



Τμήμα Ηλεκτρονικών Μηχανικών
και
Μηχανικών Υπολογιστών
ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ

Διπλωματική Εργασία με τίτλο:

*Ποσοτική Ανάλυση της Χωρικής Κατανομής των
Εντάσεων Δυναμικών Οπτικών Σημάτων:
Εφαρμογή στην Εκτίμηση του Ρίσκου Μετεξέλιξης
Νεοπλασμάτων σε Καρκίνο'*

Αρβανίτη Ευτυχία

Επιτροπή:

Επιβλέπων:

Μέλη:

Κώστας Μπάλας

Μιχάλης Ζερβάκης

Αλέξανδρος Ποταμιάνος

ΕΥΧΑΡΙΣΤΙΕΣ

Αρχικά θα ήθελα να ευχαριστήσω τον καθηγητή μου και επιβλέπων μέλος της εξεταστικής επιτροπής κύριο Κώστα Μπάλα για την συνεργασία κατά τη διάρκεια της διπλωματικής αυτής εργασίας, το πολύ ενδιαφέρον θέμα διπλωματικής που μου ανέθεσε, την καθοδήγηση, την υποστήριξη και τις ιδέες του που υλοποιήθηκαν στην εργασία αυτή. Ευχαριστώ τους καθηγητές Μιχάλη Ζερβάκη και Αλέξανδρο Ποταμιάνο για την συμμετοχή τους στην τριμελή επιτροπή της διπλωματικής αυτής εργασίας.

Πιο προσωπικά θα ήθελα να ευχαριστήσω την οικογένεια μου για την στήριξη της καθ όλη τη διάρκεια των προπτυχιακών σπουδών μου, τους πολύ καλούς μου φίλους στα Χανιά και τον Σ. Δημόπουλο για την ψυχολογική υποστήριξη τους κατά τη διάρκεια της διπλωματικής εργασίας. Τέλος όλους εσάς που διαβάσετε την εργασία αυτή.

ΠΟΣΟΤΙΚΗ ΑΝΑΛΥΣΗ ΤΗΣ ΧΩΡΙΚΗΣ ΚΑΤΑΝΟΜΗΣ ΤΩΝ
ΕΝΤΑΣΕΩΝ ΔΥΝΑΜΙΚΩΝ ΟΠΤΙΚΩΝ ΣΗΜΑΤΩΝ:
ΕΦΑΡΜΟΓΗ ΣΤΗΝ ΕΚΤΙΜΗΣΗ ΤΟΥ ΡΙΣΚΟΥ
ΜΕΤΕΞΕΛΙΞΗΣ ΝΕΟΠΛΑΣΜΑΤΩΝ ΣΕ ΚΑΡΚΙΝΟ

Copyright © October 2008 Diploma Thesis by Arvaniti Eutuxia



Πολυτεχνείο Κρήτης
Τμήμα Ηλεκτρονικών Μηχανικών και Μηχανικών Υπολογιστών
Τομέας Ηλεκτρονικής
Πολυτεχνειούπολη – Κουνουπιδιανά
73100 Χανιά – Ελλάδα

Επιτροπή:
Επιβλέπων: Κώστας Μπάλας
Μέλη: Μιχάλης Ζερβάκης
Αλέξανδρος Ποταμιάνος

Χανιά, Οκτώβριος 2008

ΠΕΡΙΛΗΨΗ

Σε αρκετές ιατρικές διαγνωστικές τεχνικές χρησιμοποιούνται διαγνωστικοί δείκτες (biomarkers) για την ενίσχυση της αντίθεσης μεταξύ φυσιολογικού και παθολογικού ιστού. Τα τελευταία χρόνια έχουν αναπτυχθεί τεχνολογίες οπτικής μοριακής απεικόνισης οι οποίες χρησιμοποιούν διαγνωστικούς δείκτες με την ιδιότητα να αλλάζουν τα χαρακτηριστικά σκέδασης ή/και φθορισμού του παθολογικού ιστού επιλεκτικά. Η *in vivo* μέτρηση των οπτικών αυτών σημάτων παρέχει πολύτιμες πληροφορίες για τις αλλαγές στην δομή και την λειτουργικότητα του ιστού, οι οποίες προκαλούνται κατά την εξέλιξη της νόσου από το προκαρκινικό στο καρκινικό στάδιο. Μέχρι σήμερα, δεν έχει προσδιοριστεί κάποια ποσοτική οπτική παράμετρος για την ποσοτικοποίηση του βαθμού οριοθέτησης (demarcation) της περιοχής που αντιδρά θετικά με τον διαγνωστικό δείκτη. Ο λόγος που αυτό κρίνεται ως ιδιαίτερα σημαντικό είναι ότι είναι δυνατόν μέσω της παραμέτρου αυτής να εκτιμηθεί η χωρική κατεύθυνση ανάπτυξης της νεοπλασίας στον ιστό και κατά συνέπεια του ρίσκου για μεταστάσεις. Ο προσδιορισμός μιας τέτοιας παραμέτρου στην περίπτωση του προκαρκίνου του τραχήλου της μήτρας, όπου αραιό διάλυμα οξικού οξέος χρησιμοποιείται ως διαγνωστικός δείκτης αποτελεί το αντικείμενο της διπλωματικής αυτής εργασίας.

Αναλύθηκαν δεδομένα από εικόνες οι οποίες έχουν ληφθεί σε χρονική διαδοχή για την καταγραφή σε κάθε εικονοστοιχείο των δυναμικών χαρακτηριστικών της οπισθοσκεδαζόμενης ακτινοβολίας. Οι παραγόμενες καμπύλες ένταση-χρόνος εκφράζουν τον ρυθμό πρόσληψης του διαγνωστικού δείκτη. Η εικόνα πριν την εφαρμογή του δείκτη αφαιρείται από αυτήν που λαμβάνεται στο peak της πρόσληψης για την παραγωγή μια νέας εικόνας η οποία εκφράζει την χωρική κατανομή της επαγόμενης από τον δείκτη μέγιστης αλλαγής των οπτικών χαρακτηριστικών. Στην συνέχεια υπολογίζεται το τρισδιάστατο προφίλ της εικόνας

αυτής. Τα μετρητικά δεδομένα του προφίλ εξομαλύνονται με την χρήση κατάλληλων αλγορίθμων προσομοίωσης και κατόπιν της επεξεργασίας αυτής υπολογίζονται αυτόματα, και σε ένα αντιπροσωπευτικό αριθμό προφίλ, η παράμετρος demarcation ως συνδυασμός κλίσεων και μεγίστων των προφίλ με συγκεκριμένα βάρη. Τελικά, χαρακτηρίζεται το συνολικό demarcation της βλάβης με μια και μοναδική παράμετρο, στη βάση της παραμετρικής ανάλυσης των επιμέρους προφίλ. Η παράμετρος αυτή υπολογίστηκε σε εκατό περιστατικά με ιστολογικά επιβεβαιωμένες χαμηλού και υψηλού βαθμού νεοπλασίες του τραχήλου της μήτρας. Βρέθηκε ότι μέσω της παραμέτρου αυτής είναι δυνατή η διάκριση μεταξύ των νεοπλασμάτων αυτών, ποσοτικά και αντικειμενικά και με ικανοποιητικό επίπεδο στατιστικής σημαντικότητας. Η δυνατότητα της νέας αυτής μεθόδου να εκτιμά επίσης τον κίνδυνο μετάστασης πρέπει να επιβεβαιωθεί με μια σειρά νέων κλινικών δοκιμών.

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΕΡΙΛΗΨΗ	01
ΚΕΦΑΛΑΙΟ 1: Κίνητρα και Αντικείμενο της εργασίας	
1.1 Εισαγωγικά	08
1.2 Δομή της εργασίας	17
ΚΕΦΑΛΑΙΟ 2: Θεωρητικό Υπόβαθρο	
2.1 Εικόνες και Ανθρώπινη Αντίληψη τους	18
2.2 Spatial Profiles	21
2.2.1 Εισαγωγικά	21
2.2.2 Επιλογή της Μεθόδου	21
2.2.3 Ανάλυση Profile – Δεδομένα – Συμπεράσματα	23
2.3 Θόρυβος εικόνας	29
ΚΕΦΑΛΑΙΟ 3: Προ-επεξεργασία Δεδομένων	
3.1 Τυχαία Δεδομένα	30
3.2 Τεχνικές Προ-επεξεργασίας Τυχαίων Δεδομένων	31
3.2.1 Smoothing	31
3.2.1.1 Moving Average	32
3.2.1.2 Median Filter	33
3.2.2 Curve Fitting	35
3.2.2.1 Regression Analysis	35
3.2.2.2 Parametric Fitting	36
3.2.2.3 Nonparametric Fitting	40
3.2.3 Διαφορές Smoothing αλγορίθμων – Fitting	41
3.3 Επιλογή Μεθόδου	42
ΚΕΦΑΛΑΙΟ 4: Επεξεργασία και Ανάλυση Δεδομένων – Αλγόριθμοι Υλοποίησης	
4.1 Βελτιστοποιήσεις	48
4.2 Ερμηνεία Χαρακτηριστικών - Εξαγωγή Παραμέτρων	49
4.3 Αλγόριθμοι Εξαγωγής Χαρακτηριστικών για κάθε Profile	53
4.4 Συνδυασμός Παραμέτρων για κάθε Profile	63

ΚΕΦΑΛΑΙΟ 5:	Αποτελέσματα	
5.1	Επιλογή Τελικών Παραμέτρων για κάθε εικόνα	66
5.2	Αποτελέσματα – Ερμηνεία Αποτελεσμάτων	69
5.3	Αποτελέσματα – Στατιστική Ανάλυση Αποτελεσμάτων	76
5.4	Συμπεράσματα	84
ΚΕΦΑΛΑΙΟ 6:	Τεχνικές Βελτίωσης – Μελλοντική Έρευνα	85
ΠΑΡΑΡΤΗΜΑ Α:	Αποτελέσματα & Εικόνες ανά ασθενή	87
	- 3D Profiles, subtracted εικόνες, χάρτης Dysis -	
ΠΑΡΑΡΤΗΜΑ Β:	Κομμάτια Κώδικα Υλοποίησης	107
ΑΝΑΦΟΡΕΣ – ΒΙΒΛΙΟΓΡΑΦΙΑ		112

Λίστα Εικόνων

- Εικόνα 1: Δομή του επιθηλιακού ιστού σε παθογόνα και μη κύτταρα
- Εικόνα 2: Dysis
- Εικόνα 3: Ένταση του οπισθοσκεδασμένου φωτός για παθογόνα και μη κύτταρα
- Εικόνα 4: Χάρτης Dysis
- Εικόνα 5: (a),(b),(c),(d),(e) Διαδικασία εξαγωγής subtracted grayscale εικόνας
- Εικόνα 6: Conditional contrast
- Εικόνα 7: Ebbinghaus illusion
- Εικόνα 8: (a) Spatial profile γραμμής (b) Εικόνα προς εξαγωγή spatial profile
- Εικόνα 9: Spatial profiles 3 τυχαίων γραμμών της εικόνας
- Εικόνα 10: Spatial profile-χρωματισμένη περιοχή που διακρίνει το ανθρώπινο μάτι
- Εικόνα 11: Spatial profile 3 τυχαίων γραμμών της εικόνας
- Εικόνα 12: Τυχαίο spatial profile με θόρυβο
- Εικόνα 13: Σύγκριση μεγέθους span. (a) Αρχικά δεδομένα (b) Smoothed δεδομένα με μέγεθος span 9 (c) Με span 31 (d) Με span 61
- Εικόνα 14: Σύγκριση μεγέθους span. (a) Αρχικά δεδομένα (b) Smoothed δεδομένα με μέγεθος span 9 (c) Με span 31 (d) Με span 61
- Εικόνα 15: Παράδειγμα fitting
- Εικόνα 16: Smoothing spatial profile
- Εικόνα 17: Smoothing spline με smoothing parameter $sp=0,34^{10-5}$, $sp=0,7^{10-4}$, $sp=0,9^{10-3}$
- Εικόνα 18: Gaussian function
- Εικόνα 19: Smoothing spline vs parametric gauss2 fitting
- Εικόνα 20: Fitted profiles με gauss2 a,b,c,d,e
- Εικόνα 21: (a) 3D αρχικής εικόνας (b) 3D με fitting όλων των profile (c) zoom βλάβης

Εικόνα 22: Θόρυβος στην εικόνα

Εικόνα 23: Profile με σχεδόν ίσες μέσες τιμές $\mu_{1,2}$ και (a) σχεδόν ίσες διασπορές $\sigma_{1,2}$ (b) διαφορετικές διασπορές $\sigma_{1,2}$

Εικόνα 24: (a) Low demarced βλάβη (b) High demarced βλάβη

Εικόνα 25: (a) Spatial profile low contrast εικόνας (profile 1) (b) Spatial profile high contrast εικόνας (profile 2)

Εικόνα 26: (a) Spatial profile με ένα τοπικό μέγιστο, διασπορές σε κοντινές τιμές (b),(c) Spatial profile με ένα τοπικό μέγιστο, διασπορές με μεγάλη διαφορά

Εικόνα 27: Spatial profile με ένα τοπικό μέγιστο, gaussian fitted συνάρτηση μιας καμπάνας

Εικόνα 28: Παράμετροι, ύπαρξη μιας peak τιμής

Εικόνα 29: Προβληματικό profile. Εύρεση νέων παραμέτρων

Εικόνα 30: (a), (b), (c) Profile με δύο T.M, (c) δύο βλάβες

Εικόνα 31: Παράμετροι για profile με 2 τοπικά μέγιστα

Εικόνα 32: (a) Profile με δύο T.M και μικρή τιμή διασποράς σ_i (b) Διαχείριση προβληματικών profile με 2 T.M

Εικόνα 33: 2 βλάβες σε απόσταση μεγαλύτερη των 200 pixel

Εικόνα 34: Gaussian κατανομή

Εικόνα 35: Υπολογισμός των σημείων x_1 , x_2

Εικόνα 36: Επιλογή γραμμών-στηλών όπου υπολογίζονται οι τελικοί παράμετροι του demarcation

Εικόνα 37: Subtracted εικόνα με artifacts και RGB original εικόνα

Εικόνα 38: (a) Median φίλτρο 25 στην subtracted εικόνα (b) Επιλογή profile μετά την εφαρμογή median filter 25

Εικόνα 39: (a) Εικόνα $t=0$ (b) Εικόνα $t=\max(\text{acetowhitening})$ (c) Subtracted εικόνα

Εικόνα 40: (a) 3D απεικόνιση original εικόνας (b) 3D απεικόνιση εικόνας ύστερα από fitting όλων των profile (c) zoom βλάβης

Εικόνα 41: Ασθενής 16_72. 3D απεικόνιση των επιλεγμένων 'φιλτραρισμένων' & μοντελοποιημένων profile

Εικόνα 42: Ασθενής 16_72. Επιλεγμένα profile εξαγωγής τελικών παραμέτρων

Εικόνα 43: (a) Το χειρότερο demarcation της βλάβης (b) Επιθηλιακός ιστός (c) χάρτης Dysis

Εικόνα 44: (a) Επιλογή profile. Subtracted εικόνα (b) Επιθηλιακός ιστός (c) Χάρτης Dysis

Εικόνα 45: (a) True negative subtracted εικόνα (b) Χάρτης Dysis

Εικόνα 46: Δέντρο κατηγοριοποίησης

Συνομμεύσεις

AW : acetowhitening

RCI : Reids Colposcopic Index

Επεξηγήσεις

True positive : είναι τα θετικά τεστ που προέρχονται από πληθυσμό των εχόντων την ασθένεια.

True negative : είναι τα αρνητικά τεστ που προέρχονται από πληθυσμό των μη εχόντων την ασθένεια.

False positive : είναι τα θετικά τεστ που προέρχονται από πληθυσμό των μη εχόντων την ασθένεια.

False negative : είναι τα αρνητικά τεστ που προέρχονται από πληθυσμό των εχόντων την ασθένεια.

1.1 Εισαγωγικά

Ο καρκίνος του τραχήλου της μήτρας είναι η πιο γνωστή μορφή καρκίνου σε γυναίκες κάτω των 35 ετών και η δεύτερη πιο γνωστή μορφή καρκίνου σε γυναίκες όλων των ηλικιών παγκοσμίως. Η κύρια αιτία του καρκίνου του τραχήλου της μήτρας προκαλείται από τον ιό HPV (human papillomavirus), η πιο γνωστή σεξουαλικά μεταδιδόμενη ασθένεια παγκοσμίως. Το 50 με 80% του γυναικείου πληθυσμού θα έρθει αντιμέτωπη με ορισμένους τύπους του ιού κάποια στιγμή της ζωής του. Η εξέλιξη της μόλυνσης του ιού μπορεί να οδηγήσει με πιθανότητα περίπου 60% σε νεοπλασία του επιθηλιακού ιστού του τραχήλου (CIN) ακόμα και σε καρκίνο του τραχήλου της μήτρας [4].

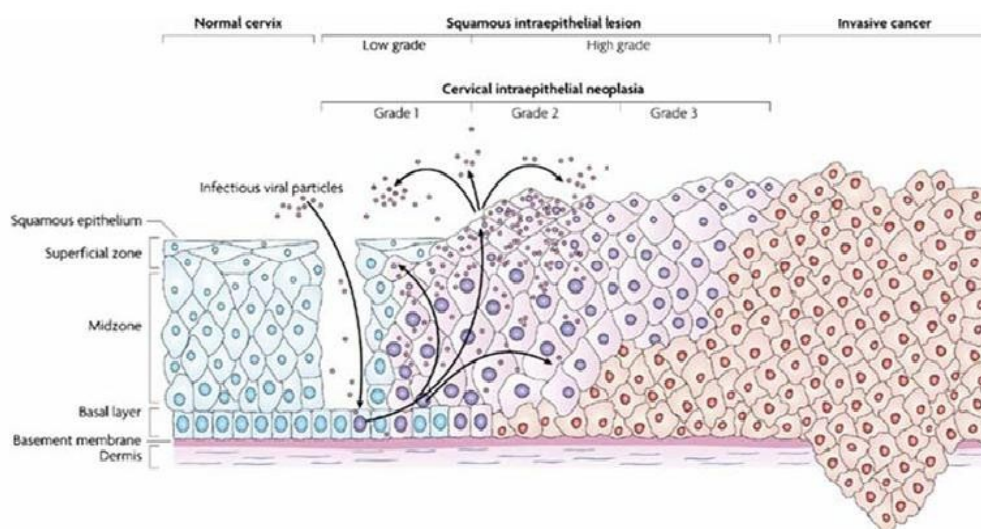
Ο επιθηλιακός ιστός ή αλλιώς επιθήλιο αποτελείται από κύτταρα στενά συνδεδεμένα μεταξύ τους με ελάχιστη μεσοκυττάρια ουσία, σχηματίζοντας εκτεταμένες επιφάνειες που καλύπτουν εξωτερικά τις διάφορες κοιλότητες και επιφάνειες του σώματος, έχοντας κυρίως προστατευτικό ρόλο για τα όργανα του οργανισμού. Ο επιθηλιακός ιστός του τραχήλου της μήτρας αποτελείται από στρώματα κυττάρων με διαφορετικά μορφολογικά χαρακτηριστικά. Η εικόνα 1 απεικονίζει τα στρώματα αυτά τα οποία σε μη παθολόγο ιστό είναι: το βασικό (basal layer) στρώμα που αποτελείται από κύτταρα κυβικού σχήματος, το μεσαίο στρώμα (midzone) και το επιφανειακό στρώμα (superficial zone). Η ωρίμανση των κυττάρων είναι η διαδικασία κατά την οποία τα κύτταρα αλλάζουν μορφολογικά και λειτουργικά κατά τη διάρκεια της ζωής. Κατά την ωρίμανση τους τα κύτταρα μετακινούνται σταδιακά από το βασικό στρώμα προς το επιφανειακό και εξελίσσονται σε κύτταρα με πιο επίπεδο σχήμα και με μεγαλύτερο όγκο [5].

Οι αλλοιώσεις (CIN) του επιθηλιακού ιστού του τραχήλου κατηγοριοποιούνται σε στάδια CIN I, II, III ανάλογα με την σοβαρότητα τους και το στρώμα των κυττάρων τα οποία έχουν επιμολυνθεί. Δεν υπάρχουν ποσοτικά κριτήρια κατηγοριοποίησης

των σταδίων αυτών και γι'αυτό η διάγνωση είναι κυρίως υποκειμενική. Σε γενικές γραμμές ο άξονας κατηγοριοποίησης των αλλοιώσεων παρουσιάζεται στον πίνακα.

Στάδιο αλλοίωσης	Ιστολογική Περιγραφή
CIN I: ήπια δυσπλασία	Δυσπλαστικά κύτταρα στο βασικό στρώμα καταλαμβάνουν το 1/3 του επιθηλιακού ιστού.
CIN II: μέτρια δυσπλασία	Δυσπλαστικά κύτταρα στο μεσαίο στρώμα καταλαμβάνουν τα 2/3 του επιθηλιακού ιστού.
CIN III: δριμύς δυσπλασία ή καρκίνωμα	Δυσπλαστικά κύτταρα σε όλο το πάχος του επιθηλίου.

Το στάδιο CIN I θεωρείται στάδιο χαμηλού κινδύνου (low-grade) και θεωρείται πιθανή η αντιμετώπιση του και από τον ίδιο τον οργανισμό, ενώ στο στάδιο CIN II, CIN III θεωρείται υψηλού κινδύνου (high-grade) που μπορεί να εξελιχθεί σε καρκίνο [5]. Στην εικόνα 1 παρουσιάζονται τα στάδια των αλλοιώσεων όπου με μπλε χρώμα είναι ο φυσιολογικός ιστός, με μοβ τα στάδια των αλλοιώσεων (προ καρκινικά κύτταρα) και με κόκκινο τα καρκινικά.



Εικόνα 1: Δομή του επιθηλιακού ιστού σε παθολόγο και μή κύτταρα [4]

Η βιοϊατρική επιστήμη στοχεύει στην εύρεση διαγνωστικών μεθόδων για έγκαιρο και έγκυρο εντοπισμό και αναγνώριση της μόλυνσης αυτής σε πρώιμα στάδια όπως CIN I [4]. Οι διαγνωστικές μέθοδοι ονομάζονται οι ιατρικοί έλεγχοι-εξετάσεις που γίνονται με σκοπό τον εντοπισμό και την πρόληψη ασθενειών, μειώνοντας όσο το

δυνατόν τα περιθώρια λανθασμένης διάγνωσης [7]. Το τεστ Παπανικολάου (κυτταρολογική εξέταση) είναι η κύρια εξέταση ανίχνευσης αλλοιώσεων χαμηλού είτε υψηλού κινδύνου και πραγματοποιείται με λήψη δείγματος κυττάρων από τον επιθηλιακό ιστό του τραχήλου και κυτταρολογική εξέταση τους στο μικροσκόπιο από ειδικό. Η ακρίβεια του τεστ περιορίζεται από σφάλματα κατά τη δειγματοληψία και την εξέταση των κυττάρων, γεγονός που οδηγεί σε χαμηλή ευαισθησία της διαγνωστικής μεθόδου (~0.59) [6]. Σε περίπτωση μη-φυσιολογικού ΠΑΠ τεστ η εξέταση ανίχνευσης αλλοιώσεων συνεχίζεται με κολποσκοπική εξέταση όπου στην κλασική μέχρι σήμερα μέθοδο εξέτασης χρησιμοποιείται low-power μικροσκόπιο και διαγνωστικοί δείκτες με την ιδιότητα να ενισχύουν την αντίθεση παθογόνου και μη παθογόνου ιστού. Η διαγνωστική μέθοδος της κολποσκόπησης βασίζεται σε οπτική παρατήρηση του τραχήλου από τον θεράποντα ιατρό και η τυχόν αλλοίωση κατηγοριοποιείται σύμφωνα με την εμπειρία του εξεταστή σε CIN I,II,III. Η κολποσκόπηση σαν διαγνωστική μέθοδος είναι αποδοτική στην ανίχνευση της θέσης της μόλυνσης αλλά ανακριβής στην ανίχνευση της σοβαρότητας της. Για το λόγο αυτό κρίνεται απαραίτητα η συνέχιση της εξέτασης ανίχνευσης αλλοιώσεων με ιστολογική εξέταση (βιοψία) από τις ύποπτες περιοχές που έχουν εντοπιστεί. Η ακρίβεια της βιοψίας περιορίζεται από σφάλματα κατά τη δειγματοληψία. Είναι μια εξέταση ποιοτική όπου δεν καθίσταται δυνατός ο εντοπισμός της παθογόνου περιοχής και μπορεί να κρύβει κινδύνους [4]. Είναι δαπανηρή διαδικασία καθώς και ψυχοφθόρα για τον ασθενή.

Δεδομένων των αναφερομένων και της συνεχόμενης αύξησης των κρουσμάτων και στις αναπτυσσόμενες χώρες, είναι εμφανής η αναγκαιότητα εύρεσης αποδοτικών μη δαπανηρών διαγνωστικών μεθόδων. Η κυτταρολογική και ιστολογική εξέταση ανίχνευσης αλλοιώσεων του τραχήλου βασίζονται σε εξέταση ιστών εκτός ζωντανού οργανισμού (*in vitro*). Τα τελευταία χρόνια, μεγάλος αριθμός διαγνωστικών μεθόδων βασίζεται σε τεχνικές οπτικής μοριακής απεικόνισης των ιστών όπου η χρήση διαγνωστικών δεικτών (biomarkers) καθιστά δυνατή την αλλαγή των οπτικών ιδιοτήτων του παθογόνου ιστού (π.χ. τα χαρακτηριστικά σκέδασης). Η *in vivo* παρακολούθηση του δυναμικού φαινομένου μπορεί να οδηγήσει σε διάγνωση καθώς και κατανόηση της δομής του παθογόνου ιστού [4].

Συγκεκριμένα, για την ανίχνευση των αλλοιώσεων στον επιθηλιακό ιστό ο διαγνωστικός δείκτης είναι το οξικό οξύ και ο ψεκάσμος του ιστού με τον δείκτη μπορεί να διαχωρίσει οπτικά τα παθογόνα κύτταρα από τα υγιή. Η εργασία αυτή ανοίγει ένα νέο κεφάλαιο εκτίμησης της σημασίας των χαρακτηριστικών των δυναμικών οπτικών σημάτων. Η σύγχρονη βιοϊατρική μελέτη στοχεύει στην κατασκευή έγκυρων διαγνωστικών εργαλείων που παρέχουν εξέταση σε ζωντανό οργανισμό (*in vivo*) και άμεση, μη χρονοβόρα εξαγωγή αποτελεσμάτων με δυνατότητα οπτικής παρακολούθησης-ανάγνωσης μέσω οθόνης. Το ιατρικό μηχάνημα DySIS παρέχει *in vivo* ανίχνευση των αλλοιώσεων του τραχήλου της μήτρας και τείνει να αντικαταστήσει την μέχρι τώρα κλασική μέθοδο κολποσκόπησης.



Εικόνα 2: DySIS [12]

Η λειτουργία του βασίζεται στην ιδιότητα του οξικού οξέος (3 – 5% acetic acid) να αντιδρά ειδικά με τα μη-φυσιολογικά επιθηλιακά κύτταρα μεταβάλλοντας προοδευτικά και αντιστρεπτά τις οπτικές ιδιότητες σκέδασης του φωτός που διαπερνά τα κύτταρα αυτά. Η αντίδραση του οξικού οξέος με παθογόνα επιθηλιακά κύτταρα παρέχει στον παρατηρητή την οπτική αντίληψη του φαινομένου αυτού μέσω του acetowhitening (AW) [12].

Το AW είναι ένα δυναμικό οπτικό φαινόμενο που γίνεται αντιληπτό από την ανθρώπινη όραση ως άσπρισμα της παθογόνου περιοχής ύστερα από τον ψεκάσμό οξικού οξέος. Κλινικές έρευνες έδειξαν ότι η διάρκεια και το ποσοστό του ασπρίσματος της παθογόνου περιοχής σχετίζεται με τον βαθμό της βλάβης (CIN) [6]. Το AW πραγματοποιείται προοδευτικά και κατά την χρονική στιγμή $t=t$ παίρνει την μέγιστη του ένταση. Έχει μελετηθεί και μοντελοποιηθεί στο Πολυτεχνείο Κρήτης το βιοφυσικό υπόβαθρο του φαινομένου του AW, σύμφωνα με την μορφολογία του επιθηλιακού ιστού [6].

Σήμερα, το επιστημονικό ενδιαφέρον του δυναμικού φαινομένου του AW είναι έντονο δεδομένης της προσπάθειας εύρεσης αποδοτικών μεθόδων διάγνωσης,

χαμηλού κόστους και χωρίς επιβλαβείς συνέπειες στην υγεία του ασθενή. Πρόσφατες έρευνες σε ευρύ σύνολο γυναικών αποδεικνύουν την αποδοτικότητα της επιθεώρησης του φαινομένου του AW ως μέθοδο διάγνωσης. Σύμφωνα με τις έρευνες αυτές όπου έλαβαν μέρος 50.000 γυναίκες, παρατηρήθηκε μείωση κατά 30% των περιστατικών του καρκίνου του τραχήλου της μήτρας μέσω οπτικής παρατήρησης του AW σε σχέση με την καθιερωμένη αλυσίδα εξετάσεων μέχρι τώρα (κυτταρική ανάλυση, κολποσκόπηση, ιστολογική εξέταση)[6]. Η οπτική επιθεώρηση του AW λαμβάνεται υπόψιν από τους ιατρούς αν και στερείται αντικειμενικότητας διότι είναι ένα καθαρά οπτικό φαινόμενο που η οπτική αντίληψη του εξαρτάται από τον παρατηρητή. Τα κύρια διαγνωστικά κριτήρια της άμεσης οπτικής επιθεώρησης του AW βασίζονται στις εξής παραμέτρους: την ταχύτητα και το μέγεθος της διαδικασίας ασπρίσματος, τον βαθμό ασπρίσματος την στιγμή που συμβαίνει το μέγιστο άσπρισμα και τον βαθμό απότομης ή μη οριοθέτησης των παθογόνων κυττάρων από τα υγιή (*demarcation*). Η διαγνωστική ακρίβεια εξαρτάται από τον τρόπο που κάθε ιατρός εφαρμόζει τις κατευθυντήριες γραμμές (στην κολποσκόπηση: δείκτες κολποσκόπησης Reid -RCI-). Εξετάζοντας τα χαρακτηριστικά: χρώμα, οριοθέτηση βλάβης και επιφανειακό σχηματισμό (*demarcation*), γειτονικά αγγεία, ιωδιούχο βάψιμο, προκύπτει η περιγραφή της εκάστοτε βλάβης όπου για κάθε χαρακτηριστικό παρέχονται αριθμητικοί δείκτες ανάλογα με την περιγραφή, το άθροισμα των οποίων δίνει τις εξής διαγνώσεις ανάλογα με την επικινδυνότητα:

0-2 : πιθανόν CIN I

3-4 : υπέρθετη βλάβη. Πιθανόν CIN I ή CIN II

5-8 : πιθανόν CIN II ή CIN III.

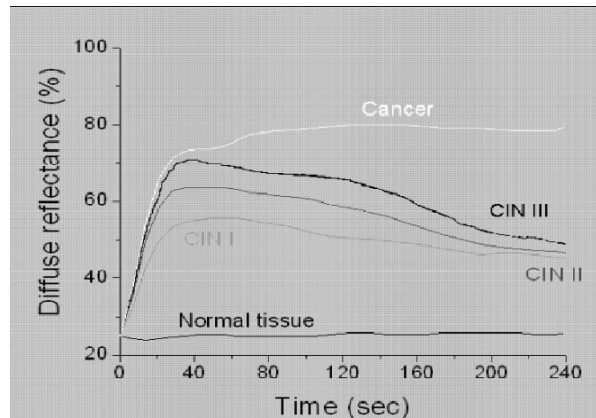
Η περιγραφική διάγνωση του ιατρού καθιστά σαφές το χρώμα και το *demarcation* της βλάβης [16]. Το *demarcation* μιας βλάβης ορίζεται ως ο βαθμός οριοθέτησης των παθογόνων κυττάρων από τα υγιή. Οι δύο 'κατηγορίες' κυττάρων διαχωρίζονται λόγω του AW και όσο πιο απότομος ο διαχωρισμός των δύο περιοχών, τόσο πιο *demarcated* η βλάβη. Το *demarcation* μέχρι τώρα δεν έχει ποσοτικοποιηθεί και αρκείται στον περιγραφικό προσδιορισμό του σύμφωνα με την μέση οπτική αντίληψη του παρατηρητή.

Η διπλωματική αυτή εργασία παρέχει για πρώτη φορά την ποσοτικοποίηση του

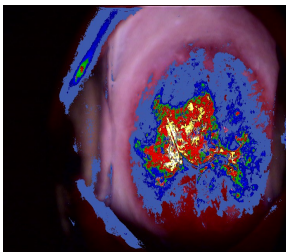
demarcation εξάγοντας ποσοτικές οπτικές παραμέτρους από εικόνες. Οι εικόνες προκύπτουν από το ιατρικό μηχάνημα DySIS. Δεν έχει αποδειχτεί ακόμα κλινικά η σημασία του demarcation και η ερμηνεία του στον επιθηλιακό ιστό. Η πρώτη προσέγγιση που παρέχει η εργασία είναι η ποσοτικοποίηση του σύμφωνα με τις μεθόδους που αναφέρονται και τα συμπεράσματα που εξάγονται από αυτήν. Το demarcation αναφέρεται στο δυναμικό φαινόμενο του AW αλλά δεν είναι μια δυναμική παράμετρος. Από την εργασία αποδεικνύεται ότι σύμφωνα με τις τελικές εξαγόμενες παραμέτρους ποσοτικοποίησης, μπορούν να διαχωριστούν βλάβες χαμηλού και υψηλού κινδύνου. Οι τελικές παράμετροι μπορούν επιπροσθέτως να λειτουργήσουν ως ποσοτικό διαγνωστικό κριτήριο της οπτικής επιθεώρησης του AW προσφέροντας αντικειμενικότητα σε ένα από τα διαγνωστικά κριτήρια, ως επιπρόσθετη διαγνωστική παράμετρος του Dysis αλλά και να παρέχει την έννοια της πρόβλεψης της κατεύθυνσης των παθογόνων κυττάρων στο επιθήλιο. Σύμφωνα με το τελευταίο μπορεί να χρησιμοποιηθεί σε ασθενείς που τα διαγνωστικά τεστ τους αποδείχτηκαν αληθώς αρνητικά (*true negative*) δίνοντας μια μελλοντική πρόβλεψη πιθανής μετάστασης. Η απόδειξη αυτού απαιτεί κλινικές και μακροχρόνιες δοκιμές.

- DySIS – Δεδομένα της Εργασίας

Το ιατρικό μηχάνημα DySIS παρέχει *in vivo* ανίχνευση αλλοιώσεων, άμεση οπτική χαρτογράφηση της βλάβης (χάρτης DySIS) και διάγνωση ενώ ταυτόχρονα καθιστά δυνατό τον ακριβή εντοπισμό (προ)-καρκινικών κυττάρων ώστε η λήψη βιοψίας ή η τυχόν εγχείρηση να πραγματοποιηθούν σε σωστά σημεία. Η λειτουργία του βασίζεται στην χρονική απορρόφηση του οξικού οξέος από τα κύτταρα που έχουν ψεκαστεί και αντιπροσωπεύονται από τις καμπύλες έντασης-χρόνου της εικόνας 3. Κάθε καμπύλη αναφέρεται σε μια περίπτωση αλλοίωσης. Κάποια χρονική στιγμή (στο μέγιστο peak) πραγματοποιείται η μέγιστη απορρόφηση. Από κάθε καμπύλη προκύπτουν δυναμικές παράμετροι (peak, ολοκλήρωμα καμπύλης) οι τιμές των οποίων ομαδοποιούνται σε χρώματα, εξάγοντας τον χάρτη του DySIS (εικόνα 4) ο οποίος οδηγεί στη διάγνωση [12].



Εικόνα 3: Ένταση του οπισθοσκεδασμένου φωτός για παθολόγια και μη κύτταρα [12]

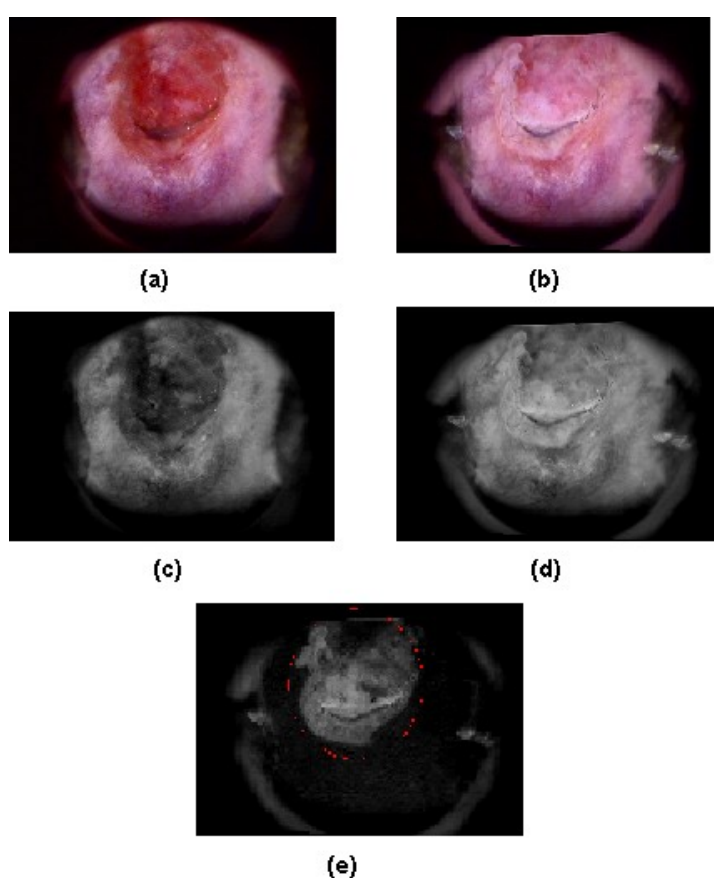


Εικόνα 4: Χάρτης Dysis

Κάθε χρώμα αντιστοιχεί σε μια περίπτωση αλλοίωσης (μπλε: CIN I, πράσινο: CIN II, κόκκινο: CIN III, άσπρο: καρκίνος). Ο χάρτης Dysis στα πλαίσια της εργασίας λειτουργεί για επιβεβαίωση της βλάβης. Το ενδεχόμενο να λειτουργούν οι παράμετροι του demarcation αποδοτικότερα από τις δυναμικές παραμέτρους θα μελετηθεί σύντομα.

Το Dysis απεικονίζει ψηφιακά τον ιστό κάθε γυναίκας που υποβάλλεται στην εξέταση από την χρονική στιγμή $t=0$ όπου δεν έχει ψεκαστεί το οξικό οξύ ανά 5 s μέχρι την χρονική στιγμή $t=240$ s. Το μέγιστο AW συμβαίνει τη χρονική στιγμή $t=\tau$ όπου οι καμπύλες δίνουν το μέγιστο peak. Το DySIS λύνει το πρόβλημα των αυθόρμητων κινήσεων των ιστών είτε της ασθενούς κάνοντας ευθυγράμμιση (registration) όλων των εικόνων που λαμβάνει. Οι εικόνες που επεξεργαζόμαστε στην εργασία είναι εικόνες που αναφέρονται στις χρονικές στιγμές $t=0$ και $t=\tau$. Έχει αποδειχτεί ότι η μέγιστη αντίθεση μεταξύ παθολόγου και μη ιστού είναι ορατή στο μήκος κύματος $\lambda=525\pm 15\text{nm}$ [4]. Για το λόγο αυτό οι RGB λαμβανόμενες εικόνες από το DySIS γίνεται grayscale στο πράσινο. Εν συνεχεία η τελική εικόνα επεξεργασίας μας (*subtracted*) προκύπτει από τη χρονική στιγμή $t=\tau$ αφαιρώντας την εικόνα την χρονική στιγμή $t=0$ με σκοπό τον αποκλεισμό άσπρων εκ γενετής περιοχών που πιθανόν να ερμηνευτούν ως AW. Η αφαίρεση γίνεται ανά pixel. Η *subtracted* εικόνα περιέχει μόνο την αλλαγή που συνέβη ύστερα από χρόνο $t=\tau$ άρα

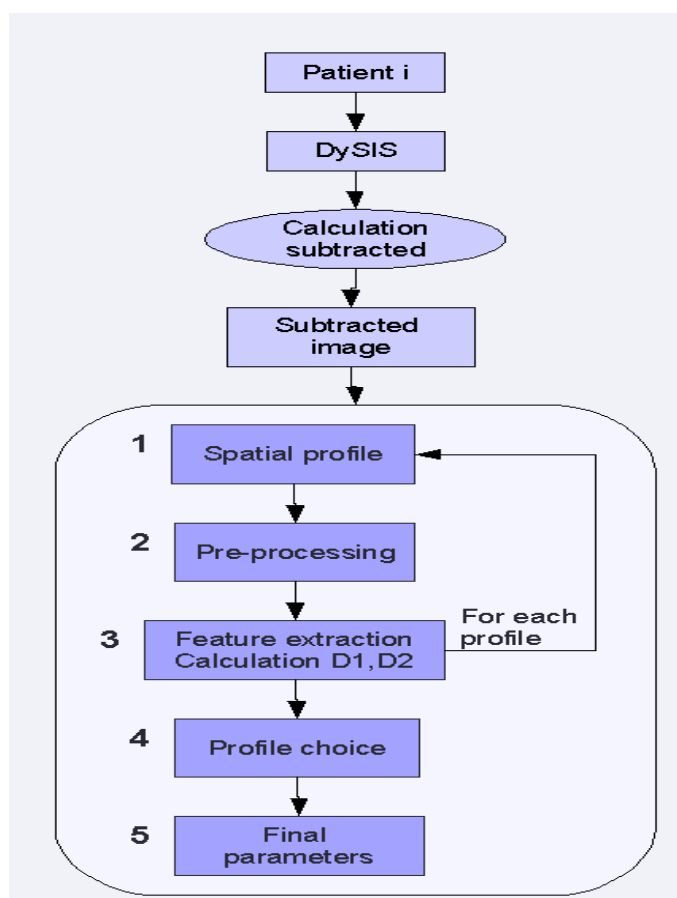
μόνο το AW μαζί με θόρυβο εικόνας ο οποίος εκτός του συνηθισμένου ψηφιακού θορύβου προέρχεται και από εργαλεία εξέτασης. Μέρος της εργασίας είναι η προσπάθεια απαλοιφής των θορύβων αυτών ώστε η επιλογή παραμέτρων να μην οδηγεί σε σφάλματα. Κάθε subtracted εικόνα αναφέρεται σε μία ασθενή. Η διαδικασία απόκτησης της subtracted αναφέρεται στην εικόνα 5. Η RGB εικόνα 5a απεικονίζει τον τράχηλο της μήτρας τη χρονική στιγμή $t=0$ ενώ η 5b την $t=\tau$ του μέγιστου AW. Η εικόνα 5c είναι η grayscale της 5a ενώ η 5d της 5b. Η subtracted είναι η εικόνα 5e όπου το AW είναι ότι φαίνεται στον κόκκινο κύκλο. Ο άσπρος κύκλος και οι κουκκίδες είναι θόρυβος από ανακλάσεις των εργαλείων εξέτασης.



