

**ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ**  
**ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΚΑΙ ΤΕΧΝΟΛΟΓΙΚΩΝ**  
**ΕΠΙΣΤΗΜΩΝ**  
**ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ**

Βιολογικά Εμπνευσμένη Αρχιτεκτονική για Αυτόνομους  
Πράκτορες Ανάκτησης Εικόνων με βάση το Περιεχόμενο  
τους. Εφαρμογή στις Μαγνητικές Τομογραφίες  
Εγκεφάλου.

Μεταπτυχιακή Εργασία

Ηράκλειο, Φεβρουάριος 2005



ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ  
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΚΑΙ ΤΕΧΝΟΛΟΓΙΚΩΝ  
ΕΠΙΣΤΗΜΩΝ  
ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

**Βιολογικά Εμπνευσμένη Αρχιτεκτονική για Αυτόνομους  
Πράκτορες Ανάκτησης Εικόνων με βάση το Περιεχόμενο  
τους. Εφαρμογή στις Μαγνητικές Τομογραφίες  
Εγκεφάλου.**

Εργασία που υποβλήθηκε από τον  
**Ιωάννη Μουστάκα**  
ως μερική εκπλήρωση των απαιτήσεων για την απόκτηση  
ΜΕΤΑΠΤΥΧΙΑΚΟΥ ΔΙΠΛΩΜΑΤΟΣ ΕΙΔΙΚΕΥΣΗΣ

Συγγραφέας:

\_\_\_\_\_  
**Ιωάννης Μουστάκας**  
Τμήμα Επιστήμης Υπολογιστών  
Πανεπιστήμιο Κρήτης

Εισηγητική Επιτροπή:

\_\_\_\_\_  
**Στέλιος Ορφανουδάκης**, Καθηγητής, Πρόεδρος

\_\_\_\_\_  
**Παναγιώτης Τσακαλίδης**, Αναπληρωτής Καθηγητής, Μέλος

\_\_\_\_\_  
**Κώστας Μαριάς**, Ερευνητής Ινστιτούτου Πληροφορικής, Ι.Τ.Ε., Μέλος

Δεκτή:

\_\_\_\_\_  
**Δημήτρης Πλεξουσάκης**, Αναπληρωτής Καθηγητής  
Πρόεδρος Επιτροπής Μεταπτυχιακών Σπουδών



# **Βιολογικά Εμπνευσμένη Αρχιτεκτονική για Αυτόνομους Πράκτορες Ανάκτησης Εικόνων με βάση το Περιεχόμενο τους. Εφαρμογή στις Μαγνητικές Τομογραφίες Εγκεφάλου.**

**Μουστάκας Ιωάννης**

Μεταπτυχιακή Εργασία

Τμήμα Επιστήμης Υπολογιστών  
Πανεπιστήμιο Κρήτης

## **ΠΕΡΙΛΗΨΗ**

Ο μεγάλος και συνεχώς αυξανόμενος όγκος των οπτικών και πολυμεσικών δεδομένων στο διαδίκτυο αποτελεί ένα σημαντικό κομμάτι πληροφορίας για το οποίο είναι αναγκαίο να δημιουργηθούν αποδοτικές μέθοδοι και εργαλεία ανάκτησης. Ειδικά, στον τομέα της Ιατρικής, ένα μεγάλο μέρος της πληροφορίας συναντάται υπό τη μορφή ιατρικών εικόνων. Η ανάκτηση παρόμοιων ιατρικών περιστατικών με βάση την οπτική πληροφορία (εξειδικευμένη σημασιολογική ομοιότητα) από μία βάση με εμπειριστατωμένα περιστατικά είναι δυνατό να αποτελέσει χρήσιμο βοήθημα στη διάγνωση.

Σε αυτή την εργασία προτείνεται μία πρωτότυπη αρχιτεκτονική για την ανάκτηση εικόνων με βάση το περιεχόμενό τους σε εξειδικευμένα πεδία εφαρμογής. Η αρχιτεκτονική ενσωματώνει στοιχεία από τη θεωρία αυτόνομων πρακτόρων, που αποτελεί μία από τις πλέον σύγχρονες μεθόδους της τεχνητής νοημοσύνης. Περαιτέρω, είναι εμπνευσμένη από τα πειραματικά δεδομένα της ψυχολογίας της οπτικής αντίληψης που υποδεικνύουν ότι η βιολογική όραση επεξεργάζεται την οπτική πληροφορία σε διάφορα στάδια. Συγκεκριμένα προτείνουμε μία αρχιτεκτονική δύο επιπέδων επεξεργασίας και ανάκτησης των εικόνων, εκ των οποίων, το πρώτο διαχειρίζεται τη γενική πληροφορία της εικόνας, ενώ το δεύτερο την πληροφορία, επιλεκτικά, των περιοχών όπου εντοπίζεται το ιδιαίτερο ενδιαφέρον του πεδίου εφαρμογής. Συμπληρωματικά των δύο επιπέδων ανάκτησης, προσφέρεται η υπηρεσία της σταθμισμένης συνδυαστικής ανάκτησης των αποτελεσμάτων τους.

Για την επίδειξη της λειτουργικότητας της προτεινόμενης αρχιτεκτονικής υλοποιήθηκε μία πλατφόρμα ανάκτησης εικόνων δύο επιπέδων, εξειδικευμένη στην εξαγωγή και σύγκριση πληροφορίας από εικόνες μαγνητικής τομογραφίας εγκεφάλου (T1, T2 και PD). Στα πλαίσια αυτής της προσπάθειας προτείνεται μία πρωτότυπη διαδικασία εντοπισμού περιοχών ενδεχόμενης παθολογίας τις οποίες χρησιμοποιεί το δεύτερο, ή επιλεκτικό επίπεδο ανάκτησης. Αρχικά αποτελέσματα αξιολόγησης της πλατφόρμας παρουσιάζονται για διάφορες φυσιολογικές κα παθολογικές εικόνες ερώτησης.

**Επόπτης:** **Στελιος Ορφανουδάκης,**  
Καθηγητής Επιστήμης Υπολογιστών,  
Πανεπιστημίου Κρήτης.



# **A Biologically Inspired Architecture using Autonomous Agents for Content-based Image Retrieval with Application to Brain Magnetic Resonance Images**

**Moustakas Ioannis**

Master of Science Thesis

Department of Computer Science

University of Crete

## **ABSTRACT**

The large and continuously increasing volume of image data in the web dictates the development of specialized techniques for accessing information such as content-based image retrieval (CBIR). One particular application is the development of specialised CBIR platforms for assisting medical diagnosis through the retrieval of relevant cases based on image content (visual semantics) from annotated clinical databases.

This thesis first proposes a novel architecture for content-based image retrieval that involves autonomous agents and is inspired from mechanisms of visual perception. The key idea underlying our work emanates from psychological studies which indicate that the human visual system processes information at several stages. More specifically, we present a two-tier architecture featuring a ‘pre-attentive’ and an ‘attentive’ level of retrieval; the first level manages generic image primitive features, while the second, ‘attentive’ level selectively describes the regions of specialized interest according to the application. Additionally, it also allows CBIR using a weighted combination of the information acquired from the two levels of retrieval.

We are then concerned with implementing a brain-MR CBIR platform, using public domain data (T1, T2 and PD scans) in order to assess the value of the proposed architecture. In order to implement the ‘attentive’ level of retrieval in this specific application, a novel method for automatically detecting asymmetrical regions of potential pathology from brain scans, is also developed. Results are presented for various, normal and pathological query-images while initial validation of the platform is also reported.

**Supervisor: Stelios Orphanoudakis**

Professor of Computer Science

University of Crete





## Ευχαριστίες

Με το πέρας της παρούσας εργασίας θα ήθελα να ευχαριστήσω τα άτομα που συντέλεσαν με οποιονδήποτε τρόπο στην επιτυχημένη ολοκλήρωσή της.

Αρχικά θα ήθελα να ευχαριστήσω τον επόπτη καθηγητή μου κ. Στέλιο Ορφανουδάκη για την επίβλεψη της εργασίας. Ο κ. Ορφανουδάκης αποτελεί στη συνείδηση μου πρότυπο επιστήμονα και ανθρώπου. Με τις ιδέες του και την καθοδήγηση του με βοήθησε να διευρύνω τον τρόπο σκέψης μου και να διαμορφώσω έναν ιδιαίτερο τρόπο αντίληψης που θα με συνοδευεί σε κάθε πτυχή της ζωής μου.

Περαιτέρω, θα ήθελα να ευχαριστήσω τον ερευνητή του Ιδρύματος Τεχνολογίας και Έρευνας (Ι.Τ.Ε) κ. Κώστα Μαρία με τον οποίο συνεργάστηκα πολύ στενά. Οι συμβουλές και η βοήθειά του ήταν καθοριστικές σε όλες τις αποφάσεις που πάρθηκαν στα πλαίσια της παρούσας εργασίας. Επίσης, θα ήθελα να ευχαριστήσω τον καθηγητή κ. Παναγιώτη Τσακαλίδη που ως μέλος της επιτροπής, με τις σωστές υποδείξεις και συμβουλές του συντέλεσε στην ορθή ολοκλήρωσή της.

