα δώσουμε το αρχικό σύνολο σημείων και ένα στιγμιότυπό του, για κάποιο σύνολο παραμέτρων μετασχηματισμού και θορύβου και εν συνεχεία το διάγραμμα με το μέσω σφάλμα υπέρθεσης. Στην περίπτωση των σημείων, το μέσο σφάλμα υπέρθεσης υπολογίζεται ως η μέση απόσταση μεταξύ των πραγματικών σημείων του συνόλου αφού εφαρμοστεί ο μετασχηματισμός υπέρθεσης και του συνόλου που προκύπτει αν εφαρμοστεί ο ίδιος μετασχηματισμός υπέρθεσης, αλλά με παραμέτρους αυτές τις οποίες υπολογίζει η μέθοδός μας. Να τονίσουμε ότι το φάσμα αυτό είναι σε μονάδες βάσης, δηλαδή εξαρτάται από τη σημασιολογία των συντεταγμένων των σημείων. Τέλος να αναφέρουμε ότι για την εξαγωγή των παρακάτω αποτελεσμάτων, θεωρήσαμε συμπαγείς μετασχηματισμούς (περιστροφή και μετατόπιση), ενώ δε συμπεριλάβαμε την κλιμάκωση. Αρχικά δοκιμάσαμε τα όρια στα οποία η μέθοδός μας δίνει ορθά αποτελέσματα. Στην περίπτωση αυτή χρησιμοποιήσαμε αυτούσιες τις εικόνες εισόδου, χωρίς να προσθέσουμε παραμόρφωση λόγω θορύβου, αλλά μεταβάλλοντας τις παραμέτρους του μετασχηματισμού. Με βάση τα πειράματα που κάναμε παρατηρήσαμε ότι για διαστήματα $\theta \in [-25, +25]$ μοίρες και διάνυσμα μετατόπισης $T \in [-50, +50] \times [-50, +50]$ μονάδες βάσης η μέθοδός μας δίνει ικανοποιητικά αποτελέσματα. Όσον

αφορά στον έλεγχο της ανοχής στο θόρυβο, κρατήσαμε σταθερή την παραμετροποίηση του μετασχηματισμού υπέρθεσης (γωνία 0 μοίρες, μετατόπιση $[0, 0]^T$) και προσθέτοντας σε κάθε στιγμιότυπο κατάλληλο θόρυβο. Ο θόρυβος που εισαγάμε στα δεδομένα προέκυψε από τον πολλαπλασιασμό των συντεταγμένων των σημείων που επιλέχθηκαν τυχαία να αλλοιωθούν με μία προκαθορισμένη μεγάλη τιμή. Για την εκτέλεση της ελαχιστοποίησης, σε κάθε περίπτωση χρησιμοποιήσαμε τον αλγόριθμο SIMPLEX. Σε όλα τα πειράματα χρησιμοποιήσαμε 100 2Δ σημεία στο αρχικό σύνολο, σχήμα 4.1, και τα οποία παράξαμε τεχνητά από ομοιόμορφη και από κανονική κατανομή στα διαστήματα $[9, 11] \times [9, 11]$, $[19, 21] \times [19, 21]$ και $[28, 32] \times [28, 32]$, ισοκατανεμημένα στο κάθε διάστημα.