Εικόνα 5: (a) $t=0$, (b) $t=\tau$, (c) $t=0$, grayscale($\lambda=525nm$) , (d) $t=\tau$, grayscale($\lambda=525nm$) , (e) subtracted ($\tau-t(0)$) grayscale($\lambda=525nm$) source: DySIS

Το DySIS σαν ιατρικό μηχάνημα εξέτασης υπάρχει στα νοσοκομεία Αλεξάνδρα, Hammersmith, StMarys και οι ασθενείς μας είναι από τα νοσοκομεία αυτά. Για κάθε μία γυναίκα που υπόκειται στην εξέταση με το DySIS εξάγονται 1. ο χάρτης DySIS

που δίνει τη διάγνωση και 2. οι εικόνες όμοιες με τις 5a,5b οι οποίες δίνουν την subtracted. Οι τελευταίες επεξεργάζονται εκμεταλλευόμενοι τα χωρικά χαρακτηριστικά τους (spatial profile). Από κάθε εικόνα εξάγονται 38 profile γραμμών και 48 στηλών ώστε να είναι δυνατή η προσέγγιση του demarcation από όλες της πλευρές της βλάβης. Κάθε spatial profile επεξεργάζεται ξεχωριστά επιλέγοντας ένα μοντέλο περιγραφής των δεδομένων (συνάρτηση παραμετρικού fitting) και εν συνεχεία εξάγονται χαρακτηριστικά ο συνδυασμός των οποίων εκφράζει το demarcation. Επιλέγονται ενδεικτικά 20 profile με τις μέγιστες τιμές peak από όπου προκύπτουν οι τελικές παράμετροι για κάθε subtracted εικόνα. Αποδείχτηκε με t-Test ότι οι δύο ομάδες low και high-grade είναι διαχωρίσιμες σύμφωνα με τις τιμές των τελικών παραμέτρων. Προτείνεται διαισθητικό σύστημα κατηγοριοποίησης μεταξύ low και high grade αλλοιώσεων σύμφωνα με τις εξαγόμενες παραμέτρους. Η διαδικασία της επεξεργασίας για εξαγωγή τελικών παραμέτρων παρουσιάζεται από το διάγραμμα:



1.2 Δομή της εργασίας

Αναλυτικότερα, στο *Κεφάλαιο 2* αναφέρεται το θεωρητικό υπόβαθρο στο οποίο στηριχθήκαμε για την μελέτη και ανάλυση της εργασίας. Αναφέρονται περιληπτικά οι έννοιες των εικόνων και της αντίληψης τους από την ανθρώπινη όραση. Αναλύεται η τεχνική των spatial profile στην οποία βασίζεται η εργασία, ο λόγος επιλογής της τεχνικής και αναφέρονται οι πληροφορίες που εξάγονται από τα profile. Ερμηνεύονται οι μορφές των spatial profiles σε σχέση με τον επιθηλιακό ιστό ερμηνεύοντας την έννοια της πρόβλεψης.

Στο *Κεφάλαιο 3* αναφέρεται αρχικά η θεωρητική ανάλυση των τεχνικών προ-επεξεργασίας δεδομένων που δοκιμάστηκαν (τεχνικές smoothing, μέθοδοι curve fitting). Αναλύονται τα βήματα και οι λόγοι της επιλογής της μεθόδου smoothing για κάθε profile. Μέσω των τεχνικών αυτών τα εξαγόμενα από την subtracted εικόνα profile γραμμών/στηλών προ-επεξεργάζονται (μοντελοποιούνται) με σκοπό την εξαγωγή χαρακτηριστικών.

Το *κεφάλαιο 4* περιλαμβάνει όλους τους αλγορίθμους επεξεργασίας ανά smoothed spatial profile. Αναφέρονται οι αλγόριθμοι επεξεργασίας των profile με αναποτελεσματικό fitting, γίνονται βελτιστοποιήσεις και παραδοχές, ερμηνεύεται η επιλογή των χαρακτηριστικών που εξάγονται από κάθε profile και ο συνδυασμός τους ώστε να εκφράζουν το demarcation της βλάβης ανά πλευρά και αναφέρονται όλοι οι αλγόριθμοι εξαγωγής των χαρακτηριστικών αυτών ανάλογα τη μορφή του profile.

Το *κεφάλαιο 5* περιλαμβάνει τον τρόπο επιλογής των 20 profile που αναφέρονται στη υποπεριοχή της εικόνας που έχει AW και όπου θα υπολογιστούν οι τελικές παράμετροι του demarcation. Αναφέρεται ο υπολογισμός των τελικών παραμέτρων του demarcation για κάθε εικόνα και ερμηνεύονται, αναλύονται τα αποτελέσματα για τρεις χαρακτηριστικές περιπτώσεις ασθενών. Υπολογίζονται οι παράμετροι για 100 ασθενείς, γίνεται στατιστική ανάλυση, πραγματοποιείται t-Test για τις low και high grade περιπτώσεις σύμφωνα με τις τιμές δύο παραμέτρων και προτείνεται ευριστικό σύστημα εκτίμησης του ρίσκου μετεξέλιξης νεοπλασμάτων σε καρκίνο.

Στο *Κεφάλαιο 6* προτείνονται τρόποι μελλοντικής έρευνας.

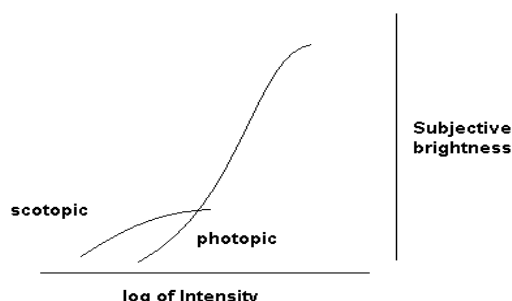
2.1 Εικόνες και Ανθρώπινη Αντίληψη τους

Ο όρος Monochrome image αναφέρεται σε μια 2-διάστατη συνάρτηση έντασης του φωτός $f(x,y)$, όπου x και y αποτελούν ζεύγος χωρικών συντεταγμένων και η τιμή της f σε κάθε σημείο (x,y) είναι ανάλογη του brightness (ή gray level) της εικόνας στο σημείο αυτό. Μια ψηφιακή εικόνα είναι μία εικόνα $f(x,y)$ που έχει διακριτές τιμές χωρικών συντεταγμένων (x,y) και διακριτές τιμές brightness. Μπορεί να θεωρηθεί ως ένας 2-διάστατος πίνακας του οποίου οι γραμμές και οι στήλες καθορίζουν ένα σημείο της εικόνας και η αντίστοιχη τιμή του στοιχείου αυτού στον πίνακα καθορίζει μια τιμή του brightness (ή gray level). Στις grayscale εικόνες τα χρώματα αγνοούνται επειδή το φως είναι μονοχρωματικό, το μόνο χαρακτηριστικό είναι η ένταση. Στις ψηφιακές εικόνες κάθε pixel έχει μόνο μία τιμή, η οποία περιγράφει την ένταση του γκρι. Η τιμή της έντασης στο 0 σημαίνει μαύρο χρώμα και στο 255 σημαίνει άσπρο, για 8-bit εικόνες.[1].

- *Brightness*: ένα από τα βασικά μεγέθη που χρησιμοποιούνται για να περιγράψουν την ποιότητα του χρωματικού φωτός. Είναι ένας υποκειμενικός δείκτης αντίληψης της εικόνας και εκφράζει την αντίληψη που εξάγεται από τη φωτεινότητα ενός αντικειμένου. Εμπεριέχει την έννοια της α-χρωματικής έντασης του φωτός και είναι ένα από τα σημεία-κλειδιά για περιγραφή της αντίληψης του χρώματος. Στις ψηφιακές μονοχρωματικές εικόνες (digital gray-level pictures) το brightness αναφέρεται στην α-χρωματική ένταση του φωτός και παίρνει διακριτές τιμές (0-255) [1,2].

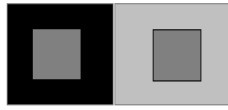
Επειδή κάθε ψηφιακή εικόνα αναφέρεται ως ένα σετ διαφορετικών τιμών brightness είναι σημαντική η ανάλυση της ικανότητας της ανθρώπινης όρασης να διαχωρίσει τις εναλλαγές του brightness. Πειραματικές μετρήσεις έχουν αποδείξει

ότι το υποκειμενικό brightness που αντιλαμβάνεται η ανθρώπινη όραση είναι λογαριθμική συνάρτηση της έντασης του φωτός που προσπίπτει στο ανθρώπινο μάτι. Το παρακάτω γράφημα αναπαριστά αυτή τη σχέση αυτή [1,2].



Το ανθρώπινο μάτι αντιλαμβάνεται τις διαφορές στις εικόνες μόνο όταν τα αντικείμενα της εικόνας διαφέρουν (διαχωρίζονται) από το background. Αν πρόκειται να αναλυθεί μια εικόνα και να βγουν συμπεράσματα από αυτήν, ο σχεδιαστής πρέπει να έχει υπόψιν του τις αρχές που διέπουν την αντίληψη της ανθρώπινης όρασης και η εξαγόμενη πληροφορία να εκφραστεί χρησιμοποιώντας μεταβλητές οι οποίες μπορούν να διακριθούν εύκολα. Αυτές οι μεταβλητές είναι ψυχο-φυσικοί παράμετροι (psycho-physical parameters) όπως το *contrast, object border, shape, texture, color* στις οποίες η ανθρώπινη αντίληψη είναι ευαίσθητη. Οι αναφερόμενες παράμετροι οδηγούν την ανθρώπινη αντίληψη όχι μόνο σε άντληση πληροφοριών αλλά και σε οφθαλμαπάτες [2].

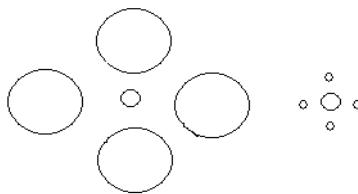
- *Contrast*: είναι η τοπική αλλαγή του brightness μιας εικόνας και ορίζεται ως η αναλογία (λόγος) της αλλαγής του brightness ενός αντικειμένου της εικόνας σε σχέση με το brightness του background. Στις υψηλές τιμές brightness απαιτείται μεγαλύτερο contrast για να γίνει αντιληπτό από την ανθρώπινη όραση. Το φαινομενικό brightness μιας περιοχής της εικόνας εξαρτάται από το brightness του γενικού background, γεγονός που προκαλεί συχνά οφθαλμαπάτες. Παράδειγμα σε αυτό αποτελεί η αντίληψη στις εικόνες παρακάτω, γνωστό ως conditional contrast:



Εικόνα 6: *Conditional contrast*[2]

Τα μικρά τετραγωνάκια έχουν το ίδιο brightness αλλά διαφορετικό background: ένα μαύρο και ένα γκρι. Το ανθρώπινο μάτι αντιλαμβάνεται το brightness των μικρών τετραγώνων ως διαφορετικό, κάτι το οποίο δεν ισχύει [1,2].

- *Acuity*: είναι η δυνατότητα ανίχνευσης λεπτομέρειας στην εικόνα και μειώνεται με την αύξηση της απόστασης από το οπτικό αντικείμενο [1,2].
- *Object Border*: η οριοθέτηση των αντικειμένων παρέχει πολλές πληροφορίες προσαρμογής του ανθρώπινου ματιού. Σαν παράδειγμα-ανάλυση μπορεί να θεωρηθεί το Ebbinghaus illusion



Εικόνα 7: *Ebbinghaus illusion*[2]

Όπου οι δύο κύκλοι με ίδια διάμετρο αλλά περιτριγυρισμένοι από κύκλους διαφορετικών διαμέτρων, μοιάζουν να έχουν διαφορετική διάμετρο [2]. Τα μεγέθη αυτά σχετίζονται με την αντίληψη του AW στις subtracted εικόνες. Το AW θεωρείται ως ένα αντικείμενο της εικόνας με ένταση $I + \Delta I$ το οποίο γίνεται αντιληπτό μέσω του contrast της εικόνας. Όσο μεγαλύτερης έντασης AW έχουν υποστεί τα κύτταρα τόσο πιο έντονο contrast. Το background και η οριοθέτηση του AW μπορεί να οδηγήσει το μάτι του ιατρού σε οφθαλμαπάτες. Η ποσοτικοποίηση οπτικών χαρακτηριστικών λύνουν το πρόβλημα της υποκειμενικής οπτικής αντίληψης.

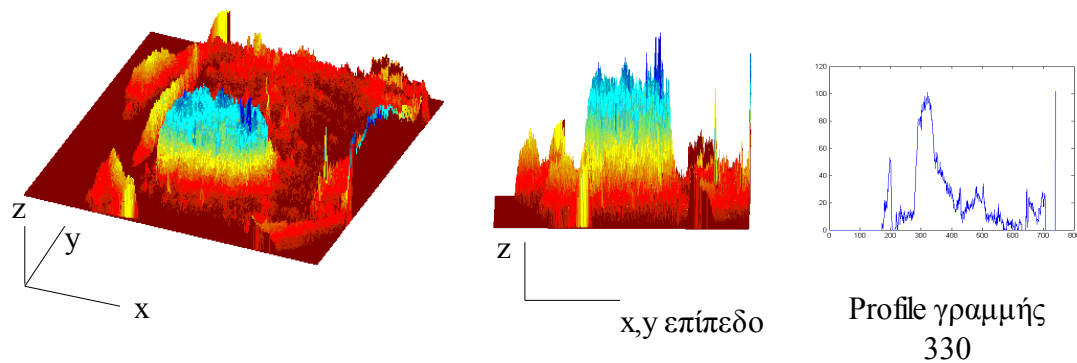
2.2 Spatial Profiles

2.2.1 Εισαγωγικά

Η μέθοδος που θα χρησιμοποιήσουμε στα πλαίσια της εργασίας με σκοπό την παραμετροποίηση του contrast και της ανθρώπινης αντίληψης των δεδομένων ιατρικών εικόνων είναι η μέθοδος χωρικής αναπαράστασης (spatial profiles) και γενικά εκφράζει την κατανομή διαφόρων μεγεθών στο χώρο. Έχει χρησιμοποιηθεί για μέτρηση της διαμέτρου στεφανιαίας αρτηρίας από αγγειόγραμμα [11]. Πιο συγκεκριμένα στα πλαίσια της εργασίας το μέγεθος, η χωρική 'κατανομή' του οποίου μας ενδιαφέρει είναι η ένταση της εικόνας. Η μέθοδος αυτή εκμεταλλεύεται τα χωρικά χαρακτηριστικά της εικόνας και την κατανομή της έντασης στις χωρικές συντεταγμένες x-y.

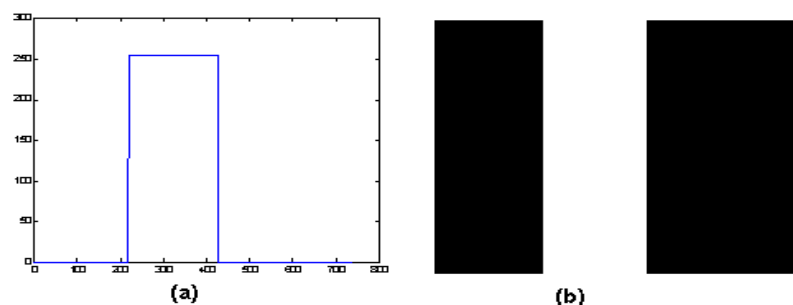
2.2.2 Επιλογή της Μεθόδου

Η μέθοδος αυτή επιλέγεται διότι αναζητείται η ποσοτικοποίηση του demarcation, χαρακτηριστικό το οποίο εξαρτάται από την χωρική κατανομή της έντασης (δηλ. του AW) στην subtracted εικόνα. Για το λόγο αυτό εκμεταλλευόμαστε τις διαφορετικές διακυμάνσεις τιμών έντασης της εικόνας στο χώρο. Θα προσπαθήσουμε να αποτυπώσουμε μαθηματικώς αυτές τις διακυμάνσεις τιμών των pixel εξάγοντας παραμέτρους και ποσοτικοποιώντας το demarcation. Κάθε spatial profile προκύπτει αν φανταστούμε την subtracted εικόνα στον τρισδιάστατο χώρο με τον z άξονα να περιέχει τις τιμές έντασης και x,y τα pixel, κοιτάζοντάς την εικόνα από πλάγια όψη ώστε τα x,y να ορίζουν ένα επίπεδο και στον z άξονα να απεικονίζονται οι εντάσεις. Εξάγοντας τα ψηφιακά αυτά σήματα ανά γραμμές και στήλες παίρνουμε τα spatial profile γραμμών και στηλών αντίστοιχα. Τα αναφερόμενα απεικονίζονται στις ακόλουθες εικόνες.



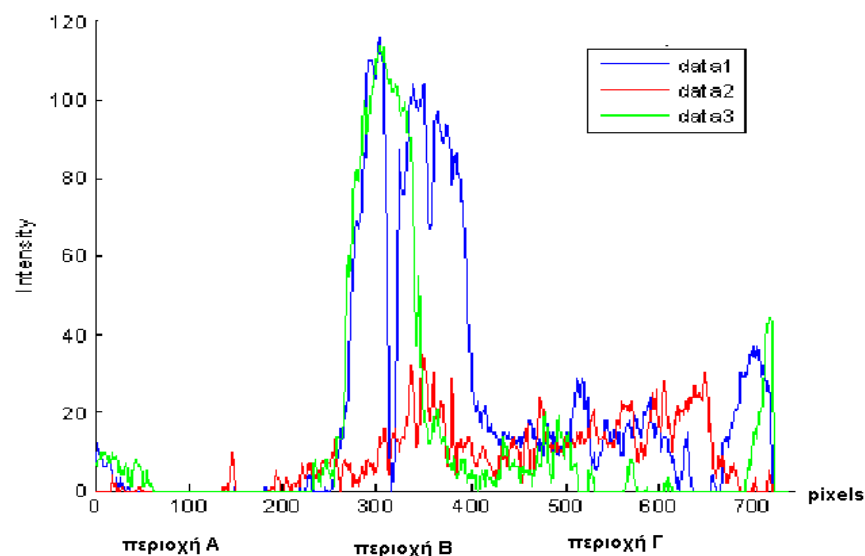
Τα spatial profiles των γραμμών/στηλών δεν είναι τίποτα άλλο παρά ψηφιακά σήματα που εκφράζουν για κάθε γραμμή/στήλη της εικόνας, την τιμή της έντασης του κάθε pixel στο 2D χώρο, όπου στον x-άξονα απεικονίζονται τα pixel της εικόνας και στον y-άξονα απεικονίζονται οι εντάσεις των αντίστοιχων pixel. Τα spatial profiles όλων των γραμμών ή στηλών δίνουν την συνολική 3D εικόνα.

Σαν παράδειγμα spatial profile αναφέρουμε ιδανικά το παρακάτω. Αν μπορούσε μια εικόνα να έχει σε όλες τις γραμμές της το προφίλ της εικόνας 8a, η εικόνα θα ήταν η 8b.



Εικόνα 8: (a) Spatial profile γραμμής (b) Εικόνα προς εξαγωγή spatial profile

Στα πλαίσια της εργασίας τα profiles που εξάγονται μπορούν να έχουν οποιαδήποτε μορφή. Στην εικόνα 9 παρουσιάζονται τρία τυχαία profiles γραμμών της subtracted εικόνας για την ασθενή 16_72 από το νοσοκομείο Αλεξάνδρα με σκοπό την ανάλυση τους.



Εικόνα 9: Spatial profiles 3 τυχαίων γραμμών της εικόνας

2.2.3 Ανάλυση Profile – Δεδομένα – Συμπεράσματα

Από το spatial profile κάθε γραμμής αντλούμε την πληροφορία της χωροταξικής τοποθέτησης της βλάβης στην εικόνα. Αναλύεται αρχικά ενδεικτικά το σήμα με το μπλε χρώμα (data 1) της εικόνας 9, μελετώντας το σε τρεις διαφορετικές περιοχές (εύρος pixel). Πιο συγκεκριμένα, παρατηρούμε ότι το σήμα αυτό από τα pixel 0 μέχρι 250 (περιοχή Α) έχει επί το πλείστον μηδενικές τιμές έντασης ενώ από τα pixel κοντά στο 250 μέχρι το pixel 400 (περιοχή Β) οι τιμές της έντασης ανεβαίνουν απότομα με μέγιστη τιμή το 118. Το ίδιο σήμα από το pixel 400 μέχρι το τέλος της εικόνας (περιοχή Γ) αρχίζει να 'σβήνει' έχοντας τιμές έντασης κοντά στο 20. Η ερμηνεία του σήματος που περιγράφηκε στην εικόνα είναι: για την συγκεκριμένη γραμμή της εικόνας στην περιοχή Α (pixels 1-250) η γραμμή της εικόνας έχει σκούρο μαύρο χρώμα. Στην περιοχή Β (pixels 250-400) η ίδια γραμμή αλλάζει απότομα χρώμα και γίνεται άσπρο ⁽¹⁾ και τέλος στην περιοχή Γ (pixels 400-740) η γραμμή γίνεται πολύ σκούρα γκρι με τιμές έντασης μικρότερες του 35. Τελικά από το spatial profile της μπλε γραμμής (data1) αντλούμε τις εξής πληροφορίες:

- i) Η παθογόνος περιοχή (περιοχή που έχει AW) εντοπίζεται στην περιοχή Β, γιατί οι

μεγάλες εντάσεις υποδηλώνουν AW και οι κλίσεις καθορίζουν τα όρια της βλάβης.

ii) Η αλλαγή της έντασης στην γραμμή αυτή της εικόνας από μαύρο σε άσπρο σαρώνοντας από αριστερά προς τα δεξιά γίνεται απότομα και αντιστρόφως.

iii) Στην περιοχή B η ένταση από πολύ μεγάλες τιμές πέφτει απότομα στο μηδέν στο pixel 320 περίπου, γεγονός που ερμηνεύεται ως ύπαρξη σε εκείνο το σημείο της εικόνας μιας μικρής μαύρης κουκκίδας που μπορεί να ερμηνευτεί σαν paper noise.

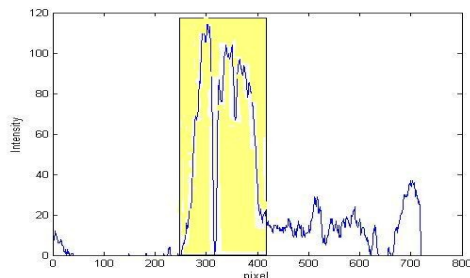
iv) Κάθε profile αποτελείται όχι μόνο από την χρήσιμη για εμάς χωρική πληροφορία 'κατανομής' της έντασης, αλλά και από θόρυβο εικόνας που απαντάται ως μικρές διακυμάνσεις τιμών έντασης. Επίσης θόρυβος που έχει προστεθεί από όργανα εξέτασης (κολποδιαστολέας) στην subtracted γίνεται φανερός στην περιοχή Γ στα pixel 700.

v) Στην περιοχή Γ παρατηρούνται μικρές τιμές έντασης με μέσο όρο το 20. Στην εικόνα τέτοιες τιμές έντασης δεν ερμηνεύονται ως AW και στον χάρτη του Dysis δεν χρωματίζονται. Γι'αυτό θεωρούμε ότι δεν δίνουν κάποια χρήσιμη πληροφορία.

vi) Σύμφωνα με την ενότητα 2.1 για το profile αυτό το contrast που θα γίνει αντιληπτό από την ανθρώπινη όραση είναι στην περιοχή B, χωρίς να αντιλαμβάνεται τις εντάσεις στην περιοχή Γ. Ειδικά στο συγκεκριμένο profile όπου το σήμα στην περιοχή B αυξάνεται απότομα και έχει αρκετά μεγάλη τιμή peak η οποία διατηρείται περίπου ίδια για την περιοχή αυτή. Η πληροφορία αυτή είναι σημαντική για την εφαρμογή fitting^(ενότητα 3.2.2).

⁽¹⁾ το 'καθαρό' άσπρο αντιστοιχεί σε τιμή έντασης 255 και η μέγιστη τιμή έντασης στο σήμα αυτό είναι το 115, άρα ο όρος 'άσπρη' για την συγκεκριμένη περίπτωση δεν είναι απόλυτα σωστός. Αλλά επειδή στις πραγματικές ιατρικές εικόνες οι παθογόνες περιοχές μετά την χρήση οξικού οξέος δεν ασπρίζουν τόσο ώστε να έχουμε τιμές έντασης 255, θα αναφέρουμε στο εξής ως άσπρες περιοχές τις περιοχές που ασπρίζουν μετά την χρήση οξικού οξέος και η έντασή τους είναι αρκετά μεγαλύτερη από του background.

Σχηματικά η ανθρώπινη όραση αντιλαμβάνεται το contrast της χρωματισμένης περιοχής και οι εντάσεις της περιοχής Γ δεν διακρίνονται εύκολα.



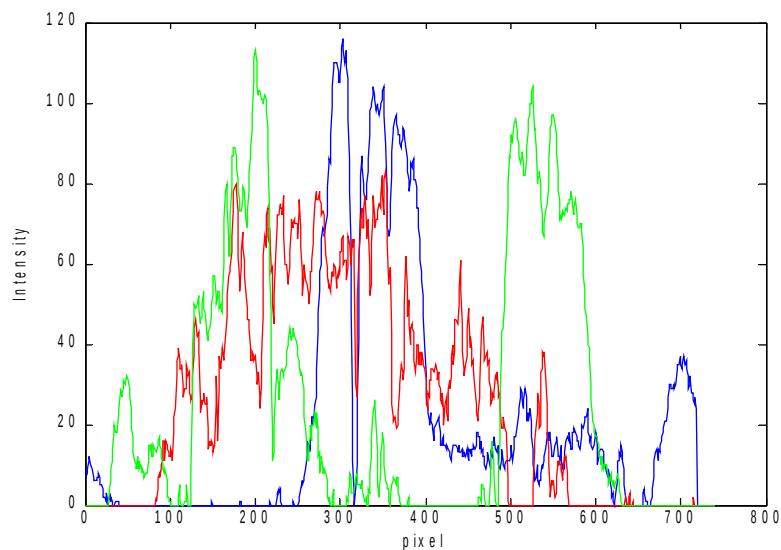
Εικόνα 10: Spatial profile-χρωματισμένη περιοχή που διακρίνει εύκολα το ανθρώπινο μάτι.

Επιπροσθέτως, ακολουθεί ανάλυση των τριών διαφορετικών spatial profiles (data1, data 2, data 3) της εικόνας 9 μεταξύ τους . Αντλούμε τις εξής πληροφορίες:

vii) Το profile με το μπλε χρώμα (data 1) λόγω των μεγάλων τιμών έντασης, φανερώνει ότι στη γραμμή αυτή υπάρχει AW ενώ το profile με το κόκκινο χρώμα (data 2) λόγω των μικρών (~20) εντάσεων του, φανερώνει ότι η γραμμή αυτή περισσότερο περιέχει θόρυβο και όχι χρήσιμη για εμάς πληροφορία.

viii) Συγκρίνοντας τα profiles με το μπλε και πράσινο χρώμα (data1 και data 3) παρατηρούμε ότι η βλάβη στα data 3 έχει μικρότερο εύρος από τα data1. Ο τρόπος με το οποίο τα σήματα αυτά παίρνουν τιμές είναι σχεδόν ίδιος, υπό την έννοια ότι οι αλλαγές από τις χαμηλές τιμές έντασης στις υψηλότερες γίνεται και στα δυο αυτά profiles απότομα. Το εύρος στο μισό του μεγίστου (FWHM) είναι μικρότερο για τα data3.

Στη συνέχεια παρουσιάζεται στην εικόνα 11 το profile με το μπλε χρώμα (data 1) συγκριτικά με άλλα δύο τυχαία profiles της ίδιας εικόνας με σκοπό την σύγκριση τους. Και τα τρία profiles έχουν AW.



Εικόνα 11: Spatial profile 3 τυχαίων γραμμών της εικόνας

Από τη σύγκριση των νέων profiles αντλούμε τις εξής πληροφορίες:

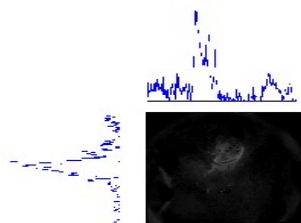
ix) Το profile με κόκκινο χρώμα παίρνει τις μεγάλες τιμές του σταδιακά σε αντίθεση με το profile με μπλε χρώμα, το οποίο παίρνει τις μεγάλες τιμές του απότομα. Η σταδιακή απόδοση τιμών στα profile ερμηνεύεται στην εικόνα ως γραμμή ή στήλη η οποία δεν έχει μεγάλο contrast και αντίστοιχα στη βλάβη ως άτονο demarcation. Η παρατήρηση αυτή αποτελεί σημαντικό κλειδί της εργασίας αυτής όπου και θα αναζητηθούν χαρακτηριστικά που το εκφράζουν.

χ) Τέλος το profile με πράσινο χρώμα παίρνει τις μέγιστες τιμές του σε δύο περιοχές άρα στην εικόνα υπάρχουν δύο περιοχές σε αρκετά μεγάλη απόσταση μεταξύ τους οι οποίες έχουν AW.

Όλες οι παραπάνω πληροφορίες θα χρησιμοποιηθούν με σκοπό την εξαγωγή χαρακτηριστικών. Για κάθε αναφερόμενη παρατήρηση προτείνονται οι λύσεις που θα υλοποιηθούν στη συνέχεια: i) διατήρηση των χαρακτηριστικών του AW ii) διατήρηση των εναλλαγών έντασης iii) τρόποι αποφυγής λαθών σε περιπτώσεις θορύβου iv) εξομάλυνση του σήματος, απαλοιφή όλων των ειδών θορύβων v) επεξεργασία μόνο χρήσιμης πληροφορίας vi) επεξεργασία profile επιλέγοντας μοντέλα που διατηρούν την πληροφορία που επιθυμούμε vii,ix) εξαγωγή

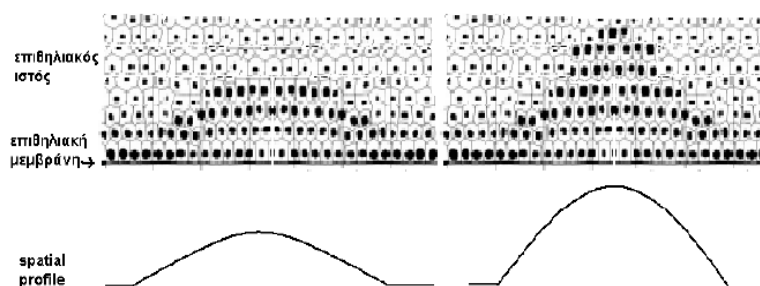
χαρακτηριστικών που εκφράζει την παρατήρηση αυτή x) ξεχωριστή επεξεργασία σε τέτοια περίπτωση βλάβης.

Για κάθε subtracted εικόνα εξάγονται profile γραμμών και στηλών ώστε να προσεγγιστεί η βλάβη από όλες τις πλευρές της. Σχηματικά:



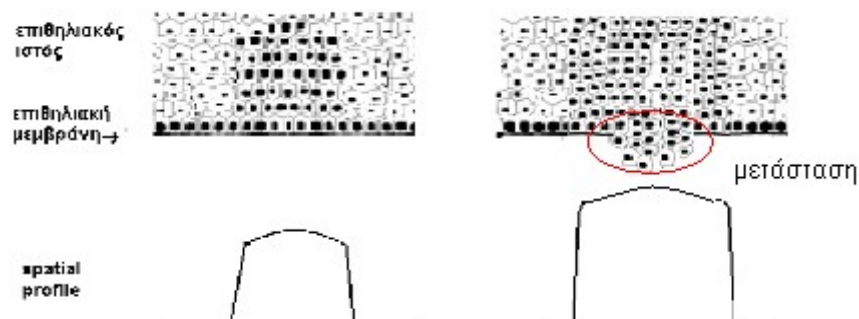
- Ερμηνεία των profile στον επιθηλιακό ιστό

Τα παθογόνα κύτταρα του επιθηλιακού ιστού ανάλογα με την ποσότητα και την τοπολογία τους στο επιθήλιο δίνουν ένα τρόπο και ένα ποσοστό ασπρίσματος (AW) μετά τον ψεκασμό του οξικού οξέος. Στη συνέχεια αναφέρεται η σχέση της μορφής των spatial profile με την τοπολογία των παθογόνων κυττάρων. Υποθέτοντας ότι τα παθογόνα κύτταρα απεικονίζονται με έντονο μαύρου χρώματος πυρήνα και ότι είναι δυνατή η απεικόνιση της πλάγιας όψης του επιθηλίου, παρακάτω παρουσιάζονται υποθετικά profile και ερμηνεύονται με βάση την κατάσταση του επιθηλίου.

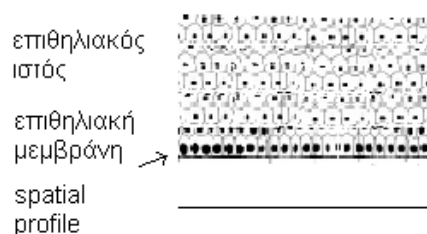


Τα παθογόνα κύτταρα στην παραπάνω εικόνα τείνουν να εξαπλωθούν στα πλάγια του επιθηλίου και η αντίδραση τους με το οξικό οξύ δίνει αχνό AW οπότε τα αντίστοιχα profile παρουσιάζουν μικρή κλίση ενώ η μέγιστη τιμή των profile εξαρτάται από το πλήθος των παθογόνων κυττάρων κατά βάθος του ιστού (ένταση

AW). Στην περίπτωση που τα παθογόνα κύτταρα εξαπλώνονται κατά βάθος του ιστού τα αντίστοιχα profile παρουσιάζονται παρακάτω με φανερά πιο μεγάλη κλίση και μέγιστη τιμή που εξαρτάται από το πλήθος των παθογόνων κυττάρων αντίστοιχα.



Τα πιο επικίνδυνα profile είναι αυτά με μεγάλη κλίση και peak τιμή διότι υποδηλώνουν πιθανή μετάσταση του καρκίνου. Ο φυσιολογικός μη-παθογόνος ιστός ιδανικά θα είχε το profile:



Τα profile αυτά είναι υποθετικά. Ο καθρεφτισμός των profile στον άξονα x αντιπροσωπεύει την τοπολογία των παθογόνων κυττάρων μέσα στο επιθήλιο. Οποιαδήποτε μορφή-τοπολογία βλάβης γίνεται αντιληπτή μέσω του AW και του demarcation λόγω του ποσού και της 'θέσης' των παθογόνων κυττάρων που αντιδρούν με το οξικό οξύ. Από τα αναφερόμενα το demarcation μπορεί να λειτουργήσει με την έννοια της πρόβλεψης της κατεύθυνσης των παθογόνων κυττάρων στο επιθήλιο. Η απόδειξη του φαινομένου αυτού απαιτεί ιατρικές και χρονοβόρες έρευνες που δεν είναι στα πλαίσια του μηχανικού.

2.3 Θόρυβος στην Εικόνα (Image Noise)

Οι πραγματικές εικόνες αλλοιώνονται συχνά από τυχαίο θόρυβο. Ο θόρυβος αυτός μπορεί να απαντάται στην εικόνα λόγω λήψης, μεταφοράς ή επεξεργασίας και μπορεί να είναι ανεξάρτητος είτε εξαρτημένος από τα περιεχόμενα της εικόνας.

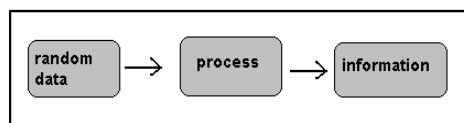
Πιο γνωστή μορφή θορύβου είναι ο Gaussian noise με κάποια μέση τιμή μ και κάποια διασπορά σ . Κατά την μεταφορά εικόνων προστίθεται (από το κανάλι) additive noise ανεξάρτητος από τα περιεχόμενα της εικόνας με την μορφή $f(x, y) = image(x, y) + n(x, y)$ όπου $n(x, y)$ θόρυβος [2]. Κατά την ψηφιοποίηση των εικόνων προστίθεται quantization noise που προκαλείται λόγω ελλειπών κατωφλίων με αποτέλεσμα αποκλεισμό πολλών τιμών σήματος. Impulsive noise απαντάται στην εικόνα όταν οι τιμές του brightness τυχαίων pixel έχουν εξαιρετικά διαφορετικές τιμές από τα γειτονικά (μαύρες και άσπρες κουκκίδες). Γνωστός και ως salt and paper θόρυβος συναντάται στις ψηφιακές εικόνες από λάθη κατά τη ψηφιοποίηση ή μεταφορά [1,2]. Αν δεν υπάρχει εκ των προτέρων γνώση των ιδιοτήτων του θορύβου, στη μηχανική όραση η απαλοιφή του περιορίζεται σε τοπική επεξεργασία ενώ σε αντίθετη περίπτωση χρησιμοποιούνται τεχνικές αποκατάστασης της εικόνας [2].

Οι subtracted εικόνες της εργασίας περιέχουν εκτός του συνηθισμένου θορύβου ψηφιακής εικόνας, θόρυβο από ανακλάσεις οργάνων γυναικολογικής εξέτασης είτε τριχών, από την αφαίρεση των εικόνων στην περίπτωση που οι αλγόριθμοι του registration εικόνων του DySIS προσπίπτουν σε μικρά σφάλματα.

3.1 Τυχαία Δεδομένα

Η διαδικασία της προ-επεξεργασίας στην *'επεξεργασία εικόνας'* περιλαμβάνει κάποιες τεχνικές και αλγορίθμους που έχουν σκοπό i) την απαλοιφή του θορύβου καθώς και ii) επαύξηση του contrast στην εικόνα iii) την απομόνωση κάποιας περιοχής που δηλώνει ότι υπάρχει πιθανότητα να δώσει αλφαριθμητικές πληροφορίες [1]. Όμως στην δική μας περίπτωση δεν θέλουμε να επέμβουμε καθόλου στην subtracted εικόνα γιατί αποτελεί πειστήριο ιατρικής εξέτασης. Για το λόγο αυτό επεξεργαζόμαστε τα spatial profile γραμμών και στηλών. Τα δεδομένα αυτά είναι τυχαία (random) και για να επεξεργαστούν θα χρειαστούν κατάλληλη προ-επεξεργασία.

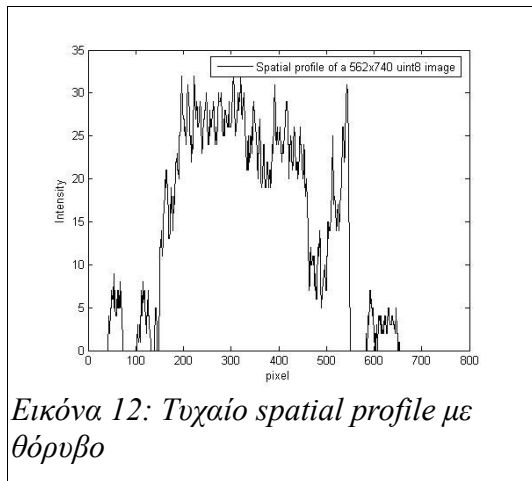
Η προ-επεξεργασία τυχαίων δεδομένων (data pre-processing) είναι μια υπολογιστική διαδικασία που μετατρέπει τα δεδομένα σε πληροφορία και γνώση. Η υπολογιστική διαδικασία θεωρείται ως μία μονάδα που έχει ως είσοδο τα τυχαία δεδομένα και ως έξοδο ερμηνεύσιμα δεδομένα που παρέχουν πληροφορία. Τα τυχαία δεδομένα ακολουθούν το διάγραμμα ροής [10,2]:



Υπάρχουν πολλές τεχνικές, αλγόριθμοι και μαθηματικά μοντέλα για το σκοπό αυτό. Στα πλαίσια της εργασίας αναλύονται και εφαρμόζονται κάποιες από αυτές.

3.2 Τεχνικές προ-επεξεργασίας Δεδομένων

Τα spatial profiles που παίρνουμε από την εικόνα έχουν την μορφή του δεξιά γραφήματος. Είναι ψηφιακά σήματα με έντονες διακυμάνσεις τιμών έντασης που ερμηνεύονται σαν θόρυβος. Η προ-επεξεργασία των profile στην εργασία αυτή έχει σκοπό την απαλοιφή του θορύβου, την ομαλοποίηση του σήματος χωρίς



Εικόνα 12: Τυχαίο spatial profile με θόρυβο

να αλλοιώνονται τα χαρακτηριστικά που εκφράζουν το demarcation και την επιλογή ενός μοντέλου περιγραφής ώστε να διευκολυνθεί η επεξεργασία.

3.2.1 Smoothing

Αν τα δεδομένα προς επεξεργασία έχουν θόρυβο τότε πρέπει να εφαρμοστεί ένας smoothing αλγόριθμος για να φανερωθούν τα χαρακτηριστικά του ψηφιακού σήματος, απαλλαγμένο από θόρυβο και να παρέχουν μία λογική αρχική προσέγγιση για επιλογή parametric fitting (ενότητα 3.2.2). Η διαδικασία του smoothing σκοπό έχει να δώσει μία προσέγγιση της μέσης κατανομής του κάθε σημείου των αρχικών δεδομένων. Η εκτίμηση αυτή βασίζεται σε καθορισμένο αριθμό γειτονικών δεδομένων. Οι smoothing αλγόριθμοι χρησιμοποιούνται κυρίως στην στατιστική, στην επεξεργασία σήματος και εικόνας.

Υπάρχουν δύο συνηθισμένοι smoothing μέθοδοι:

i) *local regression*

Οι μέθοδοι αυτοί χρησιμοποιούν *parametric linear least squares fitting* (ενότητα 3.2.2) τοπικά ανά διαστήματα με πολυωνμική fitting συνάρτηση είτε πρώτου βαθμού (*lowess*) είτε δευτέρου (*loess*). Διαφέρει από το fitting στο γεγονός ότι εφαρμόζεται σε διάστημα δεδομένων και όχι στο σύνολο δεδομένων.

- ii) *filtering*, όπου η έξοδος της μεθόδου προκύπτει άμεσα από τις τιμές των δεδομένων.

Σε κάθε μέθοδο *smooth* απαιτείται ένα διάστημα (*span*) που καθορίζει ένα παράθυρο γειτονικών σημείων που θα περιλαμβάνονται στον *smoothing* υπολογισμό για κάθε σημείο. Το παράθυρο αυτό μετακινείται δια μέσω των δεδομένων όσο υπολογίζονται οι νέες τιμές του κάθε σημείου [13]. Όσο μεγαλύτερο το *span* τόσο μεγαλύτερο το *smoothness* αλλά μειώνεται η αναλυτικότητα των δεδομένων. Το μέγεθος του *span* εξαρτάται από τα δεδομένα και την *smoothing* μέθοδο που εφαρμόζεται (χρειάζεται πειραματισμός για να βρεθεί). *Smoothing* δεδομένων μπορεί να γίνει και με μεθόδους *fitting* κυρίως με *nonparametric fitting* ^(ενότητα 3.2.2).

3.2.1.1 Moving Average Algorithm

Είναι η πιο γνωστή *filtering* μεθόδους. Ο αλγόριθμος κάνει *smooth* τα δεδομένα αντικαθιστώντας κάθε σημείο με το μέσο όρο των γειτονικών σημείων που καθορίζονται στο *span*. Η απόκριση του αλγορίθμου δίνεται από την εξίσωση

$$y_s(i) = \frac{1}{2N+1} (y(i+N) + y(i+N-1) + \dots + y(i-N+1) + y(i-N))$$

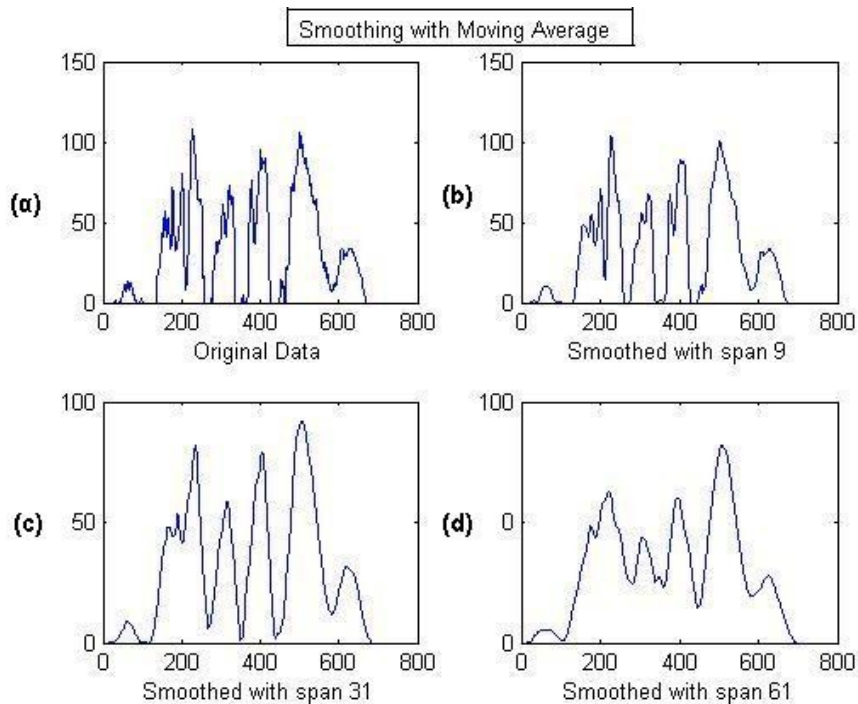
όπου $y_s(i)$ είναι η *smoothed* τιμή για το i -οστό σημείο, N είναι ο αριθμός των γειτονικών σημείων από την κάθε πλευρά του $y_s(i)$ και $2N+1$ είναι το *span*.