Θα ήταν παράλειψη να μην αναφερθώ ιδιαίτερα στο Σωκράτη Δημητριάδη ο οποίος ως παλαιότερος μεταπτυχιακός φοιτητής βοήθησε σημαντικά την προσπάθειά μου. Η μεταπτυχιακή του εργασία αποτέλεσε πηγή έμπνευσης στη παρούσα εργασία. Επίσης θα ήθελα να αναφερθώ ιδιαίτερα στους συμφοιτητές και φίλους μου Καλλιρόη Φλουρή, Μιγάλη Μανιαδάκη και Γιώργο Καζαζάκη για όποια βοήθεια μου προσέφεραν.

Κλείνοντας αυτό το σύντομο σημείωμα θα ήθελα να ευχαριστήσω το Ίδρυμα Τεχνολογίας και Έρευνας που μου παρείχε την οικονομική υποστήριξη και την υλικοτεχνική υποδομή κατά τη διάρκεια των μεταπτυχιακών μου σπουδών. Στον ιδιαίτερο χώρο του εργαστηρίου Υπολογιστικής Όρασης και Ρομποτικής συνάντησα άτομα με κοινά ενδιαφέροντα τα οποία επέδειξαν μεγάλη προθυμία να μοιραστούν τις ερευνητικές προκλήσεις της προσωπικής μου εργασίας. Θα ήθελα λοιπόν να ευχαριστήσω όλα τα μέλη του εργαστηρίου μου και να τους ευχηθώ καλή σταδιοδρομία.



# Περιεχόμενα

<b>1</b>	<b>ΕΙΣΑΓΩΓΗ</b> .....	<b>1</b>
1.1	ΠΕΡΙΕΧΟΜΕΝΟ ΕΙΚΟΝΩΝ ΚΑΙ ΣΗΜΑΣΙΟΛΟΓΙΚΗ ΑΝΑΚΤΗΣΗ .....	2
1.2	ΠΑΡΟΥΣΑ ΜΕΛΕΤΗ ΚΑΙ ΥΛΟΠΟΙΗΣΗ.....	4
1.2.1	<i>Η προτεινόμενη αρχιτεκτονική</i> .....	4
1.2.2	<i>Εφαρμογή στις μαγνητικές τομογραφίες εγκεφάλου</i> .....	7
1.3	ΟΡΓΑΝΩΣΗ ΤΟΥ ΚΕΙΜΕΝΟΥ.....	9
<b>2</b>	<b>ΒΙΟΛΟΓΙΚΗ ΏΡΑΣΗ ΚΑΙ ΕΥΦΥΪΑ</b> .....	<b>11</b>
2.1	ΝΕΥΡΟΕΠΙΣΤΗΜΕΣ ΚΑΙ ΒΙΟΛΟΓΙΚΗ ΏΡΑΣΗ .....	11
2.1.1	<i>Η πορεία της οπτικής πληροφορίας και η εξειδίκευση, μέσα στον οπτικό φλοιό</i> .....	13
2.2	ΠΕΙΡΑΜΑΤΙΚΗ ΨΥΧΟΛΟΓΙΑ ΚΑΙ ΟΠΤΙΚΗ ΑΝΤΙΛΗΨΗ.....	14
2.2.1	<i>Πρώιμο στάδιο επεξεργασίας: preattentive level of cognition</i> .....	14
2.2.2	<i>Στάδιο επιλεκτικής επεξεργασίας: attentive level of cognition</i> .....	16
2.2.3	<i>Αρχές οπτικής οργάνωσης και μη τυχαίες ιδιότητες</i> .....	18
2.3	ΒΙΟΛΟΓΙΚΗ ΏΡΑΣΗ ΚΑΙ ΕΥΦΥΪΑ .....	21
2.3.1	<i>Η έννοια της ευφύιας</i> .....	21
2.3.2	<i>Μοντελοποίηση της ευφύιας - Η θεωρία των αυτόνομων πρακτόρων</i> .....	22
2.3.3	<i>Ώραση, μία ευφύης διαδικασία</i> .....	24
<b>3</b>	<b>ΣΥΓΧΡΟΝΑ ΣΥΣΤΗΜΑΤΑ ΑΝΑΚΤΗΣΗΣ ΕΙΚΟΝΩΝ ΜΕ ΒΑΣΗ ΤΟ ΠΕΡΙΕΧΟΜΕΝΟ ..</b>	<b>27</b>
3.1	ΑΝΑΚΤΗΣΗ ΕΙΚΟΝΩΝ ΓΕΝΙΚΟΥ ΠΕΡΙΕΧΟΜΕΝΟΥ .....	27
3.2	ΑΝΑΚΤΗΣΗ ΕΙΚΟΝΩΝ ΣΤΗΝ ΙΑΤΡΙΚΗ .....	30
3.2.1	<i>Ιδιαίτερα χαρακτηριστικά ανάκτησης ιατρικών εικόνων</i> .....	31
3.2.2	<i>Χρησιμότητα του CBIR σε κλινικές εφαρμογές</i> .....	33
3.2.3	<i>Σημαντικά συστήματα ανάκτησης ιατρικών εικόνων</i> .....	35
<b>4</b>	<b>Η ΠΡΟΤΕΙΝΟΜΕΝΗ ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΑΝΑΚΤΗΣΗΣ ΕΙΚΟΝΩΝ</b> .....	<b>37</b>
4.1	ΠΑΡΟΥΣΙΑΣΗ ΤΩΝ ΒΑΣΙΚΩΝ ΛΕΙΤΟΥΡΓΙΩΝ ΤΟΥ ΣΥΣΤΗΜΑΤΟΣ.....	37
4.1.1	<i>Πράκτορες ανάλυσης και σύγκρισης οπτικών χαρακτηριστικών, η λειτουργική μονάδα του συστήματος</i> .....	39
4.1.2	<i>Ομάδες πρακτόρων για αποφάσεις οπτικής ομοιότητας</i> .....	39
4.2	ΒΙΟΛΟΓΙΚΑ ΕΜΠΝΕΥΣΜΕΝΗ ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΔΥΟ ΕΠΙΠΕΔΩΝ ΑΝΑΛΥΣΗΣ ΚΑΙ ΑΝΑΚΤΗΣΗΣ ...	40
4.2.1	<i>Το πρώιμο επίπεδο περιγραφής και ανάκτησης των εικόνων</i> .....	42
4.2.2	<i>Τμηματοποίηση των εικόνων σε εξειδικευμένες εφαρμογές</i> .....	43
4.2.3	<i>Το Επίπεδο επιλεκτικής περιγραφής και ανάκτησης Εικόνων</i> .....	45
4.2.4	<i>Ολοκλήρωση του σχήματος εξαγωγής πληροφορίας από τις εικόνες</i> .....	46
4.2.5	<i>Μορφή και αποθήκευση της οπτικής πληροφορίας από τους οπτικούς πράκτορες</i> .....	47
4.3	ΙΔΙΟΤΗΤΕΣ ΤΟΥ ΣΥΣΤΗΜΑΤΟΣ.....	48
4.4	ΜΕΛΛΟΝΤΙΚΗ ΕΡΓΑΣΙΑ ΚΑΙ ΣΥΜΠΕΡΑΣΜΑΤΑ .....	49
4.4.1	<i>Θέματα μάθησης</i> .....	49
4.4.2	<i>Θέματα κινητικότητας και επικοινωνίας</i> .....	51
4.4.3	<i>Σημσιολογικά μεταδεδομένα εικόνων και δυνατότητα επέκτασης της αρχιτεκτονικής σε γενικότερες εφαρμογές</i> .....	52
4.4.4	<i>Η προτεινόμενη αρχιτεκτονική στο πεδίο της Ιατρικής</i> .....	52
<b>5</b>	<b>ΕΞΑΓΩΓΗ ΟΠΤΙΚΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΑΠΟ ΜΑΓΝΗΤΙΚΕΣ ΤΟΜΟΓΡΑΦΙΕΣ ΕΓΚΕΦΑΛΟΥ</b> .....	<b>55</b>
5.1	ΜΑΓΝΗΤΙΚΕΣ ΤΟΜΟΓΡΑΦΙΕΣ ΕΓΚΕΦΑΛΟΥ .....	55
5.1.1	<i>Κατηγορίες μαγνητικών τομογραφιών εγκεφάλου</i> .....	55
5.1.2	<i>Η ιδιότητα της συμμετρίας στον ανθρώπινο εγκέφαλο</i> .....	56
5.2	ΕΥΡΕΣΗ ΤΟΥ ΆΞΟΝΑ ΣΥΜΜΕΤΡΙΑΣ ΚΑΙ ΕΥΘΥΓΡΑΜΜΙΣΗ ΕΙΚΟΝΑΣ .....	57