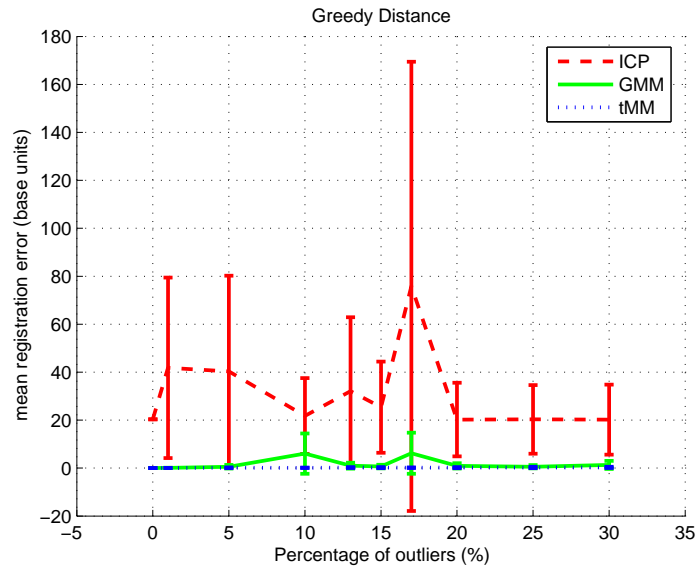


Σχήμα 4.1: Δεδομένα που χρησιμοποιήσαμε στα πειράματα ελέγχου της μεθόδου μας σε νέφη σημείων. (α) Αρχικό σύνολο σημείων (μοντέλο - σημεία που παράχθηκαν από ομοιόμορφη κατανομή). (β) Ένα στιγμιότυπο του τελικού συνόλου σημείων υπέρθεσης με προσθήκη 2% αποκλιπόντων σημείων.

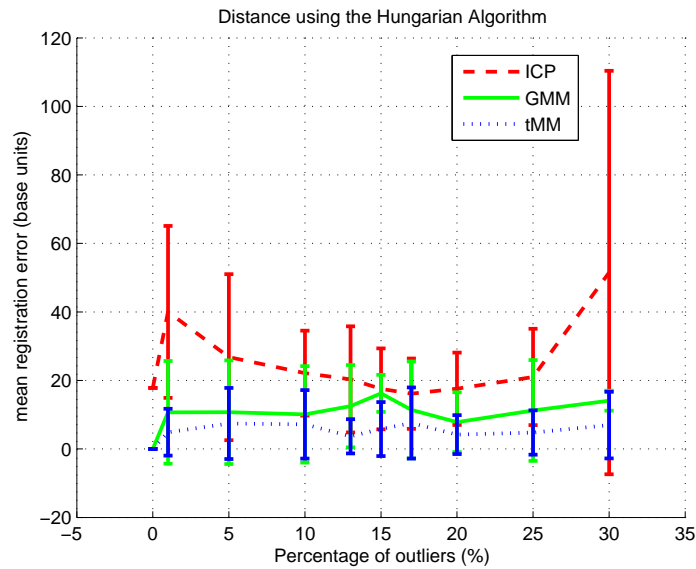
4.4 Ανάλυση αποτελεσμάτων

Ακολουθώντας την ίδια τακτική με τη μελέτη της μεθόδου μας στις εικόνες, θα κλείσουμε το κεφάλαιο παραθέτοντας κάποια σχόλια για τα πειραματικά αποτελέσματα.

Συγκρίνοντας λοιπόν τον ICP με τις μεθόδους που βασίζονται σε Μικτές κατανομές, σχήμα 4.3, διαπιστώνουμε την ανωτερότητα που επιδεικνύουν οι τελευταίες ως προς το αποτέλεσμα της υπέρθεσης. Όσον αφορά στη σύγκριση υπέρθεσης με Μικτές Κανονικές κατανομές και Μικτές Student's t -Κατανομές, είναι και εδώ εμφανές ότι οι δεύτερες παρουσιάζουν πιο σταθερή συμπεριφορά, δείχνοντας ευρωστία στο θόρυβο, γεγονός που ερμηνεύεται ως μικρότερο σφάλμα στην υπέρθεση. Η συμπεριφορά αυτή, όπως δείχνουν τα πειράματα, φαίνεται να μην επηρεάζεται από το είδος της κατανομής που έχει παράξει τα σημεία.