Ακολουθούνται οι κανόνες:

- Το μέγεθος του *span* (παράθυρο) πρέπει να είναι περιττό.
- Το σημείο που γίνεται *smoothed* πρέπει να είναι στο κέντρο του *span*.
- Τα σημεία στα άκρα του σήματος δεν λαμβάνουν μέρος στο *smooth* γιατί το *span* εκεί δεν ορίζεται.

Τα *spatial profile* (γραμμών και στηλών) προ-επεξεργάζονται ξεχωριστά, σαν τυχαία δεδομένα. Δοκιμάστηκαν πολλές μέθοδοι *smoothing* των *profile*. Αναφέρονται οι αποτελεσματικότεροι όπως η *Moving average*. Κάνοντας μια αρχική προσέγγιση της συμπεριφοράς του *spatial profile* για διάφορα *span*, εξάγουμε τις

ακόλουθες εικόνες.



Εικόνα 13: Σύγκριση μεγέθους span. (α) Αρχικά δεδομένα (β) Smoothed δεδομένα με μέγεθος span 9 (γ) Με span 31 (δ) Με span 61

Τα smoothed δεδομένα απομακρύνονται από τα αρχικά ανάλογα με το span και τις διακυμάνσεις των κοντινών τιμών σε κάθε span. Η μέθοδος με οποιαδήποτε τιμή span δίνει μεγάλο αριθμό τοπικών μεγίστων και ελαχίστων. Δεν ευνοείται η περαιτέρω επεξεργασία (εξαγωγή χαρ/κών), κατανόηση και μοντελοποίηση του smoothed profile. Δοκιμάζονται επιπλέον μέθοδοι smoothing με σκοπό την καλύτερη ομαλοποίηση του profile.

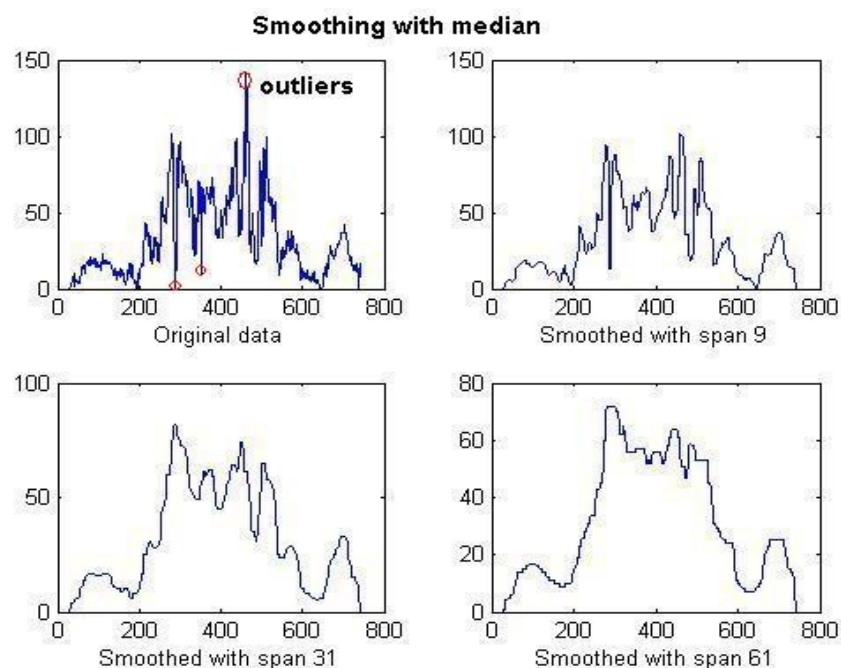
3.2.1.2 Median Algorithm

Είναι μια από τις filtering μεθόδους, ιδιαίτερα αποτελεσματική για θόρυβο ο οποίος αποτελείται από οξείες και απότομες αλλαγές μικρού εύρους. Είναι πολύ αποτελεσματικό στην απαλοιφή των outliers⁽¹⁾ διατηρώντας την ευκρίνεια του σήματος. Χρησιμοποιεί span, περιττού μεγέθους, το οποίο κινείται δια μέσω του

σήματος ακολουθώντας τα βήματα του αλγορίθμου για κάθε σημείο i :

1. Εύρεση N σημείων δεξιά του i και N σημείων αριστερά.
(Το μέγεθος του span είναι $2N+1$)
2. Ταξινόμηση των $2N+1$ στοιχείων κατά αύξουσα σειρά.
3. $i = m$ όπου ($m > N_{\text{αριστερά}}$) & ($m < N_{\text{δεξιά}}$)

Αν πρόκειται το median filter να χρησιμοποιηθεί σε εικόνα(μη γραμμικό φίλτρο), το span είναι πίνακας που καθορίζει τον αριθμό των γειτονικών pixel που θα τρέχει ο αλγόριθμος. Εφαρμόζοντας το median filter στο spatial profile με διάφορες τιμές span εξάγονται οι εικόνες.



Εικόνα 14: Σύγκριση μεγέθους span. (α) Αρχικά δεδομένα (β) Smoothed δεδομένα με μέγεθος span 9 (γ) Με span 31 (δ) Με span 61

⁽¹⁾: δεδομένα τα οποία είναι αριθμητικά απομακρυσμένα από τα υπόλοιπα δεδομένα

Τα outliers απαλείφονται ενώ το profile κρατάει τα 'εξωτερικά' χαρακτηριστικά που αφορούν την εργασία και γίνεται πιο ερμηνεύσιμο. Από την μελέτη πολλών profile που έχουν γίνει smoothed προκύπτει μια αρχική προσέγγιση των χαρακτηριστικών τους που μοιάζουν να προσεγγίζουν την gaussian mix συνάρτηση. Ο median αλγόριθμος δεν διευκολύνει την επεξεργασία οπότε η διαδικασία ομαλοποίησης των profile συνεχίζεται με μεθόδους fitting.

3.2.2 Curve Fitting (Fitting)

Μια από τις μεθόδους που χρησιμοποιούνται για ανάλυση δεδομένων αναφέρεται στην χάραξη μιας καμπύλης ανάμεσα από τα δεδομένα προς επεξεργασία. Η καμπύλη αυτή προκύπτει από *regression* τεχνικές, υπολογισμό ενός *spline* ή από *interpolation*. Σκοπός της τεχνικής αυτής είναι να παρέχει μια βαθύτερη κατανόηση των δεδομένων, καθιστώντας δυνατή:

- i) την αποδοχή ή απόρριψη ενός θεωρητικού μοντέλου
- ii) την ερμηνεία της φυσικής σημασίας των συντελεστών του fitting
- iii) την βελτίωση των τεχνικών συλλογής δεδομένων σε μελλοντικά πειράματα

Τα δεδομένα προς επεξεργασία μπορεί να είναι είτε από στατιστική έρευνα, από πειράματα, από αισθητήρες, από σήματα και ούτο καθεξής. [9,13,3]

3.2.2.1 Regression Analysis

Η Regression analysis είναι μία τεχνική επεξεργασίας και ανάλυσης αριθμητικών δεδομένων με σκοπό την εύρεση της σχέσης μεταξύ μεταβλητών. Συνήθως η εύρεση αυτή αναζητά την αιτιολογική επίδραση μιας μεταβλητής σε μια άλλη. Έτσι, έχουμε τιμές μιας εξαρτημένης μεταβλητής και ενός ή πολλών ανεξάρτητων. Στην περίπτωση που υπάρχει μία ανεξάρτητη μεταβλητή αναφερόμαστε στην *Simple Regression* ενώ όταν υπάρχουν περισσότερες του ενός ανεξάρτητες μεταβλητές στην *Multiple Regression*. Σαν παράδειγμα αναφέρεται η σχέση του εισοδήματος ενός ελεύθερου επαγγελματία με τους εξής παράγοντες: την μόρφωση, την εμπειρία, την ηλικία, το φύλο, τα κίνητρα. *Multiple Regression* θεωρείται η σχέση του εισοδήματος με όλους ή κάποιους από τους παραπάνω παράγοντες, ενώ *Simple Regression* η σχέση του εισοδήματος μόνο με την μόρφωση, αγνοώντας τους άλλους παράγοντες [8].

Η σχέση εξαρτημένη-ανεξάρτητης(-των) μεταβλητής(-ών) περιγράφεται από την regression εξίσωση (*regression equation*) η οποία μοντελοποιείται σαν συνάρτηση της/των ανεξάρτητης/-των μεταβλητής/-των, των αντίστοιχων παραμέτρων και ενός παράγοντα σφάλματος.

$$y_i = \theta(x_i) + \varepsilon_i$$

όπου x_i είναι η/οι ανεξάρτητη/-τες μεταβλητή/-ές, ε_i ο παράγοντας σφάλματος και θ η μορφή της συνάρτησης συμπεριλαμβανομένων και των παραμέτρων (συντελεστών) [8,3,9].

Ο παράγοντας σφάλματος ε_i θεωρείται τυχαία μεταβλητή που ακολουθεί κανονική κατανομή με μέση τιμή 0, $\varepsilon_i \sim N(0, \sigma)$. Οι παράμετροι εκτιμούνται με την εφαρμογή αλγορίθμων ώστε να δώσουν την συνάρτηση η οποία θα προσεγγίζει καλύτερα τα δεδομένα. Επιπροσθέτως, εκτός από την *regression analysis* (*parametric regression analysis*) υπάρχει και η *nonparametric regression analysis*, η οποία είναι μια μορφή *regression analysis* με τη διαφορά ότι το μοντέλο δεν θεωρείται δεδομένο αλλά εκτιμάται -κατασκευάζεται- από τα δεδομένα. Χρειάζεται μεγάλος όγκος δεδομένων γιατί εκτιμάται τόσο το μοντέλο όσο και οι παράμετροί του. Είναι πολύπλοκη μέθοδος και χρησιμοποιείται σε περίπτωση που δεν μας ενδιαφέρει η σχέση μεταξύ των δεδομένων αλλά μόνο η χάραξη της καμπύλης που περιγράφει τα δεδομένα [9,14].

3.2.2.2 Parametric Fitting

Το parametric fitting είναι μια μορφή curve fitting, χρησιμοποιώντας regression, η οποία αναφέρεται στην εύρεση των βέλτιστων τιμών των παραμέτρων ενός μοντέλου το οποίο έχει επιλεγθεί να κάνει fit τα δεδομένα. Το μοντέλο του οποίου οι βέλτιστοι συντελεστές αναζητούνται, επιλέγεται από τον σχεδιαστή. Κάθε δεδομένο y_i αποτελείται από έναν ντετερμινιστικό παράγοντα, ο οποίος δίνεται από τη συνάρτηση fitting f και από έναν τυχαίο παράγοντα, το σφάλμα ε_i .

$$y_i = f(x_i) + \varepsilon_i$$

$f(x_i)$: fitting συνάρτηση μεταβλητής/-ών x_i συμπεριλαμβανομένων και των συντελεστών

ε_i : παράγοντας σφάλματος [13,9].

- Αλγόριθμος

Η βασική ιδέα του αλγορίθμου είναι η αναζήτηση των συντελεστών της συνάρτησης $f(x_i)$ οι οποίοι ελαχιστοποιούν τον παράγοντα σφάλματος ε_i , ώστε τα fitted δεδομένα να υπολογίζονται από τον ντετερμινιστικό όρο $f(x_i)$. Το $f(x_i)$ εμπεριέχει την έννοια της πρόβλεψης (prediction), ενώ το ε_i είναι παράγοντας που εμπεριέχει την έννοια της αβεβαιότητας, θεωρείται τυχαία μεταβλητή και ακολουθεί κανονική (gaussian) κατανομή ($\varepsilon_i(T.M) \sim N(0, \sigma)$). Η έξοδος του αλγορίθμου δίνεται από την εξαρτημένη μεταβλητή (predictor δεδομένα) της συνάρτησης με τους βελτιστοποιημένους συντελεστές ικανοποιώντας το κριτήριο του ελάχιστου σφάλματος.

- Least square error

Ο fitting αλγόριθμος απαιτεί ένα μοντέλο που να σχετίζει τα predicted δεδομένα με τα πραγματικά ώστε να προκύπτει ο παράγοντας σφάλματος. Το μοντέλο αυτό ορίζεται ως η διαφορά r_i των πραγματικών δεδομένων από τα predicted για κάθε δεδομένο i .

$$r_i = y_i - \hat{y}_i \quad \text{ή αλλιώς} \quad \text{error} = \text{data} - \text{fitted data}.$$

(Τα fitted data προέρχονται από το ντετερμινιστικό μοντέλο $f(x_i)$)

Παίρνοντας τις διαφορές για όλα τα δεδομένα που γίνονται fit έχουμε το συνολικό σφάλμα

$$\sum_{i=1}^N r_i = \sum_{i=1}^N (y_i - \hat{y}_i)$$

όπου N : το σύνολο των δεδομένων.

Έχει επικρατήσει να παίρνουμε το άθροισμα των τετραγώνων των διαφορών r_i (square error)

$$\sum_{i=1}^N r_i^2 = \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

διότι η ανάλυση του κατά τον υπολογισμό των παραμέτρων οδηγεί σε ευκολότερους και γρηγορότερους υπολογισμούς, χωρίς να αποκλείουμε τον υπολογισμό του

συνολικού σφάλματος ως άθροισμα π.χ απολύτων τιμών

$$\sum_{i=1}^N |r_i| = \sum_{i=1}^N |y_i - \hat{y}_i|$$

[3,14,9,10] Το συνολικό square error σχετίζει τα predicted δεδομένα με τα πραγματικά και στόχος είναι η ελαχιστοποίηση του ώστε ιδανικά να συμπίπτουν. Για το λόγο αυτό το μοντέλο του σφάλματος στην fitting εξίσωση ονομάζεται ελάχιστο τετραγωνικό σφάλμα (least square error).

- Linear – Nonlinear models (equations)

Στο parametric fitting οι fitting συναρτήσεις (fitting equation) που χρησιμοποιούνται μπορεί να είναι είτε γραμμικές (*linear*) είτε μη γραμμικές (*nonlinear*).

Ως *linear* συναρτήσεις χαρακτηρίζονται οι fitting συναρτήσεις που είναι γραμμικός συνδυασμός των παραμέτρων του μοντέλου και μπορεί να είναι είτε simple είτε multiple^(ενότητα 3.2.2.1). Συναρτήσεις τέτοιες είναι οι γραμμικές, οι πολυωνυμικές [14]:

$$y_i = b_0 + b_1 x_i + \varepsilon_i, i = 1, N, \quad y_i = b_0 + b_1 x_i + b_2 x_i^2 + \varepsilon_i, i = 1, N$$

$$y_i = b_0 z_i + b_1 x_i + \varepsilon_i, i = 1, N$$

(Στο εξής θα αναφερόμαστε μόνο στις simple fitting εξισώσεις διότι αυτές χρησιμοποιούνται στην εργασία)

Οι linear συναρτήσεις χρησιμοποιούν συνήθως το linear least square error διότι ελαχιστοποιώντας το (μηδενισμός παραγώγου) καταλήγουν σε απλές μαθηματικές σχέσεις, γεγονός που τις καθιστά ως τις πιο απλές και γρήγορες fitting συναρτήσεις. Οι βέλτιστοι παράμετροι είναι συνάρτηση των x_i και y_i δεδομένων και προκύπτουν άμεσα από μαθηματικές πράξεις υπολογίζοντας παραγώγους και λύνοντας το σύστημα με απλές τεχνικές πινάκων [14,10].

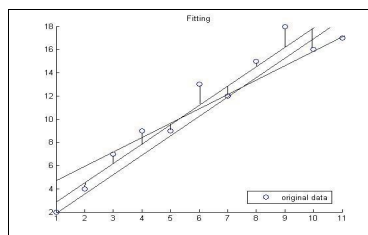
Ως *nonlinear* συναρτήσεις χαρακτηρίζονται οι fitting συναρτήσεις που είναι μη γραμμικός συνδυασμός των παραμέτρων του μοντέλου και μπορεί να είναι είτε simple είτε multiple. Συναρτήσεις τέτοιες είναι Gaussian, mix Gaussian, λόγος

πολυωνύμων, εκθετικές συναρτήσεις. Οι nonlinear συναρτήσεις χρησιμοποιούν το nonlinear least square error όπου οι συντελεστές δεν μπορούν να υπολογιστούν με επίλυση μαθηματικού συστήματος. Αντ'αυτού χρησιμοποιούνται επαναληπτικές τεχνικές ακολουθώντας τα βήματα:

1. Αρχικοποίηση των παραμέτρων της fitted συνάρτησης κάνοντας μια αρχική προσέγγιση.
2. Εύρεση predicted τιμής με τις αρχικοποιημένες παραμέτρους.
3. Ρύθμιση-αλλαγή των παραμέτρων και προσδιορισμός της βελτίωσης του fit. (Χρησιμοποιούνται αλγόριθμοι όπως *trust-region* για τον τρόπο αλλαγής των παραμέτρων)
4. Επανάληψη από το βήμα 2 μέχρι να προσεγγιστεί το κριτήριο σύγκλισης. (nonlinear least square error) [14].

- Εύρεση Best Fit model και κριτήρια Best Fit

Πρωταρχικό βήμα στο curve fitting αποτελεί η επιλογή του μοντέλου που θα κάνει fit τα δεδομένα. Σε περίπτωση που απαιτείται ερμηνεία είτε επεξεργασία των παραμέτρων της fit συνάρτησης συνήθως επιλέγεται το parametric fitting ενώ όταν απαιτείται μόνο smoothing των δεδομένων επιλέγεται το nonparametric fitting. Στο parametric fitting η επιλογή της fit συνάρτησης γίνεται κυρίως εμπειρικά από οπτική παρατήρηση των δεδομένων και της σχέσης μεταξύ τους. Έχοντας τα ακόλουθα δεδομένα (Εικόνα 15) προκύπτει από οπτική εξέταση του παρατηρητή ότι η σχέση των x, y δεδομένων είναι γραμμική δηλ. $y = ax + b$. Η επιλογή αυτή του μοντέλου fit (δηλ.: simple parametric fitting, linear equation first degree) προκύπτει από τον σχεδιαστή, γεγονός που προσθέτει ένα παράγοντα λάθους στο σύστημα έστω ε_π . Έστω ότι $\varepsilon_\pi = 0$, οι συντελεστές της ευθείας $y = ax + b$ προκύπτουν από τις μεθόδους fitting που περιγράφηκαν.



Εικόνα 15: Παράδειγμα fitting

Υπάρχουν κριτήρια που πιστοποιούν τα αποτελέσματα του fitting γραφικά και μαθηματικά. Γενικώς η οπτική παρατήρηση του σχεδιαστή μπορεί να είναι πιο ευεργετική από τα γραφικά και αριθμητικά κριτήρια διότι έχει συνολική όψη όλων των δεδομένων και είναι δυνατή η άμεση σύγκριση όλων των δεδομένων και του fit μοντέλου [8,14]. Η μαθηματική προσέγγιση περιορίζεται στην αποτύπωση όλης της 'εικόνας' σε έναν αριθμό και μπορεί να αποβεί παραπλανητική. Γραφικά ένα από τα κριτήρια είναι η μελέτη της χωρικής τοποθέτησης των αρχικών δεδομένων, τα οποία πρέπει να ισοκατανέμονται από αριστερά και δεξιά της fitting συνάρτησης [8]. Άλλο κριτήριο είναι οι διαφορές των αρχικών δεδομένων από τα fitted. Αν οι διαφορές αυτές κατανέμονται τυχαία το fitting μοντέλο λειτουργήσει αποτελεσματικά ενώ αν κατανέμονται με κάποιο σταθερό τρόπο τον μοντέλο ήταν ανεπαρκές. Μαθηματικά κριτήρια αναφέρουμε ενδεικτικά το SSE για το parametric fitting που ορίζεται ως το άθροισμα των τετραγώνων των διαφορών και τιμές του κοντά στο 0 αποδεικνύουν αριθμητικά αποτελεσματικό fitting [14].

3.2.2.3 Nonparametric Fitting

Το Nonparametric fitting χρησιμοποιείται κυρίως για smoothing δεδομένων. Διαφέρει από το parametric γιατί δεν χρησιμοποιεί καθορισμένα μοντέλα (συναρτήσεις fitting) εκπαιδευόμενα από τα αρχικά δεδομένα. Στην κατηγορία περιλαμβάνονται τα *splines*, μαθηματικά μοντέλα με πολυπλοκότητα $O(n^3)$ n:αριθμός δειγμάτων [9]. Χρησιμοποιείται μία smoothing παράμετρος *sp* (*smoothing parameter*) που καθορίζει το smoothness των δεδομένων. Αν η τιμή του *sp* είναι ίση με 1 τα fitted δεδομένα προκύπτουν από την χάραξη μιας συνεχής ευθείας γραμμής που

διέρχεται από όλα τα δεδομένα (simple interpolation) και όσο μειώνεται το sp πλησιάζοντας το 0 τόσο πιο smoothed γίνεται το σήμα [14].

- Interpolation

Αναφέρεται σε μια τεχνική η οποία σκοπεύει στην χάραξη μιας καμπύλης η οποία περνά ακριβώς πάνω από όλα τα δεδομένα, εκτιμώντας τις τιμές των σημείων ανάμεσα από αυτά. Χρησιμοποιείται κυρίως για δεδομένα που προέρχονται από δειγματοληψία. Υπάρχουν πολλοί τρόποι ένωσης των δεδομένων: γραμμική ένωση, ένωση με εκτίμηση των κοντινότερων γειτονικών σημείων, cubic spline interpolation όπου χαράζεται σε κάθε ζευγάρι γειτονικών σημείων μια ξεχωριστή πολυωνυμική καμπύλη τρίτου βαθμού [14]. Οι πρώτες δύο μέθοδοι είναι γρήγοροι αλλά δεν παρέχουν ικανοποιητικό smooth σε αντίθεση με την τρίτη μέθοδο. Το Interpolation διαφέρει από το fitting στο ότι γίνεται σε γειτονικά σημεία και η fitted καμπύλη περνά πάνω από όλα τα δεδομένα.

3.2.3 Διαφορές Smoothing αλγορίθμων – Fitting

Το smoothing σκοπό έχει να ομαλοποιήσει τα δεδομένα κάνοντας μία αρχική προσέγγιση των ωφέλιμων δεδομένων ενώ το fitting προσπαθεί να δώσει βαθύτερη ερμηνεία και γνώση αυτών. Το parametric fitting προϋποθέτει μία αρχική επιλογή μοντέλου για την περιγραφή των δεδομένων ενώ το smoothing παρέχει άμεσα αποτελέσματα από μαθηματικούς υπολογισμούς των αρχικών δεδομένων, αποφεύγοντας έτσι το ϵ_p λόγω της επιλογής μοντέλου από το σχεδιαστή. Το smoothing που θα υποστούν τα δεδομένα καθορίζεται από παραμέτρους όπως το span ενώ στο fitting αυτό δεν είναι δυνατό. Το fitting παρέχει στον σχεδιαστή παραμέτρους που μπορούν να ερμηνευτούν και να επεξεργαστούν καθώς και το fitted σήμα είναι πιο ομαλό χωρίς να παρουσιάζει έντονες εναλλαγές όπως στο smoothing.

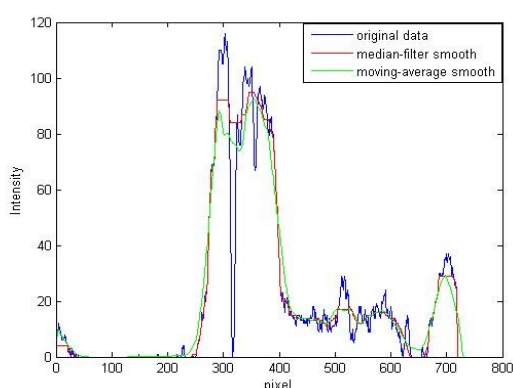
3.3 Επιλογή Μεθόδου

Είσοδος στην υπολογιστική διαδικασία της προ-επεξεργασίας είναι κάθε profile γραμμής και στήλης ξεχωριστά. Τα profile απεικονίζονται στο επίπεδο x,y σαν τυχαία δεδομένα, με διακριτές τυχαίες τιμές από 0 μέχρι 255 στον y -άξονα και από 1 μέχρι 740 ή 1 μέχρι 562 ανάλογα με το αν αναφερόμαστε σε γραμμή ή στήλη. Η πληροφορία προς εξαγωγή είναι κρυμμένη πίσω από τις τυχαίες διακυμάνσεις εντάσεων (θόρυβος). Γίνεται προ-επεξεργασία των spatial profiles ξεχωριστά εφαρμόζοντας τις τεχνικές της ενότητας 3.2.

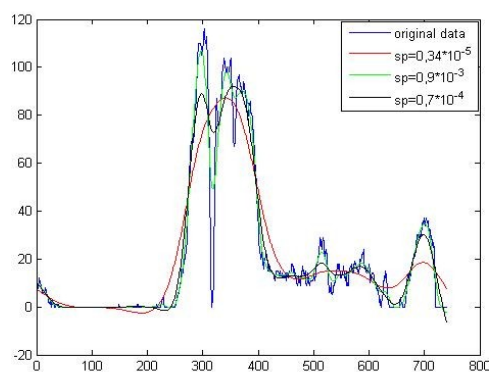
Τα profile επεξεργάζονται αρχικά με smoothing τεχνικές ώστε να γίνει μια αρχική προσέγγιση των χαρακτηριστικών τους, να ομαλοποιηθούν και να επιλεχτεί μια συνάρτηση fitting που τα μοντελοποιεί. Ακολουθεί ανάλυση για ένα τυχαίο profile (έστω γραμμής). Εφαρμόζοντας τη μέθοδο Moving Average και Median με span 41 το profile παίρνει τις μορφές της εικόνας 16. Τα smooth δεδομένα ακολουθούν τα αρχικά με μεγάλη ακρίβεια, αλλά η ύπαρξη πολλών τοπικών μεγίστων και ελαχίστων τα οποία στην εικόνα δεν γίνονται αντιληπτά από την ανθρώπινη όραση, δεν δίνει επιπλέον πληροφορία και δυσκολεύει την επεξεργασία. Είναι αναποτελεσματική η 'ομαδοποίηση' των profile τα οποία μπορούν να έχουν διαφορετικές μορφές. Δοκιμάζοντας smoothing τεχνικές σε πολλά profile παρατηρήθηκε ότι προσεγγίζουν ένα σήμα το οποίο θυμίζει την Gaussian mix συνάρτηση.

Δοκιμάζεται nonparametric fitting με smoothing spline στα profile με μεταβαλλόμενη smoothing parameter από 0 μέχρι 1. Στην εικόνα 17 απεικονίζεται ενδεικτικά το profile της εικόνας 16 και παρατηρείται ότι: όσο η smoothing parameter sp πλησιάζει την μονάδα τόσο το smoothed σήμα συμπίπτει με το αρχικό ενώ όσο μειώνεται η sp και πλησιάζει το μηδέν, τόσο πιο smoothed είναι η καμπύλη. Δίνοντας πολύ μικρή τιμή sp ($sp=0,7*10^{-4}$) το profile προσεγγίζει την Gaussian συνάρτηση με δύο κύριες κορυφές στον κεντρικό λοβό ενώ μειώνοντας την ακόμα περισσότερο ($sp=0,34*10^{-5}$) το profile προσεγγίζει την Gaussian συνάρτηση μίας κορυφής. Επειδή υπάρχουν άπειρες μορφές profile η μέθοδος του non parametric fitting καθιστά αδύνατη την ομαδοποίηση τους και συνεπώς και την εύρεση

αλγορίθμων εξαγωγής χαρακτηριστικών, οπότε απορρίπτεται. Σύμφωνα με την εφαρμογή smoothing αλγορίθμων, nonparametric fitting και τις εξαγόμενες παρατηρήσεις δοκιμάζεται parametric fitting στα profile με nonlinear fitting συνάρτηση $f(x_i)$ την Gaussian mix 2 κορυφών. Απεικονίζεται ενδεικτικά στην εικόνα 19 το fitted profile (κόκκινο χρώμα) εφαρμόζοντας την μέθοδο στο (original) profile (μπλε χρώμα) της εικόνας 16,17 .



Εικόνα 16: *Smoothing spatial profile*

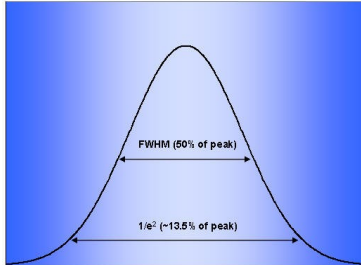


Εικόνα 17: *Smoothing spline με smoothing parameter*
 $sp=0,34*10^{-5}$, $sp=0,7*10^{-4}$, $sp=0,9*10^{-3}$

- Gaussian Συνάρτηση

Η Gaussian συνάρτηση είναι μια από τις σημαντικότερες συναρτήσεις (κατανομές) στην θεωρία των μαθηματικών και διατυπώθηκε από τον Friedrich Gauss

(1777-1855) κατά την εξαγωγή αποτελεσμάτων σε γεωδαιτικές μετρήσεις. Έχει την μορφή καμπάνας και χαρακτηρίζεται από δύο παραμέτρους: την μέση τιμή μ και την διασπορά σ^2 . Η συνάρτηση gauss έχει την γραφική παράσταση της εικόνας 18 και μαθηματική μορφή:



Εικόνα 18: Gaussian function

$$y = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

όπου σ^2 και μ παράμετροι της συνάρτησης. Η διασπορά καθορίζει το εύρος της καμπάνας ενώ η μέση τιμή θέση της μέγιστης τιμής στον x-άξονα [15].

Mixed Gaussian συνάρτηση είναι η συνάρτηση που αποτελείται από άθροισμα απλών Gaussian συναρτήσεων με μαθηματικό τύπο:

$$y = \sum_{i=1}^n \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{(x-\mu_i)^2}{2\sigma_i^2}\right)$$

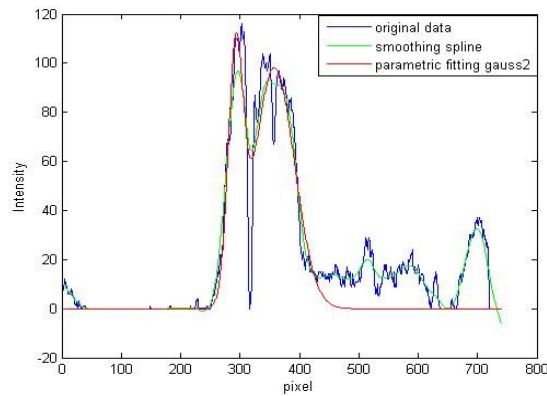
όπου n : αριθμός gaussian συναρτήσεων.

Για $n=2$ η παραπάνω εξίσωση δίνει:

$$y = \frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right) + \frac{1}{\sigma_2 \sqrt{2\pi}} \exp\left(-\frac{(x-\mu_2)^2}{2\sigma_2^2}\right) \quad (1)$$

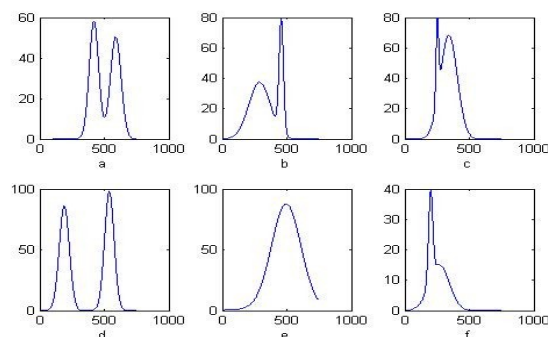
Ακολουθείται το μοντέλο του parametric fitting $y_i = f(x_i) + \varepsilon_i$ στα profile με συνάρτηση fitting $f(x_i)$ την αναφερόμενη (1) όπου για εμάς x τα pixel (x άξονας) και y η predicted έξοδος που εκφράζει την ένταση. Εκτιμούνται οι παράμετροι μ_1, σ_1 , μ_2, σ_2 που ελαχιστοποιούν το συνολικό τετραγωνικό σφάλμα $\sum \varepsilon_i$ ώστε η έξοδος της $f(x_i)$ να προσεγγίζει τα unfitted δεδομένα y_i . Όπου i αναφέρεται σε κάθε δεδομένο. Οι δύο κορυφές της gaussian fitting συνάρτησης επιλέγονται ώστε (ιδανικά) κάθε καμπάνα να 'εντοπίζει' την κάθε πλευρά της βλάβης (αν είναι profile γραμμής για την δεξιά-αριστερή πλευρά και αν είναι profile στήλης για την πάνω και κάτω πλευρά), αφού το demarcation αναφέρεται στα όρια της βλάβης. Επιθυμούμε μετά το fitting να διατηρηθούν οι peak τιμές και οι κλίσεις στα όρια της βλάβης γιατί

είναι τα χαρακτηριστικά που δίνουν την πληροφορία που επιθυμούμε. Δεν μας ενδιαφέρει η προσέγγιση του σήματος ενδιάμεσα.



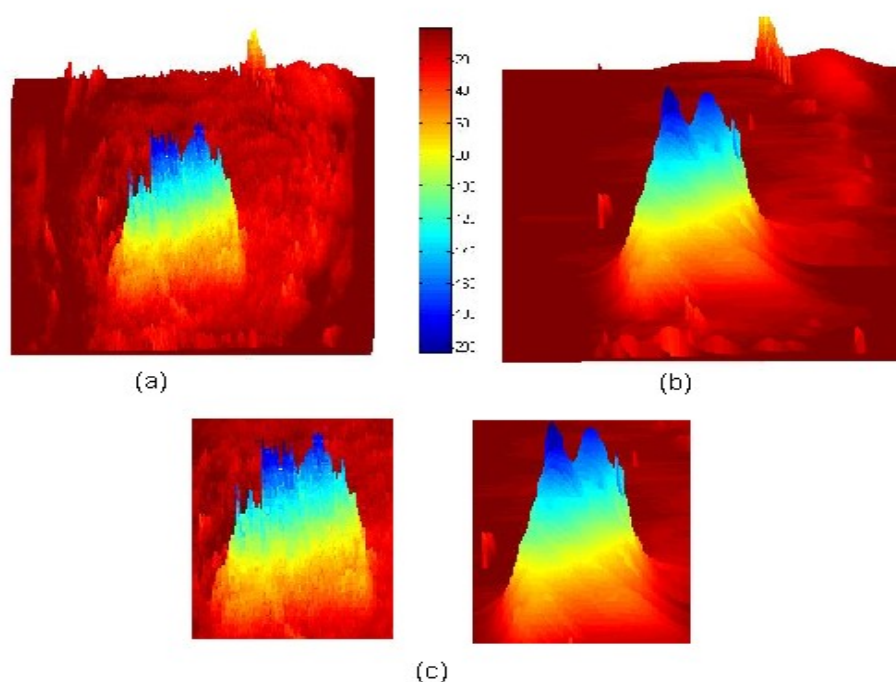
Εικόνα 19: *Smoothing spline vs parametric gauss2 fitting*

Το parametric fitting με nonlinear συνάρτηση gaussian mix με $n=2$ επιτρέπει την ανάλυση των παραμέτρων της fitting συνάρτησης, διατηρείται η απαραίτητη πληροφορία του demarcation δηλ.οι peak τιμές και οι κλίσεις ανά πλευρά^(ενότητα 4.2) και απαλείφονται οι μικρές εντάσεις που δεν δίνουν πληροφορία και δεν εκφράζουν AW καθώς και ο θόρυβος από όργανα εξέτασης που εντοπίζεται στα άκρα της εικόνας(π.χ.στα pixel 450-740)^{ενότητα 2.2.3}. Το σήμα που γίνεται αντιληπτό από την ανθρώπινη όραση και εκφράζει το AW (και έμμεσα το demarcation) στην εικόνα προσεγγίζεται πολύ ικανοποιητικά. Μπορεί να υπάρχουν θεωρητικά άπειρες μορφές fitted profiles λόγω της τυχαιότητας των εικόνων. Στην εικόνα 20 παρουσιάζεται ένα δείγμα των πιθανών fitted profile με gaussian2.



Εικόνα 20: *Fitted profiles με gauss2 a,b,c,d,e,f*

Την εκτίμηση του best fit την δίνει η οπτική εξέταση του συνόλου των δεδομένων. Αναπαρίστανται σε 3D όλες οι original subtracted εικόνες της εργασίας σε αντιστοιχία με αυτές που έχουν γίνει fitted με gaussian2 σε όλα τα profile. Η εικόνα 21 δίνει ένα παράδειγμα όπου η 21a είναι η 3D της original subtracted εικόνας και η 21b η 3D όπου έχουν γίνει fit με gaussian2 όλα τα profile (γραμμών) της εικόνας. Παρόμοια είναι και η 3D αναπαράσταση των fitted στηλών της εικόνας. Η βλάβη φαίνεται μεγενθυμένη στην εικόνα 21c. Η οπτική σύγκριση όλων των εικόνων της εργασίας οδήγησε στο συμπέρασμα ότι όλα τα χαρακτηριστικά της βλάβης διατηρούνται για την πληροφορία που εμείς επιθυμούμε και ο θόρυβος ομαλοποιείται. Οι άκρες σε όλες τις περιπτώσεις γίνονται πιο smooth από ότι στην original, κάτι το οποίο στη συνέχεια της εργασίας επιλύεται.



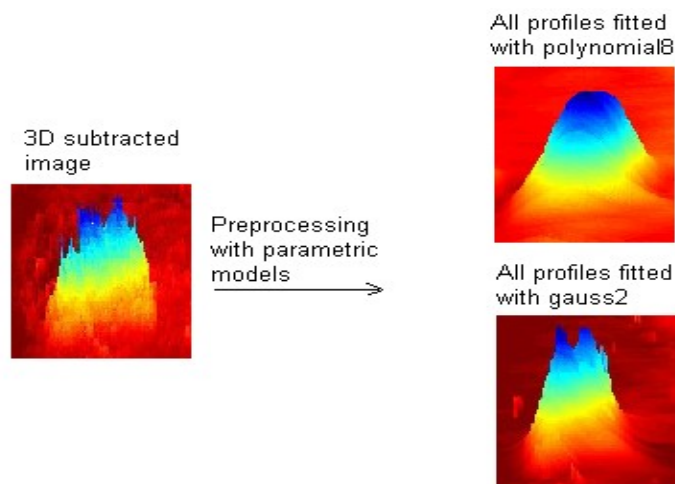
Εικόνα 21: (a) 3D αρχικής εικόνας (b) 3D εικόνα με fitting όλων των profile (c) zoom βλάβης

Κάθε εικόνα έχει 740 γραμμές και 562 στήλες άρα και 740+562 διαφορετικά profiles. Μερικά από αυτά χρειάζονται επιπλέον επεξεργασία και ερμηνεία^(κεφάλαιο 4) διότι δίνουν αναποτελεσματικό fitting και αναφέρονται στη συνέχεια. Αποφασίζεται

η εφαρμογή σε κάθε profile της μεθόδου *simple regression parametric fitting* με την *nonlinear* συνάρτηση *fitting mixed Gaussian δύο κορυφών*. Όλα τα profile γίνονται fitting με τη μέθοδο *NonlinearLeastSquares*, χωρίς robust και με τον επαναληπτικό αλγόριθμο *trust region* για την αλλαγή των παραμέτρων και την εκτίμηση του βέλτιστου fitting.

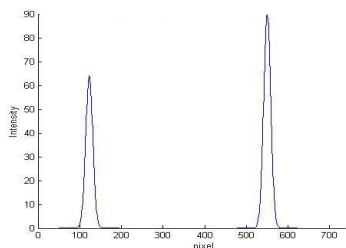
Πρέπει να ικανοποιούνται ορισμένοι περιορισμοί στις τιμές των παραμέτρων (*Coefficient Parameters*) γιατί τα όρια της εικόνας άρα και των profiles είναι περιορισμένα. Έτσι επιλέγεται ως άνω όριο των παραμέτρων $\mu_{1,2}$ και $\sigma_{1,2}$ το μέγεθος της εικόνας Upper=739 και ως κάτω όριο των $\mu_{1,2}$ και $\sigma_{1,2}$ το Lower=2. Επίσης, by default (από software) ισχύει $\sigma_{1,2} > 0$. Τα startpoints (οι αρχικές τιμές των παραμέτρων $\mu_{1,2}, \sigma_{1,2}$) προκύπτουν by default γιατί αλλάζοντάς τα δεν έδωσαν αποτελεσματικότερο fitting και προκύπτουν από τα δεδομένα κάνοντας μια αρχική βέλτιστη εκτίμηση.

Σημειώνεται ότι δεν γνωρίζουμε την μοντελοποίηση του φαινομένου του AW (αυτό ενδιαφέρει τους βιολόγους και απαιτεί κλινικές έρευνες) και ούτε το gaussian επιλέγεται με αυτό το σκεπτικό. Οδηγηθήκαμε στην επιλογή του μοντέλου αυτού γιατί αυτό έδωσε την βέλτιστη προσέγγιση για την πληροφορία που επιθυμούμε σύμφωνα με οπτική παρατήρηση όλων των δεδομένων και των εικόνων της εργασίας. Τέλος, δοκιμάστηκαν και και άλλα παραμετρικά μοντέλα fitting τα οποία κρίθηκαν αναποτελεσματικά διότι αλλοίωναν πολύ τα χαρακτηριστικά της βλάβης. Ένα οπτικό παράδειγμα αναφέρεται παρακάτω όπου η εφαρμογή parametric fitting με πολυωνυμική συνάρτηση 8ου βαθμού αλλοιώνει πολύ τα χαρακτηριστικά προς εξαγωγή που μας ενδιαφέρουν.



4.1 Βελτιστοποιήσεις

Το fitting μπορεί σε ορισμένες περιπτώσεις να είναι αναποτελεσματικό δεδομένου ότι μπορεί να επηρεαστεί από outliers ή salt and paper θόρυβο. Τέτοιου είδους θόρυβος στην εικόνα έχει υψηλές εντάσεις αλλά μικρό εύρος. Το μικρό εύρος εξηγείται γιατί η διαστάσεις του θορύβου αυτού στην εικόνα είναι μικρές σε σχέση με το AW. Το gaussian2 fitting σε τέτοιες περιπτώσεις θα δώσει καμπύλες οι οποίες έχουν πολύ μικρή διασπορά. Διακρίνονται δύο περιπτώσεις: η ύπαρξη μικρής τιμής διασποράς και στις δύο καμπάνες του mixed gaussian2 fitted profile και η ύπαρξη μικρής τιμής διασποράς σε μια από τις δύο.



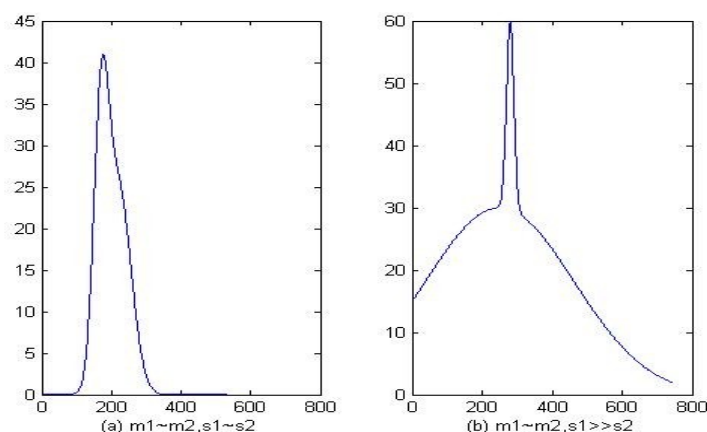
Εικόνα 22: Θόρυβος στην εικόνα

Στην πρώτη περίπτωση το profile θα έχει την μορφή του σχήματος της εικόνας 22 και σε όποια απόσταση και να βρίσκονται οι δύο καμπάνες μεταξύ τους, θεωρείται ως θόρυβος λόγω μικρής τιμής διασποράς και δεν επεξεργάζεται. Η δεύτερη περίπτωση χειρίζεται με διαφορετικό τρόπο διότι

θεωρείται ότι εμπεριέχεται μέσα σε χρήσιμη πληροφορία (ενότητα 4.3).