5.2.1	<i>Χρήση της θεωρίας των ροπών αδράνειας – moments</i>	57
5.2.2	<i>Χρήση διαδοχικών δοκιμών εταιροσυσχέτισης</i>	59
5.3	<b>ΕΞΑΓΩΓΗ ΠΛΗΡΟΦΟΡΙΑΣ ΓΙΑ ΤΗ ΓΕΝΙΚΗ ΜΟΡΦΟΛΟΓΙΑ</b>	60
5.3.1	<i>Χρωματική Πληροφορία – ιστόγραμμα</i>	61
5.3.2	<i>Πληροφορία υφής – texture Gabor</i>	62
5.3.3	<i>Αμοιβαία πληροφορία – mutual information</i>	65
5.3.4	<i>Τοπογραφικός χάρτης ασυμμετρίας</i>	67
5.3.5	<i>Σχήμα ως έλλειψη– best fitting ellipse</i>	70
5.4	<b>ΤΜΗΜΑΤΟΠΟΙΗΣΗ ΜΑΓΝΗΤΙΚΩΝ ΤΟΜΟΓΡΑΦΙΩΝ ΕΓΚΕΦΑΛΟΥ</b>	70
5.4.1	<i>Χρωματικός διαχωρισμός με τον αλγόριθμο kmeans</i>	71
5.4.2	<i>Τμηματοποίηση με βάση την ασυμμετρία</i>	72
5.5	<b>ΕΞΑΓΩΓΗ ΠΛΗΡΟΦΟΡΙΑΣ ΓΙΑ ΤΙΣ ΠΕΡΙΟΧΕΣ ΕΝΔΙΑΦΕΡΟΝΤΟΣ</b>	77
5.5.1	<i>Χρωματική πληροφορία – ιστόγραμμα</i>	77
5.5.2	<i>Σχετική θέση</i>	77
5.5.3	<i>Μέγεθος</i>	77
5.5.4	<i>Μορφή</i>	78
5.6	<b>ΣΥΜΠΕΡΑΣΜΑΤΑ</b>	79
<b>6</b>	<b>ΕΦΑΡΜΟΓΗ ΤΗΣ ΑΡΧΙΤΕΚΤΟΝΙΚΗΣ ΑΝΑΚΤΗΣΗΣ ΣΕ ΜΑΓΝΗΤΙΚΕΣ ΤΟΜΟΓΡΑΦΙΕΣ ΕΓΚΕΦΑΛΟΥ</b>	<b>81</b>
6.1	<b>Το ΣΥΣΤΗΜΑ ΠΟΥ ΥΛΟΠΟΙΗΘΗΚΕ</b>	81
6.2	<b>ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΚΤΗΣΗΣ</b>	83
6.2.1	<i>Αποτελέσματα με κριτήριο τη γενική μορφολογία</i>	84
6.2.2	<i>Αποτελέσματα με κριτήριο τις περιοχές ενδιαφέροντος</i>	86
6.2.3	<i>Συνδυαστικά αποτελέσματα</i>	88
<b>7</b>	<b>ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΕΠΙΛΟΓΟΣ</b>	<b>93</b>
<b>8</b>	<b>ΒΙΒΛΙΟΓΡΑΦΙΑ</b>	<b>95</b>

# 1 Εισαγωγή

Ο μεγάλος και συνεχώς αυξανόμενος όγκος των οπτικών και πολυμεσικών δεδομένων στο διαδίκτυο αποτελεί ένα σημαντικό κομμάτι πληροφορίας για το οποίο είναι αναγκαίο να δημιουργηθούν αποδοτικές μέθοδοι και εργαλεία ανάκτησης. Οι ευρέως διαδεδομένες λεκτικές (textual) μέθοδοι ανάκτησης, στις περισσότερες περιπτώσεις, δεν αποτελούν ικανοποιητική λύση, καθώς η περιγραφή μίας εικόνας από συγκεκριμένες λέξεις κλειδιά αδυνατεί να εκφράσει το ουσιαστικό περιεχόμενό της. Στον αντίποδα αυτής της προσέγγισης, η ανάκτηση εικόνων με βάση το οπτικό περιεχόμενό τους (CBIR-Content Based Image Retrieval) φαίνεται να κερδίζει το ενδιαφέρον της ερευνητικής κοινότητας, αποτελώντας τα τελευταία δέκα χρόνια μία από τις περισσότερο ενεργές ερευνητικές περιοχές στο πεδίο της υπολογιστικής όρασης. Περαιτέρω, υβριδικά συστήματα των δύο προσεγγίσεων αποτελούν μία αρκετά ελκυστική λύση.

Το πρόβλημα του CBIR ουσιαστικά διατυπώνεται ως εξής: δοσμένης μίας εικόνας εισόδου αναζητούνται «παρόμοιες» στο περιεχόμενο εικόνες από κάποια πηγή εικόνων. Ωστόσο, *το σημασιολογικό περιεχόμενο των εικόνων είναι υποκειμενικό με αποτέλεσμα η κατανόησή τους να απαιτεί εν γένη κάποια πρότερη γνώση ή ένα σημείο αναφοράς*. Επιπρόσθετα σύμφωνα με τις πλέον σύγχρονες θεωρήσεις περί της έννοιας της ευφυΐας, η αντίληψη του οπτικού περιεχομένου αποτελεί μία ευφυή διαδικασία. Η κατανόηση του περιεχομένου των εικόνων είναι ένα δυσεπίλυτο πρόβλημα, ωστόσο στη φύση, οι βιολογικοί οργανισμοί ανταποκρίνονται με εξαιρετικές επιδόσεις στις απαιτήσεις ενός δυναμικού και σύνθετου οπτικού περιβάλλοντος. Κατά συνέπεια, στην προσπάθεια μοντελοποίησης υπολογιστικών συστημάτων αντίληψης της οπτικής πληροφορίας, αντλείται έμπνευση από τους μηχανισμούς της βιολογικής όρασης.

Πέρα από ερευνητική προσέγγιση της ανάκτησης εικόνων γενικού περιεχομένου που σχετίζεται με την αντίληψη και την κατανόηση της οπτικής πληροφορίας, πολλά είναι τα εξειδικευμένα πεδία εφαρμογής της, όπως η κατηγοριοποίηση εικόνων (image indexing and classification), η ρομποτική πλοήγηση, η ιατρική διαγνωστική, η αναγνώριση προσώπων, η αναζήτηση έργων τέχνης μέσω του ιστού κ.α. Όπως θα γίνει φανερό, συστήματα εξειδικευμένα στην ανάκτηση εικόνων από ένα συγκεκριμένο πεδίο εφαρμογής κάνουν αποδοτική χρήση της πρότερης γνώσης από το συγκεκριμένο πεδίο εικόνων. Επιπρόσθετα τέτοια συστήματα έχουν τη δυνατότητα να προσδιορίσουν με μεγαλύτερη ακρίβεια τους στόχους τους και το βαθμό επίτευξής αυτών

Σε αυτή την εργασία προτείνεται μία νέα αρχιτεκτονική για εξειδικευμένες πλατφόρμες CBIR. Η αρχιτεκτονική ενσωματώνει στοιχεία από τη θεωρία αυτόνομων πρακτόρων, που αποτελεί μία από τις πλέον σύγχρονες μεθόδους της τεχνητής νοημοσύνης. Επιπλέον, είναι εμπνευσμένη από τα πειραματικά δεδομένα της ψυχολογίας για την οπτική αντίληψη. Η επεξεργασία και ανάκτηση των εικόνων λαμβάνει χώρα σε δύο επίπεδα, εκ των οποίων, το πρώτο σχετίζεται με τη γενική πληροφορία της εικόνας ενώ το δεύτερο με τις περιοχές όπου εντοπίζεται το ιδιαίτερο ενδιαφέρον του πεδίου εφαρμογής. Πέρα από τα δύο αυτά επίπεδα ανάκτησης προσφέρεται η υπηρεσία της σταθμισμένης συνδυαστική ανάκτησης των αποτελεσμάτων παρουσιάζει ενδιαφέρουσες ιδιότητες. Για την επίδειξη της λειτουργικότητας της προτεινόμενης αρχιτεκτονικής υλοποιήθηκε μία πλατφόρμα CBIR, εξειδικευμένη στην εξαγωγή πληροφορίας από μαγνητικές τομογραφίες εγκεφάλου (MRI- Magnetic Resonance Images). Αποτελέσματα

παρουσιάζονται τόσο σε επίπεδο ανάκτησης εικόνων όσο και σε επίπεδο ανάλυσης του περιεχομένου των μαγνητικών τομογραφιών του εγκεφάλου.

Σε τούτο το εισαγωγικό κεφάλαιο, αρχικά γίνεται μία σύντομη ανάλυση του προβλήματος της κατανόησης του ουσιαστικού περιεχομένου των εικόνων. Στη συνέχεια, γίνεται μία σύντομη αλλά περιεκτική περιγραφή της προτεινόμενης αρχιτεκτονικής, συνοδευόμενη από χαρακτηριστικά αποτελέσματα της πλατφόρμας για MRIs εγκεφάλου. Τέλος για τη διευκόλυνση του αναγνώστη παρουσιάζεται η οργάνωση που ακολουθεί το υπόλοιπο κείμενο.