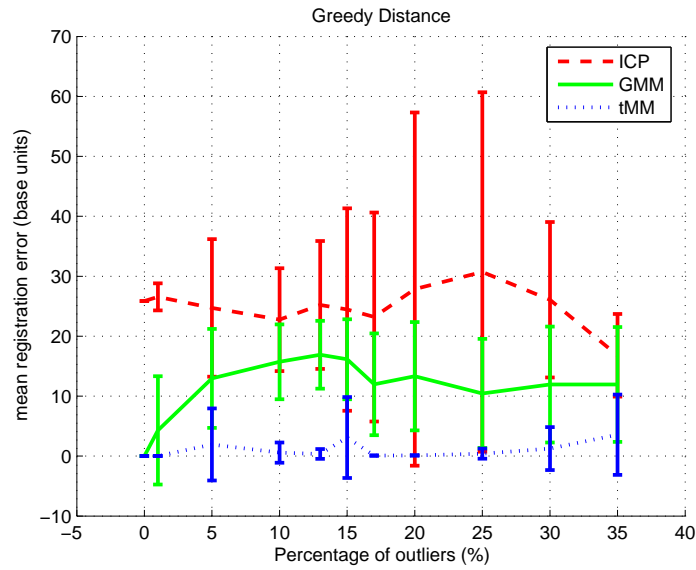


(α)

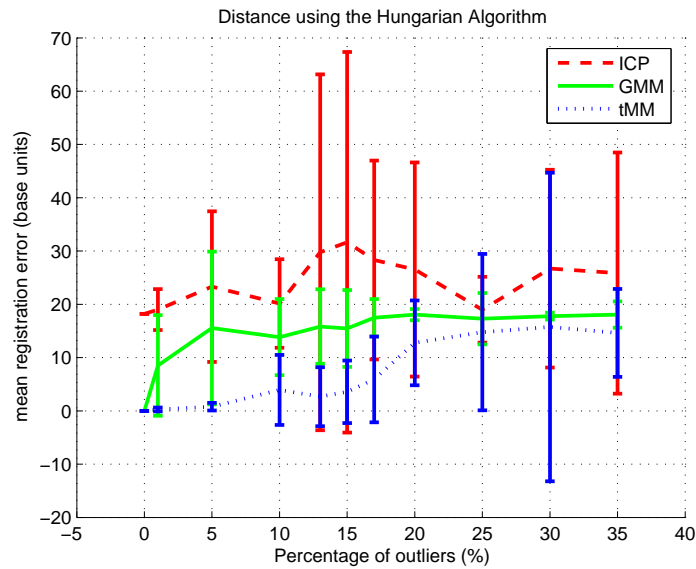


(β)

Σχήμα 4.2: Πειραματικά αποτελέσματα για τα νέφη σημείων που περιγράφονται στο σχήμα 4.1 (α) Διάγραμμα σφάλματος υπέρθεσης (Γκαουσιανή κατανομή + αποκλίνοντα σημεία). Σύγκριση Μικτών κατανομών χωρίς αντιστοιχίες. (β) Διάγραμμα σφάλματος υπέρθεσης (Γκαουσιανή κατανομή + αποκλίνοντα σημεία). Σύγκριση Μικτών κατανομών με αντιστοιχίες (Hungarian αλγόριθμος).



(α)



(β)

Σχήμα 4.3: Πειραματικά αποτελέσματα για τα νέφη σημείων που περιγράφονται στο σχήμα 4.1 (α) Διάγραμμα σφάλματος υπέρθεσης (Ομοιόμορφη κατανομή + αποκλίνοντα σημεία). Σύγκριση Μικτών κατανομών χωρίς αντιστοιχίες. (β) Διάγραμμα σφάλματος υπέρθεσης (Ομοιόμορφη κατανομή + αποκλίνοντα σημεία). Σύγκριση Μικτών κατανομών με αντιστοιχίες (Hungarian αλγόριθμος).

ΚΕΦΑΛΑΙΟ 5

ΕΠΙΛΟΓΟΣ

Στην παρούσα μεταπτυχιακή εργασία ασχοληθήκαμε με μία μέθοδο υπέρθεσης εικόνας. Το πρόβλημα της υπέρθεσης δύο εικόνων εστιάζεται στον υπολογισμό εκείνων των παραμέτρων των γεωμετρικών μετασχηματισμών, που αν εφαρμοστούν στη μία εικόνα, θα τη μετασχηματίσουν γεωμετρικά έτσι ώστε να ταυτιστεί νοηματικά με την άλλη. Με άλλα λόγια στην υπέρθεση εικόνας στόχος μας είναι η γεωμετρική στοίχιση δομών της μίας εικόνας με αυτών της άλλης με κριτήριο μία νοηματική αντιχτοίχιση μεταξύ των δομών. Για παράδειγμα στην περίπτωση υπέρθεσης δύο δορυφορικών φωτογραφιών, επιθυμούμε να ν' αζητήσουμε εκείνους τους μετασχηματισμούς (τύπος και παράμετροι) έτσι ώστε περιοχές της μίας που εικονίζουν κάποια γεωγραφική αναφορά (π.χ. μία πεδιάδα) να συμπέσουν, άμα τη εφαρμογή του μετασχηματισμού, με τις αντίστοιχές τους στην άλλη εικόνα. Είναι προφανές ότι η υπέρθεση μπορεί να αναφέρεται σε σύνολο εικόνων.