Υπάρχουν profiles τα οποία δεν δίνουν την απαραίτητη πληροφορία για την ανάλυση, λόγω των πολύ μικρών τιμών εντάσεων. Αν η μέγιστη τιμή ενός τέτοιου profile είναι μικρότερη του 20 δεν λαμβάνεται υπόψιν στην ανάλυση και αγνοείται. Περιοχές της εικόνας με ένταση μικρότερη του 20 έχουν επιβεβαιωθεί και από το Dysis ότι δεν χρωματίζονται στον χάρτη άρα δεν υποδηλώνουν βλάβη.

Επιπροσθέτως, αν οι μέσες τιμές των δύο Gaussian καμπυλών ορισμένων profile είναι περίπου ίσες μπορούν να έχουν και τις μορφές:



Εικόνα 23: Profile με σχεδόν ίσες μέσες τιμές $\mu_{1,2}$ και (a) σχεδόν ίσες διασπορές $\sigma_{1,2}$ (b) διαφορετικές διασπορές $\sigma_{1,2}$

Στο profile της εικόνας 23b το fitting δεν έδωσε ικανοποιητικά αποτελέσματα. Κάνοντας βελτιστοποίηση τα profile με κοντινές μέσες τιμές γίνονται fit με Gaussian συνάρτηση μιας καμπάνας ($n=1$) με τους αντίστοιχους περιορισμούς fitting, προσπαθώντας να προσεγγιστούν καλύτερα, αποφεύγοντας τυχόν σφάλματα κατά την ανάλυση. Τα υπόλοιπα profile διαχειρίζονται χωρίς να υποστούν αλλαγές γιατί θεωρείται ότι περιέχουν την πληροφορία που επιθυμούμε, χωρίς επιπλέον θόρυβο.

4.2 Ερμηνεία Χαρακτηριστικών και Εξαγωγή Παραμέτρων

Στα profile αναζητούνται χαρακτηριστικά που εκφράζουν το demarcation της βλάβης. Εξάγονται χαρακτηριστικά (feature extraction) σε κάθε profile που αποδίδουν με μαθηματικούς αριθμούς αυτό που η ανθρώπινη όραση αντιλαμβάνεται ως demarcation στην εικόνα. Η διαδικασία αυτή γίνεται για την ποσοτικοποίηση του demarcation της βλάβης.

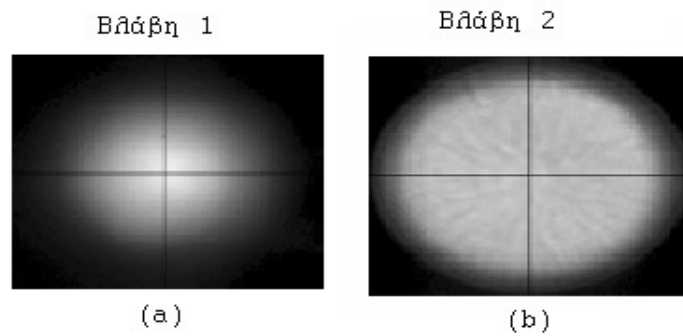
Κάθε υποπεριοχή της εικόνας που ασπρίζει μετά τον ψεκασμό του οξικού οξέος αποτυπώνεται στην εικόνα ως περιοχή υψηλής έντασης. Όσο περισσότερο ασπρίζει μια υποπεριοχή τόσο πιο μεγάλη η ένταση της εικόνας στην υποπεριοχή αυτή και τόσο πιο επικίνδυνη η εξέλιξη της πάθησης του ασθενή. Οι εντάσεις στην subtracted εικόνα σε σχέση με τον χάρτη του Dysis ύστερα από μετρήσεις δείχνει ότι, εντάσεις

με τιμές 0-20 δεν χρωματίζονται στον χάρτη Dysis, 20-50 είναι μπλε ανοιχτό, 50-60 αρχίζουν να γίνονται σκούρο μπλε, 60-75 πράσινο, 75-95 κόκκινο και τιμές από 95 και πάνω χρωματίζονται με άσπρο. Η αντιστοιχία των χρωμάτων στον χάρτη με τις αλλοιώσεις αναφέρθηκαν στο κεφάλαιο 1.

Υπάρχει περίπτωση η άσπρη υποπεριοχή (αν υπάρχει) στον χάρτη του Dysis να περιστοιχίζεται από τα υπόλοιπα χρώματα όπως κόκκινο, πράσινο και μπλε σκούρο (άρα από μικρότερες εντάσεις) σε μεγάλη έκταση. Σε αυτή την περίπτωση το contrast της subtracted εικόνας δεν είναι πολύ μεγάλο γιατί το άσπρο κάποιας υποπεριοχής της σβήνει ομαλά καθώς απομακρυνόμαστε από αυτήν. Το demarcation της βλάβης σε τέτοιες περιπτώσεις είναι μικρό και πιθανολογούμε ότι τα παθογόνα κύτταρα έχουν την τάση εξέλιξης-μετάδοσης κατά μήκος του επιθηλιακού ιστού. Μικρή τιμή του demarcation ερμηνεύεται ως πιθανή τάση κατεύθυνσης των παθογόνων κυττάρων προς τα 'πλάγια' του ιστού, παρατήρηση που έχει την έννοια της πρόβλεψης.

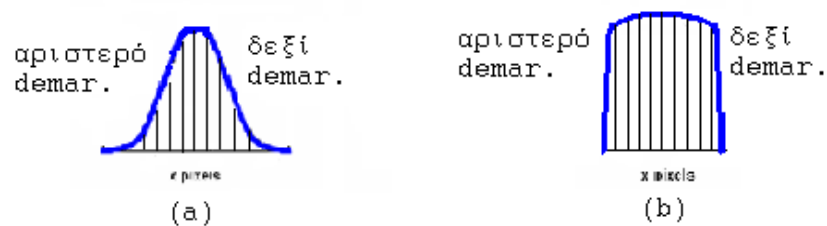
Στην αντίθετη περίπτωση η οποία θεωρείται και πιο επικίνδυνη, η άσπρη υποπεριοχή (αν υπάρχει) στον χάρτη DySIS δεν περιστοιχίζεται από άλλα χρώματα σε μεγάλη έκταση ή από χρώματα μικρών εντάσεων όπως μπλε ή μαύρο. Το contrast της subtracted εικόνας στην περίπτωση αυτή θα είναι μεγαλύτερο. Το demarcation της βλάβης σε τέτοιες περιπτώσεις είναι μεγάλο και πιθανολογείται ότι τα παθογόνα κύτταρα έχουν την τάση εξέλιξης-μετάδοσης κατά βάθος του επιθηλιακού ιστού -περίπτωση μετάστασης-, η οποία δεν ανιχνεύεται με γυμνό μάτι. Οι περιπτώσεις αυτές είναι και οι πιο επικίνδυνες από ιατρικής πλευράς διότι μπορεί η ιατρική εγχείρηση να μην φτάσει στο κατάλληλο βάθος του επιθηλίου ώστε να αφαιρεθούν τα καρκινικά κύτταρα τα οποία εξελίσσονται κατά την κατεύθυνση αυτή και να προκαλέσουν πιθανώς μετάσταση στο μέλλον σπάζοντας τον επιθηλιακό ιστό.

Οι μαθηματικοί αριθμοί που θα προκύψουν από την εξαγωγή χαρακτηριστικών (feature extraction) 'μετρούν' το demarcation των subtracted εικόνων το οποίο εν συνεχεία παρέχει και την πρόβλεψη κατεύθυνσης των παθογόνων κυττάρων όπως αναφέρθηκε. Τα παραπάνω αναλύονται μέσω των spatial profile. Για την ανάλυση παρουσιάζονται ιδανικά παραδείγματα εικόνων. Έστω οι subtracted εικόνες 24a, 24b οι οποίες έχουν AW ίσης έντασης στο κέντρο τους.



Εικόνα 24: (a) Low demarced βλάβη (b) High demarced βλάβη

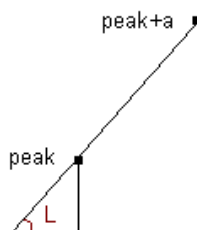
Η εικόνα 24a αναφέρεται στην πρώτη περίπτωση όπου το demarcation της βλάβης είναι μικρότερο και οι εντάσεις 'σβήνουν' σταδιακά στα άκρα ενώ η βλάβη της εικόνας 24b έχει μεγαλύτερο demarcation και η ένταση της μεταβάλλεται απότομα στα όρια παθογόνων – υγιών κυττάρων. Αναζητούνται χαρακτηριστικά τα οποία το εκφράζουν. Εξάγοντας τα spatial profile της γραμμής στο κέντρο της κάθε εικόνας προκύπτει:



Εικόνα 25: (a) Spatial profile low contrast εικόνας (profile 1)
(b) Spatial profile high contrast εικόνας (profile 2)

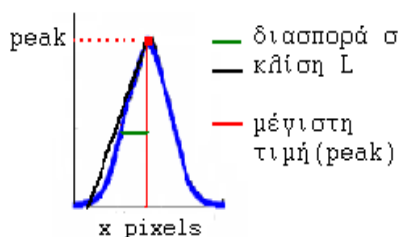
Το profile 1 είναι μια καμπάνα που παίρνει την μέγιστη τιμή της (peak) στο κέντρο και με διασπορά που καθορίζει πόσο απότομα μηδενίζεται η ένταση. Στο profile 2 η μέγιστη τιμή διατηρείται για κάποιο διάστημα και παίρνει επίσης μια peak τιμή έντασης η οποία μηδενίζεται πολύ απότομα στα άκρα της βλάβης. Στα πραγματικά δεδομένα της εργασίας τέτοιου είδους profile σαν το 2 δεν υπάρχουν λόγω του mix gaussian2 fitting. Το AW προσδιορίζεται από το εμβαδό των καμπυλών (γραμμοσκιασμένη περιοχή) των profile ενώ το αριστερό demarcation της βλάβης εκφράζεται από την αριστερή κλίση της καμπύλης και το δεξί αντιστοίχως από την δεξιά για τα profile που έχουν ίδια μέγιστη τιμή έντασης (peak) και

αναφέρονται σε profile γραμμής. Το demarcation εξαρτάται όχι μόνο από τις κλίσεις του εκάστοτε profile οι οποίες εκφράζουν την απότομη ή μη αλλαγή έντασης, αλλά και από την μέγιστη τιμή της έντασης (peak) του διότι ως οπτικό χαρακτηριστικό της εικόνας το οποίο γίνεται αντιληπτό μέσω του contrast, όσο μεγαλύτερο το ΔI της υποπεριοχής με ένταση $I + \Delta I$, με I να εκφράζει την ένταση του γενικού background, τόσο πιο αντιληπτό το contrast της εικόνας. Μπορεί διαφορετικά profile να έχουν ίση τιμή κλίσης αλλά οπτικά η βλάβη και το demarcation δεν γίνονται το ίδιο αντιληπτά. Σχηματικά:



Εξάγονται τα ακόλουθα χαρακτηριστικά ανά profile από τα οποία εξαρτάται το demarcation: i) η peak τιμή του profile η οποία εκφράζει το βαθμό ασπρίσματος που γίνεται αντιληπτό, ii) η γεωμετρική κλίση από τις μικρές τιμές έντασης μέχρι την(τις) peak τιμή(-ες λόγω του gaussian2 fitting) στα όρια της βλάβης. Η βλάβη οριοθετείται από τις κλίσεις αυτές, οι τιμές των οποίων εκφράζουν την απότομη ή μη αλλαγή της έντασης. Κλίση με μεγάλη τιμή αποδεικνύει και απότομη αλλαγή της έντασης άρα στην ιατρική εικόνα ερμηνεύεται ως έντονο demarcation της βλάβης δίνοντας την πρόβλεψη για κατεύθυνση των καρκινικών κυττάρων προς το βάθος του ιστού, ενώ κλίση με μικρότερη τιμή σημαίνει μικρότερο demarcation της βλάβης με πρόβλεψη για κατεύθυνση των καρκινικών κυττάρων στα πλάγια του ιστού. Επιπλέον, εξάγεται και iii) η διασπορά της κάθε καμπάνας που πιθανολογούμε ότι μπορεί να χρησιμοποιηθεί.

Τελικώς τα χαρακτηριστικά που εξάγονται από κάθε profile (έστω μίας gaussian καμπάνας) παρουσιάζονται στο επόμενο διάγραμμα είναι: η peak τιμή, η κλίση και η διασπορά. Τα χαρακτηριστικά αυτά πρέπει να συνδυαστούν κατάλληλα ώστε να δώσουν το demarcation του profile ανά πλευρά βλάβης.



Μπορεί να εξαχθεί και το FWHM που είναι το εύρος της καμπάνας στο μισό του μεγίστου αλλά παρόμοια πληροφορία μας δίνει και η διασπορά, οπότε αυτή η εκδοχή απορρίπτεται.

4.3 Αλγόριθμοι Εξαγωγής Χαρακτηριστικών για κάθε πιθανή μορφή profile

- Αρχικά βήματα αλγορίθμου

Για κάθε εξαγόμενο profile της εικόνας υπολογίζονται τα χαρακτηριστικά που αναφέρθηκαν στην ενότητα 4.2. Τα δεδομένα μας είναι οι ψηφιακές subtracted grayscale στο πράσινο εικόνες bitmap μεγέθους 407KB διαστάσεων 562x740 που έχουν ληφθεί από το Dysis κατά τη διάρκεια γυναικολογικής εξέτασης, προκύπτουν από όσα αναφέρθηκαν και περιέχουν το μέγιστο AW σε όλα τα pixel.

Η εικόνα σαρώνεται αρχικά από πάνω προς τα κάτω ανά 15 γραμμές κατά τον x-άξονα της ώστε να υπάρχει κάποιο περιθώριο ουσιαστικής αλλαγής των profile και στη συνέχεια σαρώνεται ανά 15 στήλες κατά τον y-άξονα της ακολουθώντας την ίδια διαδικασία. Κάθε φορά φορτώνεται ένα profile το οποίο προ-επεξεργάζεται (εφαρμογή parametric fitting με gaussian 2 κορυφών με τους περιορισμούς και τα χαρακτηριστικά που αναφέρθηκαν στην ενότητα 3.3) και εξάγονται τα χαρακτηριστικά. Οι τιμές μικρότερες του 0.9 στρογγυλοποιούνται στο 0. Κρατείται η μέγιστη τιμή του fitted profile. Κάθε fitted profile επεξεργάζεται ξεχωριστά.

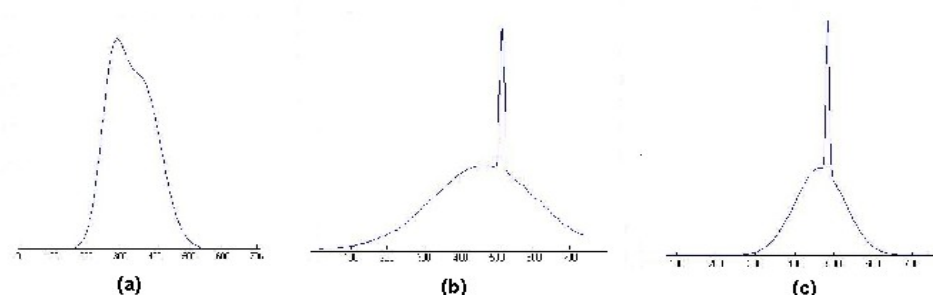
Εφαρμόζονται οι βελτιστοποιήσεις και οι έλεγχοι για τα 'προβληματικά' profile που αναφέρθηκαν στην ενότητα 4.1 με βάση τις τιμές των παραμέτρων $\sigma_{1,2}$, $\mu_{1,2}$ του gaussian fitting: Αν $\max(\text{peak}) < 20$ το profile δεν επεξεργάζεται επιπλέον για εξαγωγή χαρακτηριστικών λόγω ελλιπούς πληροφορίας. Αν οι διασπορές και των δύο κορυφών του gaussian fitted profile είναι πολύ μικρές (< 8.5) το profile θεωρείται μόνο θόρυβος και δεν επεξεργάζεται. Η περίπτωση μικρής τιμής διασποράς (< 7.1)

μόνο στη μια gaussian καμπάνα αναφέρεται στη συνέχεια. (Σημείωση: οι μικρές τιμές διασπορών δεν αποκλείουν την ύπαρξη μικροσκοπικής βλάβης στα σημεία αυτά και για αυτό τα κατώφλια αποφασίζονται με επιφύλαξη. Με τόσο μικρές τιμές διασπορών που επιλέχτηκαν θεωρούμε ότι πιθανή βλάβη δεν είναι ορατή μέσω του AW). Αν οι μέσες τιμές απέχουν κατά 25 pixel γίνεται fitting με μία καμπάνα Gauss ($n=1$) και πάνω στο profile αυτό γίνεται η εξαγωγή χαρακτηριστικών. Η εξαγωγή χαρακτηριστικών της τελευταίας αναφέρεται στην εικόνα 27 γιατί υπάρχει μόνο μια peak τιμή. Σε οποιαδήποτε άλλα profiles ακολουθείται η διαδικασία παρακάτω όπου η επεξεργασία εξαρτάται από την τιμή της παραγώγου:

- Αλγόριθμος επεξεργασίας βασικών profile

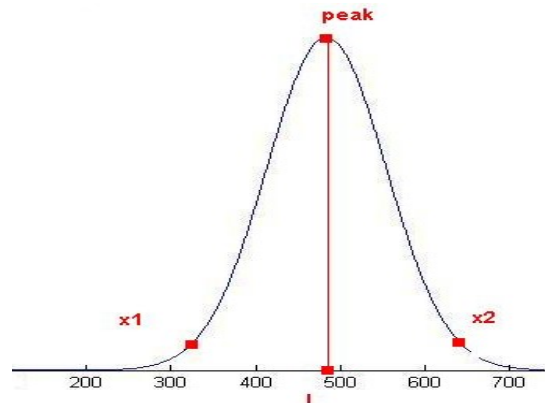
Τα 'βασικά' profile που επεξεργάζονται είναι όσα profile δεν έχουν εξαιρεθεί από τα αναφερόμενα. Ακολουθείται η διαδικασία: Εύρεση της παραγώγου σε κάθε σημείο του spatial profile. Εύρεση τοπικών μεγίστων και ελαχίστων ανάλογα με την αλλαγή προσήμου της παραγώγου και αποθήκευση τους. Σε αυτή τη φάση γίνεται ένας διαχωρισμός στην επεξεργασία των profile.

- 1) Αν κάποιο profile δεν παρουσιάζει κανένα τοπικό μέγιστο θεωρείται προβληματικό και αγνοείται.
- 2) Αν παρουσιάζει ένα μόνο τοπικό μέγιστο και κανένα τοπικό ελάχιστο, κάποιες ενδεικτικές μορφές τους φαίνονται στην εικόνα 26:



Εικόνα 26: (a) Spatial profile με ένα τοπικό μέγιστο, διασπορές σε κοντινές τιμές (b),(c) Spatial profile με ένα τοπικό μέγιστο, διασπορές με μεγάλη διαφορά

καθώς και στη εικόνα 27 έχοντας τη μορφή μίας μονής Gaussian καμπάνας στην περίπτωση που έχει γίνει fit με μια Gaussian συνάρτηση (optimization ενότητα 4.1).



Εικόνα 27: Spatial profile με ένα τοπικό μέγιστο, gaussian fitted συνάρτησης μιας καμπάνας

Στις περιπτώσεις με ένα τοπικό μέγιστο και κανένα τοπικό ελάχιστο η παθογόνος περιοχή είναι η περιοχή που περικλείεται από το εμβαδό της(των) Gaussian καμπάνας(ών). Υπολογίζεται η peak τιμή και εντοπίζεται η θέση της I στον x -άξονα. Υπολογίζονται τα σημεία $x1, x2$ με έλεγχο ανάλογα με την σειρά των παραμέτρων $\sigma_{1,2}$, $\mu_{1,2}$ που έχουν γίνει fit τα δεδομένα, (αναλύεται στην παράγραφο αλγόριθμος εύρεσης των $x1, x2$). Σε περίπτωση που τα $x1, x2$ ξεφεύγουν από τα όρια της εικόνας, παίρνουν τις τιμές $x1=1$ και $x2=\text{μήκος γραμμής/στήλης}$. Στη συνέχεια υπολογίζονται οι κλίσεις $L1$ από την αριστερή πλευρά της βλάβης και $L2$ από την δεξιά αν πρόκειται για profile γραμμής ενώ αν πρόκειται για profile στήλης οι $L1$ αναφέρεται στην πάνω πλευρά και η $L2$ στην κάτω. Η κλίση ευθύγραμμου τμήματος ως προς τον άξονα x ορίζεται ως

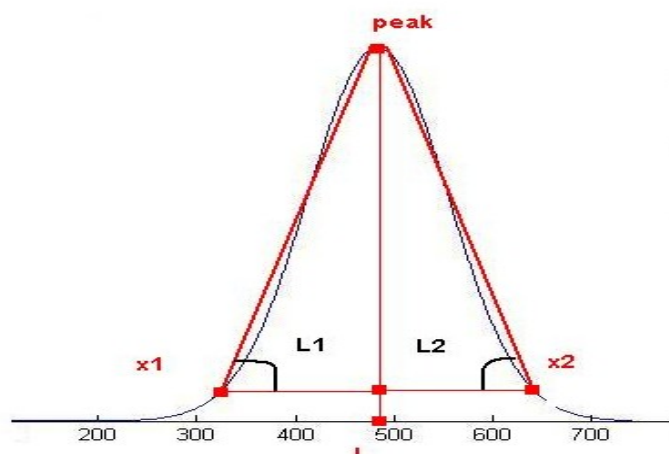
$$L = \frac{\Delta y}{\Delta x}$$

Συγκεκριμένα για την εργασία είναι:

$$L1 = \left| \frac{\text{fitted}(\text{peak}) - \text{fitted}(x1)}{I - x1} \right| \quad \text{και} \quad L2 = \left| \frac{\text{fitted}(\text{peak}) - \text{fitted}(x2)}{I - x2} \right|$$

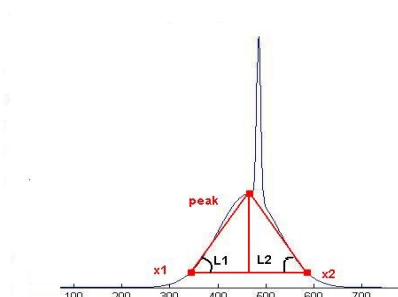
με $\text{fitted}(\cdot)$ η καμπύλη που προέκυψε από την διαδικασία του fitting.

Σχηματικά:



Εικόνα 28: Παράμετροι, ύπαρξη μιας peak τιμής

Στην περίπτωση ενός τοπικού μεγίστου και μικρής τιμής διασποράς σε έναν από τους δύο όρους της gaussian2 fitting συνάρτησης και χωρίς να ικανοποιούν τα κριτήρια της ενότητας 4.1, προκύπτουν profile όμοια με της εικόνας 26b, 26c. Είναι η περίπτωση αναποτελεσματικού fitting στα profile που επηρεάζεται από outliers και εμπεριέχει χρήσιμη πληροφορία για την εικόνα. Το fitting μοντέλο για τέτοια profile δίνει μεγάλες τιμές κλίσεων που δεν ανταποκρίνονται στο AW ($\varepsilon_{\pi} \neq 0$). Προσπαθώντας να λυθεί αυτό οι κλίσεις υπολογίζονται χωρίς να λαμβάνεται υπόψιν η gaussian καμπάνα με την μικρή τιμή διασποράς δηλ. ένας από τους όρους της Gaussian mix συνάρτησης. Υπολογίζονται εκ νέου τα σημεία $x1, x2$ από τον 'ωφέλιμο' όρο, γίνεται έλεγχος και αντικατάσταση των ορίων τους στην εικόνα και η peak τιμή του profile είναι η max τιμή του 'ωφέλιμου' όρου της fitting συνάρτησης. Οι κλίσεις L1, L2 προκύπτουν από τις νέες τιμές των $x1, x2, peak$.



Εικόνα 29: Προβληματικό profile. Εύρεση νέων παραμέτρων

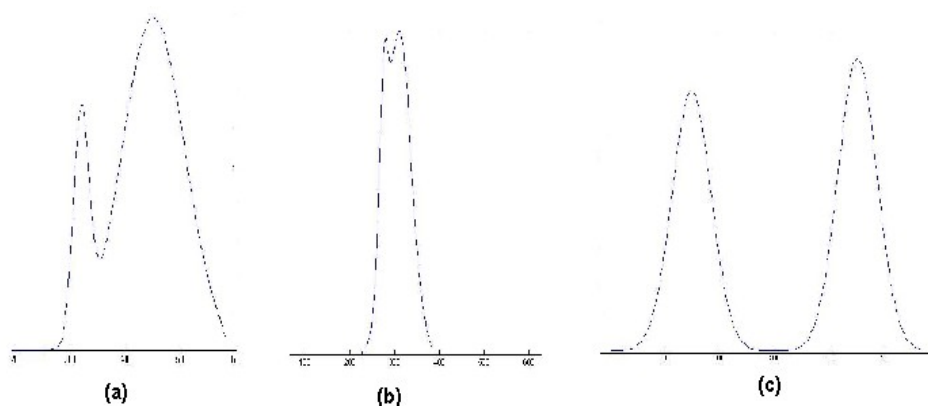
Αριστερά παρουσιάζονται οι νέες κλίσεις L1, L2 σε περίπτωση που η διασπορά ενός όρου της gaussian mix συνάρτησης έχει τιμή μικρότερη του 10. Εμπειρικά τέτοια profiles οδηγούν σε υπολογισμούς κλίσεων που αν δεν αποκλειστούν θα προέκυπταν σφάλματα.

Στη συνέχεια κρατούνται οι διασπορές των gaussian καμπανών (στη συνέχεια το χαρακτηριστικό αυτό δεν χρησιμοποιείται γιατί δεν μπορεί να συσχετιστεί)

Τέλος, για την περίπτωση 2 (ύπαρξη ενός τοπικού μεγίστου) αποθηκεύονται όλοι οι παραπάνω παράμετροι: L1,L2,διασπορές, peak, τιμή για να επεξεργαστούν στη συνέχεια.

Διευκρίνηση: για την αριστερή πλευρά της βλάβης οι παράμετροι είναι L1,s1 ενώ για την δεξιά πλευρά τα L2,s2. Το peak εδώ είναι κοινό.

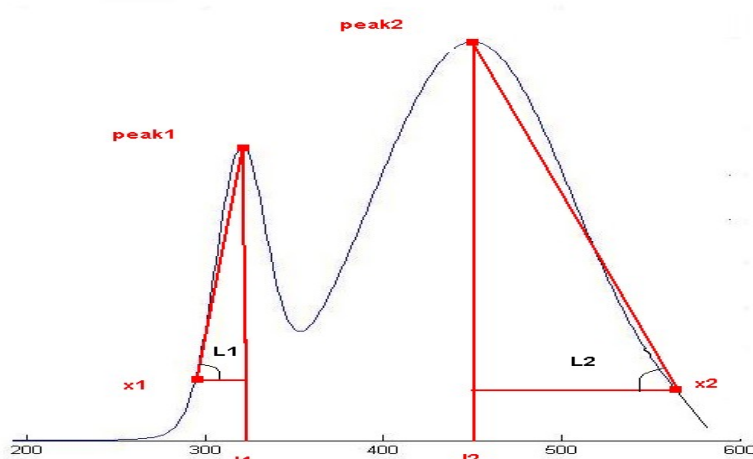
3) Αν παρουσιάζονται δύο τοπικά μέγιστα στο profile είναι δεδομένη η ύπαρξη ενός τοπικού ελαχίστου. Αν το τοπικό ελάχιστο είναι ίσο με 0, το profile θα έχει την μορφή της εικόνας 30c. Στην εικόνα 30 παρουσιάζονται ενδεικτικά ορισμένες μορφές profile με δύο τοπικά μέγιστα.



Εικόνα 30: (a),(b),(c) Profile με δυο T.M , (c) δύο βλάβες

Στην περίπτωση των δύο τοπικών μεγίστων η περιοχή που έχει υποστεί AW είναι η περιοχή που περικλείεται ανάμεσα στο εμβαδό των δύο Gaussian καμπυλών. Υπάρχουν δύο peak τιμές (peak1,peak2) και η βλάβη οριοθετείται από αριστερά από την αριστερή κλίση L1 μέχρι την peak1 ενώ η δεξιά πλευρά από την δεξιά κλίση L2 μέχρι την peak2 αν πρόκειται για profile γραμμής και για profile στήλης πάνω και κάτω αντίστοιχα. Επικεντρωνόμαστε στα άκρα της βλάβης γιατί αναζητούμε το demarcation στα άκρα της βλάβης. Η τιμή που έχει σε κάθε profile το τοπικό

ελάχιστο δεν μελετάται εκτός αν έχει μηδενική τιμή που ερμηνεύεται ως ύπαρξη δύο ξεχωριστών βλαβών. Ο αλγόριθμος για την περίπτωση 3) είναι: Υπολογίζονται οι δύο peak τιμές του profile καθώς και η θέση τους $I1, I2$ στον x-άξονα. Υπολογίζονται τα σημεία $x1, x2$, το $x1$ υπολογίζεται από τις παραμέτρους της αριστερά gaussian καμπύλης ενώ το $x2$ από της δεξιά αν πρόκειται για profile γραμμής(αντίστοιχοι έλεγχοι σύμφωνα με την σειρά που έχει προκύψει από το fitting). Γίνεται έλεγχος ώστε να μην ξεπερνάνε τα $x1, x2$ τα όρια της εικόνας και πιθανή αντικατάσταση $x1=1$ και $x2=μῆκος γραμμής/στήλης$. Υπολογίζονται οι κλίσεις $L1, L2$ σύμφωνα με τους ακόλουθους τύπους. Στην εικόνα 31 παρουσιάζονται σχηματικά τα αναφερόμενα σε ένα τυχαίο profile με δύο T.M και ένα T.E και οι κλίσεις όπως υπολογίζονται:



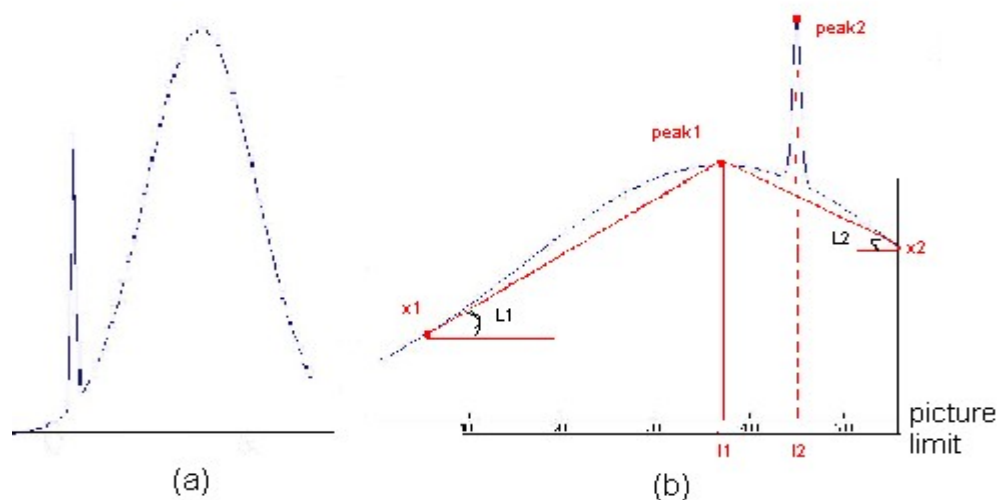
Εικόνα 31: Παράμετροι για profile με 2 τοπικά μέγιστα

Υπολογίζοντας τις κλίσεις:

$$L1 = \left| \frac{fitted(peak1) - fitted(x1)}{I1 - x1} \right| \quad \text{και} \quad L2 = \left| \frac{fitted(peak2) - fitted(x2)}{I2 - x2} \right|$$

με $fitted(.)$ η καμπύλη που προέκυψε από την διαδικασία του fitting.

Γίνεται έλεγχος για ύπαρξη fitted profile μη απαλλαγμένο από θόρυβο (ύπαρξη μικρής τιμής διασποράς σε μια από τις δύο καμπάνες) το οποίο δεν υποστηρίζεται από τις προαναφερθείσες περιπτώσεις και εμπεριέχει χρήσιμη πληροφορία για το demarcation. Παρουσιάζονται πιθανές μορφές τέτοιων profile:

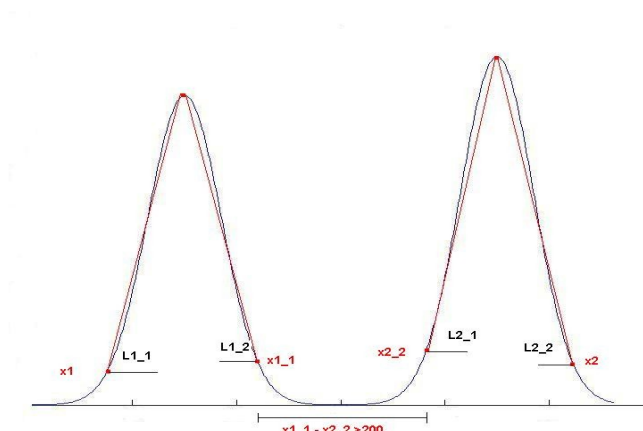


Εικόνα 32: (a) Profile με δύο T.M και μικρή τιμή διασποράς σ_i (b) Διαχείριση προβληματικών profile με 2 T.M

Στην περίπτωση αυτή ο gauss όρος της fitting συνάρτησης με την μικρή τιμή διασποράς (<10) αγνοείται και λαμβάνεται υπόψιν μόνο ο ένας από τους όρους της fitting συνάρτησης. Υπολογίζονται νέες τιμές για τα x_1, x_2 σύμφωνα με τις παραμέτρους του 'οφέλιμου' όρου της Gaussian mix συνάρτησης και οι κλίσεις L_1, L_2 προκύπτουν από την peak τιμή του όρου αυτού και τα καινούρια σημεία x_1, x_2 όπως δείχνεται στην εικόνα 32b. Σημείωση: το x_2 της εικόνας 32b επειδή πέφτει έξω από τα όρια της εικόνας τοποθετείται στο σημείο $x_2 = \text{μήκος profile}$, κάτι που γίνεται σε οποιαδήποτε παρόμοια περίπτωση.

Ελέγχεται η περίπτωση όπου το τοπικό ελάχιστο είναι 0 σε profile με δύο τοπικά μέγιστα, γεγονός που ερμηνεύεται ως ύπαρξη δύο βλαβών, αρκεί να είναι σε κάποια απόσταση μεταξύ τους, αλλιώς επεξεργάζονται όπως παραπάνω (εικόνα 33). Αναζητούνται τα σημεία $x_1, x_2, x_{1_1}, x_{2_2}$ όπως απεικονίζονται στην εικόνα 33 τα οποία προκύπτουν από τις παραμέτρους του mixed Gaussian fitting. Τα σημεία x_{1_1} και x_{2_2} δεν πρέπει να ξεφεύγουν από τα όρια της εικόνας -έλεγχος και πιθανή αντικατάσταση με 1 και μήκος profile-. Αν τα σημεία x_{1_1} και x_{2_2} απέχουν κατά 200 pixel θα πρέπει να θεωρηθούν σαν ξεχωριστές βλάβες και να υπολογιστούν νέοι παράμετροι. Υπολογίζονται οι κλίσεις $L_{1_1}, L_{1_2}, L_{2_1}, L_{2_2}$ που φαίνονται στην

εικόνα και κρατούνται οι δύο μεγαλύτερες τιμές τους, προσπαθώντας να αποδοθεί η χειρότερη περίπτωση του demarcation ($L1 = \max(L1_1, L1_2)$, $L2 = \max(L2_1, L2_2)$).



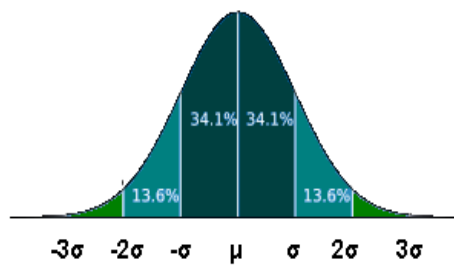
Εικόνα 33: 2 βλάβες σε απόσταση μεγαλύτερη των 200 pixel

Κρατούνται επίσης οι διασπορές των gaussian καμπανών. Οι κλίσεις $L1_1, L1_2$ και $L2_1, L2_2$ δεν έχουν μεγάλες διαφορές τιμών λόγω της μικρής επιρροής του mix gaussian στα σημεία $x1_1, x2_2$. Για κάθε profile στην περίπτωση 3 (ύπαρξη 2 τοπικών μεγίστων) υπολογίζονται και αποθηκεύονται οι παράμετροι $L1, peak1$ και $L2, peak2$ από την δεξιά.

Τέλος για κάθε profile ελέγχονται οι τιμές των $L1, L2$ και αν υπερβαίνουν την μέγιστη θεωρητική τιμή που έχει υπολογιστεί και είναι 15 αντικαθίσταται από το 1/3 της.

- Αλγόριθμος εύρεσης των σημείων $x1, x2$ για κάθε profile

Η κανονική (gaussian) κατανομή έχει το χαρακτηριστικό να περιέχει το 68,2% της πληροφορίας της στην περιοχή από $-σ$ μέχρι $σ$, το 95,4% της πληροφορίας της στην περιοχή $-2σ$ μέχρι $2σ$ και το 99,6% στην περιοχή $-3σ$ μέχρι $3σ$ [15], όπως παρουσιάζεται στην ακόλουθη εικόνα:



Εικόνα 34: Gaussian κατανομή

Εφαρμόζοντας fitting σε όλα τα profiles κάθε subtracted εικόνας και κάνοντας την 3D αναπαράσταση σε όλες τις εικόνες της εργασίας, ύστερα από σύγκριση με την 3D αναπαράσταση των αρχικών εικόνων με τα unfitted profile, παρατηρήθηκε οπτικά ότι οι άκρες των fitted profile γίνονται πιο smooth από ότι στα unfitted. Τα fitted profile στην περιοχή (-2σ) μέχρι (-3σ) είναι πιο smoothed από ότι αυτά στην subtracted unfitted εικόνα και η βλάβη δείχνει πιο 'απλωμένη'. Η εικόνα 21 φανερώνει την παρατήρηση αυτή. Για να αποφευχθεί αυτό γίνεται ένα είδος φιλτραρίσματος κάθε φορά που εξάγονται χαρακτηριστικά σε κάθε profile ώστε να 'κόβονται' οι περιοχές 3σ μέχρι 2σ . Στόχος είναι να προσεγγίζουν τα fitted profile όσο το δυνατόν περισσότερο τα unfitted και η βλάβη να μην αλλοιώνει τα χαρακτηριστικά της. Για το λόγο αυτό τα σημεία x_1, x_2 , από όπου υπολογίζονται οι κλίσεις L_1, L_2 , τοποθετούνται στα σημεία -2σ και 2σ αντίστοιχα έχοντας υπόψιν την παραπάνω παρατήρηση και την πληροφορία που δίνει η εικόνα 34. Αγνοώντας μόνο το 4,6% της πληροφορίας των profile, οι κλίσεις υπολογίζονται με μεγαλύτερη ακρίβεια. Συγκεκριμένα, σε profile γραμμής το σημείο x_1 τοποθετείται στο σημείο -2σ σύμφωνα με τον όρο του gaussian mix fitting που δίνει την αριστερή gaussian καμπάνα ενώ το x_2 στο σημείο 2σ που προκύπτει από την δεξιά gaussian καμπάνα. Ομοίως σε profile στήλης, πάνω και κάτω αντίστοιχα.

Για να βρεθούν τα σημεία x_1, x_2 απαιτείται i) η εύρεση της θέσης των δύο gaussian καμπανών στον x άξονα ανάλογα με την μέση τιμή τους και ii) εύρεση της σχέσης των παραμέτρων μεταξύ της συνάρτησης Gauss (ενότητα 3.3) και της mix gaussian fitting συνάρτησης που προκύπτει από το software. Η τελευταία είναι:

$$fitted_{gauss2} = a1 * \exp\left(-\frac{x-b1}{c1}\right)^2 + a2 * \exp\left(-\frac{x-b2}{c2}\right)^2$$

άρα

$$a_{1,2} = \frac{1}{\sigma_{1,2} \sqrt{2\pi}}$$

$$b1 = \mu_1 \quad \text{και} \quad b2 = \mu_2$$

$$c1 = \sqrt{2} * \sigma_1 \quad \text{και} \quad c2 = \sqrt{2} * \sigma_2$$

είναι η αντιστοιχία των παραμέτρων της gaussian fitting συνάρτησης $fitted_{gauss2}$ και της συνάρτησης Gauss. Πρέπει να διασφαλιστεί ότι τα σημεία $x1$ και $x2$ τοποθετούνται πάντα στον x άξονα ικανοποιώντας τη σχέση $x1 < x2$. Αυτό επιτυγχάνεται με ελέγχους των $b1, b2$. Υπολογίζοντας για κάθε περίπτωση την απόσταση 2σ είναι:

$$2\sigma_i = \frac{2}{\sqrt{2}} ci = \sqrt{2} * ci, i = 1, 2$$

Συγκεκριμένα:

$$2\sigma_2 = \frac{2}{\sqrt{2}} c2 = \sqrt{2} * c2 \quad \text{και} \quad 2\sigma_1 = \frac{2}{\sqrt{2}} c1 = \sqrt{2} * c1$$

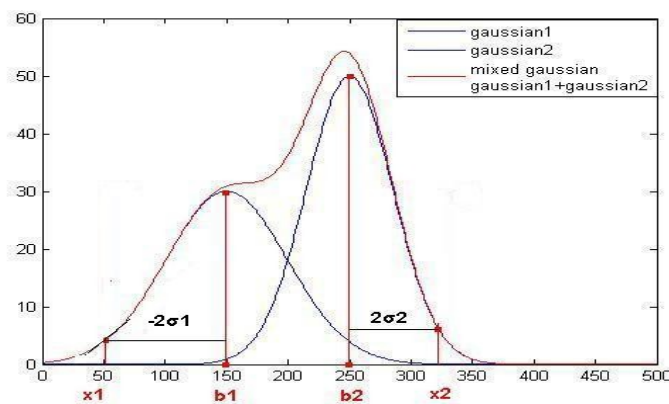
Υπολογίζοντας τα σημεία $x1, x2$ στην περίπτωση που $b1 < b2$ είναι

$$x1 = \mu_1 - 2\sigma_1 = b1 - \sqrt{2} * c1 \quad \text{και} \quad x2 = \mu_2 + 2\sigma_2 = b2 + \sqrt{2} * c2$$

ενώ αν $b1 > b2$ προκύπτουν:

$$x1 = \mu_2 - 2\sigma_2 = b2 - \sqrt{2} * c2 \quad \text{και} \quad x2 = \mu_1 + 2\sigma_1 = b1 + \sqrt{2} * c1$$

Γραφικά απεικονίζεται η περίπτωση $b1 < b2$:



Εικόνα 35: Υπολογισμός των σημείων $x1, x2$

Σε περίπτωση που υπάρχουν πολύ μεγάλες μέσες ή πολύ μικρές τιμές των gaussian1,2 συναρτήσεων, τα σημεία x_1, x_2 υπάρχει περίπτωση να πέσουν έξω από τα όρια της εικόνας. Αν συμβεί αυτό, αντικαθίστανται από τις τιμές $x_1=1$ και $x_2=\text{μήκος profile}$. Η διαδικασία για τον υπολογισμό των x_1, x_2 γίνεται τόσο για τα profile του x άξονα της εικόνας όσο και για τον y.

4.4 Συνδυασμός Παραμέτρων για κάθε profile

Η εικόνα σαρώνεται αρχικά ανά 15 γραμμές κατά τον x άξονα της και εξάγονται 38 profile. Κρατούνται σε πίνακες οι τιμές των εξαγόμενων χαρακτηριστικών και για τα 38 profile:

Γραμμή εικόνας	Γραμμή εικόνας
Peak 1 – x axis	Peak 2 – x axis
Κλίση L1 – x axis	Κλίση L2 – x axis

για την αριστερή πλευρά της βλάβης | για την δεξιά πλευρά της βλάβης

Οι τιμές των διασπορών που έχουν κρατηθεί αποτέλεσαν τελικώς έναν δείκτη επιβεβαίωσης των κλίσεων L1,L2 και δεν θα ληφθούν υπόψη στην διαδικασία του συνδυασμού των χαρακτηριστικών που να εκφράζουν το demarcation.