## 1.1 Περιεχόμενο Εικόνων και Σημασιολογική Ανάκτηση

Το σημασιολογικό περιεχόμενο των εικόνων είναι σύνθετο και υποκειμενικό. Για παράδειγμα, στον παρακάτω ζωγραφικό πίνακα (Εικόνα 1) είναι πολύ δύσκολο να κατανοήσει κανείς με ακρίβεια της προθέσεις του δημιουργού του. Ωστόσο, υπάρχουν μερικοί κανόνες και γενικές αρχές, διατυπωμένες κυρίως από τη ψυχολογία της οπτικής αντίληψης, βάσει των οποίων είναι δυνατό να προσδιορισθεί το πλαίσιο κατανόησης του οπτικού περιεχομένου.



Εικόνα 1: Παράδειγμα σύνθετου σημασιολογικού περιεχομένου σε ζωγραφικό πίνακα.

Σε ένα αρχικό στάδιο της βιολογικής όρασης η οπτική πληροφορία συναντάται υπό τη μορφή των πρωτογενών οπτικών χαρακτηριστικών (χρώμα, μορφή κίνηση κ.α.). Ωστόσο, το ουσιαστικό νόημα του οπτικού σήματος, εμπεριέχεται στις οργανωμένες διατάξεις των χαρακτηριστικών αυτών, οι οποίες αποτελούν τα αντικείμενα του φυσικού περιβάλλοντος (π.χ. άνθρωπος, βέλος, κένταυρος, κτήριο, πόλη κτλ). Ομοίως, η σημασιολογική ανάκτηση εικόνων, περνά από τον εντοπισμό και σύγκριση της σημασιολογίας των οντοτήτων αυτών. Οι μηχανισμοί οργάνωσης που χρησιμοποιεί η βιολογική όραση παραμένουν σε μεγάλο βαθμό αδιαφανείς, σε επίπεδο νευροεπιστημών και νευροφυσιολογίας. Επιπρόσθετα, το υποκείμενο δεν έχει συνειδητή πρόσβαση σε αυτούς τους μηχανισμούς. Κατά συνέπεια, οι επικρατέστερες απόψεις για τη λειτουργία τους στηρίζονται κυρίως στα πειραματικά δεδομένα και στις αρχές της ψυχολογίας της οπτικής αντίληψης.

Η υπεραρχή της Prögnanz διατυπώνεται ως εξής: «οι απλές και σταθερές ερμηνείες ευνοούνται». Επομένως, οι ιδιότητες που διέπουν συχνότερα τις οργανωμένες διατάξεις στο φυσικό κόσμο, αποτελούν τους κανόνες ομαδοποίησης των πρωτογενών χαρακτηριστικών μίας εικόνας. Τέτοιες ιδιότητες ονομάζονται μη τυχαίες ιδιότητες. Χαρακτηριστικά παραδείγματα μη τυχαίων ιδιοτήτων είναι οι κανόνες Gestalt όπως η ομοιότητα η εγγύτητα η συμμετρία κ.α. Επιπρόσθετα, ένα αντικείμενο που συναντάται συχνά με συγκεκριμένο νόημα είναι πολύ πιθανό και στην εξεταζόμενη εικόνα να διατηρεί την ίδια σημασιολογία (π.χ. «αρχαίος πολεμιστής με βέλος στη φτέρνα»= «Αχιλλέας»). Ωστόσο, το νόημα μίας εικόνας μπορεί να διαφοροποιείται σημαντικά ανάλογα με τη γωνία παρατήρησης. Για την επίδειξη αυτού του προβλήματος στο [1] επιλέχθηκε η παρακάτω πολύ χαρακτηριστική εικόνα (Εικόνα 2).



**Εικόνα 2: Παράδειγμα διαφοροποίησης της σημασιολογία των εικόνων ανάλογα με τη γωνία παρατήρησης**

Τρία είναι τα βασικά προβλήματα που καλείται να επιλύσει ένα υπολογιστικό σύστημα για την κατανόηση του περιεχομένου μίας εικόνας. Πρώτον, η τμηματοποίηση των αντικειμένων που απεικονίζονται, δεύτερον η κατανόηση της σημασιολογίας τους και τρίτον ο συσχετισμός τους με το συνολικό νόημα της εικόνας. Μολονότι τα τρία αυτά προβλήματα είναι εξαιρετικά δυσεπίλυτα απουσία ενός πλαισίου αναφοράς, σε εξειδικευμένες εφαρμογές είναι δυνατόν να δοθούν ικανοποιητικές λύσεις. Σε εξειδικευμένες εφαρμογές υπάρχει η δυνατότητα καθορισμού συγκεκριμένων στόχων. Επιπρόσθετα, μπορούν να καθοριστούν οι ιδιότητες που διέπουν τις περιοχές όπου εντοπίζεται το ειδικό ενδιαφέρον της εφαρμογής. Κατά συνέπεια, αντί για την τμηματοποίηση και κατανόηση του συνολικής πληροφορίας μίας εικόνας, είναι δυνατό να χρησιμοποιηθούν απ' ευθείας διαδικασίες εντοπισμού μόνο των αντικειμένων εκείνων που εμφανίζουν τις προκαθορισμένες ιδιότητες. Εν τέλει, είναι δυνατό να προσδιορισθεί ο βαθμός επίτευξης των αντικειμενικών πλέον στόχων της εφαρμογής και να γίνει η αξιολόγηση του αποτελέσματος. Στον αντίποδα, λόγω της υποκειμενικότητας του οπτικού περιεχομένου, είναι δύσκολο να καθοριστούν αντικειμενικά κριτήρια αξιολόγησης των συστημάτων ανάκτησης εικόνων γενικού περιεχομένου. Γενικότερα, *συστήματα όπου το αποτέλεσμα είναι υποκειμενικό πρέπει να είναι ευφυή ή αλληλεπιδραστικά, ενώ συστήματα με συγκεκριμένο στόχο πρέπει να είναι αποδοτικά.* Στη συγκεκριμένη περίπτωση ως ευφύς νοείται το σύστημα που μέσω της

προσαρμοστικότητας του απέναντι σε ένα δυναμικό οπτικό περιβάλλον καταφέρνει να αντιληφθεί την ύπαρξη του ουσιαστικού υποκείμενου νοήματος. Επίσης με τον όρο αλληλεπίδραση, γίνεται αναφορά στην ουσιαστική επικοινωνία με το χρήστη με στόχο την προσαρμογή του αποτελέσματος ανάκτησης στις εκάστοτε προτιμήσεις και στόχους του. Κατά συνέπεια, η αποδοτική ανάκτηση εικόνων γενικού περιεχομένου είναι ένα εξαιρετικά δυσεπίλυτο πρόβλημα, απουσία τέτοιας αλληλεπίδρασης, ενώ περισσότερο ρεαλιστικά είναι τα εξειδικευμένα σε συγκεκριμένο πεδίο συστήματα ανάκτησης.

## **1.2 Παρούσα Μελέτη και Υλοποίηση**

Η παρούσα εργασία αφορά το σχεδιασμό μίας πρωτότυπης αρχιτεκτονικής για συστήματα ανάκτησης εικόνων και την υλοποίηση της αντίστοιχης πλατφόρμας για εξειδικευμένα πεδία εικόνων. Καθώς ο εντοπισμός και η κατανόηση του ουσιαστικού περιεχομένου των εικόνων απαιτούν ευφυή αντιμετώπιση, μελετήθηκαν οι σύγχρονες μέθοδοι τεχνητής νοημοσύνης. Περαιτέρω, μελετήθηκαν οι μηχανισμοί της βιολογικής όρασης και ενσωματώθηκε μέρος των δεδομένων από τις επιστήμες της νευροφυσιολογίας και της ψυχολογίας της οπτικής αντίληψης.

### **1.2.1 Η προτεινόμενη αρχιτεκτονική**

Η πλέον ελπιδοφόρα κατεύθυνση της τεχνητής νοημοσύνης, οι αυτόνομοι πράκτορες, κάνουν λόγο για αυθύπαρκτες οντότητες που προσπαθούν να εκπληρώσουν ένα σύνολο από στόχους σε ένα δυναμικό περιβάλλον [2]-[6]. Στο ανώτατο επίπεδο περιγραφής οι βασικότερες παράμετροι της προτεινόμενης αρχιτεκτονικής οι ακόλουθες:

- Καθορισμός ενός δυναμικού περιβάλλοντος εικόνων
- Συνεχής-δια βίου αναζήτηση εικόνων του πεδίου εφαρμογής στις δυναμικές πηγές εικόνων από έναν αυτόνομο πράκτορα αναζήτησης. Ο ίδιος πράκτορας είναι υπεύθυνος και για την εξυπηρέτηση αιτήσεων ανάκτησης εικόνων από απομακρυσμένους πελάτες και συχνά θα ονομάζεται υπέρ-πράκτορας
- Ανάλυση των εικόνων και αποθήκευση της οπτικής πληροφορίας τους από ένα σύνολο ανεξάρτητων οπτικών υπο-πρακτόρων. Οι εικόνες αυτές αποτελούν τη βάση εικόνων του υπέρ-πράκτορα και το σύνολο των προς ανάκτηση εικόνων σε κάθε αίτηση ανάκτησης. Η δια-βίου εισαγωγή (αναζήτηση- ανάλυση- αποθήκευση) των εικόνων γίνεται σε νεκρό χρόνο και εγκυάται την αποδοτική εξυπηρέτηση των αιτήσεων ανάκτησης
- Δυναμική διεπαφή χρήστη. Οι ανακτηθείσες εικόνες επιστρέφονται στη διεπαφή του πελάτη μία προς μία, τοποθετούμενες στη σωστή θέση σύμφωνα με φθίνουσα σειρά ομοιότητας ως προς την εικόνα ερώτησης. Η διαδικασία τερματίζεται όταν εξετασθούν όλες οι εικόνες στη βάση του πράκτορα είτε με εντολή του χρήστη.
- Συνεχής προσαρμογή στα χαρακτηριστικά του πεδίου εικόνων της εφαρμογής