Το πρόβλημα της υπέρθεσης έχει απασχολήσει εκτενώς τη διεθνή επιστημονική κοινότητα, και αυτό φαίνεται από το μεγάλο πλήθος εργασιών που έχουν δημοσιευτεί. Αξίζει να αναφερθεί ότι τη δεκαετία 1993 - 2003, περισσότερες από χίλιες δημοσιεύσεις παρουσιάστηκαν για το θέμα αυτό [4], αξιοποιώντας διάφορες τεχνικές. Μία από τις μεθόδους που έχει γνωρίσει ευρεία αποδοχή για την επίλυση του προβλήματος αυτού, και δει πολυτροπικών εικόνων (*multimodal images*), είναι αυτές που βασίζονται στη μεγιστοποίηση της αμοιβαίας πληροφορίας. Παραπέμπουμε τον αναγνώστη στις εργασίες [26], [15] για λεπτομερή περιγραφή της μεθόδου. Η προσέγγιση που προτείνεται στην εργασία μας ακολουθεί τις μιστές κατανομές [7],[8] για την εύρεση των παραμέτρων των μετασχηματισμών υπέρθεσης.

Η μέθοδος μας βασίζεται στην παρατήρηση ότι αν διαμερίσουμε τη μία εικόνα με κάποιο τρόπο σε ομάδες (εξ' ου και η χρήση των μιστών κατανομών) και υπολογίσουμε την κατανομή της κάθε ομάδας τότε αν προβάλουμε την κάθε ομάδα στην άλλη εικόνα, υπολογίσουμε την κατανομή της φωτεινότητας επί της προβολής και τέλος υπολογίσουμε την ολική απόσταση μεταξύ των κατανομών, τότε η βέλτιστη τιμή των ζητούμενων παραμέτρων θα προκύψει στην περίπτωση της ελάχιστης ολικής απόστασης μεταξύ των κατανομών. Ως βασική κατανομή στο μοντέλο μιστών κατανομών, χρησιμοποιήσαμε τόσο τις Κανονικές όσο και τις Student's *t*-Κατανομές. Τα αποτελέσματα ήταν ικανοποιητικά, ειδικά στην περίπτωση της δεύτερης, ενώ υπερτερούσαν κατά πολύ εκείνων που υπολογίσαμε με βάση την αμοιβαία πληροφορία.

Με βάση τα πειραματικά αποτελέσματα, μπορούμε να ισχυριστούμε ότι η προτεινόμενη μέθοδος χρήζει περεταίρω μελέτης, ενώ είναι μία πολλά υποσχόμενη προσέγγιση στην υπέρθεση εικόνας.

Η ευρωστία που επιδεικνύεται παρουσία θορύβου είναι ένας παράγοντας που συνηγορεί υπέρ του αλγορίθμου μας.

Στην εργασία αυτή τέθηκαν οι βάσεις για την παρουσίαση μίας νέας μεθόδου υπέρθεσης εικόνων. Μένουν όμως αρκετές πτυχές που πρέπει να ερευνηθούν, όπως για παράδειγμα η επιλογή εκείνων των χαρακτηριστικών που θα παίζουν ρόλο στην εκπαίδευση των μικτών κατανομών αλλά και στον μετέπειτα υπολογισμό της ολικής απόστασης μεταξύ των κατανομών. Στα πειράματά μας χρησιμοποιήσαμε μόνο την τιμή της φωτεινότητας κάθε εικονοστοιχείου.