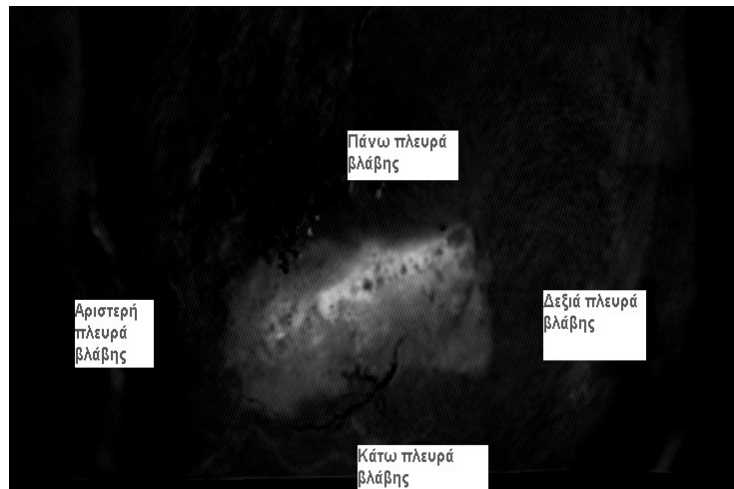
Στη συνέχεια γίνεται σάρωση της subtracted εικόνας κατά τον y άξονα της, αλλάζοντάς τις διαστάσεις και τους περιορισμούς του gaussian fitting για τα όρια Upper και Lower τα οποία γίνονται: μέση τιμή μ και διασπορά σ Upper =560 και Lower=2 αντίστοιχα. Η σάρωση γίνεται ανά 15 στήλες και φορτώνονται 48 profile. Κρατούνται σε πίνακες οι τιμές και για τα 48 profile κατά τον y άξονα:

Στήλη εικόνας	Στήλη εικόνας
Peak 1 – y axis	Peak 2 – y axis
Κλίση L1 – y axis	Κλίση L2 – y axis

για την πάνω πλευρά της βλάβης | για την κάτω πλευρά της βλάβης

Σε περίπτωση ενός τοπικού μεγίστου η $\text{peak1}=\text{peak2}=\text{peak}$. Όταν αναφέρεται η αριστερή, δεξιά, πάνω και κάτω πλευρά της βλάβης στην εικόνα, σχηματικά αναφερόμαστε στα πεδία της εικόνας, προσπαθώντας να οριοθετήσουμε τα όρια του

AW:



Το demarcation εξαρτάται όχι μόνο από την απότομη ή μη αλλαγή έντασης (κλίση) στα όρια της βλάβης αλλά και από την μέγιστη τιμή (peak). Αν και τα δύο αυτά μεγέθη έχουν μεγάλες τιμές το demarcation της βλάβης αναμένεται μεγάλο. Για το λόγο αυτό επιθυμούμε μια γραμμική σχέση μεταξύ των peak τιμών και των κλίσεων. Βασιζόμενοι σε αυτό υπολογίζεται η παράμετρος D_i που εκφράζει το demarcation ανά πλευρά της βλάβης και είναι:

$$D_i = 50\% \cdot peak_i / 100 + 50\% \cdot L_i$$

Η τιμή του peak είναι απαραίτητο να κανονικοποιηθεί για να είναι συγκρίσιμα και προσθέσιμα τα δύο αυτά μεγέθη. Διαιρείται με 100 που είναι περίπου ο μέσος όρος όλων των peak σε όλες τις εικόνες. Έγιναν δοκιμές για το βάρος που θα δοθεί σε κάθε παράμετρο. Η τελική τιμή βαρών που προέκυψε είναι 50% ανά χαρακτηριστικό για δύο λόγους. Διότι η αντίληψη του demarcation εξαρτάται εξίσου από τα δύο αυτά χαρακτηριστικά με τα μέχρι τώρα δεδομένα μελέτης και το ποσοστό 50% μπορεί να διαχωρίσει καλύτερα τις positive (high grade) και negative (low grade) βλάβες. Στους πίνακες προστίθενται οι παράμετροι D_1, D_2 για κάθε πλευρά της βλάβης:

Γραμμή εικόνας	Γραμμή εικόνας	Στήλη εικόνας	Στήλη εικόνας
Peak 1 – x axis	Peak 2 – x axis	Peak 1 – y axis	Peak 2 – y axis
Κλίση L1 – x axis	Κλίση L2 – x axis	Κλίση L1 – y axis	Κλίση L2 – y axis
D1	D2	D1	D2

Για την αριστερή και την πάνω πλευρά της βλάβης ανά άξονα οι παράμετροι είναι:

$$D1 = 50 \% \cdot peak1 / 100 + 50 \% \cdot L1$$

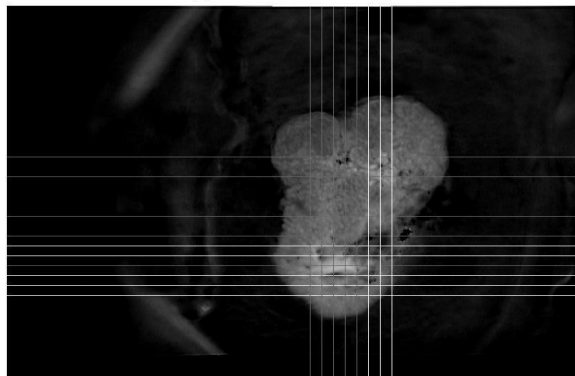
ενώ για την δεξιά και κάτω πλευρά της βλάβης ανά άξονα:

$$D2 = 50 \% \cdot peak2 / 100 + 50 \% \cdot L2$$

Από κάθε profile (γραμμών είτε στηλών) εξάγονται δύο τελικές παραμέτρους D1,D2 που εκφράζουν το demarcation ανά πλευρά. Η σάρωση της εικόνας (στην υποπεριοχή που εντοπίζεται το AW) κατά τον x και y άξονα της και η εξαγωγή των D1,D1 ανά άξονα δίνει τις παραμέτρους του demarcation για όλες τις πλευρές της βλάβης. Το AW εντοπίζεται μόνο σε μια υποπεριοχή της subtracted εικόνας και οι τελικές παράμετροι ανά ασθενή πρέπει να αναφέρονται στο AW και όχι στα υπόλοιπα πεδία της εικόνας. Στα επόμενα αναφέρεται η επιλογή των profile που 'εντοπίζουν' το AW και ο τρόπος που συνδυάζονται τα D1,D2 για εξαγωγή των τελικών παραμέτρων ώστε το demarcation της βλάβης κάθε ασθενή να περιγράφεται από 2 βασικές τελικές αριθμητικές παραμέτρους.

5.1 Επιλογή Τελικών Παραμέτρων για κάθε εικόνα

Σκοπός της εργασίας είναι να αποδίδονται σε κάθε ασθενή (σε κάθε subtracted εικόνα) τιμές τελικών παραμέτρων που προκύπτουν συναρτήσει των D1,D2 που να εκφράζουν το demarcation της βλάβης τους. Επόμενο βήμα της εργασίας είναι ο εντοπισμός των profile που αναφέρονται στην υποπεριοχή που έχει υποστεί AW. Μια προσέγγιση εντοπισμού του AW στην εικόνα είναι η επιλογή των profile με τις μεγαλύτερες τιμές peak, ώστε να προσεγγιστεί η 'χειρότερη περίπτωση' ασπρίσματος μετά τον ψεκασμό οξικού οξέως. Επιλέγονται συνολικά 20 profile (10 profile γραμμών και 10 στηλών) με τις μέγιστες τιμή peak όπου θα υπολογιστούν οι τελικοί παράμετροι της βλάβης. Συγκεκριμένα, επιλέγονται 5 profile γραμμών όπου παρουσιάζονται οι 5 μεγαλύτερες τιμές peak1 από την αριστερή πλευρά της βλάβης, 5 profile γραμμών με τις 5 μεγαλύτερες τιμές peak2 από την δεξιά πλευρά της, 5 profile στηλών με τις μεγαλύτερες τιμές peak1 από την πάνω πλευρά και 5 profile στηλών με τις 5 μεγαλύτερες τιμές peak2 από την κάτω. Τα 5 profile για την κάθε πλευρά της βλάβης λαμβάνονται για να υπάρχει μεγαλύτερη ανοχή σε θόρυβο διότι υπάρχει πιθανότητα να επιλεγεί profile θορύβου που δεν έχει αποφευχθεί από την επεξεργασία που αναφέρθηκε. Σαν παράδειγμα αναφέρεται η ασθενής 84_208 του νοσοκομείου Αλεξάνδρα όπου επιλέγονται 20 profile με τις μέγιστες 20 τιμές peak που αναφέρθηκαν και οι οποίες δείχνονται στην εικόνα 36:

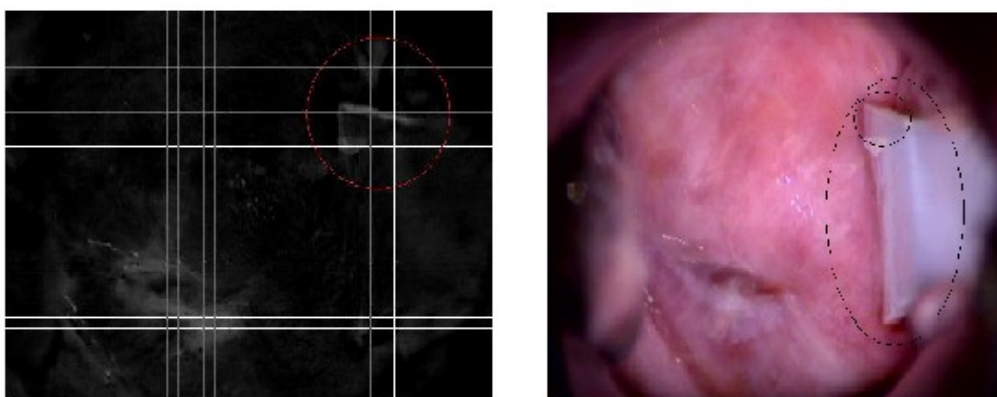


Εικόνα 36: Επιλογή γραμμών-στηλών όπου υπολογίζονται οι τελικοί παράμετροι του demarcation

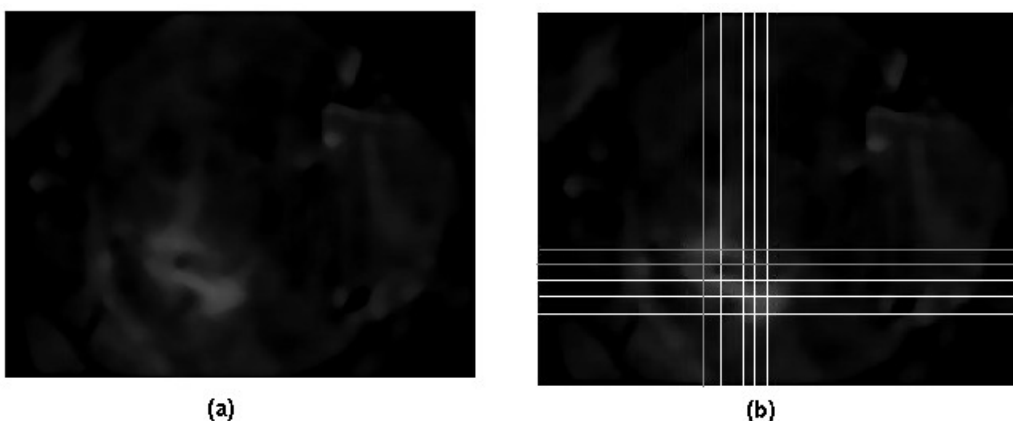
Με άσπρο χρώμα διακρίνονται τα profile γραμμών και στηλών που επιλέγονται για την αριστερή και πάνω πλευρά του AW αντίστοιχα και με γκρι χρώμα από την δεξιά και κάτω πλευρά. Οι στήλες στη συγκεκριμένη περίπτωση δεν είναι 10 γιατί συμπεφτουν από την πάνω και κάτω πλευρά. Αυτό συνήθως γίνεται διότι το fitting δίνει συνάρτηση με ένα τοπικό μέγιστο άρα $peak1=peak2$ αλλά δεν αποκλείεται και η ταυτόχρονη ύπαρξη μεγάλων τιμών $peak1$ και $peak2$.

Με βάση την επιλογή των max τιμών των peak στην εικόνα εντοπίζονται τα πεδία με τις μεγαλύτερες εντάσεις. Αυτό δεν αποκλείει την επιλογή κάποιου profile το οποίο δεν αναφέρεται σε AW αλλά σε θόρυβο εικόνας. Ο θόρυβος μπορεί να υπάρχει λόγω μη αποτελεσματικού registration των εικόνων από το DySIS που δεν έχει αποφευχθεί από τα αναφερόμενα. Στην εικόνα 37 απεικονίζεται μια τέτοια περίπτωση subtracted εικόνας με τα επιλεγμένα από το software της εργασίας 20 profile των max peak. Εδώ η απόκτηση της subtracted δεν έδωσε ικανοποιητικά αποτελέσματα και ο θόρυβος του οργάνου εξέτασης παρέμεινε στην εικόνα σαν άσπρα σημάδια. Η RGB εικόνα επιβεβαιώνει την ύπαρξη οργάνου εξέτασης. Για να αποφευχθούν τέτοιου είδους σφάλματα, προτείνεται η εφαρμογή median φίλτρου στην εικόνα. Η επιλογή των max peak από την filtered subtracted εικόνα λύνει το πρόβλημα ικανοποιητικά. Το median filter εφαρμόζεται MONO για την επιλογή των profile στα οποία θα υπολογιστούν οι τελικές παράμετροι. Οι παράμετροι D1,D2 για

τα profile αυτά έχουν υπολογιστεί και αποθηκευτεί στους πίνακες που αναφέρθηκαν από την unfiltered subtracted εικόνα με τα fitted profile αφού η εξαγωγή των D1,D2 πρέπει να αναφέρεται στην πραγματική-ανεπεξεργαστη ιατρική εικόνα. Μετά την εφαρμογή του median φίλτρου 25 γειτόνων στην subtracted της εικόνας 37, προκύπτει η εικόνα 38a και αναζητώντας στη filtered εικόνα τα max peak επιλέγονται τελικά τα profile της εικόνας 38b. Αποφεύγεται η λανθασμένη προηγούμενη επιλογή των profile.



Εικόνα 37: Subtracted εικόνα με artifacts και RGB original εικόνα



Εικόνα 38: (a)Median φίλτρο 25 στην εικόνα (b)Επιλογή profile μετά την εφαρμογή median filter 25

Έχοντας εξασφαλίσει εγκυρότητα για την σωστή επιλογή των 20 profile που αναφέρονται στο AW εξάγονται οι τελικές παράμετροι για κάθε subtracted συναρτήσε των παραμέτρων D1,D2:

$overall = mean(D1,D2)$ και των 20 επιλεγμένων profile. Λαμβάνουν μέρος στον

υπολογισμό οι τιμές των 5 D1(αριστερή πλευρά), 5 D2(δεξιά πλευρά), 5 D1(πάνω πλευρά), 5 D2(κάτω πλευρά). Η παράμετρος αυτή εκφράζει το συνολικό demarcation της βλάβης. Οι τιμές της κυμαίνονται για τα δεδομένα που δοκιμάστηκαν στα πλαίσια της εργασίας από 0,20 μέχρι 3.

$worst_case_x = MAX (mean(5 D1(αριστερή πλευρά)) , mean(5 D2(δεξιά πλευρά)))$
για τα profile γραμμών και εκφράζει την χειρότερη περίπτωση demarcation από αριστερά ή δεξιά της βλάβης (x διάσταση της εικόνας). Ομοίως για τις στήλες:

$worst_case_y = MAX (mean(5 D1(πάνω πλευρά)) , mean(5 D2(κάτω πλευρά)))$ και εκφράζει το χειρότερο demarcation από την πάνω και κάτω πλευρά της βλάβης (y διάσταση της εικόνας). Οι τιμές τους κυμαίνονται επίσης μεταξύ 0,20 και 3.

$EA = MAX (worst_case_x, worst_case_y)$ εκφράζοντας το χειρότερο demarcation όλης της βλάβης.

Οι παραπάνω παράμετροι έχουν ποσοτικοποιήσει πλέον το demarcation κάθε ασθενούς.

5.2 Αποτελέσματα – Ερμηνεία αποτελεσμάτων

Η χρήση οξικού οξέος εντοπίζει τα (προ)-καρκινικά κύτταρα μέσω του AW και η οπτική επιθεώρηση οδηγεί σε διάγνωση του σταδίου αλλοίωσης CIN I,II,III ή καρκίνου (κολποσκόπηση). Πειράματα στην Κίνα το 1997 απέδειξαν ότι η οπτική επιθεώρηση του AW εξισώνεται στον παράγοντα ευαισθησίας (sensitivity) με την κυτταρολογία.[17]. Ένα από τα χαρακτηριστικά της οπτικής επιθεώρησης του AW (σύμφωνα και με το RCI) είναι και το demarcation το οποίο για πρώτη φορά μετριέται με μαθηματικούς δείκτες στα πλαίσια αυτής της διπλωματικής εργασίας. Οι τελικές παράμετροι *overall*, *worst_case_y*, *worst_case_x*, *EA* εκφράζουν αριθμητικά το demarcation κάθε subtracted εικόνας που προκύπτει από το DySIS. Για κάθε ασθενή που υπόκειται στην γυναικολογική εξέταση με το DySIS μπορούν να εξάγονται οι παράμετροι αυτοί του demarcation εξασφαλίζοντας αντικειμενικότητα του χαρακτηριστικού και παρέχοντας την έννοια της πρόβλεψης των παθογόνων κυττάρων στο επιθήλιο.

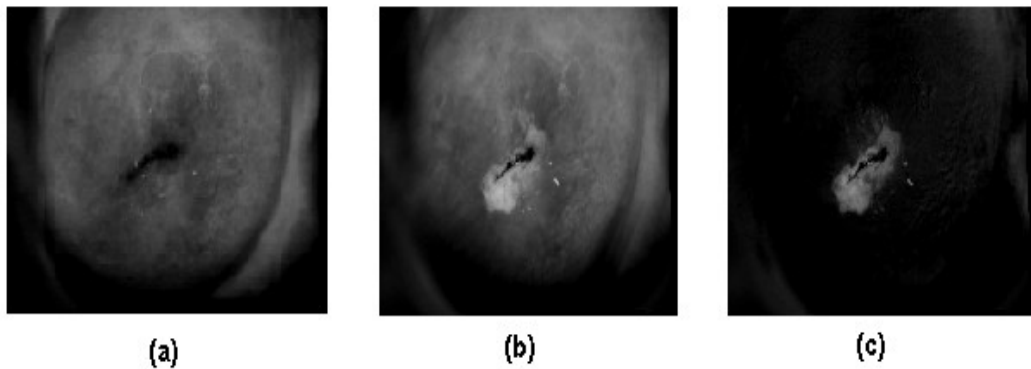
Η μαθηματική παράμετρος *overall* δίνει μία γενική εκτίμηση του demarcation

της βλάβης. Αναμένονται μεγάλες τιμές της σε εικόνες με γενικά έντονο demarcation (high-grade βλάβες) ενώ μικρές τιμές σε χαμηλό (low-grade βλάβες). Με βάση τις τιμές της παραμέτρου αυτής γίνεται στη συνέχεια διαχωρισμός low και high grade βλαβών. Η overall δίνει μια πρώτη γενική εικόνα του demarcation ενώ οι άλλες παράμετροι αναφέρονται σε πιο τοπικό επίπεδο. Η παράμετροι *worst_case_x*, *worst_case_y* εκφράζουν το χειρότερο demarcation της βλάβης ανά διάσταση της subtracted εικόνας και η *EA* το χειρότερο demarcation όλης της βλάβης. Αναμένονται μεγαλύτερες τιμές τους σε σχέση με την *overall* στην ίδια εικόνα γιατί εκφράζουν το χειρότερο demarcation της βλάβης. Εκμεταλλευόμενοι τα χωρικά χαρακτηριστικά των παραμέτρων μπορεί να προσδιοριστεί σαν μελλοντική δουλειά η θέση του χειρότερου demarcation στην εικόνα. Όλες οι παράμετροι δίνουν μία πιθανή πρόβλεψη της κατεύθυνσης των (προ)-καρκινικών κυττάρων στο επιθήλιο τόσο στις low-grade όσο και στις high-grade βλάβες. Γενικότερα, όσο μεγαλύτερες οι τιμές των παραμέτρων σε κάθε εικόνα, τόσο πιο επικίνδυνη είναι η βλάβη.

Τα αποτελέσματα που ακολουθούν προέκυψαν από εικόνες ασθενών από το Νοσοκομείο Αλεξάνδρα και Hammersmith που υποβλήθηκαν στην εξέταση με το ιατρικό μηχάνημα Dysis. Αναλύονται 3 επιβεβαιωμένες και με βιοψία περιπτώσεις ασθενών: δύο υψηλού κινδύνου νεοπλασίες -high grade- (true positive) και μια χαμηλού -low grade- (true negative). Λαμβάνοντας σαν είσοδο τις προ-επεξεργασμένες εικόνες *subtracted.bmp*^(κεφάλαιο1) και τρέχοντας το software όπου υλοποιούνται οι αναφερόμενοι αλγόριθμοι και αποθηκεύοντας όλες τις παραμέτρους και εξαγόμενες εικόνες, προκύπτουν τα παρακάτω.

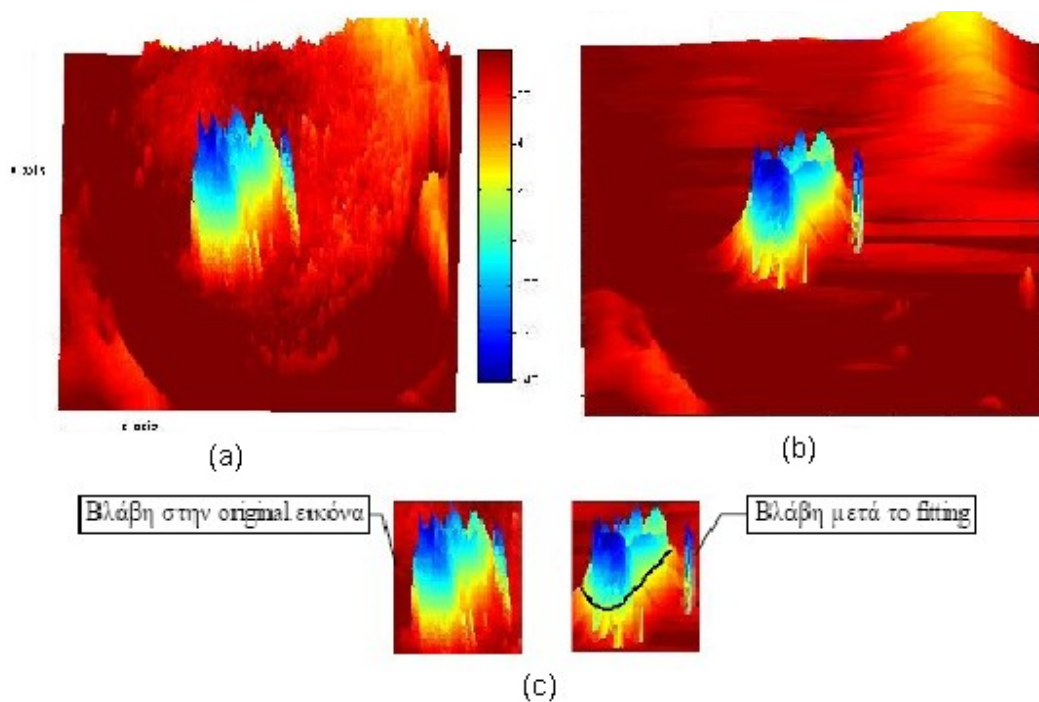
- Για την ασθενή 16_72 του νοσοκομείου Αλεξάνδρα:

Η subtracted εικόνα είναι η εικόνα 39c και προκύπτει από την αφαίρεση των εικόνων 39a και 39b θεωρώντας ότι σε κάθε pixel υπάρχει το μέγιστο AW:



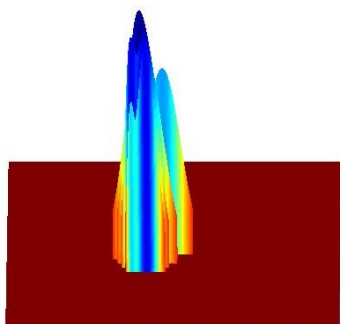
Εικόνα 39: Ασθενής 16_72 (a) Grayscale εικόνα $t=0$ (b) Εικόνα $t=\max(\text{acetowhitening})$ σε όλα τα pixel (c) subtracted εικόνα

Η 3D αναπαράσταση της subtracted εικόνας 39c που ακολουθεί (40a) δίνει μια καλύτερη οπτική αντίληψη της βλάβης στο χώρο, του demarcation και του θορύβου. Η εικόνα 40b αναπαριστά σε 3D την subtracted εικόνα ύστερα από parametric fitting με mixed Gaussian2 σε όλες τις γραμμές με σκοπό τη σύγκριση.

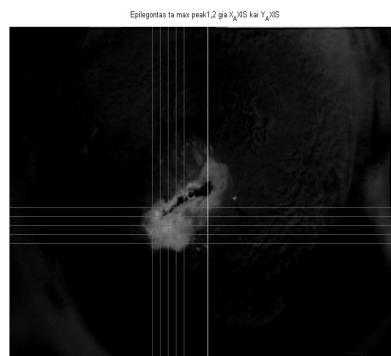


Εικόνα 40: Ασθενής 16_72 (a) 3D απεικόνιση original εικόνας (b) 3D απεικόνιση εικόνας ύστερα από fitting όλων των profile (c) zoom βλάβης

Οι άκρες της βλάβης γίνονται πιο smoothed από ότι στην original εικόνα. Για το λόγο αυτό η εξαγωγή παραμέτρων γίνεται στα fitted profile στην περιοχή που είναι πάνω από την μαύρη γραμμή της εικόνας 40c. Τα σημεία x_1, x_2 ^(ενότητα 4.3) από όπου υπολογίζονται οι κλίσεις L1, L2 βρίσκονται πάνω στην μαύρη γραμμή 'φιλτράροντας' έτσι τις smoothed άκρες των περιοχών 3σ μέχρι 2σ της fitting συνάρτησης. Σκοπός της επεξεργασίας και της μοντελοποίησης της βλάβης είναι η καλύτερη προσέγγιση της original εικόνας 40a. Σε κάθε subtracted εικόνα η μέθοδος της εργασίας 'κατασκευάζεται το μοντέλο της βλάβης' και ομοίως τα σημεία x_1, x_2 'φιλτράρουν' τις smoothed άκρες εξάγοντας τις τελικές παραμέτρους στην περιοχή του AW. Για την ασθενή 16_72, η εικόνα 41 απεικονίζει ενδεικτικά τα 5 αριστερά μοντελοποιημένα και 'φιλτραρισμένα' profile που επιλέγονται για εξαγωγή των D1αριστερά. Τα 20 επιλεγμένα profile με τα μέγιστα peak όπου υπολογίζονται οι τελικές παράμετροι του demarcation στην subtracted εικόνα (για την ασθενή 16_72) φαίνονται στην εικόνα 42 καθώς και οι τιμές των παραμέτρων στον ακόλουθο πίνακα. Η εικόνα 43a απεικονίζει το χειρότερο demarcation που αποτυπώνεται στην παράμετρο worst_case_y, 43b ερμηνεύει την κατάσταση των παθογόνων κυττάρων στο επιθήλιο, 43c γίνεται επιβεβαίωση από τον χάρτη DySIS.



Εικόνα 41: Ασθενής 16_72. 3D απεικόνιση των (5) επιλεγμένων 'φιλτραρισμένων' & μοντελοποιημένων profile

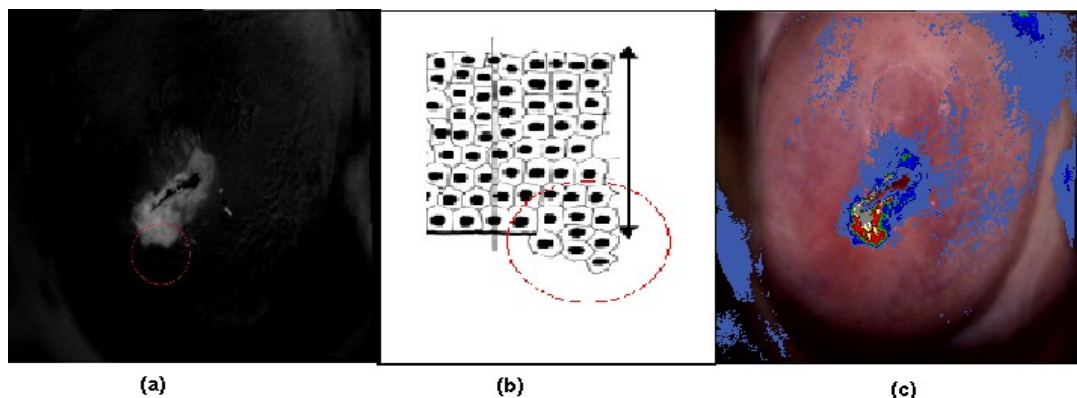


Εικόνα 42: Ασθενής 16_72. Επιλεγμένα profile εξαγωγής τελικών παραμέτρων

Αποτελέσματα

Patient_16_72:

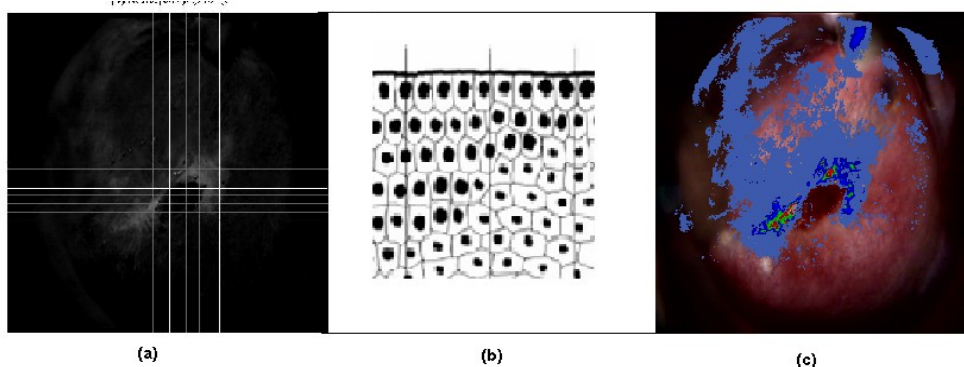
overall:	1,72
worst_case_x:	1,90
worst_case_y:	2,09



Εικόνα 43: Ασθενής_16_72 (a) Το χειρότερο demarcation της βλάβης (b) Επιθηλιακός ιστός (c) χάρτης Dysis

Οι τιμές όλων των παραμέτρων για την παραπάνω ασθενή είναι υψηλές. Η τιμή της παραμέτρου overall ίση με 1.72 αποτυπώνει αριθμητικά το απότομο demarcation του AW της συνολικής βλάβης. Το χειρότερο demarcation παρουσιάζεται κατά τον y άξονα της εικόνας με τιμή worst_case_y 2.09 όπως φαίνεται στην εικόνα 43a όπου από μαύρη περιοχή μηδενικής έντασης υπάρχει απότομη αλλαγή σε AW πολύ μεγάλης έντασης. Η υποκειμενική και περιγραφική αντίληψη του demarcation αντικαθίσταται σύμφωνα με την εργασία από την ποσοτική έκφραση 'AW με γενικό demarcation 1.72 και τοπικά 2.09 (εκφράζοντας το χειρότερο ενδεχόμενο της βλάβης)'. Σύμφωνα με αυτό, η αντίληψη του demarcation είναι πιο αντικειμενική και οι εξαγόμενες τελικές παράμετροι μπορεί να χρησιμοποιηθούν για κατηγοριοποίηση. Η πρόβλεψη για την βλάβη αυτή λόγω των μεγάλων τιμών όλων των παραμέτρων υποδηλώνει (προς το παρόν εμπειρικά) ότι τα παθογόνα κύτταρα έχουν τάση κατεύθυνσης προς το βάθος του επιθηλιακού ιστού, τείνοντας να σπάσουν την επιθηλιακή μεμβράνη προκαλώντας πιθανή μετάσταση (εικόνα 44b). Παρουσιάζεται ο χάρτης Dysis στην εικόνα 43c για επιβεβαίωση.

- Αναλύοντας μια διαφορετική περίπτωση και συγκεκριμένα της ασθενή 54_139 του νοσοκομείου Αλεξάνδρα, προκύπτουν τα ακόλουθα:



Εικόνα 44: Ασθενής 54_139 (a) Επιλογή Profile. Subtracted εικόνα (b) Επιθηλιακός ιστός (c) Χάρτης Dysis

Αποτελέσματα:

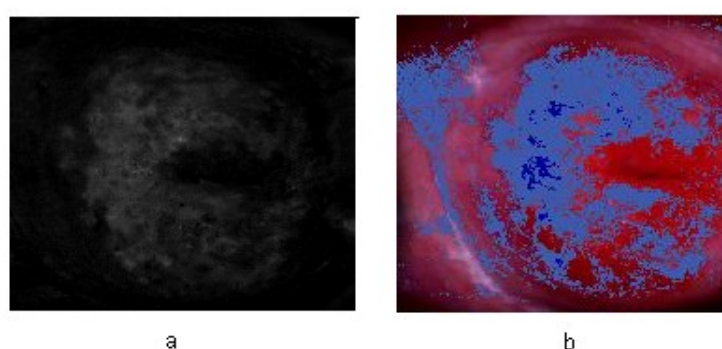
Patient_54_139:

overall:	0,86
worst_case_x:	0,88
worst_case_y:	1,40

Οι τιμές των παραμέτρων στην περίπτωση αυτή είναι μικρότερες σε σχέση με την ασθενή 16_72, αναμενόμενο σύμφωνα με την μέση ανθρώπινη οπτική αντίληψη του demarcation της βλάβης. Το AW έχει μικρότερη ένταση από την 16_72 ενώ είναι ομαλά οριοθετημένο δίνοντας γενική παράμετρο του demarcation overall ίση με 0.86 και χειρότερη περίπτωση κατά την y διάσταση ίση με 1,40 όπως αποδεικνύει και η οπτική αντίληψη της εικόνας 44a. Ο χάρτης Dysis της εικόνας 44c επιβεβαιώνει την ύπαρξη μικρότερων εντάσεων σε σχέση με την ασθενή 16_72 λόγω της έλλειψης άσπρου και εκτεταμένου κόκκινου χρώματος. Επίσης το μπλε χρώμα εδώ είναι απλωμένο σε μεγάλη έκταση γύρω από τη βλάβη. Δεδομένου ότι όσο μεγαλύτερες οι τιμές των παραμέτρων τόσο πιο επικίνδυνη η βλάβη, των αναφερομένων και της μέσης οπτικής αντίληψης των εικόνων, οι παράμετροι ανταποκρίνονται ικανοποιητικά. Η ποσοτική έκφραση του demarcation σύμφωνα με την εργασία είναι

'AW με γενικό demarcation 0.86 και τοπικά η χειρότερη περίπτωση demarcation 1.40' ενώ η πρόβλεψη λόγω των ενδιάμεσων τιμών των παραμέτρων και της μεγάλης τιμής του worst_case_y υποδηλώνει επικίνδυνη τάση εξέλιξης των (προ)-καρκινικών κυττάρων κατά μήκος του επιθηλιακού ιστού (εικόνα 44b) με υποψία εξέλιξης σε ορισμένα σημεία κατά τον y άξονα της εικόνας κατά βάθος του επιθηλίου.

- Για true-negative περιπτώσεις και συγκεκριμένα της ασθενούς 114_250 του Αλεξάνδρα προκύπτουν τα ακόλουθα:



Εικόνα 45: Ασθενής 114_250 (a) True negative subtracted εικόνα (b) Χάρτης Dysia

Αποτελέσματα:

Patient_114_250:

overall:	0,52
worst_case_x:	0,51
worst_case_y:	0,76

Οι τιμές όλων των παραμέτρων της ασθενούς 114_250 είναι μικρές. Το demarcation είναι μικρό σύμφωνα με την μέση οπτική αντίληψη. Λόγω των μικρών τιμών η βλάβη μπορεί να χαρακτηριστεί ως low-grade γεγονός που αποδεικνύεται και από τον χάρτη (εικόνα 45b). Η ποσοτική έκφραση του demarcation είναι 'AW με demarcation 0.52 και στη χειρότερη περίπτωση 0.76' και με πρόβλεψη εξέλιξης των παθογόνων κυττάρων κατά πλάτος του ιστού.

Παρατηρείται ότι οι περισσότερες true negative εικόνες δίνουν τιμές παραμέτρων μικρότερες του 0,80 ενώ οι τιμές στις true positive συνήθως είναι μεγαλύτερες του 1. Ακολουθείται στατιστική ανάλυση των αποτελεσμάτων για

καλύτερη ερμηνεία. Επειδή αναζητούμε το χειρότερο demarcation της εκάστοτε βλάβης η παράμετρος EA είναι αυτή που χρησιμοποιείται συχνότερα.

5.3 Αποτελέσματα – Στατιστική Ανάλυση Αποτελεσμάτων

Επιλέγονται 100 ασθενείς (40 True Positive και 60 True Negative) από το νοσοκομείο Αλεξάνδρα και Hammersmith που υποβλήθηκαν στην εξέταση αρχικά με το DySIS και στη συνέχεια σε βιοψία. Για κάθε subtracted εικόνα για τις 100 ασθενείς εξάγονται οι ακόλουθες τιμές τελικών παραμέτρων του demarcation. Αρχικά παρουσιάζονται οι ασθενείς με χαμηλού κινδύνου νεοπλασίες (low grade):

True_Negative/ PATIENT No.	overall	worst_case_x	worst_case_y	EA
59_153	0,72	0,88	0,77	0,88
90_219	0,77	0,59	0,66	0,66
93_224	0,66	0,61	0,43	0,61
174_325	0,68	0,48	1,04	1,04
169_317	0,66	0,90	0,57	0,90
167_312	0,82	1,08	0,92	1,08
164_308	0,76	1,00	0,70	1,00
158_301	0,81	0,99	0,89	0,99
121_257	1,15	1,20	1,20	1,20
108_243	0,65	0,50	1,10	1,10
107_242	0,56	0,79	0,47	0,79
14_70	0,97	1,00	1,20	1,20
35_117	0,50	0,63	0,47	0,63
39_121	0,37	0,47	0,39	0,47
99_231	1,10	0,65	1,80	1,80
166_310	0,64	1,10	0,63	1,10
9_65	0,80	1,00	1,30	1,30
126_262	1,00	1,30	0,98	1,30
178_333	0,83	1,00	0,90	1,00
119_255	0,78	1,00	0,72	1,00
112_248	0,94	0,90	1,20	1,20
123_259	0,71	1,00	0,58	1,00
28_107	0,48	0,50	0,60	0,60

<i>10_66</i>	0,33	0,47	0,28	0,47
<i>14_70</i>	0,90	1,00	1,30	1,30
<i>28_107</i>	0,40	0,50	0,50	0,50
<i>31_112</i>	1,20	1,40	1,30	1,40
<i>36_118</i>	0,50	0,51	0,63	0,63
<i>40_122</i>	1,00	1,00	1,20	1,20
<i>45_128</i>	0,55	0,77	0,44	0,77
<i>47_130</i>	0,80	0,90	1,00	1,00
<i>56_150</i>	0,40	0,50	0,60	0,60
<i>68_167</i>	1,00	0,90	1,40	1,40
<i>69_168</i>	0,48	0,46	0,70	0,70
<i>71_170</i>	0,71	0,70	0,70	0,70
<i>72_172</i>	0,60	0,61	0,83	0,83
<i>73_191</i>	0,83	1,00	1,10	1,10
<i>190_316</i>	0,52	0,48	0,64	0,64
<i>182_306</i>	0,91	1,09	0,91	1,09
<i>162_269</i>	0,68	1,04	0,88	1,04
<i>28_90</i>	0,55	0,66	0,83	0,83
<i>29_91</i>	1,42	1,04	1,93	1,93
<i>30_92</i>	0,70	0,50	1,00	1,00
<i>32_94</i>	0,59	0,68	0,64	0,68
<i>33_95</i>	0,56	0,50	0,73	0,73
<i>36_98</i>	0,71	0,68	0,71	0,71
<i>39_101</i>	0,81	1,00	0,90	1,00
<i>41_104</i>	1,09	0,91	1,46	1,46
<i>42_105</i>	0,72	0,79	0,71	0,79
<i>43_106</i>	0,64	0,64	0,99	0,99
<i>52_115</i>	0,66	0,54	0,94	0,94
<i>54_117</i>	0,72	1,06	0,55	1,06
<i>60_137</i>	0,78	0,78	1,00	1,00
<i>61_138</i>	0,56	0,56	0,59	0,59
<i>62_139</i>	0,96	0,98	1,00	1,00
<i>65_142</i>	0,75	0,84	0,80	0,84
<i>66_143</i>	0,59	0,66	0,81	0,81
<i>70_147</i>	0,80	0,83	1,00	1,00
<i>74_151</i>	0,63	0,90	0,59	0,90
<i>80_158</i>	1,12	0,96	1,58	1,58

Οι low grade περιπτώσεις που έχουν παραμέτρους κοντά και μεγαλύτερες του 1 στις περισσότερες περιπτώσεις έχουν έντονο demarcation και χάρτη κόκκινου χρώματος. Τα παραπάνω αποτελέσματα χωρίς την οπτική παρατήρηση των εικόνων είναι εν μέρη 'ανούσια' από ιατρικής πλευράς αλλά χρησιμοποιούνται για στατιστική προσέγγιση. Στο παράρτημα Α παρουσιάζονται οι σχετικές εικόνες από ορισμένους εκ των ασθενών. Για τις ασθενείς με υψηλού κινδύνου νεοπλασίες (high grade) προέκυψαν οι παράμετροι του demarcation:

True Positive / PATIENT No.	overall	worst_case_x	worst_case_y	EA
<i>16_72</i>	1,72	1,90	2,09	2,09
<i>54_139</i>	0,86	0,88	1,40	1,40
<i>57_151</i>	1,40	1,59	1,71	1,71
<i>80_199</i>	1,03	1,24	1,19	1,24
<i>84_208</i>	1,25	1,54	1,34	1,54
<i>85_210</i>	0,65	0,73	0,71	0,73
<i>136_274</i>	1,07	1,17	1,03	1,17
<i>13_69</i>	1,06	1,20	1,20	1,20
<i>63_160</i>	1,62	1,17	2,80	2,80
<i>120_256</i>	0,86	0,93	1,12	1,12
<i>122_258</i>	1,32	1,07	1,81	1,81
<i>66_164</i>	1,06	1,38	1,01	1,38
<i>94_225</i>	0,97	0,96	1,19	1,19
<i>118_254</i>	1,01	1,22	1,37	1,37
<i>176_329</i>	0,60	0,77	0,69	0,77
<i>34_86</i>	0,92	0,59	1,69	1,69
<i>35_87</i>	1,36	1,69	1,59	1,69
<i>44_98</i>	1,10	1,30	1,33	1,33
<i>44_103</i>	1,05	1,22	1,29	1,29
<i>52_110</i>	1,09	1,19	1,24	1,24
<i>69_131</i>	1,43	1,53	1,95	1,95
<i>70_132</i>	1,01	1,11	0,99	1,11
<i>72_134</i>	1,26	1,21	1,78	1,78
<i>76_139</i>	1,63	1,75	2,21	2,21
<i>77_140</i>	0,84	0,94	0,84	0,94
<i>78_141</i>	0,96	1,00	1,40	1,40

<i>81_144</i>	1,49	1,22	2,29	2,29
<i>124_213</i>	0,96	1,06	1,14	1,14
<i>140_236</i>	1,29	1,20	1,85	1,85
<i>154_259</i>	0,67	0,73	0,81	0,81
<i>180_304</i>	0,95	1,09	0,93	1,09
<i>38_100</i>	1,08	0,99	1,55	1,55
<i>88_168</i>	1,08	1,20	1,32	1,32
<i>143_248</i>	1,35	1,59	1,31	1,59
<i>155_261</i>	1,28	1,36	1,45	1,45
<i>184_308</i>	1,07	1,00	1,15	1,15
<i>185_309</i>	1,21	1,42	1,29	1,42
<i>187_312</i>	1,57	1,88	2,85	2,85
<i>188_313</i>	1,22	1,73	1,34	1,73
<i>111_197</i>	0,90	1,32	1,19	1,32

- Στατιστική ανάλυση αποτελεσμάτων

Η ανάλυση γίνεται για την κατανόηση των τιμών των παραμέτρων σε γνωστές περιπτώσεις ασθενών. Σύμφωνα με τους παραπάνω πίνακες, οι μέσες τιμές μ και τυπικές αποκλίσεις σ των παραμέτρων overall, worst_case_x, worst_case_y, EA στις high και low grade βλάβες είναι:

<i>High grade</i>	<i>overall</i>	<i>worst_case_x</i>	<i>worst_case_y</i>	<i>EA</i>
<i>mean(μ)</i>	<i>1,13</i>	<i>1,22</i>	<i>1,44</i>	<i>1,49</i>
<i>STDEV(σ)</i>	<i>0,26</i>	<i>0,31</i>	<i>0,50</i>	<i>0,48</i>
<i>Low grade</i>	<i>overall</i>	<i>worst_case_x</i>	<i>worst_case_y</i>	<i>EA</i>
<i>mean(μ)</i>	<i>0,74</i>	<i>0,81</i>	<i>0,88</i>	<i>0,97</i>
<i>STDEV(σ)</i>	<i>0,22</i>	<i>0,24</i>	<i>0,34</i>	<i>0,30</i>

Οι μέσες τιμές των low grade περιπτώσεων όλων των παραμέτρων είναι μικρότερες σε σχέση με τις high grade. Για να αποφασιστεί αν είναι στατιστικά σημαντική διαφορά πραγματοποιείται T-test μεταξύ των low(true positive) και high(true negative) grade νεοπλασιών για τις τελικές παραμέτρους overall και EA που είναι

και οι δύο κύριες παράμετροι ποσοτικοποίησης. Άλλωστε (EA= max(worst_case_x,worst_case_y)).