Οι ιδιότητες που χαρακτηρίζουν έναν πράκτορα έχουν προκύψει από τη μελέτη των μηχανισμών και των ιδιοτήτων που παρουσιάζουν οι βιολογικοί οργανισμοί. Ο Brooks στην αρχιτεκτονική *subsumption architecture*, πρότεινε ένα σύστημα παράλληλων, ασύγχρονων, ελαφρά συνδεδεμένων διεργασιών για τη μοντελοποίηση των πρώιμων



σταδίων στις σύνθετες βιολογικές διεργασίες. Τέτοια συστήματα είναι ευρέως γνωστά ως συστήματα πρακτόρων (multi-agent systems). Κάθε διεργασία σε αυτό το μοντέλο αναλαμβάνει ένα μικρό ανεξάρτητο υπολογιστικό κομμάτι του συνολικού προβλήματος. Σύμφωνα με τα παραπάνω, σε επίπεδο οπτικής επεξεργασίας, καθορίστηκαν τα ακόλουθα:

- Η υπολογιστική μονάδα διαχείρισης της οπτικής πληροφορίας είναι οι οπτικοί πράκτορες. Κάθε τέτοιος πράκτορας αναλαμβάνει τη διαχείριση (εξαγωγή, αποθήκευση) ενός συγκεκριμένου χαρακτηριστικού της εικόνας ή μίας υποπεριοχής (π.χ. χρώμα, υφή, συμμετρία κτλ). Επίσης, κατά την εξυπηρέτηση μίας αίτησης ανάκτησης αναλαμβάνει την ανεξάρτητη σύγκριση του χαρακτηριστικού αυτού ανάμεσα σε δύο εικόνες ή περιοχές των εικόνων.
- Ομάδες οπτικών πρακτόρων είναι επιφορτισμένες με τη συνολική περιγραφή της πληροφορίας μίας εικόνας ή υποπεριοχής της. Σε μία τέτοια ομάδα οι πράκτορες αποτελούν ανεξάρτητες παράλληλες διεργασίες. Κατά την εξυπηρέτηση μίας αίτησης ανάκτησης οι ομάδες πρακτόρων αποφασίζουν για τη συνολική ομοιότητα ανάμεσα σε δύο εικόνες ή δυο υποπεριοχές τους. Για αυτό το σκοπό τα αποτελέσματα από τους επιμέρους πράκτορες συμψηφίζονται.
- Η ψηφοφορία απαιτεί την κανονικοποίηση των ανεξάρτητων τιμών σύγκρισης από τους επιμέρους οπτικούς πράκτορες. Για το σκοπό αυτό κάθε πράκτορας διαθέτει ενσωματωμένη δια βίου διαδικασίες απόκτησης στατιστικής εμπειρίας από τα αποτελέσματα των συγκρίσεων που διενεργεί. Καθώς το σύστημα αποκτά εμπειρία μέσω ανακτήσεων, τα μεγέθη αυτά σταθεροποιούνται και αντικατοπτρίζουν την κλίμακα των τιμών του κάθε πράκτορα για εικόνες της συγκεκριμένη εφαρμογής.
- Κάθε οπτικός πράκτορας σε μία ομάδα πρακτόρων έχει συγκεκριμένη συνεισφορά στο αποτέλεσμα της ομοιότητας, η οποία καθορίζεται εφ' άπαξ κατά την εκκίνηση του υπερ-πράκτορα και σχετίζεται με τους εξειδικευμένους στόχους του πεδίου εφαρμογής. Ο δυναμικός προσδιορισμός των βαρών αυτών ως αποτέλεσμα της απόκτησης εμπειρίας και σε ένα ανώτερο επίπεδο προσαρμογής των ομάδων στα δεδομένα του πεδίου εφαρμογής είναι ένα από τα βασικότερα θέματα που αφήνονται για μελλοντική εργασία.

Ο εμπλουτισμός της αρχιτεκτονικής με επιπρόσθετα στοιχεία ευφυΐας που αφήνονται για μελλοντική εργασία καθώς και τα θέματα επέκτασης και ολοκλήρωσης της πλατφόρμας με το εκάστοτε πεδίο εφαρμογής της, σχολιάζονται στην παράγραφο 4.4.

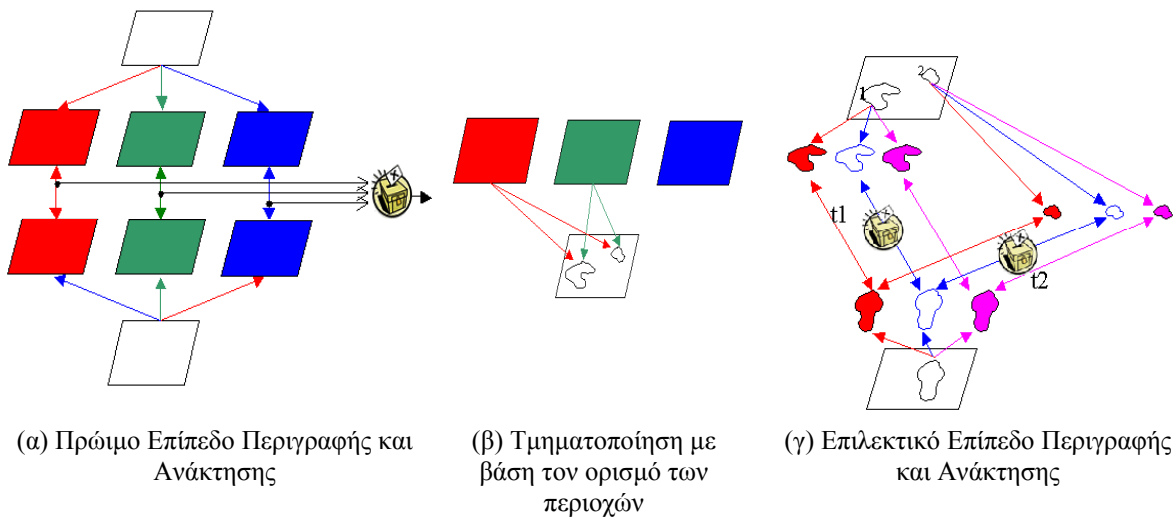
Στην προτεινόμενη αρχιτεκτονική η περιγραφή της οπτικής πληροφορίας από τις ομάδες πρακτόρων λαμβάνει χώρα σε δύο επίπεδα σύμφωνα με δεδομένα από τη μελέτη της βιολογικής όρασης:

- Στα πρώιμα στάδια της όρασης τα διάφορα πρωτογενή χαρακτηριστικά της οπτικής σκηνής (χρώμα, κίνηση, βάθος, ακμές κ.α.) συναντώνται με τη μορφή ανεξάρτητων τοπογραφικών χαρτών. Τα διάφορα χαρακτηριστικά εξετάζονται σχεδόν ταυτόχρονα και παράλληλα σε κάθε χάρτη, ενώ η επεξεργασία γίνεται ταχύτερα και σε σταθερό χρόνο.
- Τα πρωτογενή από τους ανεξάρτητους χάρτες με κάποιο τρόπο οργανώνονται ώστε να επανασυνθέσουν τις σημασιολογικές οντότητες της οπτικής σκηνής.

- Η αντίληψη και κατανόηση της οπτικής σκηνής βασίζεται κυρίως στην επιλεκτική και σειριακή εξέταση των σημασιολογικών αντικειμένων που «τραβούν» την προσοχή του υποκειμένου.