Ένα ακόμα θέμα για μελλοντική ενασχόληση είναι και αυτό της εκτέλεσης της ελαχιστοποίησης. Ο αλγόριθμος SIMPLEX που χρησιμοποιήσαμε να μεν ελαχιστοποιεί την ενέργεια αλλά είναι ένας αλγόριθμος τοπικής ελαχιστοποίησης, και επομένως η αρχικοποίησή του καθορίζει σημαντικά το αποτέλεσμα. Η αξιοποίηση ενός αλγορίθμου ολικής ελαχιστοποίησης θα παρακάμψει τα εμπόδια αυτά. Να σημειωθεί στο σημείο αυτό ότι η προτεινόμενη μέθοδος ευνοεί τη χρήση καθολικής ελαχιστοποίησης, σε σχέση με την αμοιβαία πληροφορία, αφού ο υπολογισμός της αντικειμενικής συνάρτησης είναι αρκετά απλός (εφαρμογή των τρεχουσών παραμέτρων του μετασχηματισμού και υπολογισμός της ολικής απόστασης μεταξύ των κατανομών), έναντι αυτού της αμοιβαίας πληροφορίας, όπου απαιτείται ο υπολογισμός της απο κοινού κατανομής σε κάθε βήμα του αλγορίθμου ελαχιστοποίησης και άρα την εύρεση του ιστογράμματος της από κοινού κατανομής.

Ένα επίσης βασικό ζήτημα για παραπάνω έρευνα είναι και αυτό της ενσωμάτωσης τεχνικών αυτόματης εξαγωγής του πλήθους των ομάδων, που ισοδυναμεί με την εύρεση του πλήθους των συνιστωσών του μικτού μοντέλου. Στη βιβλιογραφία υπάρχουν αρκετές εργασίες επί του θέματος. Ενδεικτικά αναφέρουμε [27],[2],[23],[5]. Τέλος μία ακόμα πτυχή που πιστεύουμε ότι πρέπει άμεσα να μελετηθεί είναι η παραλληλοποίηση της μεθόδου, ειδικά σε μία εποχή όπου η γενικότερη τάση στον ερευνητικό τομέα είναι ο παράλληλος υπολογισμός για καλύτερες αποδόσεις των διαφόρων αλγορίθμων.

Κλείνοντας θα θέλαμε να επισημάνουμε ότι προσπαθήσαμε να επεκτείνουμε τη χρήση της μεθόδου και σε υπέρθεση νεφών σημείων. Συγκρίναμε τα αποτελέσματά μας με τον παραδοσιακό και συχνά χρησιμοποιούμενο ICP αλγόριθμο. Τα αποτελέσματα ήταν εξίσου ενθαρρυντικά. Αναφορικά με τις ερευνητικές πτυχές που εμφανίζονται στην περίπτωση των σημείων μπορούμε να πούμε ότι ισχύουν όσο αναφέρθηκαν προηγουμένως. Να σημειωθεί όμως ότι καθώς η έννοια της αντιστοίχισης των επιμέρους περιοχών των διαμερίσεων δεν είναι αυτονόητη στο εν λόγω πρόβλημα υπέρθεσης, θεωρούμε ότι αξίζει μελέτης η εύρεση ενός εύρωστου και αξιόπιστου τρόπου εντοπισμού των τελευταίων.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] A. Goshtasby and G. C. Stockman. Point pattern using convex hull edges. *IEEE Transactions on Systems, Man and Cybernetics*, 15:631–637, 1985.
- [2] A. Likas and N. Vlassis and J. J. Verbeek. The global k-means clustering algorithm. *Pattern Recognition*, 36:451–461, 2002.
- [3] B. Jian and B. C. Vemuri. A robust algorithm for point set registration using mixture of Gaussians. In *Proceedings of the 10th International Conference on Computer Vision (ICCV'05)*, pages 1246–1251, Beijing, China, 2005.
- [4] B. Zitova and J. Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21(11):1–36, 2003.
- [5] C. Constantinopoulos and M. K. Titsias and A. Likas. Bayesian Feature and Model Selection for Gaussian Mixture Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(6):1013–1018, 2006.
- [6] C. H. Papadimitriou and K. Steiglitz. *Combinatorial optimization: algorithms and complexity*. Dover, 1998.
- [7] C. M. Bishop. *Neural networks for pattern recognition*. Oxford University Press Inc., Oxford, UK, 1995.
- [8] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.
- [9] C. Nikou and F. Heitz and J. P. Armspach. Robust voxel similarity metrics for the registration of dissimilar single and multimodal images. *Pattern Recognition*, 32:1351–1368, 1999.
- [10] C. Studholme and D. L. G. Hill and D. J. Hawkes. An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognition*, 32:71–86, 1998.
- [11] D. A. Forsyth and J. Ponce. *Computer Vision. A modern approach*. Pearson Educational International, 2003.
- [12] D. Peel and G. J. McLachlan. Robust mixture modeling using the t -distribution. *Statistics and Computing*, 10:339–348, 2000.
- [13] D.I. Barnea and H.F. Silverman. A class of algorithms for fast digital image registration. *IEEE Transactions on Computing*, 21:179–186, 1972.