- T-test

Πραγματοποιείται Two Samples T-test το οποίο χρησιμοποιείται όταν εξετάζουμε τη διαφορά μεταξύ δύο ομάδων και επιθυμούμε να γνωρίζουμε αν αυτή είναι στατιστικώς σημαντική (significantly different) ή αν προέκυψε τυχαία. Το τεστ γίνεται με βάση έναν βαθμό ελευθερίας (degree of freedom) ο οποίος εξαρτάται από τον αριθμό του πληθυσμού του δείγματος και χρησιμοποιείται για να καθοριστεί η critical value στον πίνακα Student's t-Distribution. Η critical value είναι η τιμή πάνω από την οποία αποφασίζεται ότι το τεστ απέρριψε την μηδενική υπόθεση (null hypothesis). Τα βήματα του τεστ είναι:

1)Καθορισμός της μηδενικής υπόθεσης (null hypothesis) H_0 η οποία ορίζει μία αρχική σχέση μεταξύ των μέσων τιμών. Καθορισμός της υπόθεσης που αναζητείται $H_1 : \mu_1 \neq \mu_2$.

2)Καθορισμός του επιπέδου σημαντικότητας (significance level) alpha: καθορίζει την πιθανότητα τα δεδομένα να προέρχονται από την ίδια ομάδα με ίδια μέση τιμή. Καθορισμός της critical value και του degree of freedom df ($df = n_1 + n_2 - 2$) σύμφωνα με τον πίνακα Student's t-Distribution.

3)Υπολογισμός των στατιστικών παραμέτρων t του τεστ σύμφωνα με τα δείγματα n_1, n_2 των δύο ομάδων. Για την περίπτωση του Two Samples T-test ισχύει

$$t = \frac{mean_1 - mean_2}{sX - X} \quad \text{όπου} \quad sX - X = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}.$$

Απαραίτητη η γνώση των μέσων τιμών $mean_1, mean_2$ των δύο ομάδων, των διασπορών τους s_1^2, s_2^2 και του αριθμού των δειγμάτων n_1, n_2 για κάθε ομάδα.

4) Απόφαση απόρριψης της μηδενικής υπόθεσης H_0 σύμφωνα με την τιμή της t και την critical value σύμφωνα με τον πίνακα Student's t-Distribution και το βήμα 2. Αν $t > \text{critical value}$ απορρίπτεται η H_0 και γίνεται δεκτή η H_1 .

Το T-test για τις παραμέτρους overall και EA για τις δύο ομάδες low και high grade με $H_0: \mu_1 - \mu_2 = 0$ και $H_1: \mu_1 \neq \mu_2$ έδωσε:

parameter:overall High grade(group1) Low grade(group2)

<i>mean</i>	<i>1,13</i>	<i>0,74</i>
<i>σ / s^2</i>	<i>0,26 / 0,0676</i>	<i>0,22 / 0,0484</i>
<i>n</i>	<i>40</i>	<i>60</i>
<i>Alpha Level</i>	<i>0.001</i>	
<i>df</i>	<i>98</i>	
<i>t</i>	<i>7.8</i>	
<i>Διαχωρίσιμα?</i>	<i>NAI</i>	

Επειδή $t = 7.8$, τιμή μεγαλύτερη από 3.390 σύμφωνα με τον πίνακα Student's t-Distribution για $df=98$ και $\alpha=0.001$, εξάγεται το συμπέρασμα ότι για την παράμετρο overall η διαφορά των μέσων τιμών των high grade (group1) και low grade (group2) είναι στατιστικώς σημαντική. Για την παράμετρο EA:

parameter:EA High grade(group1) Low grade(group2)

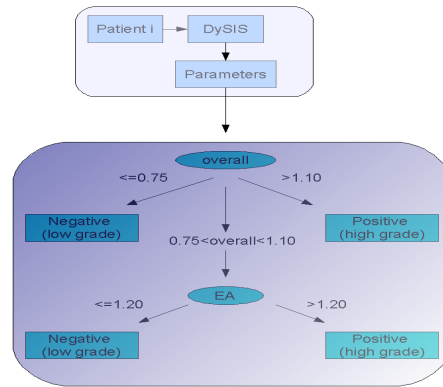
<i>mean</i>	<i>1,49</i>	<i>0,97</i>
<i>σ / s^2</i>	<i>0,48 / 0,2304</i>	<i>0,30 / 0,09</i>
<i>n</i>	<i>40</i>	<i>60</i>
<i>Alpha Level</i>	<i>0.001</i>	
<i>df</i>	<i>98</i>	
<i>t</i>	<i>6.1</i>	
<i>Διαχωρίσιμα?</i>	<i>NAI</i>	

Ομοίως, για την παράμετρο EA η διαφορά των μέσων τιμών των high και των low grade νεοπλασιών είναι στατιστικώς σημαντική (με πολύ καλή ακρίβεια $\alpha=0.001$) επειδή $t = 6.1$ μεγαλύτερο του 3.390.

Γενικό συμπέρασμα: οι low και high grade περιπτώσεις νεοπλασιών είναι διαχωρίσιμες σύμφωνα με τις τιμές των παραμέτρων overall και EA.

- Ευριστικό σύστημα

Στόχος της εργασίας δεν είναι η κατασκευή ταξινομητή που κατηγοριοποιεί low και high grade βλάβες σύμφωνα με το demarcation αφού είναι μόνο ένα από τα χαρακτηριστικά παρατήρησης για τη διάγνωση. Προτείνεται ένα διαισθητικό σύστημα για να εκτιμηθεί το ρίσκο μετεξέλιξης νεοπλασμάτων σε καρκίνο καθώς και η αποδοτικότητα της μεθόδου της εργασίας. Χρησιμοποιούνται οι μισές εικόνες (30 true negative, 20 true positive) όπου η μέση τιμή της παραμέτρου overall των low grade είναι 0,72 ενώ των high grade 1,16 και της EA 0,97 και 1,54 αντίστοιχα. Το σύστημα αποφασίζει low grade για overall μικρότερες ή ίσες του 0,75 και high-grade για overall μεγαλύτερες του 1,10. Σε περιπτώσεις που η overall παράμετρος παίρνει τιμές στη ζώνη μεταξύ 0,75 και 1,10 επειδή το overall demarcation στο διάστημα αυτό θεωρείται 'μεσσαιό', ο έλεγχος γίνεται με βάση το χειρότερο demarcation της βλάβης άρα από την τιμή της παραμέτρου EA. Αν EA παίρνει τιμές μικρότερες ή ίσες του 1,2 αποφασίζεται low-grade βλάβη ενώ αν παίρνει τιμές μεγαλύτερες του 1,2 αποφασίζεται high-grade ($\text{mean}(0.97, 1.54) \approx 1.2$). Γίνεται αρχικά εκτίμηση με βάση το γενικό (overall) demarcation της βλάβης εκτιμώντας ότι έντονο overall demarcation ($>1,10$) υποδηλώνει high grade βλάβη ενώ άτονο ($<0,75$) low grade. Η ενδιάμεση κατάσταση του overall demarcation καθορίζεται από την χειρότερη περίπτωση του demarcation (EA) σύμφωνα με την οποία αποφασίζεται η επικινδυνότητα της βλάβης. Το δέντρο απόφασης του συστήματος είναι:



Εικόνα 46: Δέντρο κατηγοριοποίησης

Σε σύνολο 100 ασθενών (40 true positive και 60 true negative γνωστές και επιβεβαιωμένες με βιοψία) το σύστημα αυτό έδωσε : 11 false negative και 11 false positive. Υπολογίζοντας την ειδικότητα (specificity) και την ευαισθησία (sensitivity) προέκυψε:

$$sensitivity = \frac{\# \text{ατόμων με θετικό τεστ και νόσο}}{\# \text{ατόμων με νόσο}} = \frac{29}{40} = 0,725$$

$$specificity = \frac{\# \text{ατόμων με αρνητικό τεστ και χωρίς νόσο}}{\# \text{ατόμων χωρίς νόσο}} = \frac{49}{60} \approx 0.82$$

Η ευαισθησία ενός διαγνωστικού τεστ είναι η πιθανότητα το τεστ να είναι θετικό δεδομένου ότι ο ασθενής έχει τη νόσο που εξετάζουμε και η ειδικότητα είναι η πιθανότητα το διαγνωστικού τεστ να είναι αρνητικό δεδομένου ότι ο ασθενής δεν έχει τη νόσο που εξετάζουμε. Για να είναι αξιόπιστο το διαγνωστικό τεστ πρέπει οι δύο αυτές πιθανότητες να είναι υψηλές και να προσεγγίζουν το 1 [7]. Σε περίπτωση που αποδειχτεί κλινικά η σημασία του demarcation και ο συσχετισμός του με την κατεύθυνση των παθογόνων κυττάρων, οι εξαγόμενοι παράμετροι της εργασίας έχουν αποδείξει ότι μπορεί να γίνει κατηγοριοποίηση με ικανοποιητικά αποτελέσματα μεταξύ low και high grade αλλοιώσεων.

Μπορούν να κατασκευαστούν πολλά εναλλακτικά συστήματα απόφασης. Το κόστος είναι ένας παράγοντας που πρέπει να μελετηθεί δεδομένου ότι μια high grade

βλάβη που δεν θα εντοπιστεί θεωρείται ότι έχει μεγαλύτερο κόστος στην υγεία του ασθενή. Από την άλλη πλευρά οι κοινωνικές, ψυχολογικές και οικονομικές επιπτώσεις που μπορεί να έχει μια λανθασμένη high grade διάγνωση σε υγιή γυναίκα είναι σημαντικές και πρέπει να μελετηθούν.

5.4 Συμπεράσματα

Στα προηγούμενα έχουν εξαχθεί συμπεράσματα για όσα αφορούν την εργασία αυτή. Συνοψίζοντας αναφέρεται ότι: έχουν βρεθεί για πρώτη φορά ποσοτικές οπτικές παράμετροι που σχετίζονται με το demarcation, οι παράμετροι της εργασίας ανταποκρίνονται ικανοποιητικά στην μέση οπτική αντίληψη του demarcation. Οι αναμενόμενες τιμές των παραμέτρων επαληθεύονται, δίνοντας σε βλάβες με έντονο demarcation μεγαλύτερες τιμές σε σχέση με όσες έχουν άτονο demarcation. Το κύριο εξαγόμενο συμπέρασμα της εργασίας για την *διαχωριστικότητα (διάκριση) μεταξύ low και high grade νεοπλασιών σύμφωνα με τις παραμέτρους του demarcation (overall,EA)* ανοίγει σειρά ερευνητικών διεργασιών. Πιθανή μελλοντική κλινική απόδειξη της σχέσης του demarcation με την εκτίμηση του ρίσκου μετεξέλιξης νεοπλασμάτων σε καρκίνο μπορεί να οδηγήσει σε ιατρικές καινοτομίες. Η εργασία παρέχει την πρώτη προσέγγιση της διαδικασίας. Επιπρόσθετα, η παραμετροποίηση του demarcation έχει επιτευχθεί και είναι 'αλγοριθμικώς αντικειμενική' για κάθε βλάβη, με την έννοια ότι πάντα θα δίνει την ίδια τιμή χωρίς να εξαρτάται από την υποκειμενική οπτική αντίληψη του παρατηρητή-ιατρού (συχνά διαφωνεί και με την αρχική του εκτίμηση). Οι παράμετροι overall και EA για κάθε βλάβη μπορούν να λειτουργήσουν σαν επιπρόσθετο αριθμητικό διαγνωστικό κριτήριο της άμεσης οπτικής επιθεώρησης του AW για την ανίχνευση των αλλοιώσεων του τραχήλου της μήτρας, αφού ποσοτικοποιούν μια οπτική παράμετρο διάγνωσης. Μπορεί να προστεθεί μια αντικειμενική, αριθμητική παράμετρος που περιγράφει το demarcation στις κατευθυντήριες γραμμές διάγνωσης (RCI).

Τεχνικές Βελτίωσης – Μελλοντική Έρευνα

Η εργασία αυτή παρέχει μια πρώτη μέτρηση του demarcation των εικόνων των ασθενών. Τεχνικές βελτίωσης και μελλοντικής έρευνας θεωρούμαι ότι θα δώσουν ικανοποιητικότερα αποτελέσματα και πιθανόν αποτελέσουν μέρος ελέγχου αλλοιώσεων (CIN) σε real-time ιατρικές εφαρμογές. Η αποδοτικότητα της μεθόδου παραμετροποίησης του demarcation σε σχέση με την άμεση οπτική επιθεώρηση απαιτεί μακροχρόνιους ιατρικούς ελέγχους. Θεωρώντας ότι η αποδοτικότητα αποδειχτεί πιο αξιόπιστη, η ενσωμάτωση της μεθόδου σε ολοκληρωμένο κύκλωμα σε υψηλής ανάλυσης ιατρικά μηχανήματα που χρησιμοποιούν ψηφιακές μεθόδους απεικόνισης μπορεί να οδηγήσει σε άμεση παραμετροποίηση του demarcation της βλάβης και να συμπληρώνει τη διάγνωση σύμφωνα με όσα αναφέρθηκαν.

Μελλοντικές έρευνες αποσκοπούν σε:

- i) Κλινικές δοκιμές για την επιβεβαίωση της προγνωστικής αξίας του demarcation στην εκτίμηση του κινδύνου μετάστασης
- ii) Εκτίμηση της δυνατότητας αύξηση της διαγνωστικής πληροφορίας με τον συνδυασμό δυναμικών οπτικών παραμέτρων και demarcation
- iii) Μελέτη από ιατρικής πλευράς των διακυμάνσεων των εντάσεων των unfitted profile συσχετίζοντας τις διακυμάνσεις αυτές με την μορφολογία των παθογόνων κυττάρων στο επιθήλιο

Επιπροσθέτως, αναφέρεται ότι οι subtracted εικόνες που είναι σαν είσοδοι στο σύστημα θεωρείται ότι δίνουν το μέγιστο AW τη χρονική στιγμή t σε όλα τα pixel της εικόνας. Το ζήτημα αυτό είναι υπό μελέτη στο Πολυτεχνείο Κρήτης διότι ιδανικά θεωρώντας ότι κάθε pixel αναφέρεται σε ένα κύτταρο του επιθηλίου η χρονική στιγμή t που συμβαίνει το μέγιστο AW μπορεί να μην είναι ίδια για όλα τα σημεία ή κύτταρα του επιθηλίου. Υπάρχει το ενδεχόμενο κάποια σημεία του επιθηλίου να

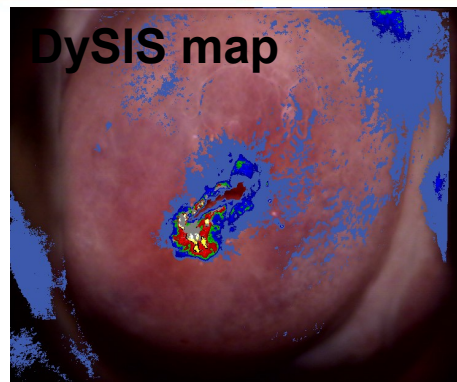
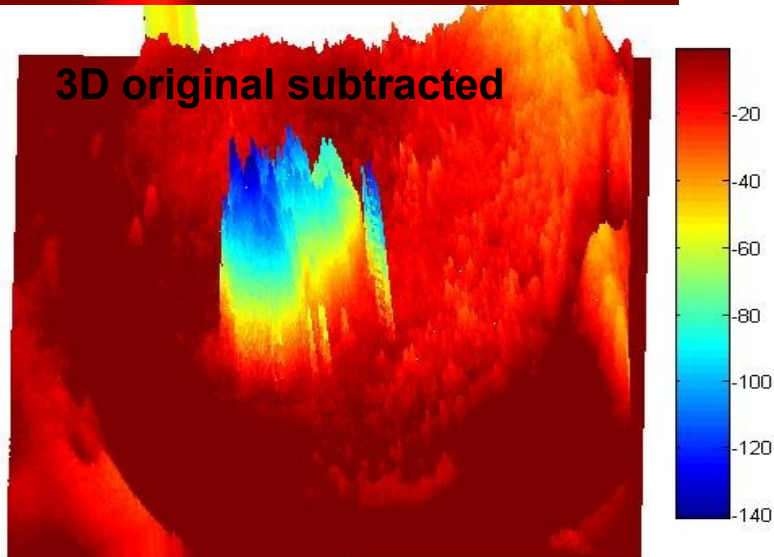
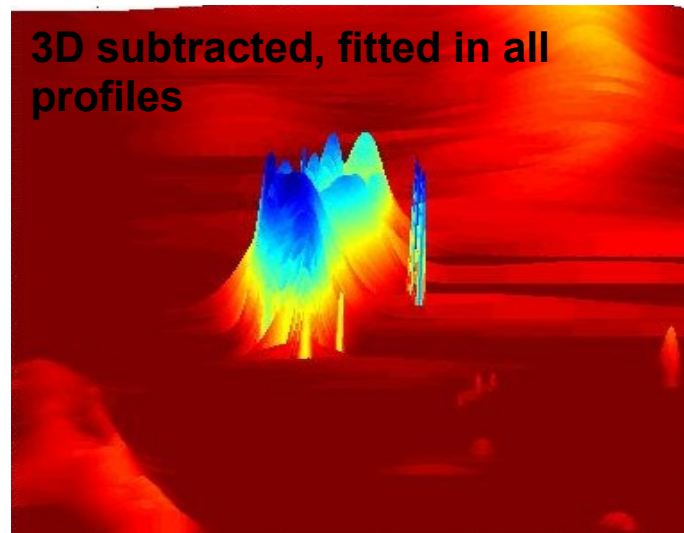
δώσουν το μέγιστο AW τη χρονική στιγμή $t' > t$ (ή $t' < t$) διότι το οξικό οξύ δεν έχει απορροφηθεί σε όλο το βάθος του επιθηλίου όπου υπάρχουν παθογόνα κύτταρα τη χρονική στιγμή t . Στόχος είναι η 'κατασκευή' της subtracted εικόνας από τις δυναμικές παραμέτρους που περιέχει σε κάθε pixel το μέγιστο AW. Τέλος, αναφέρονται ως τεχνικές βελτίωσης η κατάλληλη επεξεργασία εικόνας ή εφαρμογής τεχνικών μηχανικής όρασης ώστε να εντοπίζεται το AW χωρίς να αλλοιώνεται η πληροφορία του και να απαλείφεται ο θόρυβος από όργανα εξέτασης στην υπόλοιπη εικόνα. Ο πειραματισμός με αλλαγές των ποσοστών-βαρών της κλίσης και του peak (30%-70%, 40%-60%...) των παραμέτρων D1, D2 έχει δοκιμαστεί στην εργασία αυτή αλλά για την αποδοχή του αντικειμενικότερου ποσοστού είναι απαραίτητα ιατρικά, εργαστηριακά και χρονοβόρα πειράματα που ξεφεύγουν τα πλαίσια μιας διπλωματικής εργασίας. Θεωρούμε ότι το demarcation εξαρτάται εξίσου από την κλίση και την peak τιμή τα οποία εκφράζουν διαφορετικά μεγέθη αντίληψής του (ποσοστό 50%).

Στο παράρτημα Α παρουσιάζεται η οπτική αντίληψη των παραμέτρων του demarcation για 21 ασθενείς (15 True Positive και 6 True Negative) επιβεβαιωμένες με βιοψία. Γίνεται 3D αναπαράσταση των subtracted εικόνων για καλύτερη οπτική αντίληψη της βλάβης. Για κάθε ασθενή παρουσιάζονται οι εικόνες:

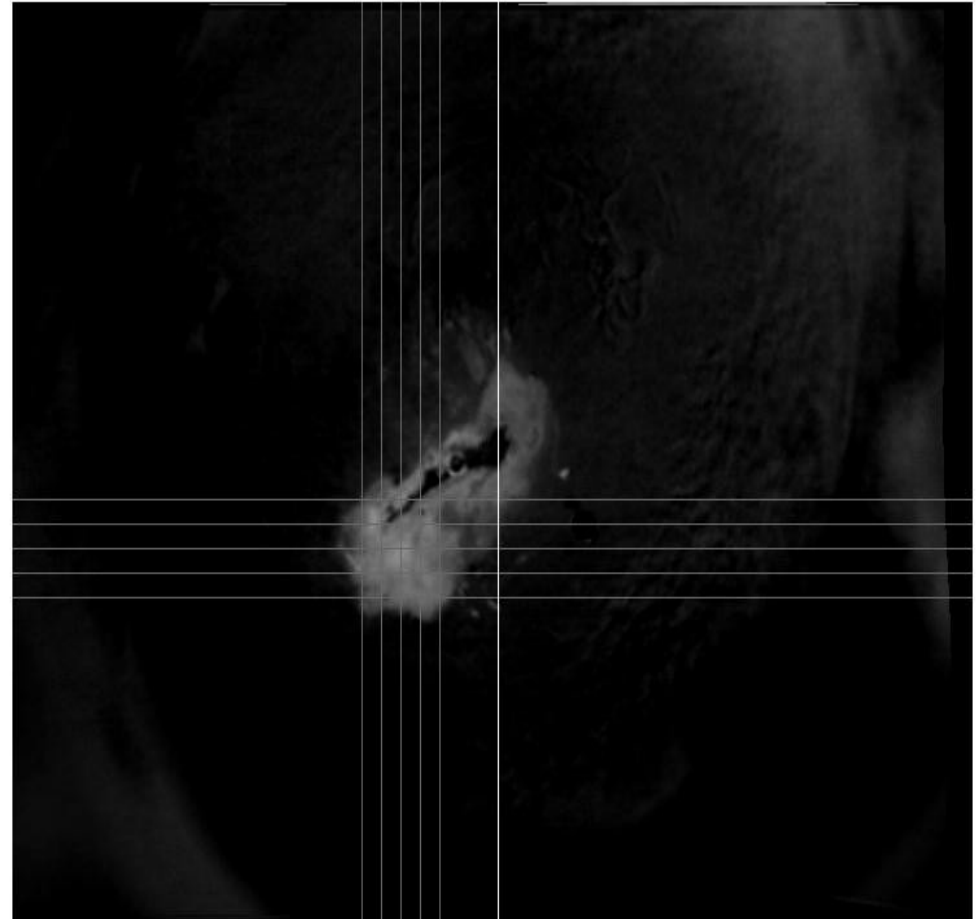
- i) 3D subtracted εικόνα ύστερα από fitting με mix gauss2 σε όλες τις γραμμές της με σκοπό την σύγκριση και την εκτίμηση του μοντέλου fitting (πάνω αριστερά)
- ii) Η unfitted, original subtracted εικόνα (αριστερά)
- iii) Subtracted εικόνα με τα profile (γραμμές και στήλες) που προσδιορίζουν το AW και χρησιμοποιούνται για εξαγωγή τελικών παραμέτρων overall, worst_case_x, worst_case_y. ($EA = \max(\text{worst_case_x}, \text{worst_case_y})$)
- iv) Ο χάρτης του Dysis (κάτω αριστερά)

Στις τελευταίες 4 περιπτώσεις ασθενών true negative δεν παρουσιάζεται η 3D αναπαράσταση διότι οι εντάσεις έχουν μικρές τιμές. Όλες οι εικόνες είναι ευθυγραμμισμένες δηλ. κοιτάζονται από την ίδια γωνία παρατήρησης.

Patient_16_72:



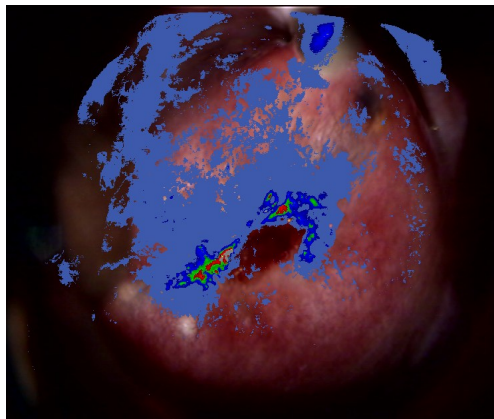
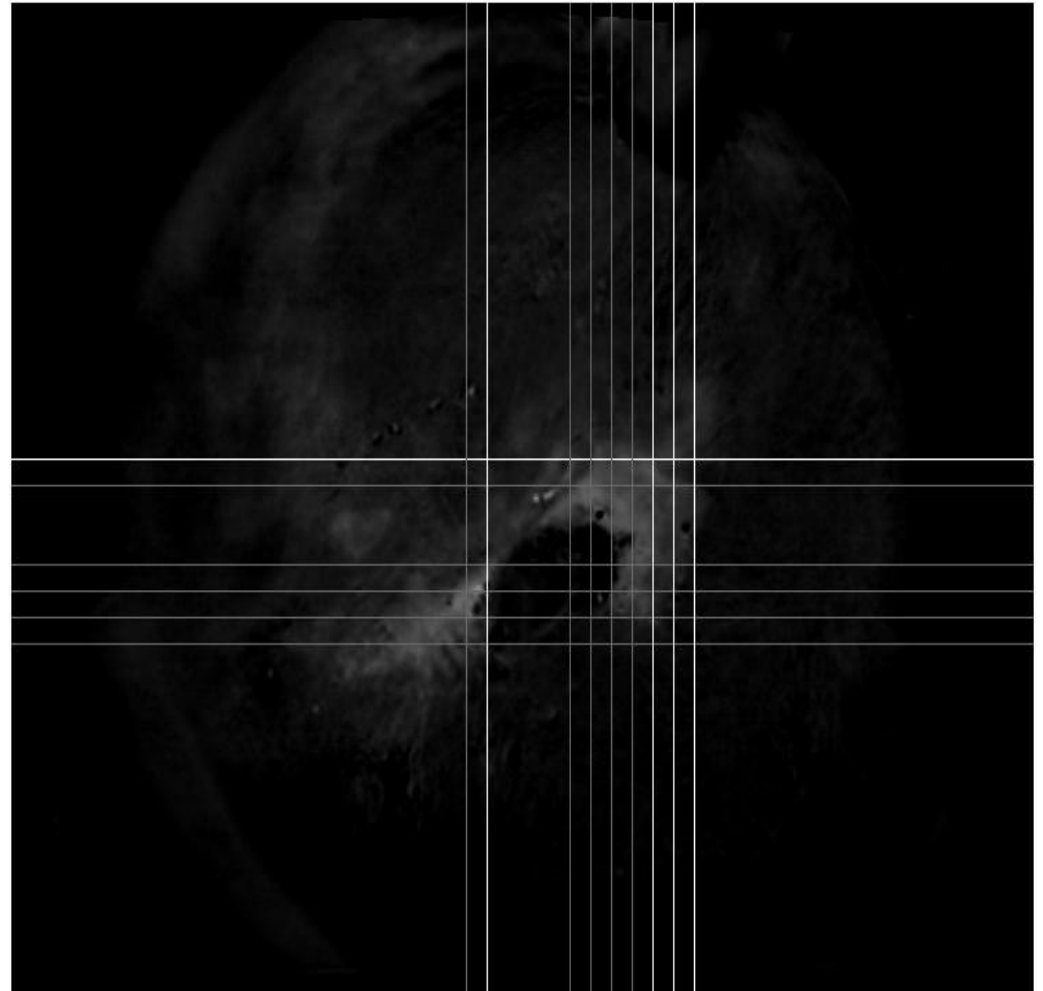
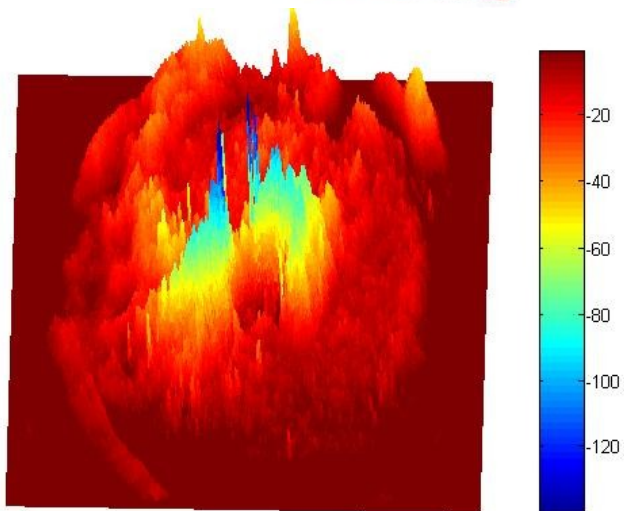
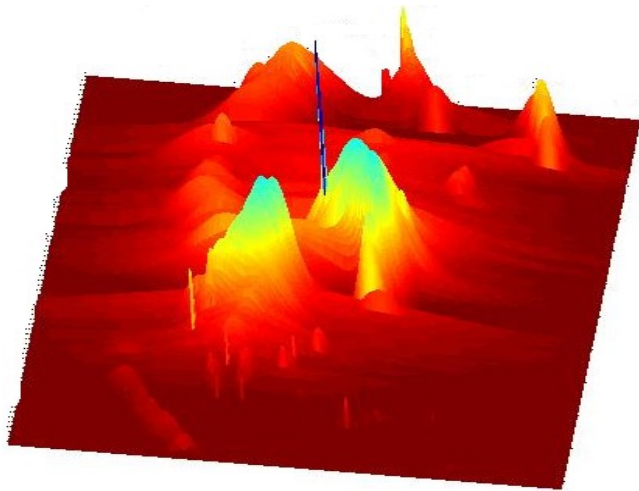
Epilegontas ta max peak1,2 gia X_A XIS kai Y_A XIS **Subtracted**



Patient_16_72:
overall=1.7
worst_case_x=1.9
worst_case_y=2.09

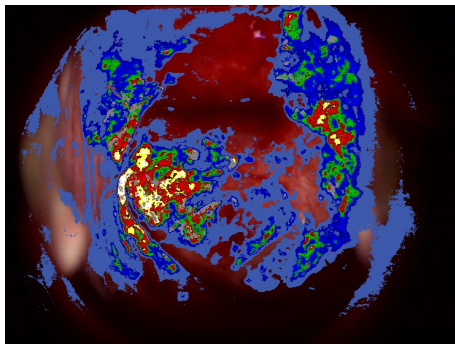
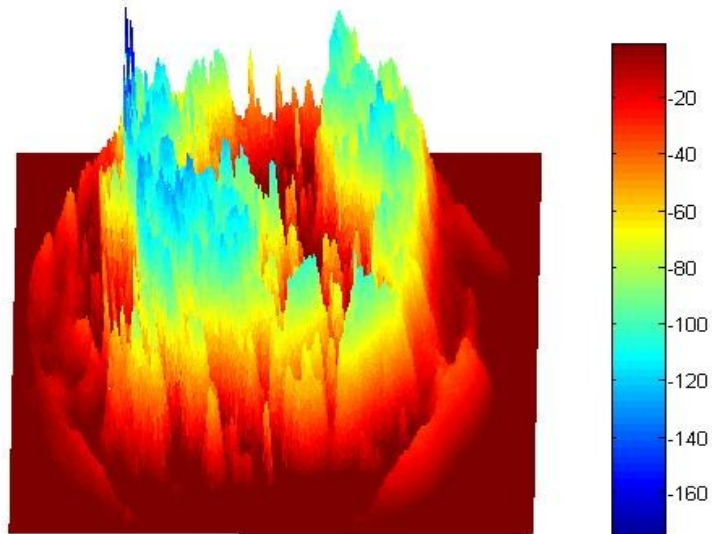
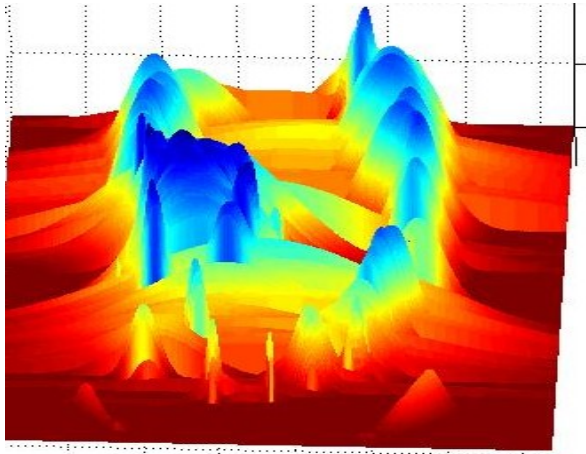
Patient_54_139:

Epilėgontas ta max peak1,2 gja X_A XIS kai Y_A XIS

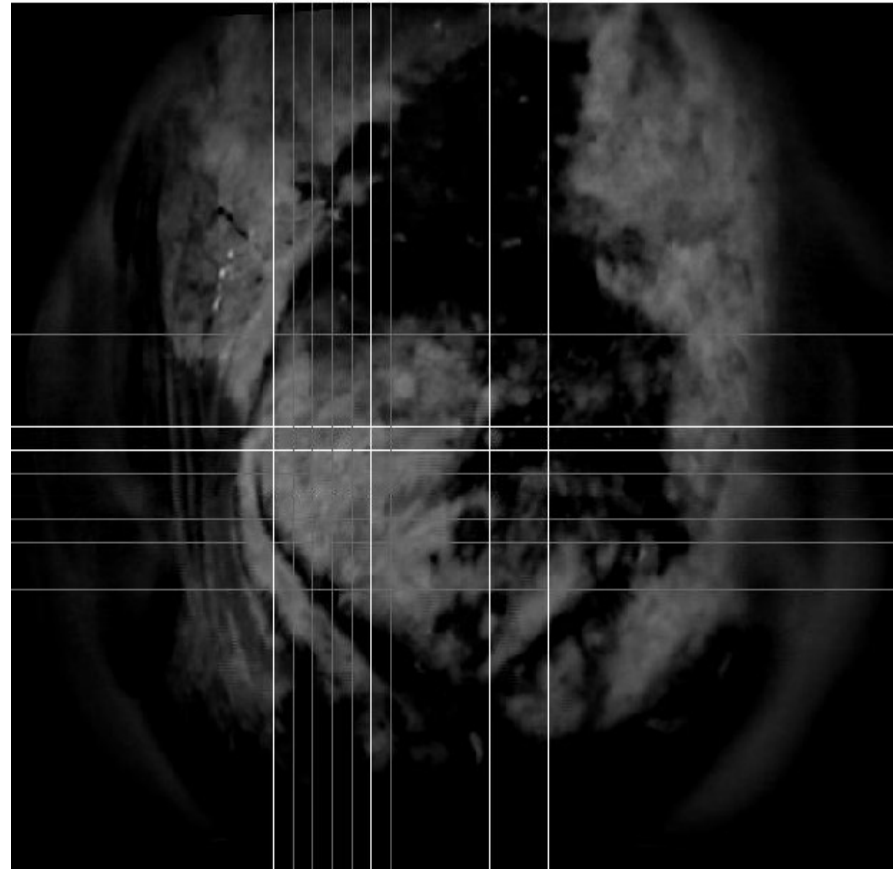


Patient_54_139:
overall=0.86
worst_case_x=0.88
worst_case_y=1.4

Patient_57_151



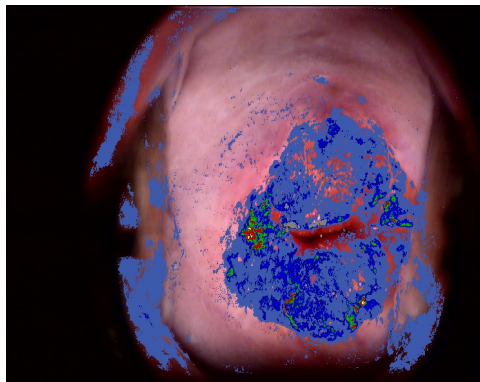
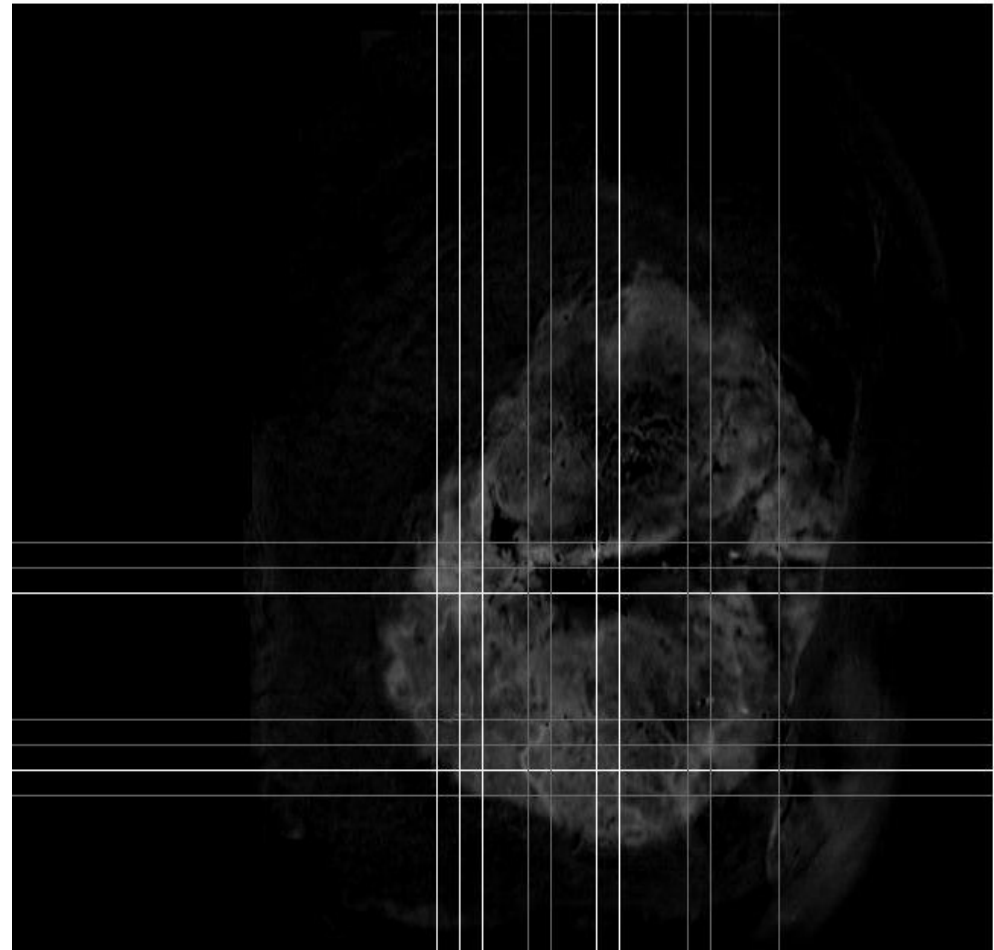
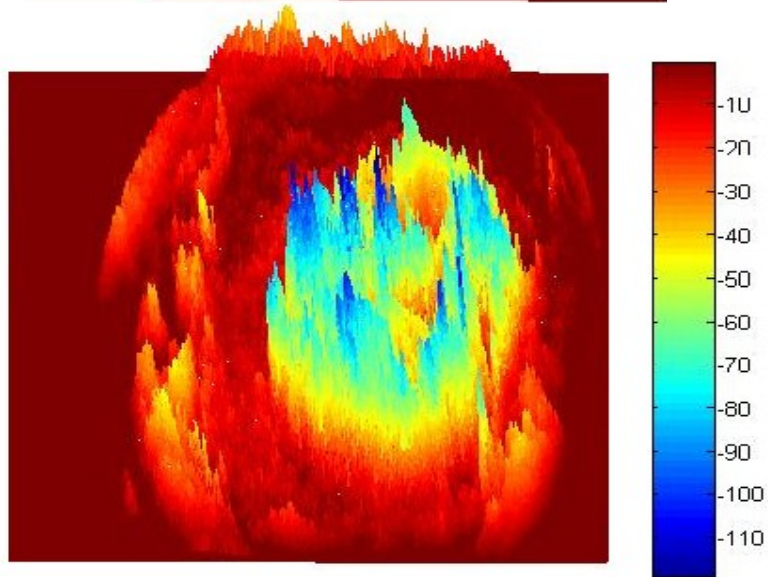
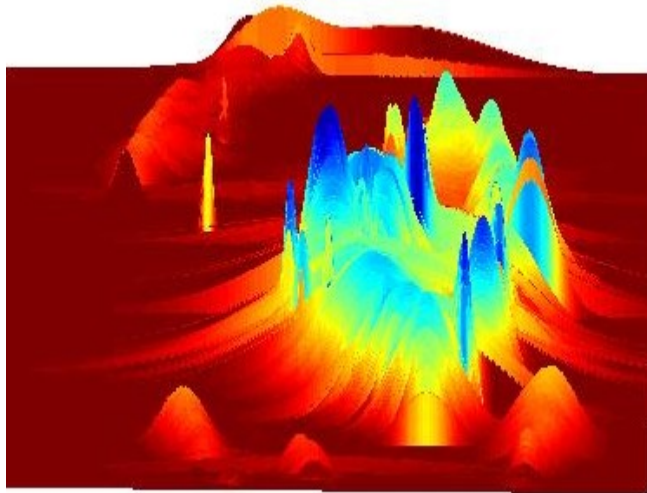
Epi egontae ta max peak', 2 gia X_{Δ} XIS ka Y_{Δ} XIS



Patient_57_151:
overall=1.40
worst_case_x=1.59
worst_case_y=1.71

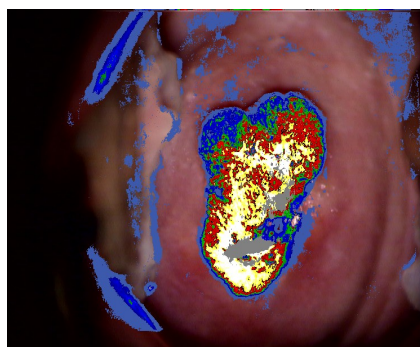
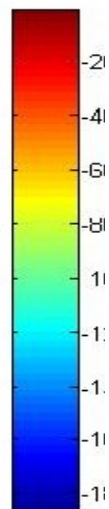
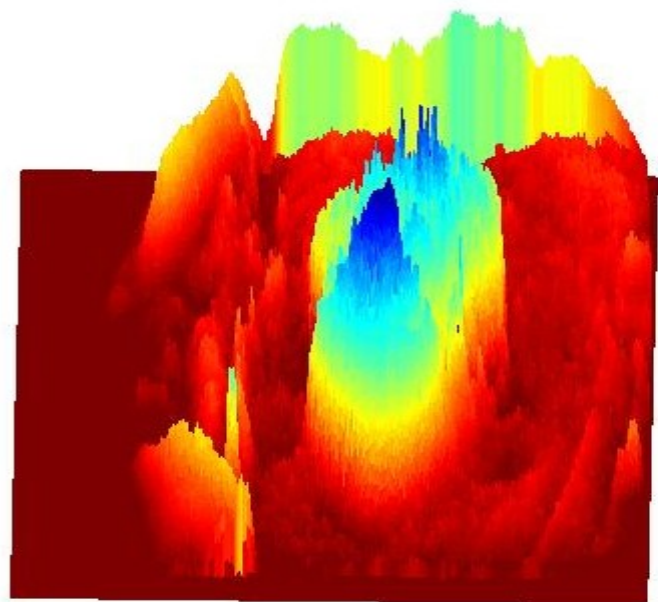
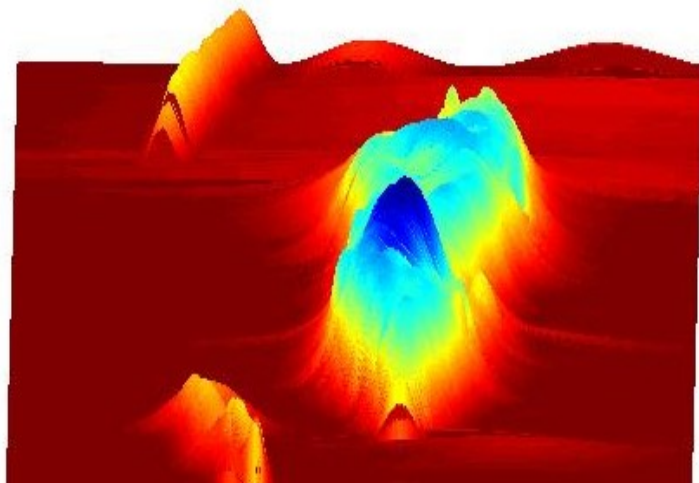
Patient_80_199:

Epilēgontas ta max peak1,2 gia $X_A XIS$ kai $Y_A XIS$

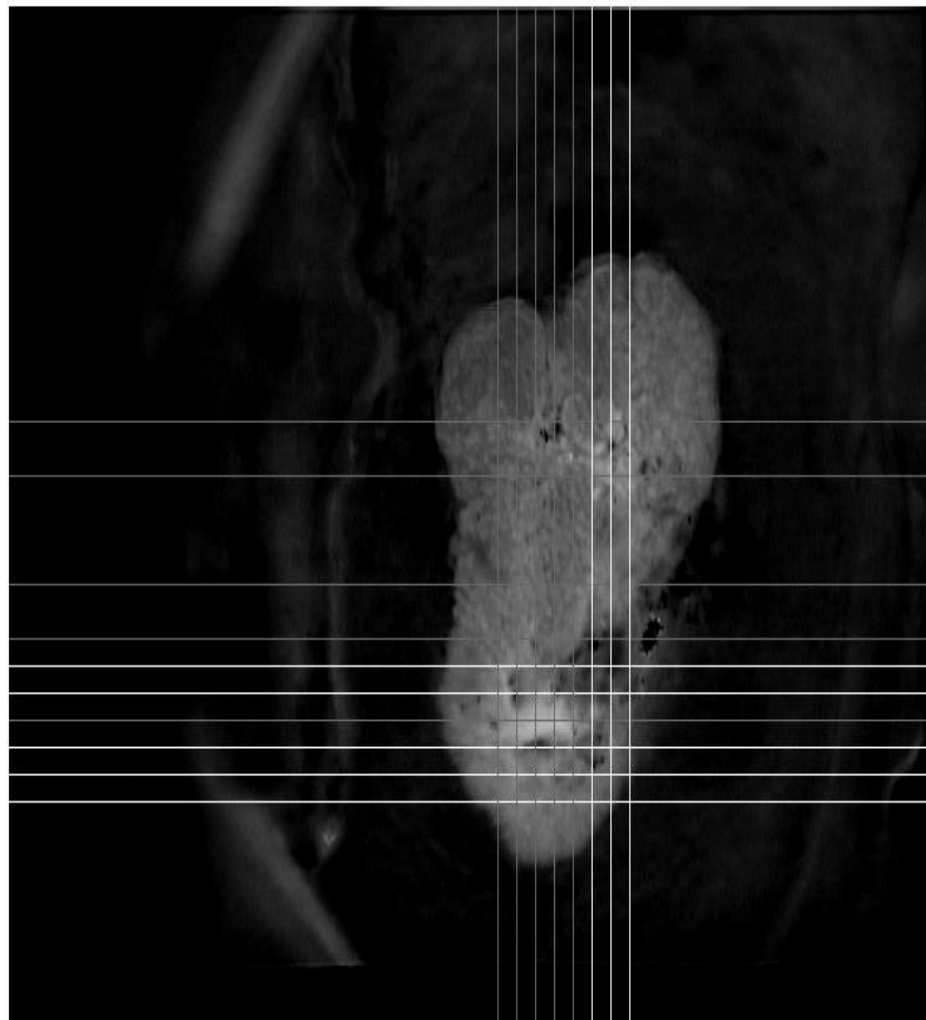


Patient_80_199:
overall=1.03
worst_case_x=1.24
worst_case_y=1.19

Patient_84_208

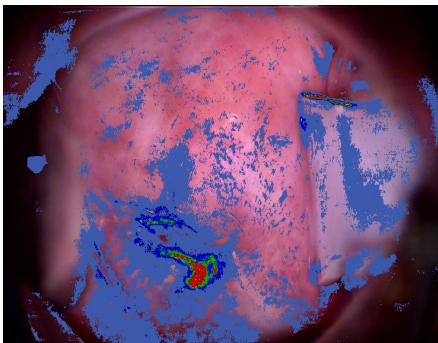
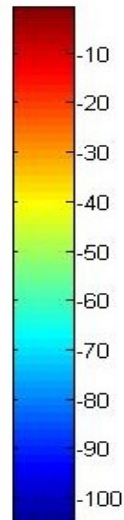
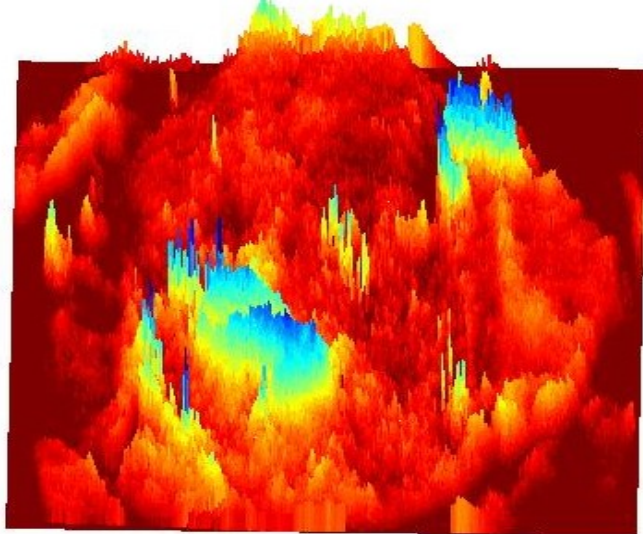
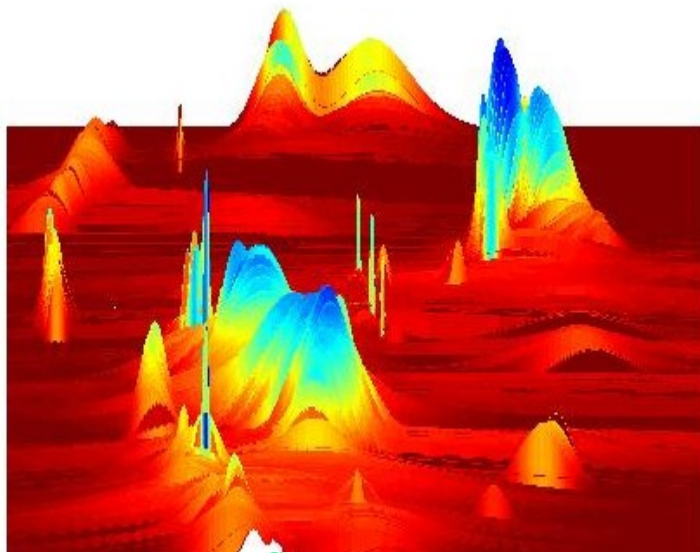


Epilęontas ta max peak1,2 ğia $X_A X_S$ kai $Y_A X_S$

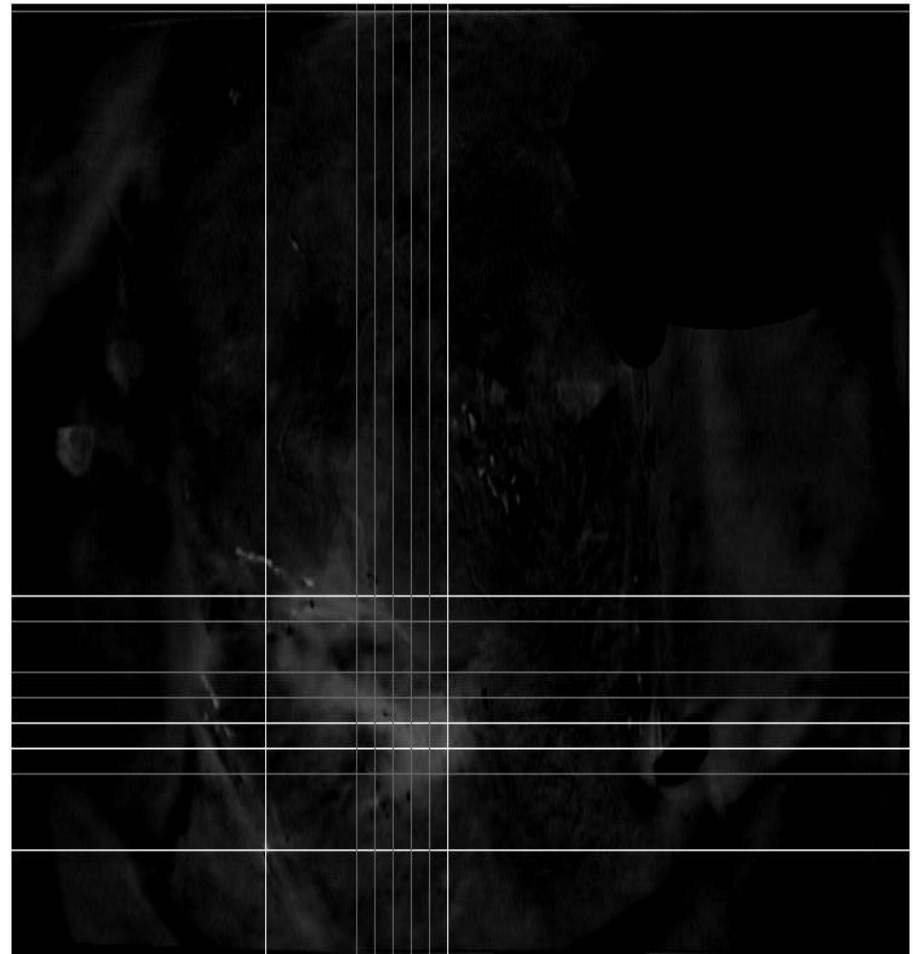


Patient_84_208:
overall=1.25
worst_case_x=1.54
worst_case_y=1.34

Patient_85_210:

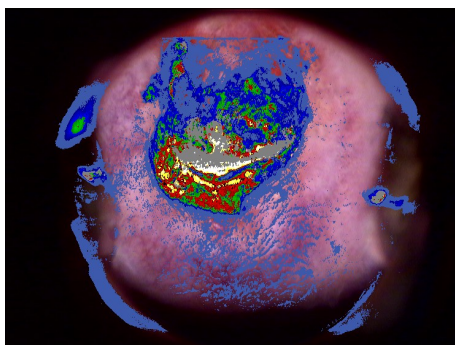
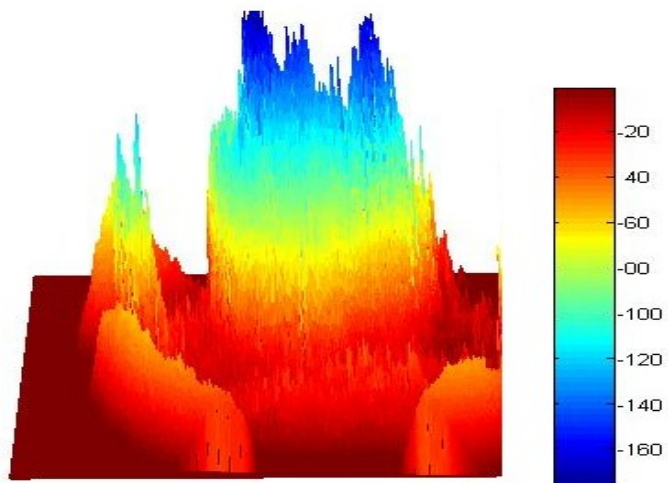
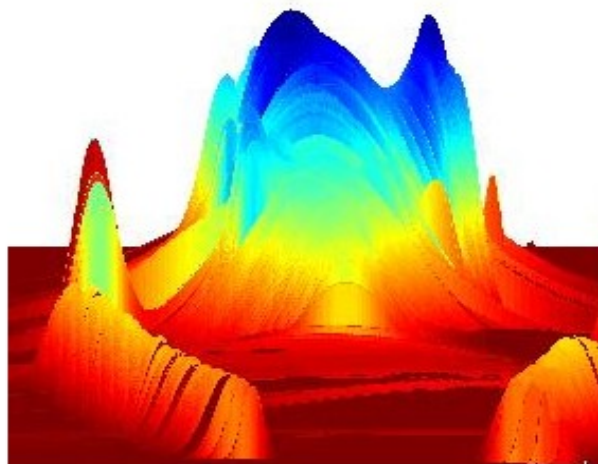


Epiliegontas ta max peak1,2 gia X_A XIS kai Y_A XIS

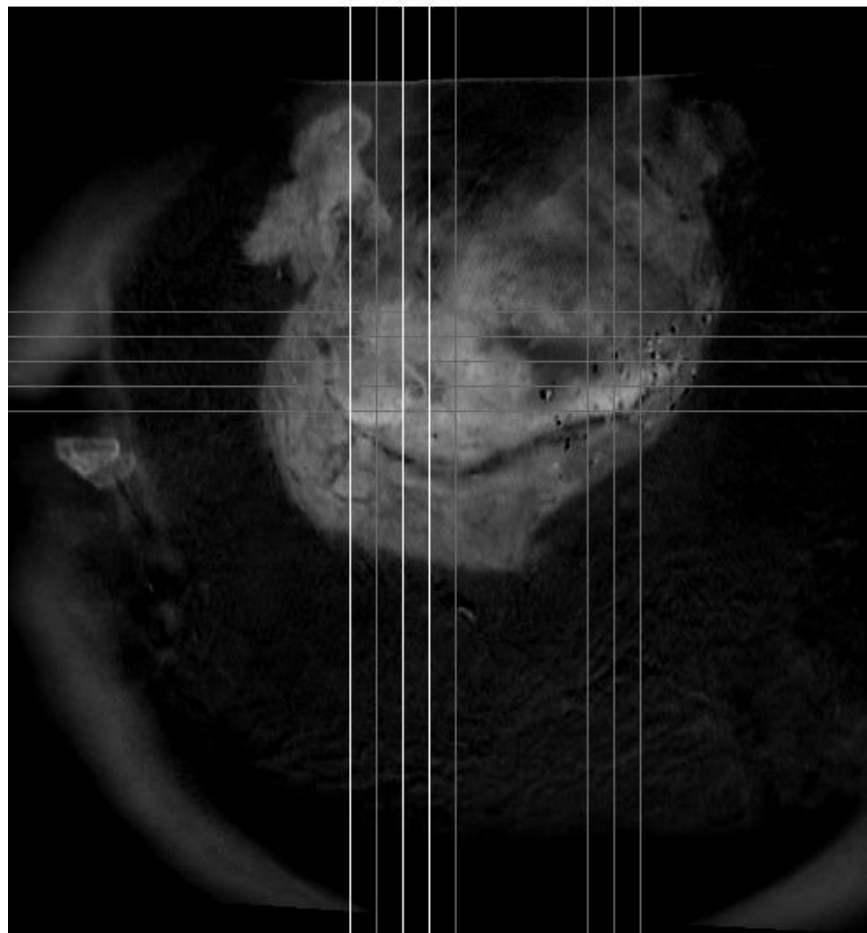


Patient_85_210:
overall=0.65
worst_case_x=0.73
worst_case_y=0.71

Patient_136_274



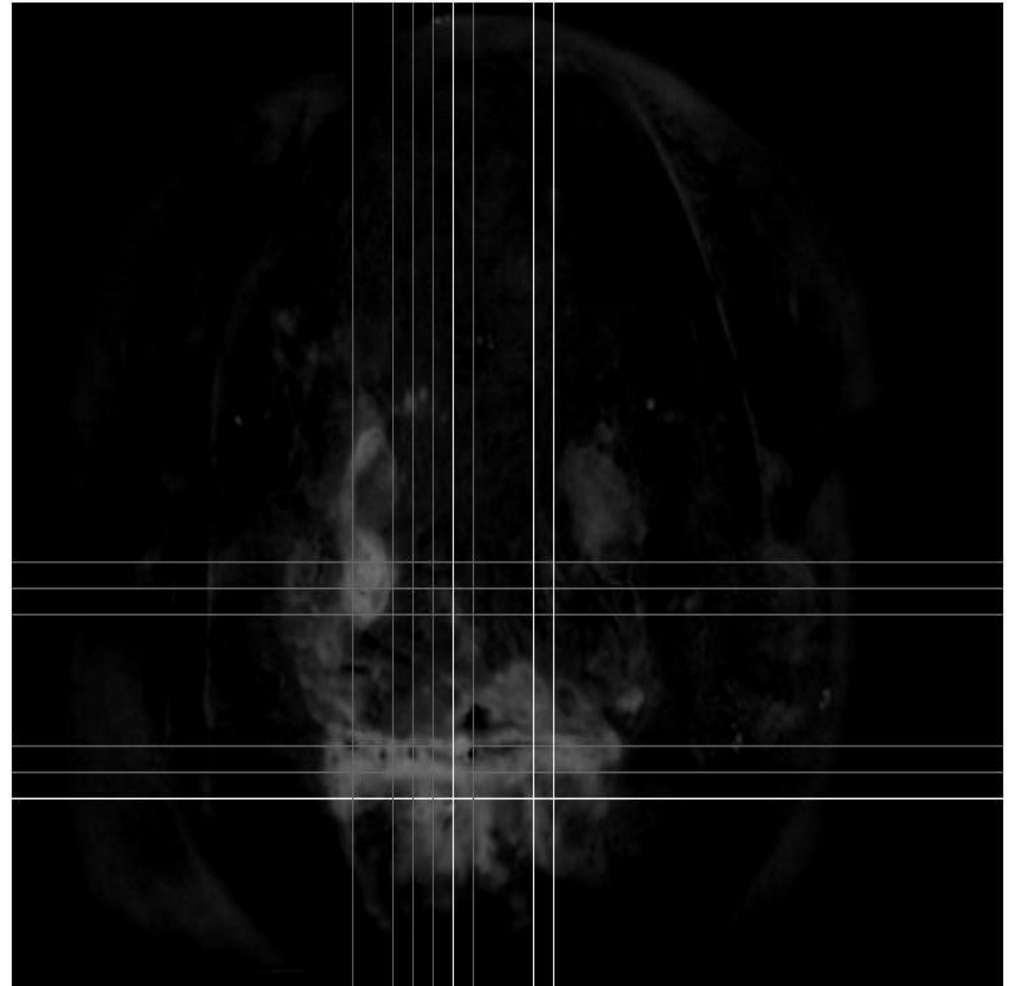
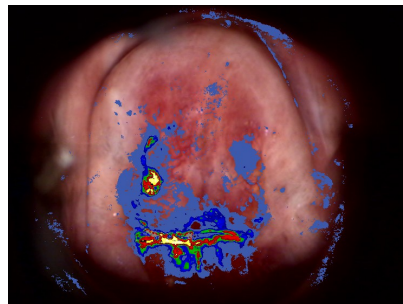
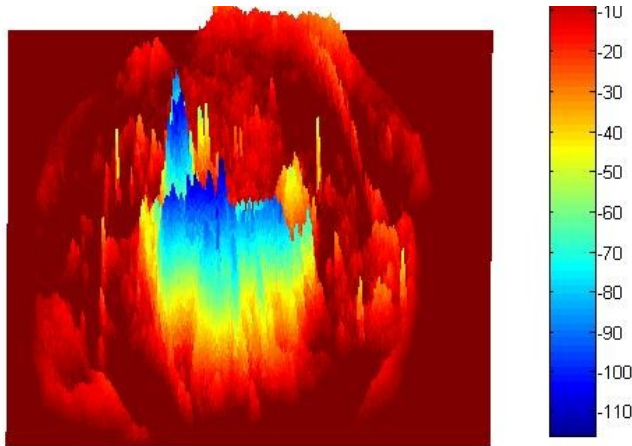
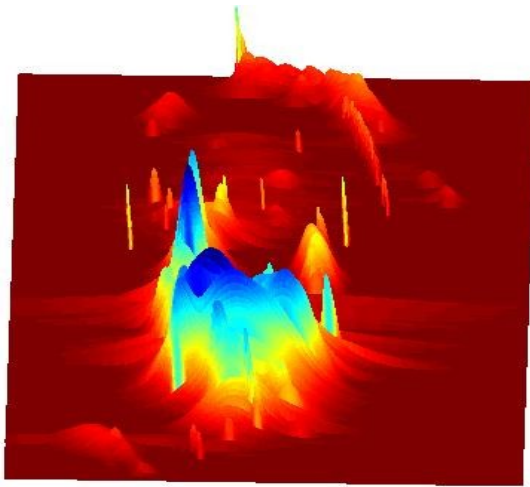
Εpilegontas ta max peak1,2 gia X_A XIS kai Y_A XIS



Patient_136_274:
overall=1.07
worst_case_x=1.17
worst_case_y=1.03

patient_13_69:

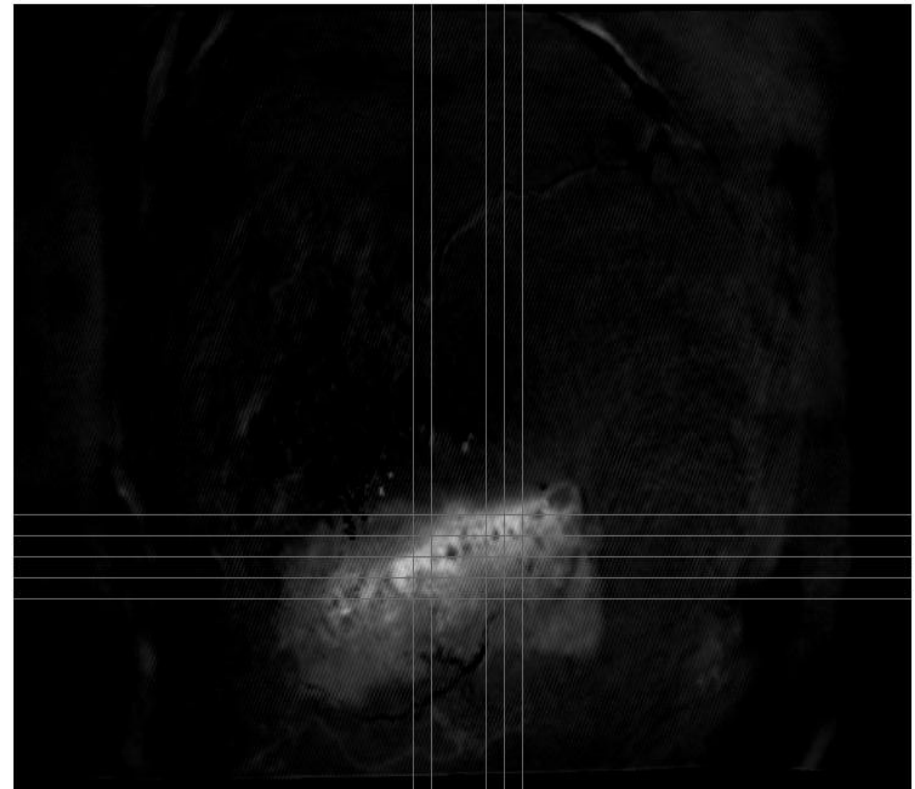
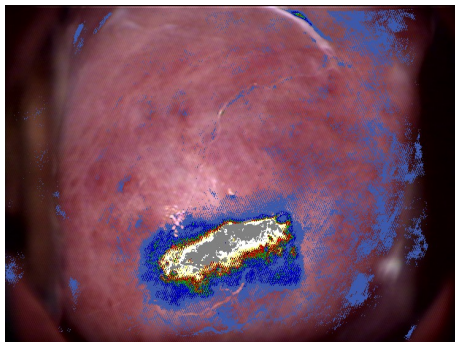
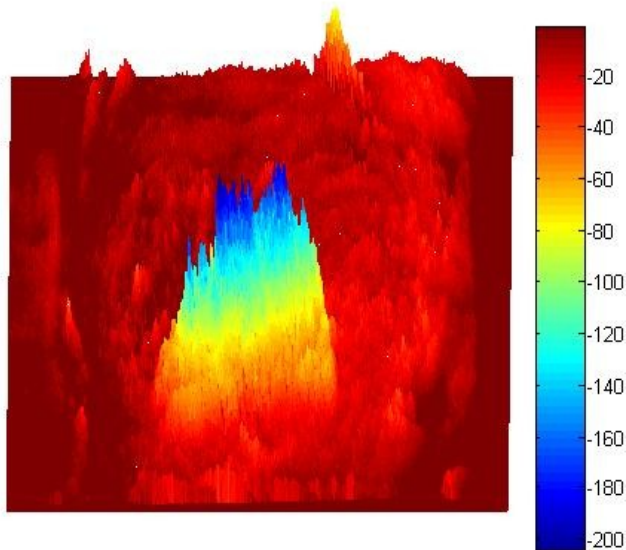
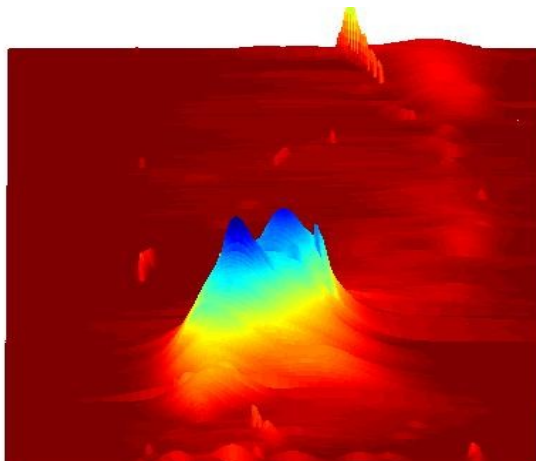
Epilėgontas ta max peak1,2 gja X_A XIS kai Y_A XIS



Patient_13_69:
overall=1.06
worst_case_x=1.2
worst_case_y=1.2

patient_63_160:

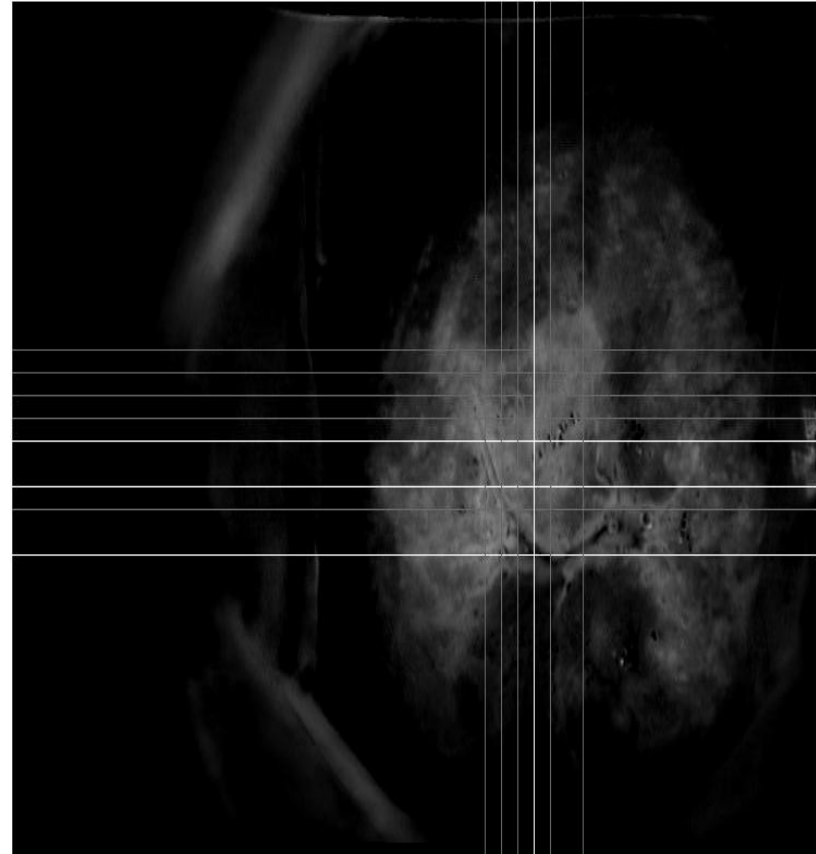
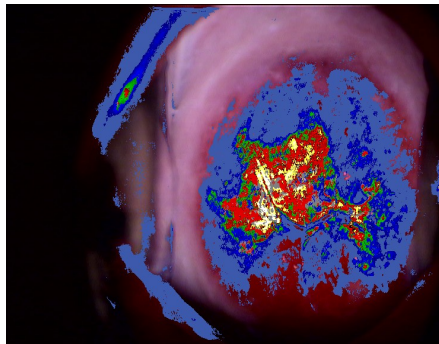
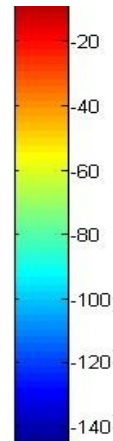
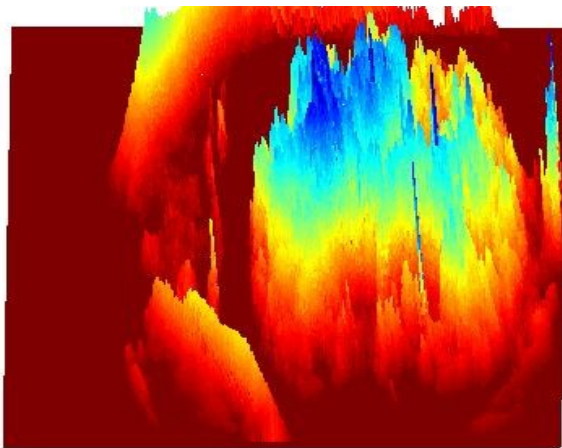
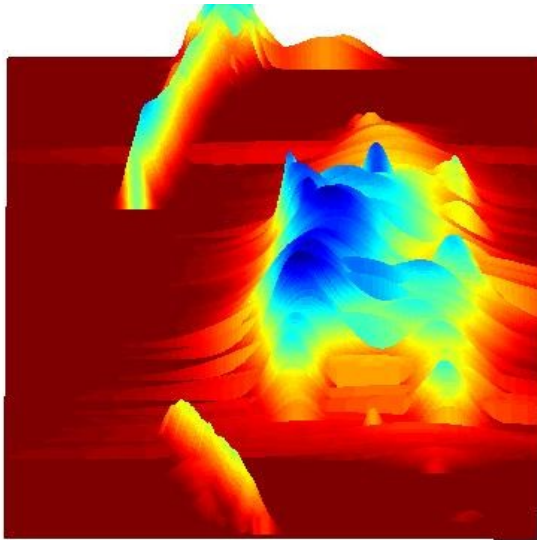
Epiliegontas ta max peak1,2 gia X_A XIS kai Y_A XIS



Patient_63_160:
overall=1.62
worst_case_x=1.17
worst_case_y=2.8

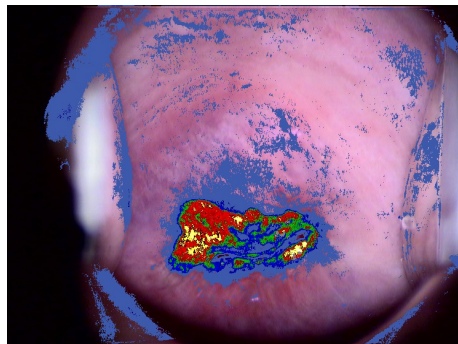
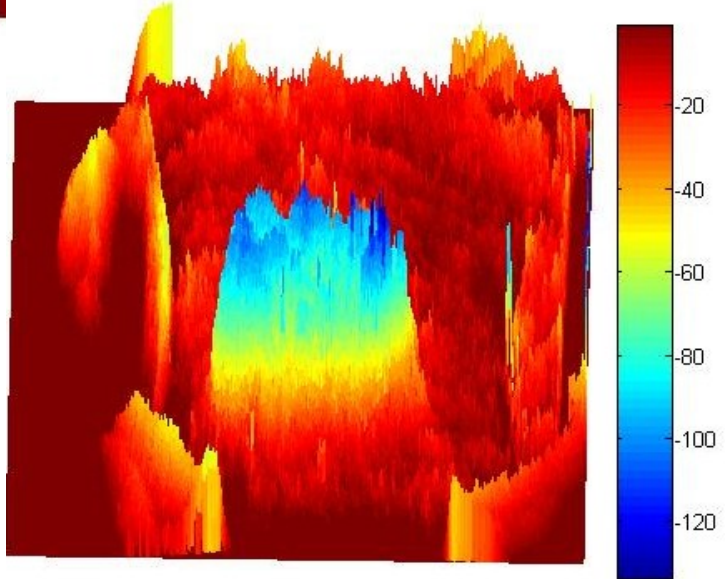
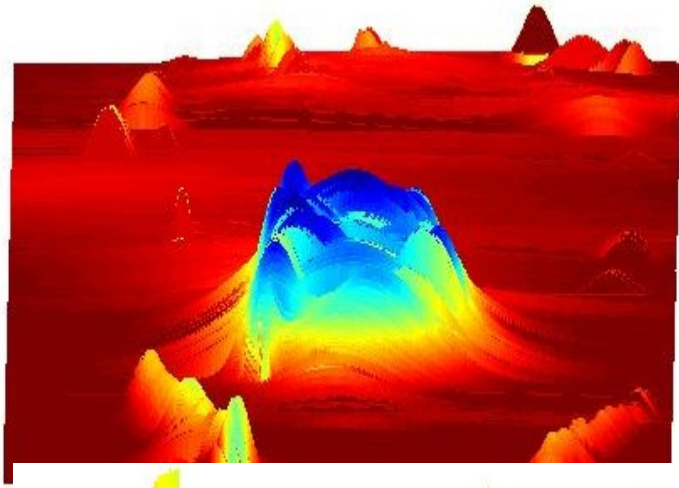
patient_120_256:

Επιλογεται το max peak1,2 για X_A, XIS και Y_A, YIS



Patient_120_256:
overall=0.86
worst_case_x=0.93
worst_case_y=1.12

patient_122_258:



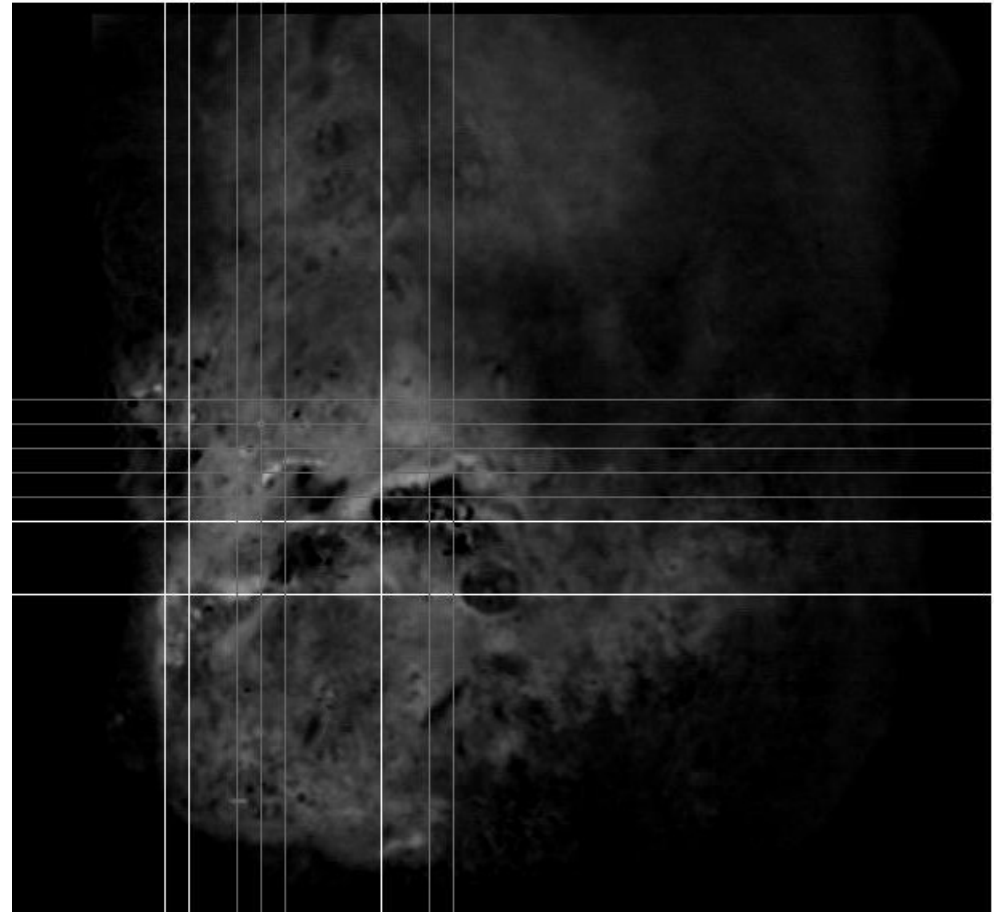
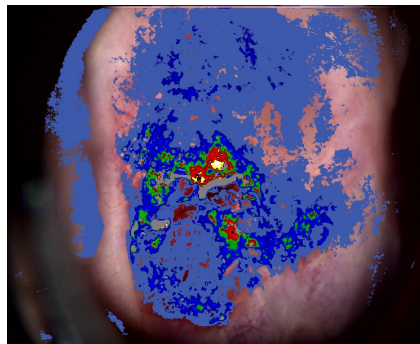
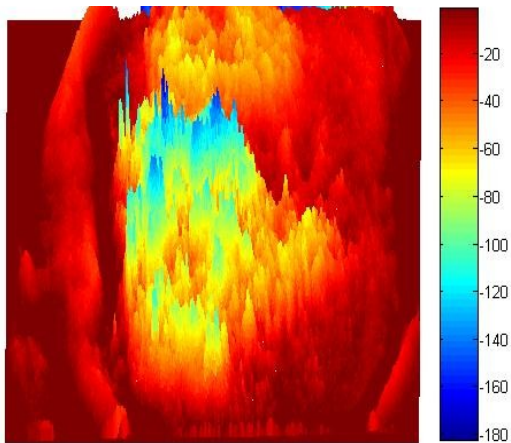
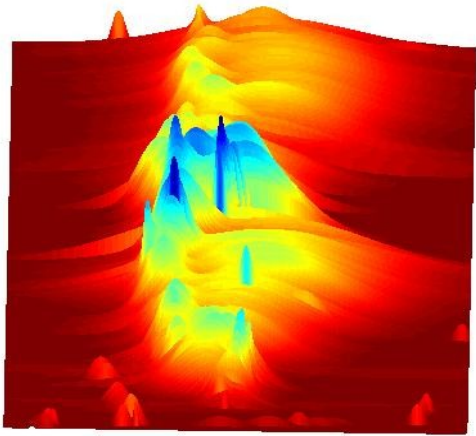
Epilegontas ta max peak1,2 gia X_A XIS kai Y_A XIS



Patient_122_258:
overall=1.32
worst_case_x=1.07
worst_case_y=1.81

patient_66_164

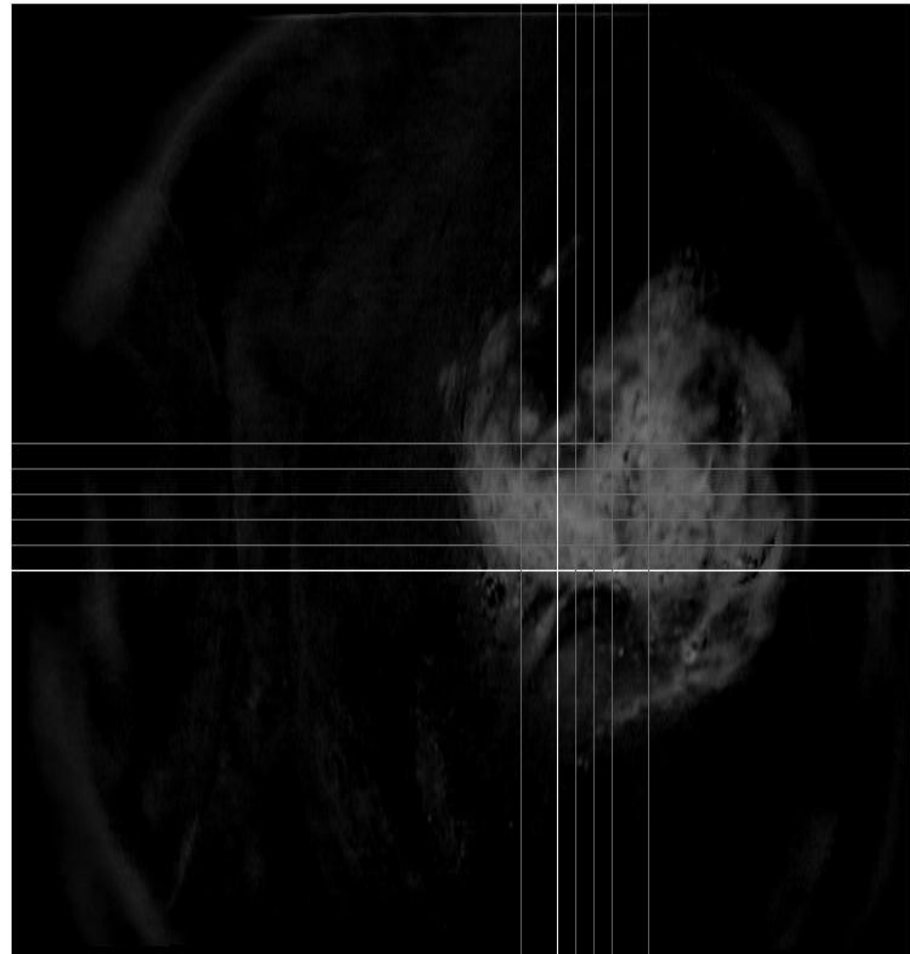
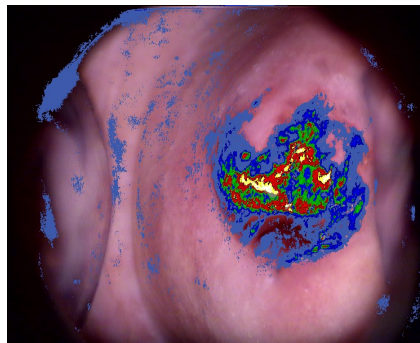
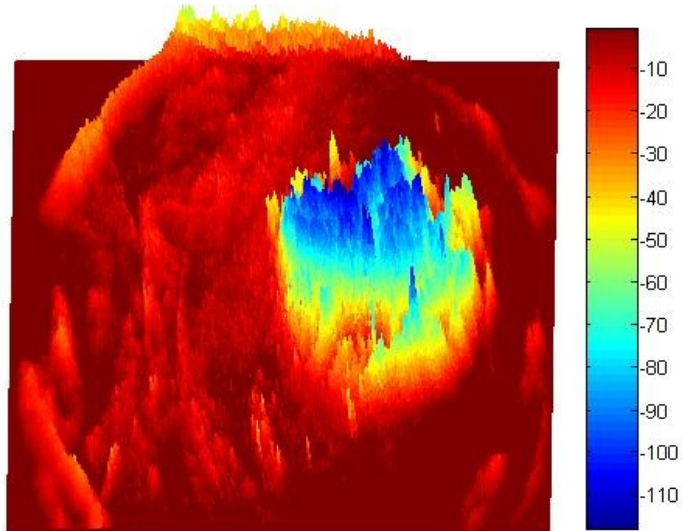
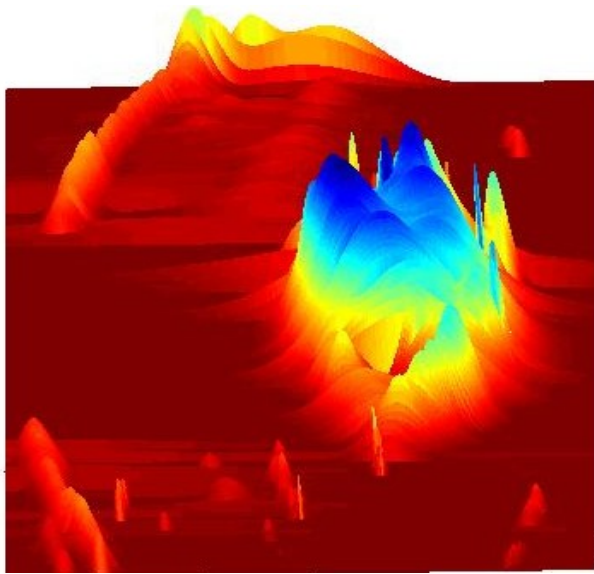
Εpilegontas ta max peak1,2 gia X_A XIS kai Y_A XIS