Σύμφωνα με τα παραπάνω ένα μοντέλο δύο ή περισσότερων επιπέδων επεξεργασίας αρχίζει να γίνεται αποδεκτό από την κοινωνία της ψυχολογίας της οπτικής αντίληψης. Τα δύο επίπεδα που μέχρι στιγμής διαχωρίζονται είναι το πρώιμο επίπεδο (preattentive level) με τους παράλληλους χάρτες και το επιλεκτικό (attentive level) της σειριακής επιλογής και προσεκτικής των περιοχών ενδιαφέροντος. Κατά συνέπεια στην προτεινόμενη αρχιτεκτονική κρίθηκαν απαραίτητα δύο τουλάχιστον επίπεδα ανάλυσης και ανάκτησης των εικόνων, το πρώιμο και το επιλεκτικό:

- Στο πρώιμο επίπεδο (Εικόνα 3 (α)), μία ομάδα οπτικών πρακτόρων αναλαμβάνει την γενική περιγραφή των εικόνων και την σύγκριση τους σε επίπεδο γενικής ομοιότητας
- Στο επιλεκτικό επίπεδο (Εικόνα 3γ)), μία δεύτερη ομάδα οπτικών πρακτόρων αναλαμβάνει την περιγραφή των περιοχών ενδιαφέροντος των εικόνων και την σύγκριση τους κατά την εξυπηρέτηση αιτήσεων ανάκτησης
- Η χρήση του συστήματος σε εξειδικευμένες εφαρμογές, ώστε να είναι διαθέσιμος ο ορισμός των περιοχών όπου περιέχεται το ουσιαστικό περιεχόμενο της εικόνας, επιτρέπει τον αποδοτικό εντοπισμό τους. Με βάση τον ορισμό αυτό, κριτήρια (Εικόνα 3β)) εφαρμόζονται πάνω στους πρώιμους τοπογραφικούς χάρτες



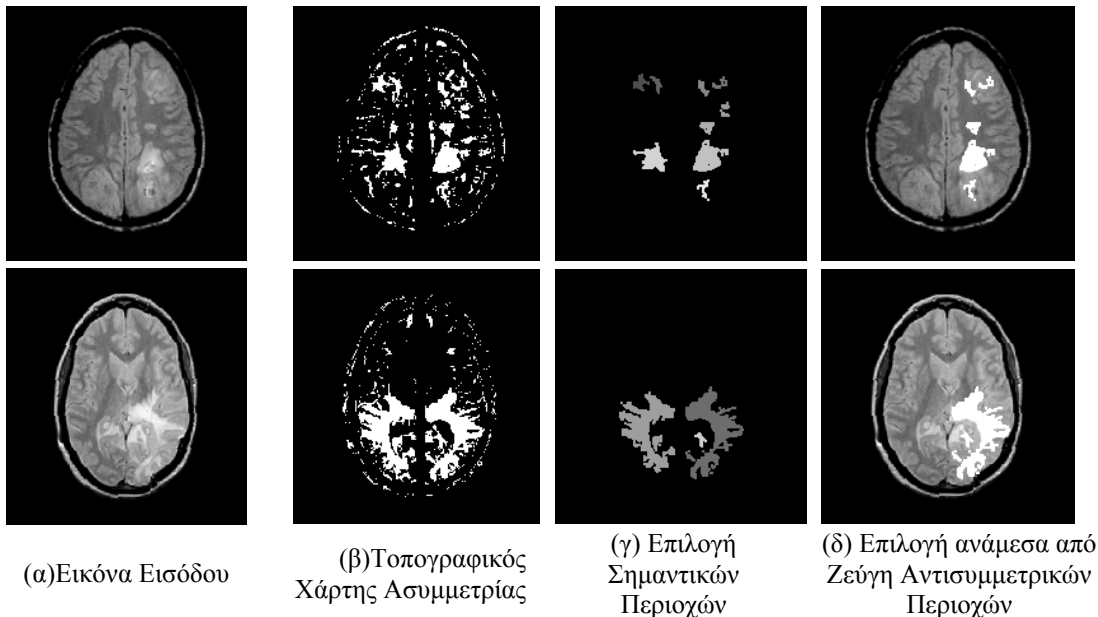
**Εικόνα 3: Η προτεινόμενη αρχιτεκτονική δύο επιπέδων περιγραφής και ανάκτησης των εικόνων**

Ο χρήστης μπορεί να διεξάγει ανάκτηση σε κάθε επίπεδο ξεχωριστά ενώ το σύστημα προβλέπει και επιλογή σταθμισμένο συνδυασμού στα αποτελέσματα των δύο επιπέδων. Οι τρεις αυτές υπηρεσίες της πλατφόρμας παρουσιάζουν ενδιαφέρουσες ιδιότητες ως προς την πληροφορία ανάκτησης. Η αρχιτεκτονική είναι απόλυτα επεκτάσιμη ως προς την προσθαφαίρεση οπτικών πρακτόρων και ομάδων πρακτόρων. Περαιτέρω, έχοντας την πεποίθηση ότι *το πρόβλημα έχει τεθεί στις σωστές βάσεις αυτονομίας, είναι εύκολο να εμπλουτιστεί με επιπρόσθετα στοιχεία συνεχούς μάθησης και προσαρμογής όπως και με ιδιότητες από τη θεωρία αυτόνομων πρακτόρων.*

## 1.2.2 Εφαρμογή στις μαγνητικές τομογραφίες εγκεφάλου

Το πεδίο εικόνων που επιλέχθηκε για την εφαρμογή, επίδειξη και αξιολόγηση της συγκεκριμένης αρχιτεκτονικής CBIR είναι οι μαγνητικές τομογραφίες εγκεφάλου και συγκεκριμένα αυτές που απεικονίζουν την συμμετρία του ανθρώπινου εγκεφάλου. Δύο είναι οι βασικότεροι λόγοι αυτής της επιλογής. Πρώτον, η χρησιμότητα και κοινωνική απήχηση τέτοιων συστημάτων στην ιατρική και κυρίτερα στη ιατρική διαγνωστική, την εκπαίδευση και την έρευνα. Δεύτερον, η συστηματική δουλειά που έχει γίνει στον ιατρικό τομέα, όσον αφορά την ιατρική οντολογία (UMLS, MeSH), τα σημασιολογικά μεταδεδομένα που συνοδεύουν τις εικόνες (DICOM) και άλλα ιδιαίτερα χαρακτηριστικά που όπως γίνει κατανοητό μας επιτρέπουν να επεκτείνουμε την εφαρμογή από ένα συγκεκριμένο πεδίο εφαρμογής (MR εγκεφάλου) στο γενικότερο πεδίο της Ιατρικής. Ένας από τους βασικούς μας στόχους είναι η ενσωμάτωση του προτεινόμενης πλατφόρμας στα σύγχρονα ολοκληρωμένα ιατρικά συστήματα.

Μετά από μια αρχική επεξεργασία που αφορά την ευθυγράμμιση των εικόνων ως προς το κέντρο βάρους και τον άξονα συμμετρίας του υποκείμενου εγκεφάλου, αναζητούνται χαρακτηριστικά για την εκφραστική πληροφορία τόσο της συνολικής μορφολογίας-ανατομίας του εγκεφάλου όσο και των επιμέρους περιοχών ενδιαφέροντος. Στο πρώιμο επίπεδο, πέρα από τα κλασσικά πρωτογενή χαρακτηριστικά για σύγκριση των εικόνων (αποχρώσεις του γκρι, υφή, σχήμα φιγούρας εγκεφάλου, αμοιβαία πληροφορία κ.α. ) παρουσιάζεται ένας πρωτότυπος τοπογραφικός χάρτης απεικόνισης των περιοχών που προκαλούν ασυμμετρία. (Εικόνα 4(β)). Η καλή συμμετρία αποτελεί μία χαρακτηριστική ιδιότητα του υγιούς ανθρώπινου εγκεφάλου. Συνεπαγωγικά, περιοχές που προκαλούν ασυμμετρία αποτελούν πιθανές περιοχές παθολογίας, για αρκετές από τις γνωστές παθήσεις (π.χ. καρκίνος, εγκεφαλικό, κ.α.). Οι τοπογραφικοί χάρτες ασυμμετρίας, με κατάλληλη επεξεργασία (Εικόνα 4(γ)), αναγάγουν το πρόβλημα του εντοπισμού των περιοχών ενδιαφέροντος στην επιλογή ανάμεσα από ζεύγη αντισυμμετρικών περιοχών.