- [14] E. D. Castro and C. Morandi. Registration of translated and rotated images using finite Fourier transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9:679–698, 1987.
- [15] F. Maes and A. Collignon and D. Vandermeulen and G. Marchal and P. Suetens. Multi-modality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16(2):187–198, 1997.
- [16] G. P. Penney and J. Weese and J. A. Little and P. Desment and D. L. G. Hill and D. J. Hawkes. A comparison of similarity measures for use in 3D-3D medical image registration. *IEEE Transactions on Medical Imaging*, 17:586–595, 1998.
- [17] G. Stockman and S. Kopstein and S. Benet. Matching images to models for registration and object detection via clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 4:229–241, 1982.
- [18] G. Wolberg and S. Zokai. Robust image registration using log-polar transform. In *Proceedings of the IEEE International Conference on Image Processing*, Canada, 2000.
- [19] J. B. A. Maintz and M. A. Viergever. A survey of medical image registration techniques. *Medical Image Analysis*, 2(1):1–36, 1998.
- [20] J. P. W. Pluim and J. B. A. Maintz and M. A. Viergever. Interpolation Artefacts in Mutual Information-Based Image Registration. *Computer Vision and Image Understanding*, 77:211–232, 1999.
- [21] L. G. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–376, 1992.
- [22] N. Ritter and R. Owens and J. Cooper and R. H. Eikelboom and P. P. van Saarloos. Registration of stereo and temporal images of the retina. *IEEE Transactions on Medical Imaging*, 18:404–418, 1999.
- [23] N. Vlassis and A. Likas. A Greedy EM Algorithm for Gaussian Mixture Learning. *Neural Processing Letters*, 15:77–87, 2002.
- [24] P. Besl and N. McKay. A method for registration of 3D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2), 1992.
- [25] P. Thévenaz and M. Unser. Optimization of mutual information for multiresolution image registration. *IEEE Transactions on Image Processing*, 9(12):2083–2099, 2000.
- [26] P. Viola and W. Wells III. Alignment by maximization of mutual information. *International Journal of Computer Vision*, 24(2):137–154, 1997.
- [27] S. Chen and H. Wang and B. Luo. Greedy EM algorithm for robust *t*-mixture modeling. In *Proceedings of the 3rd IEEE International Conference on Imaging and Graphics (ICIG'04)*, Hong-Kong, China, 2004.
- [28] S. Umeyama. Least-Squares Estimation of Transformation Parameters Between Point Patterns. *IEEE Transactions on pattern analysis and machine intelligence*, 13:376–380, 1991.

- [29] Szymon Rusinkiewicz. Efficient Variants of the ICP Algorithm. In *Proceedings of the 3rd IEEE International Conference on 3-D Digital Imaging and Modeling*, Quebec City, Canada, May 2001.

ΠΑΡΑΡΤΗΜΑ: Ο HUNGARIAN ΑΛΓΟΡΙΘΜΟΣ

Ο Hungarian αλγόριθμος είναι ένας αλγόριθμος συνδυαστικής βελτιστοποίησης, ο οποίος επιλύει το πρόβλημα της 1-με-1 ανάθεσης σε πολυωνυμικό χρόνο, $O(n^3)$. Ο Hungarian αλγόριθμος διατυπώθηκε από τον Harold Kuhn το 1955 και βελτιώθηκε από τον James Munkres το 1957.