Patient_66_164:
overall=1.06
worst_case_x=1.38
worst_case_y=0.87

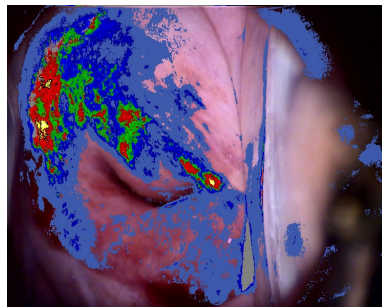
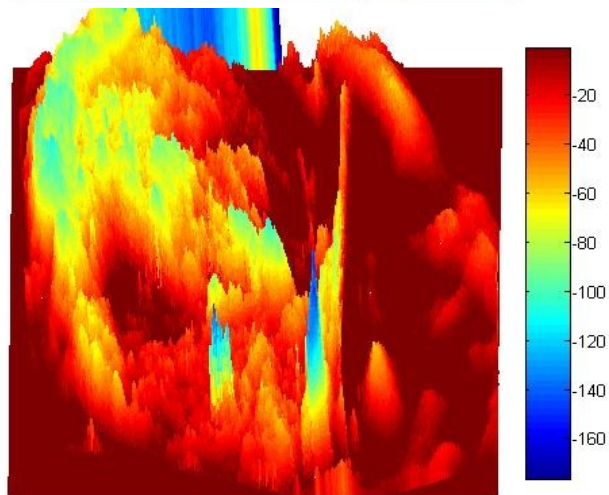
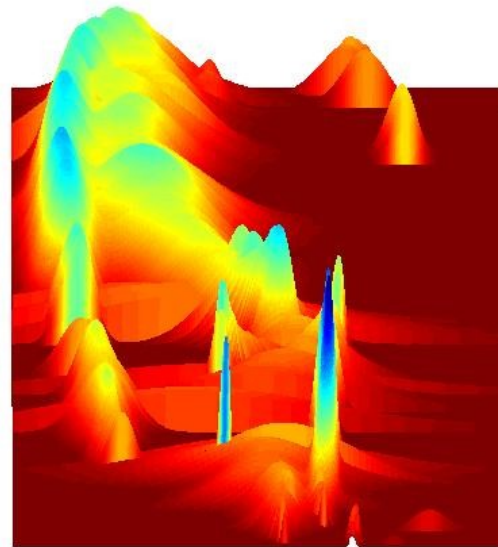
patient_94_225:

Epilegontas ta max peak1,2 gia X_A XIS kai Y_A XIS

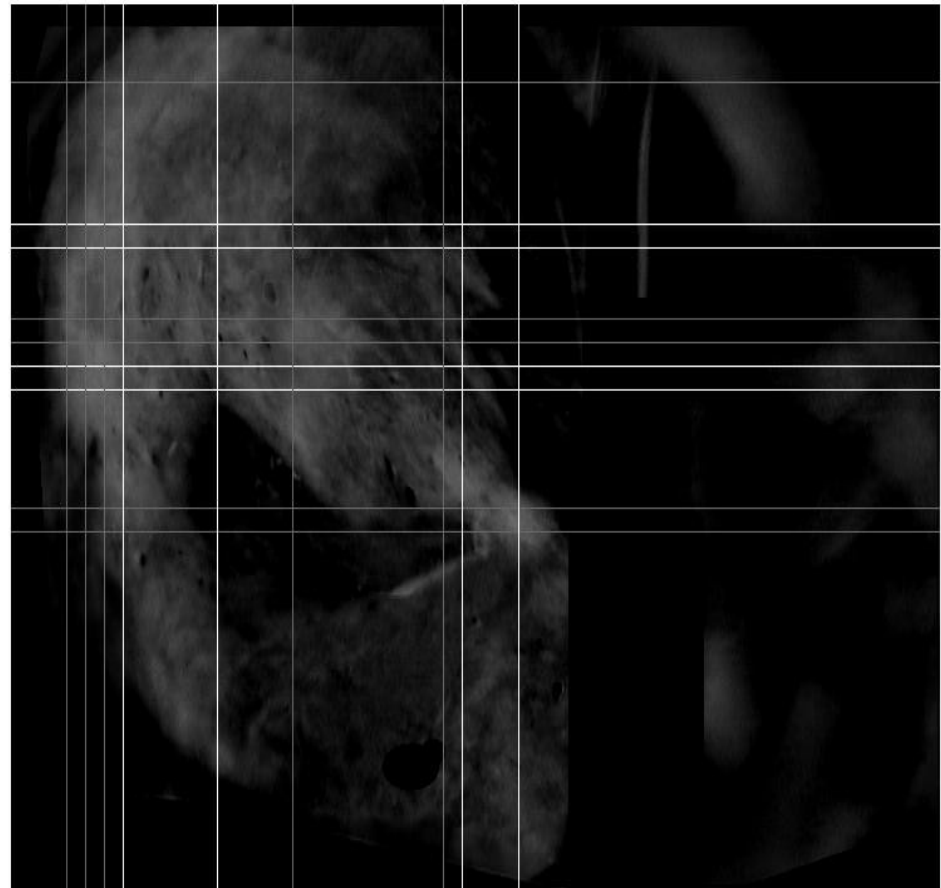


Patient_94_225:
overall=0.97
worst_case_x=0.97
worst_case_y=1.19

patient_118_254:

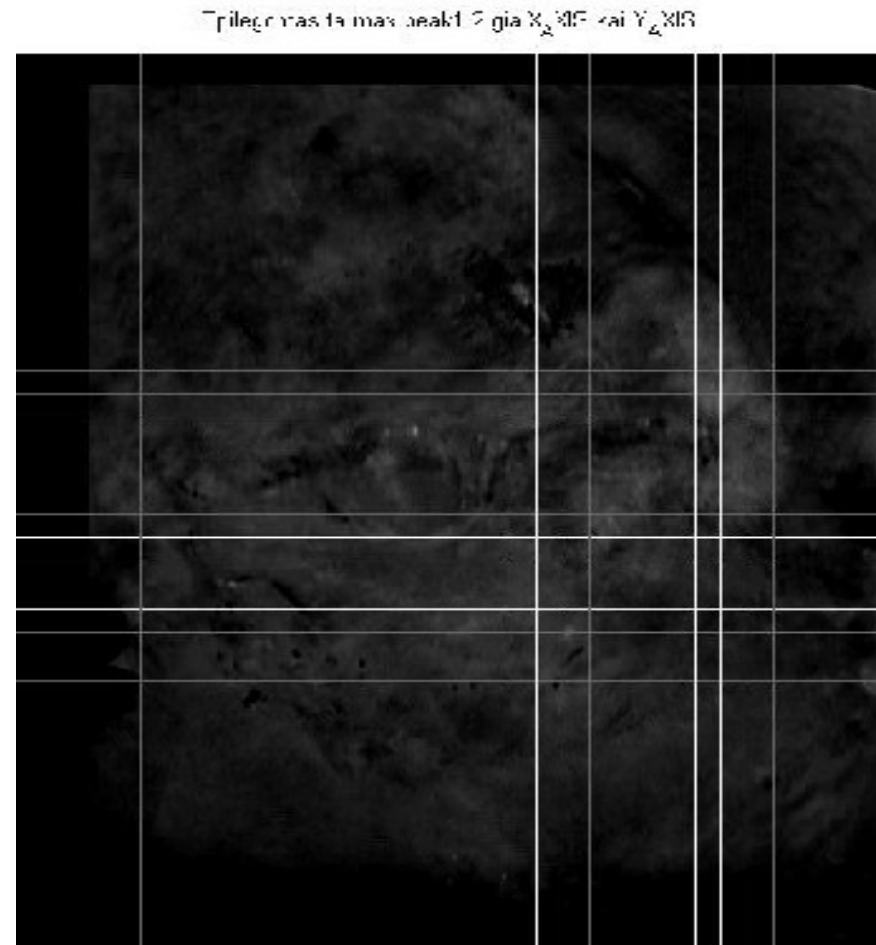
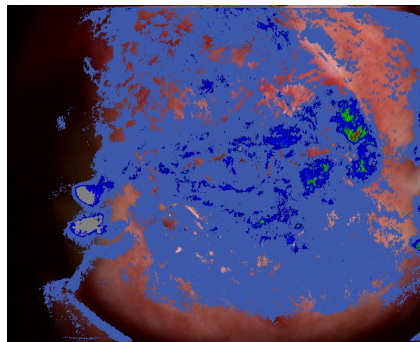
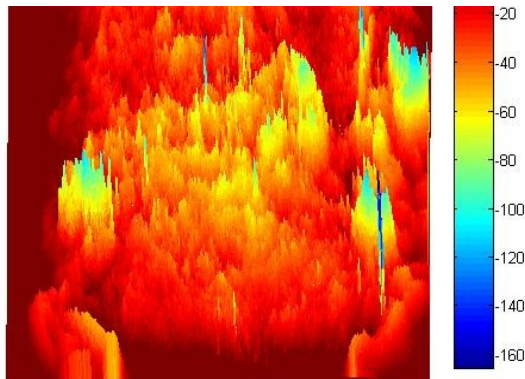
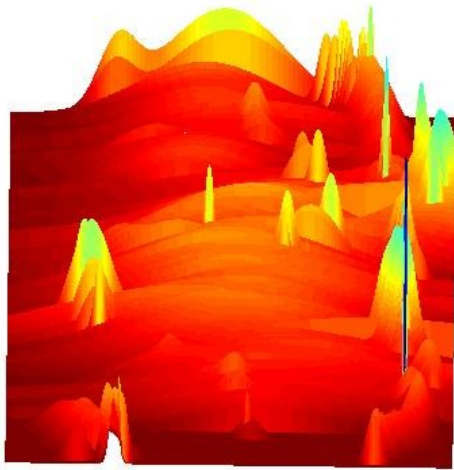


Epilegontas ta max peak1,2 gia X_A, XIS kai Y_A, YIS



Patient_118_254:
overall=1.01
worst_case_x=1.22
worst_case_y=1.37

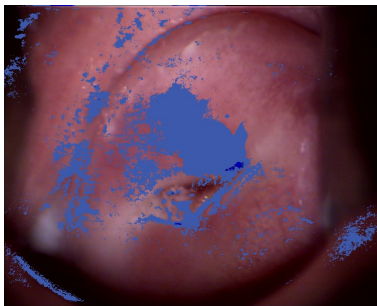
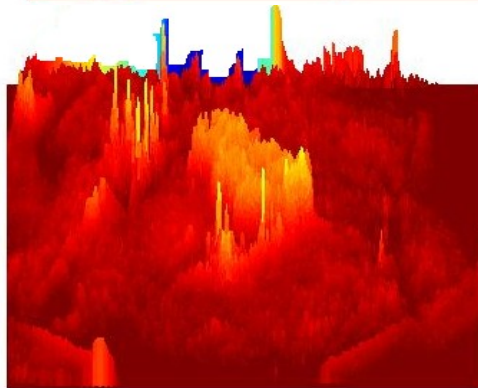
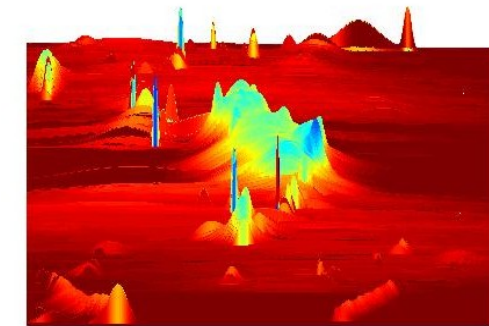
patient_176_329:



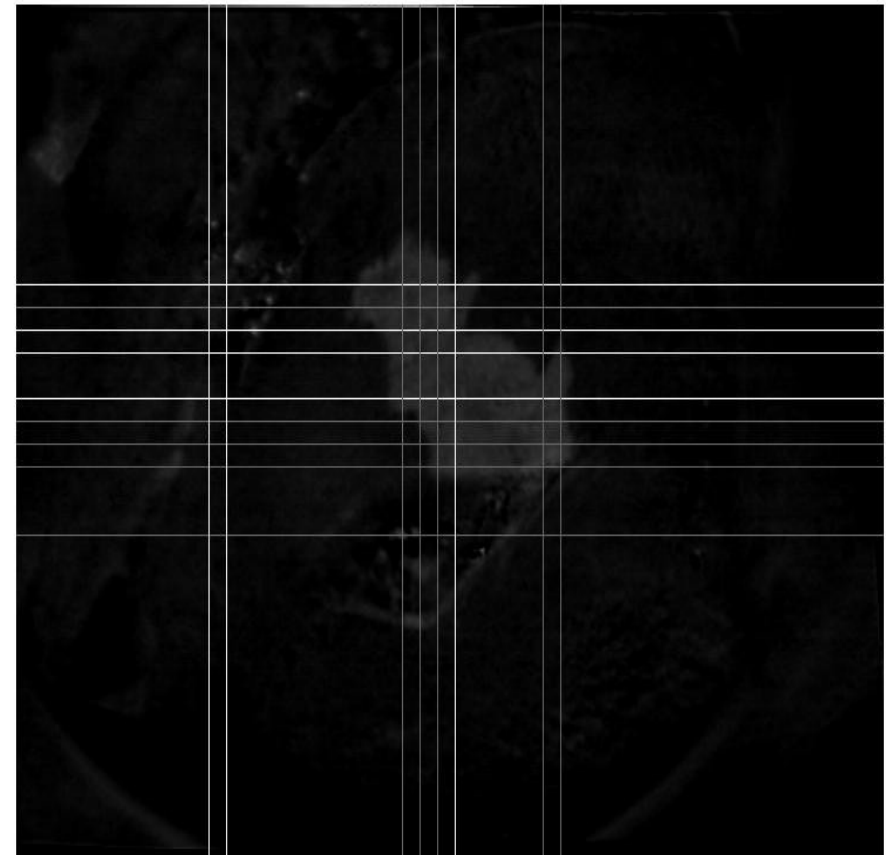
Patient_176_329:
overall=0.60
worst_case_x=0.77
worst_case_y=0.69

TRUE NEGATIVE

Patient_35_117:



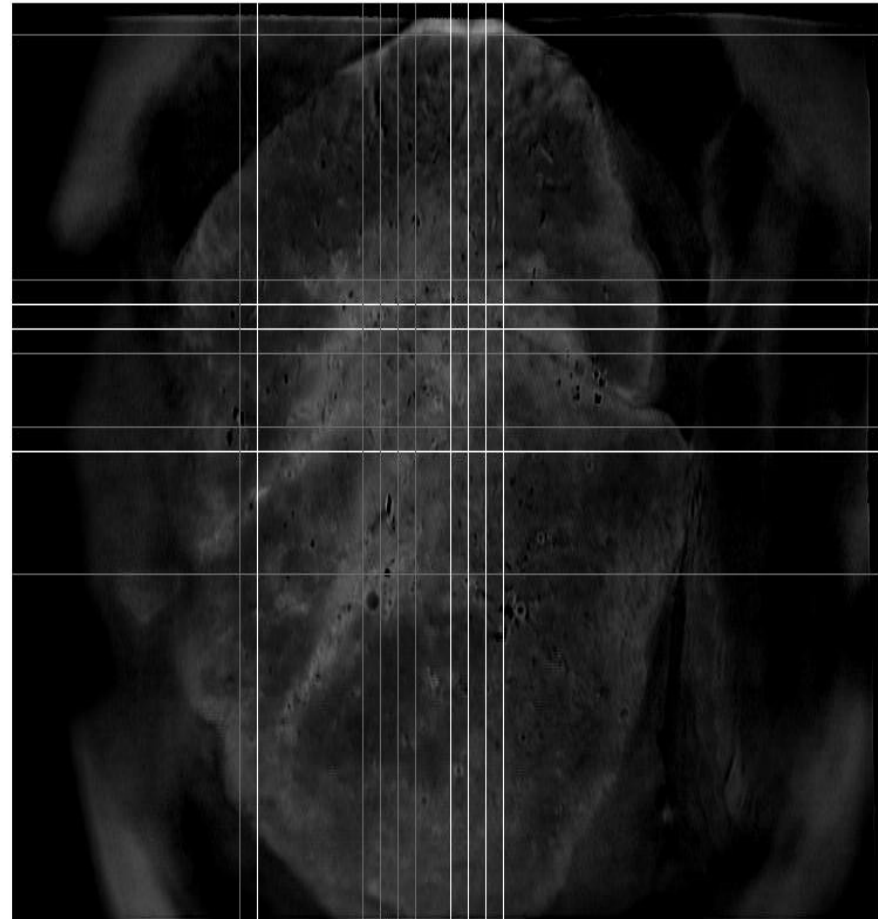
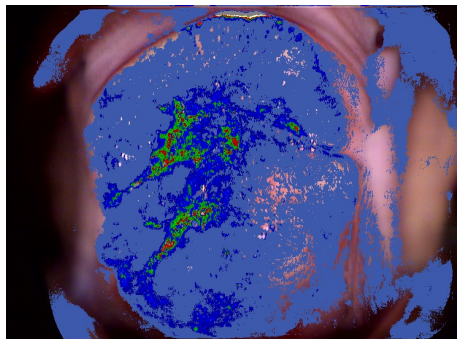
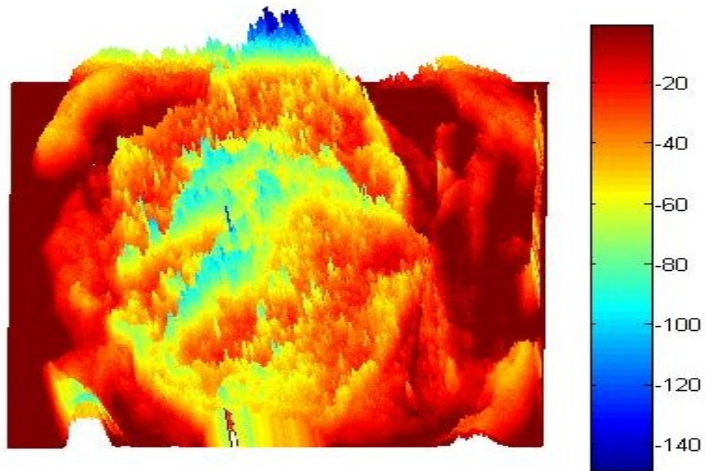
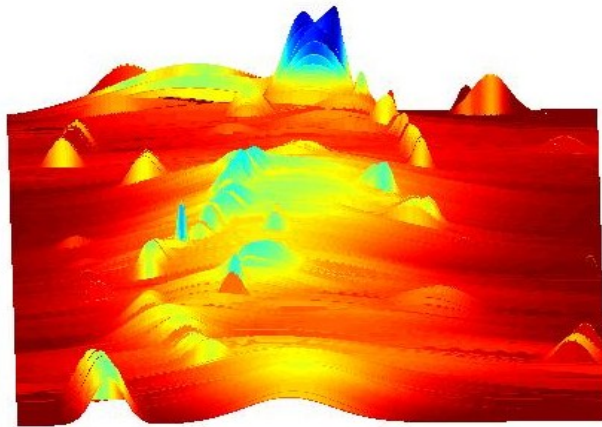
Epilegontas ta max peak1,2 gia X_A XIS kai Y_A XIS



Patient_35_117:
overall=0.50
worst_case_x=0.63
worst_case_y=0.47

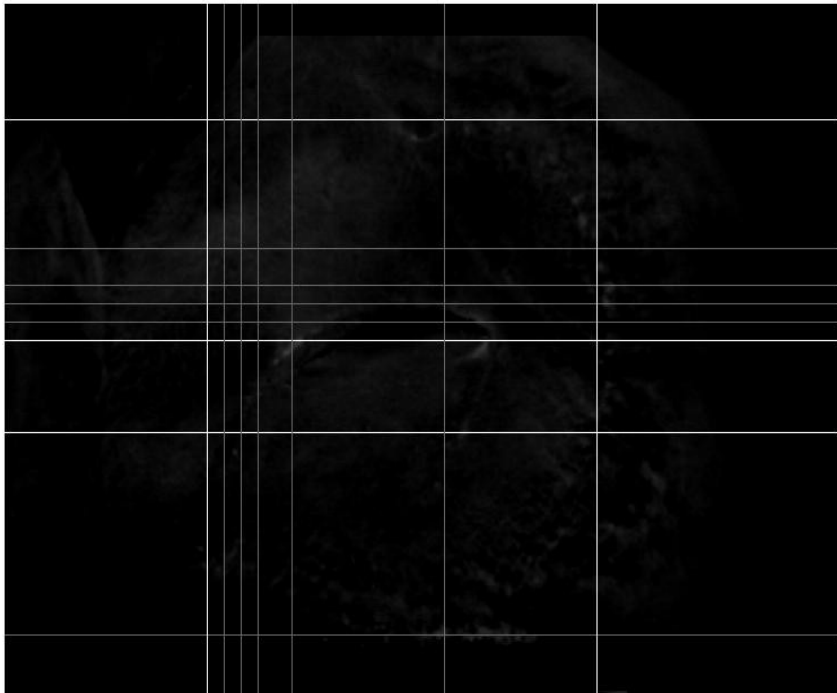
patient_93_224:

Epilegontas ta max peak1,2 gia X_A, XIS kai Y_A, YIS

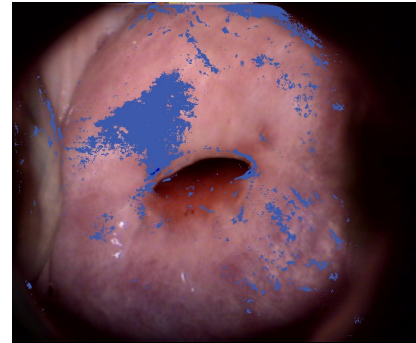


Patient_93_224:
overall=0.66
worst_case_x=0.61
worst_case_y=0.43

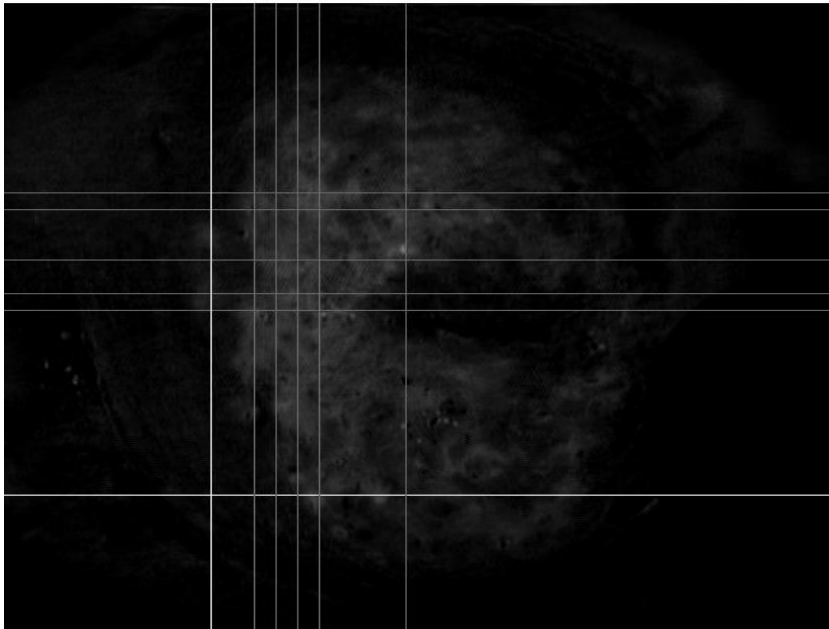
Epilegontas ta max peak1,2 gia X_A XIS kai Y_A XIS



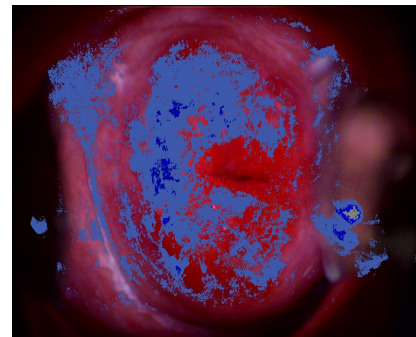
Patient_10_66:
overall=0.33
worst_case_x=0.47
worst_case_y=0.28



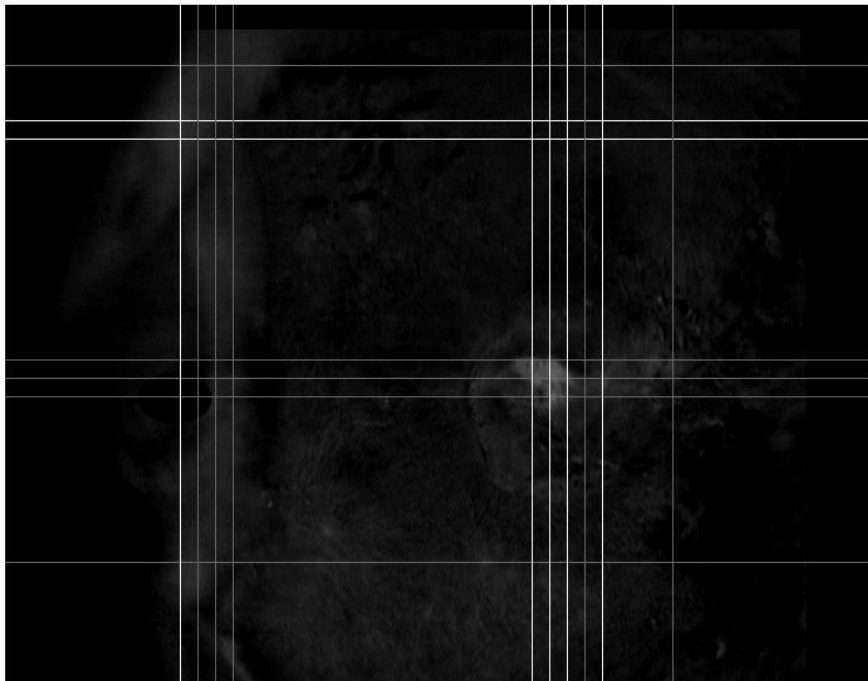
Epilegontas ta max peak1,2 gia X_A XIS kai Y_A XIS



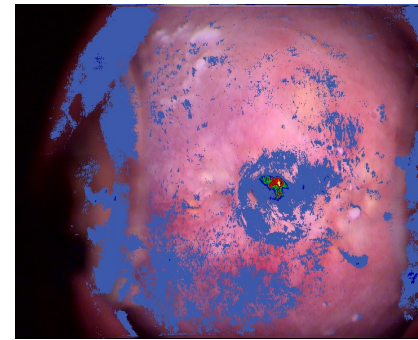
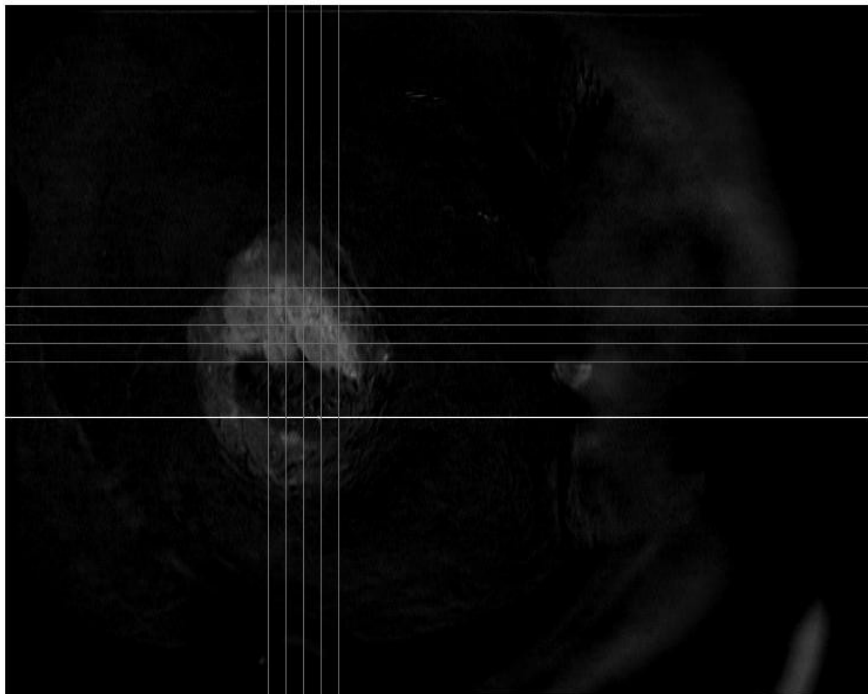
Patient_114_250:
overall=0.53
worst_case_x=0.51
worst_case_y=0.77



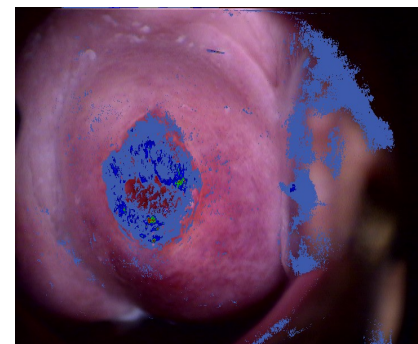
Epilegontas ta max peak1,2 gia X_A XIS kai Y_A XIS



Epilegontas ta max peak1,2 gia X_A XIS kai Y_A XIS



Patient_108_243:
overall=0.65
worst_case_x=0.50
worst_case_y=1.10



Patient_121_257:
overall=1.15
worst_case_x=1.20
worst_case_y=1.20

Αναφέρονται τα σημαντικότερα κομμάτια του κώδικα υλοποίησης.

Extract profile of the picture.

```

z=imread('subtracted.bmp'); column=double(z(:,k));%k=15:15:720

line=double(z(k,:))%k=5:15:560

...

for every profile of lines and column(eg.for a line)

    %fitting with gauss2 and restrictions

    t = 1:length(line);
    opts = fitoptions('gauss2');
    opts.Lower=[2 2 2 2 2];
    opts.Upper=[100 738 739 100 738 739];
    fitted_line=fit(t',line,'gauss2',opts);
    efil= fitted_line(t);
    ...

    % fitted_line is an object concluding the fitting parameters
    % efil contains the fitted data-profile

correct values of every profile

    ... if (efil(i)<0.9) ... value(profile)=0

hold max-min values and I

    [maxvalue,I]=max(efil);
    minvalue=min(efil);
    % I is the x position of the maxvalue

main processing of every profile(eg.line)

    %check for problem in profile -optimazation
    %exception No1
    if(maxvalue<20)
        L1=0;L2=0; %null all parameters eg.peak1,peak2 ...
        disp('This line gives no information.Exception No1');

    %exception No2
    elseif((fitted_line.c1<12) & (fitted_line.c2<12))
        L1=0;L2=0; %null all parameters ...
        disp('THIS PROFILE IS NOISE.Exception No1');

    %correction No1
    elseif(abs(fitted_line.b1-fitted_line.b2)<25)
        disp('NEAR MAXVALUES!FIT WITH GAUSS1 AND FEATURE EXTRACTION.Correction
No1')
        opts.Lower=[0 2 2];
        opts.Upper=[100 738 739];
        fitted_line=fit(t',line,'gauss1',opts);
        x1=round(fitted_line.b1-(sqrt(2)*fitted_line.c1));
        x2=round(fitted_line.b1+(sqrt(2)*fitted_line.c1));
        %check if x1,x2 throws of the picture
        if(x1<=1){ x1=1;} end
        if(x2>=length(efil)) {x2=length(efil);}end
        %ascend calculation
        L1=abs(efil(round(fitted_line.b1))-efil(x1))/abs(x1-round(fitted_line.b1));
        L2=abs(efil(round(fitted_line.b1))-efil(x2))/abs(x2-round(fitted_line.b1));

```

```

    peak1=efil(round(fitted_line.b1));
    peak2=efil(round(fitted_line.b1));
    D1=0.5*L1+0.5*peak1/100;
    D2=0.5*L2+0.5*peak2/100;
    ... ...
else
    continuing the main processing,excluding the above cases

%1st derivative calculation
    ...
    for(i=2:length(diff_efil)-1)
        if(diff_efil(i)<0)
            if(diff_efil(i-1)>0)
                topika_megista(counter3)=i-1;
                counter3=counter3+1;
            end
            if(diff_efil(i+1)>0)
                topika_elaxista(counter4)=i;
                counter4=counter4+1;
            end
        end
    end
    ... ...

separating the processing of every profile per local maximum (L.M)
    %case with none L.M
    %exception No3
    if(length(topika_megista)==0)
        disp('PROBLEM IN FITTING.IGNORE!Exception No3');    L1=0;    L2=0;    ...
    %case with 1 L.M (text processing page 57)
elseif(length(topika_megista)==1)
    %find x1,x2 for every case.The order of fitted parameters depends on
    %the trust-region algorithm.We need to know the left and right gauss
    %bell in order to demarcate the AW,so the following checkisnecessary
    if(fitted_line.b1<fitted_line.b2)
        x1=round(fitted_line.b1-(sqrt(2)*fitted_line.c1));
        x2=round(fitted_line.b2+(sqrt(2)*fitted_line.c2));
    else
        x1=round(fitted_line.b2-(sqrt(2)*fitted_line.c2));
        x2=round(fitted_line.b1+(sqrt(2)*fitted_line.c1));
    end
    %check if x1,x2 throws of the picture
    if(x1<=1)... end
    L1=abs(maxvalue-efil(x1))/abs(x1-I);
    L2=abs(maxvalue-efil(x2))/abs(x2-I);

    ... ...
    %exclude the noisy profiles containing useful information
    %(mentioned in the text in page 59).Calculate new x1,x2 and right
    %L1,L2
    %correction No2
    if(fitted_line.c1<10)
        x1=round(fitted_line.b2-(sqrt(2)*fitted_line.c2));
        x2=round(fitted_line.b2+(sqrt(2)*fitted_line.c2));
        %check x1,x2
        if(x1<=1)... end
        L1=abs(efil(round(fitted_line.b2))-efil(x1))/abs(x1-
            fitted_line.b2);
        L2=abs(efil(round(fitted_line.b2))-efil(x2))/abs(x2-
            fitted_line.b2);
        disp('PROBLEM IN PROFILE.CORRECTION No2');
    end
    %correction No3
    same for the other gaussian bell.
    if(fitted_line.c2<10)
        x1=round(fitted_line.b1-(sqrt(2)*fitted_line.c1));
        x2=round(fitted_line.b1+(sqrt(2)*fitted_line.c1));
        ... ... end

%new calculation of those parameters in case of the noisy profile with the

```

```

%right L1,L2
peak1=efil(topika_megista(1)); D1=0.5*L1+0.5*peak1/100;
peak2=... D2=...
... ..
disp(sprintf('Maxvalue of line %d is %d',k,int32(maxvalue)));

%case with 2 L.M
elseif(length(topika_megista)>1)
y1=topika_megista(1);
y2=topika_megista(length(topika_megista));
    %find x1,x2 for every case.Depends on the order
    if(fitted_line.b1<fitted_line.b2)
        x1=round(fitted_line.b1-(sqrt(2)*fitted_line.c1));
        x2=round(fitted_line.b2+(sqrt(2)*fitted_line.c2));
    else
        x1=round(fitted_line.b2-(sqrt(2)*fitted_line.c2));
        x2=round(fitted_line.b1+(sqrt(2)*fitted_line.c1));
    end
%check x1,x2
if(x1<=1)...end

L1=abs(efil(y1)-efil(x1))/abs(x1-y1);
L2=abs(efil(y2)-efil(x2))/abs(x2-y2);

%exclude the noisy profiles containing useful information(page 62)
%correction No4
if(fitted_line.c1<10)
    x1=round(fitted_line.b2-(sqrt(2)*fitted_line.c2));
    x2=round(fitted_line.b2+(sqrt(2)*fitted_line.c2));
    %check x1,x2
    if(x1<=1)... end

    L1=abs(efil(round(fitted_line.b2))-efil(x1))/abs(x1-fitted_line.b2);
    L2=abs(efil(round(fitted_line.b2))-efil(x2))/abs(x2-fitted_line.b2);
    disp('PROBLEM IN PROFILE.CORRECTION No4');
end

same for the other gaussian bell.
%correction No5
if(fitter_line.c2<10)
    x1=round(fitted_line.b1-(sqrt(2)*fitted_line.c1));
    x2=round(fitted_line.b1+(sqrt(2)*fitted_line.c1));
    %check x1,x2
    if(x1<=1)... end

    L1=abs(efil(round(fitted_line.b1))-efil(x1))/abs(x1-fitted_line.b1);
    L2=abs(efil(round(fitted_line.b1))-efil(x2))/abs(x2-fitted_line.b1);
    disp('PROBLEM IN PROFILE.CORRECTION No5');
end

... ..

%case 2 injured areas.Find new,right slopes L1,L2(page 63)
%correction No6
if(length(topika_elaxista)==0)
    if(fitted_line.b1<fitted_line.b2)
        x1_1=round(fitted_line.b1+(sqrt(2)*fitted_line.c1));
        x2_2=round(fitted_line.b2-(sqrt(2)*fitted_line.c2));
    else
        x1_1=round(fitted_line.b2+(sqrt(2)*fitted_line.c2));
        x2_2=round(fitted_line.b1-(sqrt(2)*fitted_line.c1));
    end
%now check x1_1,x2_2
    if(x1_1<=1)... end
    ...
    if(abs(x1_1 - x2_2)>400)
        L1_1=abs(efil(y1)-efil(x1))/abs(x1-y1);
        L1_2=...
        L2_1=...

```

```

        L2_2=...
        L1=max(L1_1,L1_2);
        L2=max(L2_2,L2_1);
disp('There is no local minimum.There are two injured
      areas.Correction No6');
    end
end
... ..
D1... D2...

%end of processing per L.M
end

for every calculation of L1,L2 inside the above code
check if L1,L2 greater than theoretical values
if(L1>18) {L1=L1/6;} ...

Changing the limits,almost same calculation for column profiles
(process every profile. Feature extraction- Feature combination)
... ..
... ..

for every extracted profile hold the profile and the parameters
L1,L2,peak1,peak2,line number,column number
total_gauss2_profiles(counter1,:)=efil;

... ..
slopes_L1(counter1)=L1; slopes_L2(counter1)=L2;
myline(1,counter1)=k;
myline(2,counter1)=counter1;

... ..

Save in tables for every profile the following data:
Number of line/column,L1,L2,peak1,peak2,D1,D2(same as table page 68)
EA1_all(1,:)=myline(1,:);
EA1_all(2,:)=slopes_L1;
EA1_all(3,:)=peak1_all;
EA1_all(4,:)=D1_all;
EA2_all(1,:)=myline(1,:);
EA2_all(2,:)=slopes_L2;
EA2_all(3,:)=peak2_all;
EA2_all(4,:)=D2_all;
... ..
for choosing the profiles referring to AW area sort per peak1 and peak2
[A_basi_peak1,index]=sort(EA1_all,2);
A_basi_peak1=EA1_all(:,index(3,:));
[B_basi_peak2,index]=sort(EA2_all,2);
B_basi_peak2=EA2_all(:,index(3,:)); (3 is the line position of peaks)
... ..
and choose the max peak1-x profiles,peak2-x profiles,peak1-y profiles,peak2-y
profiles
... ..
... ..
In general: calculate overall,worst_case_x,worst_case_y
OVERALL=mean([lines_10,ylines_10]);
worst_case_x=max(D1_mean,D2_mean);
worst_case_y=max(yD1_mean,yD2_mean);

if problem in parameters use median filter to find AW
im=imread('subtracted.bmp');
L = medfilt2(im,[25 25]);
counter1=1;
for(i=5:15:555)
    reg(1,counter1)=max(L(i,:));
    reg(2,counter1)=i;
    counter1=counter1+1;
end
[sorted_reg,index]=sort(reg,2);

```

```

sorted_reg=reg(:,index(1,:));
%and calculate parameters of the 20 max peak profiles

3D visualization(mesh,view)
figure; mesh(-z);
view(0,-90);

t = 1:length(z(1,:));
for (i=5:560)
    opts = fitoptions('gauss2');
    opts.Lower=[0 2 2 0 2 2];
    opts.Upper=[100 738 739 100 738 739];
    temp=fit(t',z(i,:)','gauss2',opts);
    weby(i,:)=temp(t);
end
figure; h=mesh(-weby); view(0,-90);

Testing smoothingspline and the other smoothing methods
    opts = fitoptions('Method','Smooth','SmoothingParam',7.051872e-04);
    test1=fit(t',line,'smoothingspline',opts);
    plot(t,efi6(t),t,line,'r',t,efi1,'g');
    test2=smooth(line,29,'moving');
    test3=smooth(line,19,'loess');
    test4=smooth(line,'sgolay',2);
    test5=smooth(line,19,'rloess');
    test6=medfilt1(line,20);

About x1,x2:transformation pages 64,65

```

ΑΝΑΦΟΡΕΣ - ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] Rafael C. Gonzalez, Richard E. Woods, “Digital Image Processing”, Addison-Wesley Publishing Company, 1993.
- [2] Milan Sonka, Vaclav Hlavac, Roger Boyle, “Image Processing, Analysis and Machine Vision” , Brooks/Cole Publishing Company, 1999, second edition.
- [3] Gerald Recktenwald, “Least Squares Fitting of Data to a Curve”, Portland State University, Department of Mechanical Engineering ,2007, second edition.
- [4] Costas Balas, “A Novel Optical imaging Method for the Early Detection, Quantitative Grading, and Mapping of Cancerous and Precancerous Lesion of Cervix” , IEEE Transaction on Biomedical Engineering, vol.48, 2001.
- [5] D C Walker, B H Brown, A D Blackett, J Tidy, R H Smallwood, “A study of the morphological parameters of cervical squamous epithelium”, Institute of Physics Publishing, p.121-124, 2003.
- [6] Costas Balas, George Papoutsoglou, Antonis Potirakis, “In Vivo Molecular Neoplasia Using Acetic Acid as Biomarker”, IEEE Journal of Selected Topics in Quantum Electronics Vol.14, January/February 2008.
- [7] Ι. Ντζούφρας, “Βιοστατιστική ΙΙ, Μάθημα 6: Διαγνωστικοί έλεγχοι”, Οικονομικό Πανεπιστήμιο Αθηνών, 2006.
- [8] Alan O. Sykes, “ An Introduction to Regression Analysis”, University of Chicago , Foundation Press, 2000.
- [9] Richard A. Berk, “Regression Analysis: A Constructive Critique”, Sage Publications, 2004.
- [10] D. Kleinbaum, L. Kupper, K. Muller , “Applied Regression Analysis and Other Multivariable Methods”, PWS-Kent Publications, 1988.
- [11] R. Shimamoto, J. Suzuki, J.Nishikawa, F. Nakamura, T.Tomaru, T.Toyo-

- oka, “Measuring the Diameter of Coronary Arteries on MR Angiograms Using Spatial Profile Curves”, American Roentgen Ray Society, 1997.
- [12] Forth Photonics Company, www.forth-photonics.gr, Costas Balas: Founder – Chief Technology Officer.
- [13] C. Goodall, “A Survey of Smoothing Techniques”, Modern Methods of Data Analysis, Sage Publications, 1990.
- [14] N.R Draper, H. Smith, “Applied Regression Analysis”, John Wiley & Sons, 1998, third edition.
- [15] A. Papoulis, “Πιθανότητες Τυχαίες Μεταβλητές και Στοχαστικές Διαδικασίες”, Εκδόσεις Τζιόλα, 2002, 3η έκδοση.
- [16] J.W Sellors, R.Sankaranarayanan, “Colposcopy and Treatment of Cervical Intraepithelial Neoplasia: A Beginner's Manual”, IARC, 2003.
- [17] J.L Belinson, R.G.Pretorius, W.Zhang, “Cervical Cancer Screening by Simple Visual Inspection After Acetic Acid”, Elsevier Science Inc., Σεπτέμβριος 2001