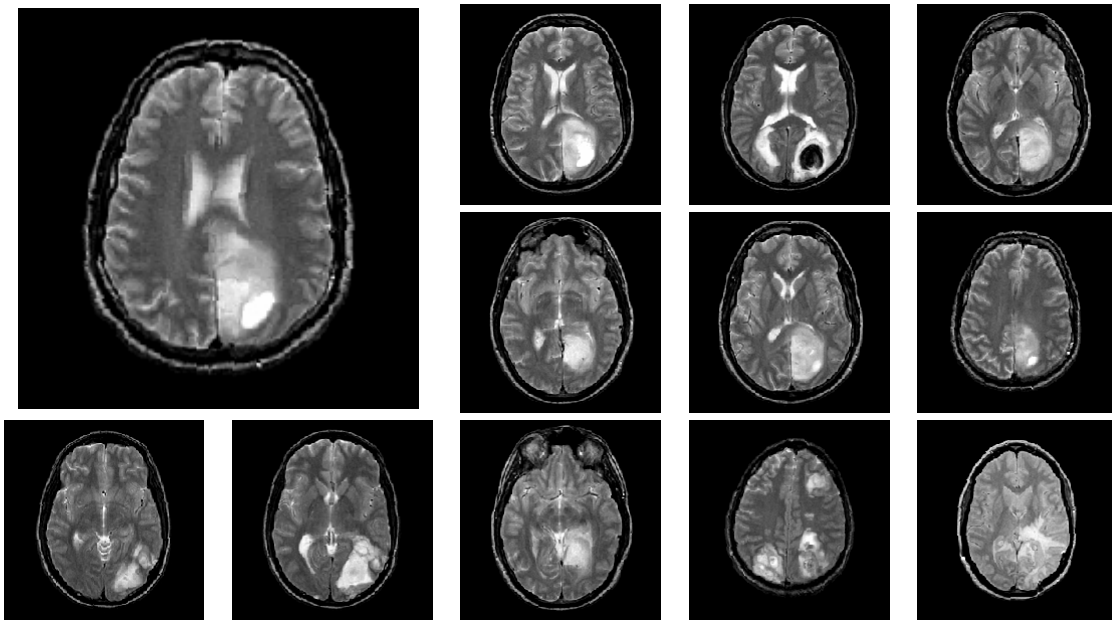


**Εικόνα 4: Χρήση τοπογραφικών χαρτών ασυμμετρίας για την τμηματοποίηση MRIs εγκεφάλου**

Η τελική επιλογή ανάμεσα από δύο αντισυμμετρικές περιοχές περιλαμβάνει κριτήρια εσωτερικής και εξωτερικής ομοιογένειας. Περιοχές με περισσότερο ξεκάθαρο περίγραμμα όπως και περιοχές με καλύτερη εσωτερική ομοιογένεια ευνοούνται. Με αυτόν τον τρόπο όπως δείχνουν τα χαρακτηριστικά παραδείγματα (Εικόνα 4), επιτυγχάνεται η τμηματοποίηση των περιοχών που προκαλούν την ασυμμετρία. Διάφορες παραλλαγές της προτεινόμενης τμηματοποίησης των MRIs εγκεφάλου είναι δυνατό να βασιστούν στη χρήση των τοπογραφικών χαρτών ασυμμετρίας.

Έχοντας υλοποιήσει μία ικανοποιητική και πρωτότυπη διαδικασία εντοπισμού των πιθανών παθήσεων, αναζητούνται κατάλληλα χαρακτηριστικά για το επίπεδο της επιλεκτικής περιγραφής. Επιλέγονται χαρακτηριστικά όπως η φωτεινότητα, το απλό σχήμα, το μέγεθος και η σχετική θέση της κάθε περιοχής μέσα στον εγκέφαλο. Ωστόσο για μία ουσιαστικότερη αντιμετώπιση κρίνεται αναγκαία η συνεργασία με εξειδικευμένους ιατρούς και ο εμπλουτισμός της εφαρμογής με περισσότερο εκφραστικούς περιγραφείς.

Στη συνέχεια ακολουθεί μία σειρά πειραμάτων ανάκτησης μαγνητικών τομογραφιών εγκεφάλου. Διεξάγονται πειράματα στο πρώτο επίπεδο γενικής μορφολογίας, στο δεύτερο επίπεδο ομοιότητας των περιοχών ενδιαφέροντος και σε επίπεδο του σταθμισμένου συνδυασμού των δύο πρώτων. Χαρακτηριστικά αποτελέσματα παρουσιάζονται στην παράγραφο 6.2 καθώς και μία στατιστική επεξεργασία σημαντικών μετρήσεων. (Εικόνα 5).



**Εικόνα 5:** Αποτελέσματα συνδυαστικής ανάκτησης (50-50) για την εικόνα εισόδου (επάνω αριστερά). 10 καλύτερες εικόνες από σύνολο 100 MRIs (67 εμφανίζουν μορφώματα, 33 χωρίς ευρήματα)

### **1.3 Οργάνωση του Κειμένου**

Το υπόλοιπο κείμενο οργανώνεται ως εξής:

Στο κεφάλαιο 2 γίνεται μια σύντομη αναφορά στους εσωτερικούς μηχανισμούς των πρώιμων σταδίων της βιολογικής όρασης, όπως προκύπτουν από τα επιστημονικά δεδομένα της νευροφυσιολογίας. Στη συνέχεια περιγράφονται πειραματικές προσπάθειες κατανόησης αυτών των μηχανισμών, δανεισμένες από τη ψυχολογία της όρασης. Απώτερος, σκοπός της μελέτης είναι η άντληση έμπνευσης για το σχεδιασμό ενός μοντέλου ανάλυσης, κατανόησης και ανάκτησης εικόνων με βάση το περιεχόμενο τους. Σε αυτή τη προσπάθεια επιβεβαιώνεται πως η όραση αποτελεί μια ευφυή διαδικασία. Συνεπώς, οποιαδήποτε προσπάθεια μοντελοποίησής της περνά μέσα από τις σύγχρονες κατευθύνσεις της τεχνητής νοημοσύνης. Μέσω μίας φιλοσοφικής αναζήτησης της έννοιας της ευφυΐας καταλήγουμε σε μία από τις πλέον σύγχρονες και αποδεκτές θεωρίες τεχνητής ευφυΐας τη θεωρία των αυτόνομων ή ευφύων πρακτόρων.

Στο 3<sup>ο</sup> κεφάλαιο παρουσιάζονται τα σημαντικότερα συστήματα ανάκτησης εικόνων που συναντά κανείς στη βιβλιογραφία. Επίσης, γίνεται μία σύντομη αναφορά στα βασικότερα προβλήματα που παρουσιάζουν και τις διάφορες μεθόδους αντιμετώπισης τους. Καθώς το πεδίο εφαρμογής της παρούσας εργασίας είναι οι ιατρικές εικόνες, γίνεται ιδιαίτερος λόγος για τα συστήματα ανάκτησης ιατρικών εικόνων και για τα ιδιαίτερα χαρακτηριστικά αυτού του πεδίου.

Στο κεφάλαιο 4 γίνεται η παρουσίαση της προτεινόμενης, βιολογικά έμπνευσμένης αρχιτεκτονικής για εξειδικευμένα συστήματα ανάκτησης εικόνων, που θέτει τις κατάλληλες βάσεις για μια περισσότερο ευφυή προσέγγιση με βάση τη θεωρία των αυτόνομων πρακτόρων.

Το 5<sup>ο</sup> κεφάλαιο είναι αφιερωμένο στην μελέτη των μαγνητικών τομογραφιών εγκεφάλου και συγκεκριμένα στην εξαγωγή οπτικής πληροφορίας με νόημα από αυτές. Οι διάφοροι αλγόριθμοι που παρουσιάζονται είναι κατάλληλοι για χρήση από την προτεινόμενη αρχιτεκτονική ανάκτησης εικόνων.

Στο κεφάλαιο 6 γίνεται η παρουσίαση του πλατφόρμας που υλοποιήθηκε για την ανάκτηση των μαγνητικών τομογραφιών εγκεφάλου. Αναφέρονται οι βασικές παράμετροι του συστήματος και επιδεικνύονται τα βασικά τμήματα της διεπαφής του. Τέλος παρουσιάζονται και σχολιάζονται διάφορα χαρακτηριστικά πειραματικά αποτελέσματα ανακτήσεων που επιδεικνύουν την λειτουργικότητα και απόδοση του συστήματος.



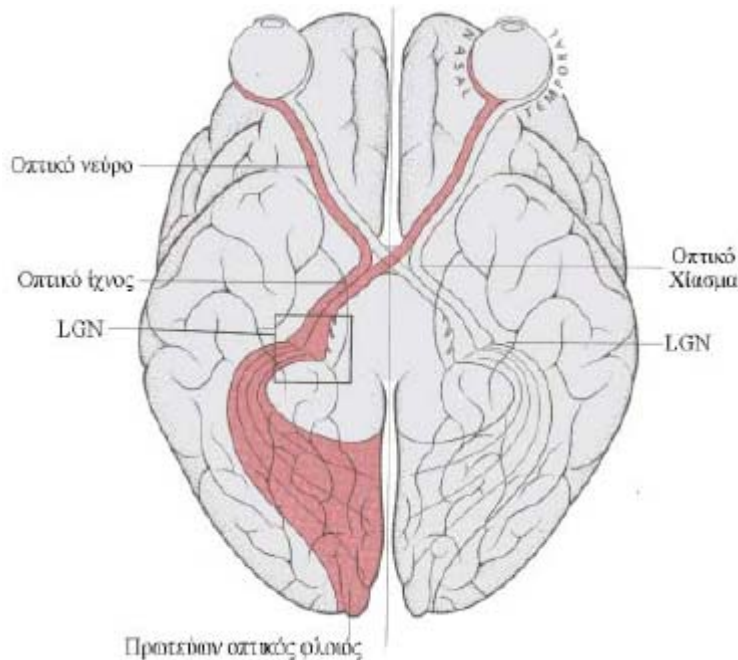
## 2 Βιολογική Όραση και Ευφυΐα

Το φως εισέρχεται στα μάτια μας ως ένα συνεχές πλέγμα από σημεία φωτεινότητας, προερχόμενο ως επί το πλείστο μέσω ανακλάσεων από αντικείμενα και επιφάνειες της οπτικής σκηνής, μεταφέροντας με αυτόν τον τρόπο την τεράστια πληροφορία για τη δομή του περιβάλλοντος χώρου. Η αναγνώριση και αντιστοίχιση των οργανωμένων προτύπων του φωτός σε συμπαγή σημασιολογικά αντικείμενα αποτελεί το ουσιαστικότερο καθήκον της βιολογικής όρασης. Ωστόσο, η αντίληψη σημασιολογικού περιεχομένου από το οπτικό ερέθισμα είναι μία ευφυής και σύνθετη διαδικασία, στους μηχανισμούς της οποίας το υποκείμενο δεν έχει συνειδητή πρόσβαση. Αν και στις μέρες μας η συστηματική μελέτη για την κατανόηση αυτών των μηχανισμών έχει ενταθεί, με διάφορους κλάδους της επιστήμης (νευροεπιστήμες, ψυχολογία της όρασης, γνωσιακή επιστήμη κ.α.) να ολοκληρώνουν τα επιστημονικά δεδομένα και τις ενδείξεις τους, το μεγαλύτερο και το πιο ουσιαστικό κομμάτι της βιολογικής όρασης παραμένει αδιαφανές.