Ο αλγόριθμος μοντελοποιεί το πρόβλημα της ανάθεσης ως έναν πίνακα κόστους $C : m \times n$ του οποίου το στοιχείο C_{ij} αναπαριστά το κόστος ανάθεσης της εργασίας j στον εργάτη i . Στην περίπτωση ελαχιστοποίησης το πρόβλημα μπορεί να επιλυθεί με εφαρμόζοντας Γκαουσιανή απαλοιφή. Στην περίπτωση της μεγιστοποίησης τα πράγματα γίνονται λίγο δυσκολότερα, αφού θα πρέπει να εφαρμοστεί πρώτα ένας μετασχηματισμός έτσι ώστε η ελαχιστοποίηση των τροποποιημένων στοιχείων να οδηγήσει τελικά σε μεγιστοποίηση.

- ΒΗΜΑ 0 (προαιρετικό)
Στην περίπτωση μεγιστοποίησης τροποποιούμε τον πίνακα κόστους C ως εξής:
 $M = \max\{C\}$, $C_{ij} \leftarrow M - C_{ij}$
- ΒΗΜΑ 1:
 - Για κάθε σειρά, βρίσκω το ελάχιστο στοιχείο της και το αφαιρώ από όλα τα στοιχεία της.
 - Για κάθε στήλη, βρίσκω το ελάχιστο στοιχείο της και το αφαιρώ από όλα τα στοιχεία της.
- ΒΗΜΑ 2:
Σχεδίαση γραμμών επί του πίνακα κόστους C έτσι ώστε να καλυφθούν όλα τα μηδενικά που έχουν εμφανιστεί ενώ παράλληλα το πλήθος των γραμμών είναι το ελάχιστο.
- ΒΗΜΑ 3:
Εάν το πλήθος των γραμμών του προηγούμενου βήματος είναι ίσο με το πλήθος των σειρών του πίνακα κόστους C , τότε έχουμε βρει τη βέλτιστη ανάθεση, ΤΕΛΟΣ.
- ΒΗΜΑ 4:
 - Βρίσκουμε το στοιχείο των σειρών που δεν καλύπτονται από τις γραμμές του βήματος 2 με την ελάχιστη τιμή. Έστω M η τιμή αυτή.
 - Αφαιρώ το M από τις σειρές και στήλες που δεν καλύπτονται από τις γραμμές που αναφέρονται στο βήμα 2.
 - Προσθέτω το M στις σειρές και στήλες που καλύπτονται από τις γραμμές που αναφέρονται στο βήμα 2.
 - Πήγαινε στο βήμα 3.

Αλγόριθμος 9: Ο Hungarian αλγόριθμος.

ΔΗΜΟΣΙΕΥΣΕΙΣ

- D. Gerogiannis, C. Nikou and A. Likas. Robust image registration using mixtures of t-distributions. 8th IEEE Computer Society Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA'07), in conjunction with ICCV'07, 14-20 October 2007, Rio de Janeiro, Brazil.
- D. Gerogiannis, C. Nikou and A. Likas. Rigid image registration based on pixel grouping. In Proceedings of the 14th International Conference on Image Analysis and Processing (ICIAP'07). 10-14 September 2007, Modena, Italy.

ΒΙΟΓΡΑΦΙΚΟ

Γεννήθηκα στα Ιωάννινα το 1983. Αποφοίτησα το 2000 από το 1ο Ενιαίο Λύκειο Ιωαννίνων, και το 2001 εισάγομαι στο τμήμα Πληροφορικής του Πανεπιστημίου Ιωαννίνων, από όπου και παίρνω το πτυχίο μου το 2005. Την ίδια χρονιά εισάγομαι στο Μεταπτυχιακό Πρόγραμμα της ίδιας σχολής. Στα ενδιαφέροντά μου συγκαταλέγεται η Υπολογιστική Όραση, η υπολογιστική νοημοσύνη και η βελτιστοποίηση.