Σε τούτο το κεφάλαιο αρχικά γίνεται μια σύντομη αναφορά στους εσωτερικούς μηχανισμούς των πρώιμων σταδίων της βιολογικής όρασης όπως προκύπτουν από τα επιστημονικά δεδομένα της νευροφυσιολογίας [7]. Στη συνέχεια περιγράφονται πειραματικές προσπάθειες κατανόησης αυτών των μηχανισμών, δανεισμένες από τη ψυχολογία της όρασης. [8]-[15] Απώτερος σκοπός της μελέτης είναι η άντληση έμπνευσης για το σχεδιασμό ενός μοντέλου ανάλυσης, κατανόησης και ανάκτησης εικόνων με βάση το ουσιαστικό περιεχόμενό τους. Σε αυτή τη προσπάθεια επιβεβαιώνεται πως η όραση αποτελεί μια ευφυή διαδικασία. Συνεπώς, οποιαδήποτε προσπάθεια μοντελοποίησής της περνά μέσα από τις σύγχρονες κατευθύνσεις της τεχνητής νοημοσύνης. Μέσω μίας φιλοσοφικής αναζήτησης της έννοιας της ευφυΐας καταλήγουμε σε μία από τις πλέον σύγχρονες και αποδεκτές θεωρίες τεχνητής ευφυΐας τη θεωρία των αυτόνομων ή ευφυών πρακτόρων.

### 2.1 Νευροεπιστήμες και Βιολογική Όραση

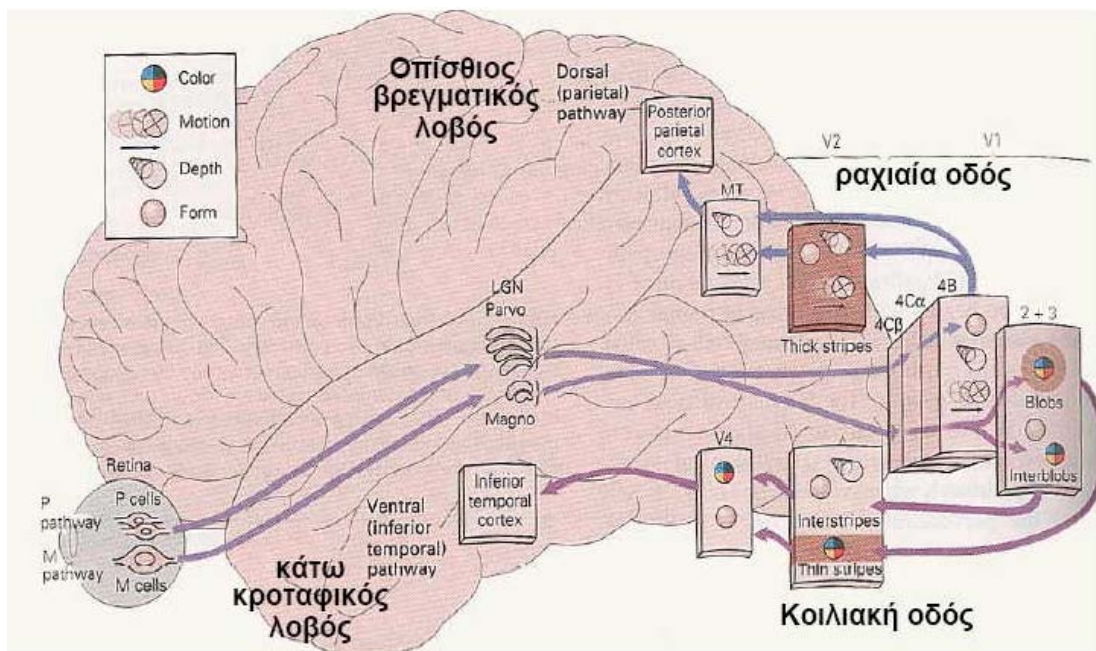
Ο εγκεφαλικός φλοιός του ανθρώπου είναι διαχωρισμένος σε περιοχές που παρουσιάζουν εξειδίκευση όσον αφορά τη λειτουργικότητά τους. Είναι βέβαιο πως η όραση (όπως και η ακοή), αποτελεί μία από αυτές τις περιοχές. Μάλιστα, είναι εντοπισμένη με μεγάλη ακρίβεια η θέση του οπτικού φλοιού (occipital lobe-ινιακός λοβός), στον εγκέφαλο η οποία παρουσιάζεται στην Εικόνα 6. Ωστόσο, υπάρχουν και άλλες ανεξάρτητες περιοχές που επιδρούν με κάποιο τρόπο στην όραση, γεγονός που επιβεβαιώνει το σύνθετο της διαδικασίας της οπτικής αντίληψης. Περαιτέρω, είναι πλέον αποδεκτό πως η εξειδίκευση της λειτουργικότητας συνεχίζεται σε υποπεριοχές μέσα στον ίδιο το μηχανισμό της όρασης.



**Εικόνα 6: Εγκεφαλική τομή σε κάτοψη όπου είναι εντοπισμένη η περιοχή του έξω γονατώδη πυρήνα**

Στην Εικόνα 7, παρουσιάζεται η πορεία του οπτικού σήματος από τη στιγμή που αυτό μέσω του οπτικού νεύρου φτάνει στον οπτικό φλοιό και συγκεκριμένα στον έξω γονατώδη πυρήνα (LGN) [7]. Ο LGN είναι η κύρια υποφλοιική θέση επεξεργασίας οπτικών πληροφοριών. Οι ίνες από το αριστερό ημιμόριο του αμφιβληστροειδή καταλήγουν στον δεξιό LGN και μεταφέρουν μία πλήρη αντιπροσώπευση του αριστερού οπτικού ημιπεδίου. Η προβολή των νευροαξόνων στον LGN γίνεται με τοπογραφική οργάνωση, με αποτέλεσμα ο τελευταίος να παραλαμβάνει ένα τοπογραφικό χάρτη του αμφιβληστροειδή. Ο χάρτης αυτός δεν αποτελεί σημείο προς σημείο αντιπροσώπευση της οπτικής σκηνής καθώς περισσότερες πληροφορίες-νευρώνες ξεκινούν από το κεντρικό τμήμα-βοθρίο του αμφιβληστροειδούς. Η περαιτέρω επεξεργασία και κατανόηση αυτού του σήματος-εικόνας αποτελεί το θέμα αυτής της παραγράφου. Ωστόσο, για την καλύτερη κατανόηση των μηχανισμών της όρασης και εξαγωγή γενικότερων συμπερασμάτων, προτείνεται η μελέτη ολόκληρου του οπτικού συστήματος συμπεριλαμβανομένης και της πορείας από τον αμφιβληστροειδή έως τα γαγγλιακά κύτταρα που τροφοδοτούν το οπτικό νεύρο. Μία τέτοια ολοκληρωμένη μελέτη είναι ικανή να βοηθήσει στην κατανόηση του τρόπου με τον οποίο λειτουργεί η φύση καθώς εξελικτικά προσπαθεί να τροφοδοτήσει τους βιολογικούς οργανισμούς με δυνατότητες ισχυρές και σύνθετες. Απλοί κανόνες εξελικτικής εξειδίκευσης και δια βίου προσαρμογής στο οργανωμένο περιβάλλον, κρύβονται πίσω από αυτές τις «εξ' αποστάσεως» ευφυείς διαδικασίες.





Εικόνα 7: Περιοχές του φλοιού των εγκεφαλικών ημισφαιρίων που εξυπηρετούν την όραση

### 2.1.1 Η πορεία της οπτικής πληροφορίας και η εξειδίκευση, μέσα στον οπτικό φλοιό

Δύο παράλληλα, ασύγχρονα μονοπάτια P, M μεταφέρουν την οπτική πληροφορία από τον έξω γονατώδη πυρήνα στο πρώτο στάδιο επεξεργασίας του οπτικού φλοιού, V1. Το P μονοπάτι διαχειρίζεται κυρίως την πληροφορία του χρώματος, ενώ το M μονοπάτι την πληροφορία του βάθους από τη στερεοσκοπική όραση και της κίνησης. Ωστόσο, όπως είναι γνωστό, στις περιοχές όπου το μάτι εστιάζει, το χρώμα είναι η κύρια πληροφορία για την αναγνώριση μορφών. Περαιτέρω, το χρώμα εμπεριέχει κάποια πληροφορία για το βάθος. Επομένως, πέρα από την πληροφορία του χρώματος, το P μονοπάτι μεταφέρει πληροφορία χρήσιμη για την αναγνώριση μορφών και βάθους. Από την άλλη, η πληροφορία της κίνησης από το M μονοπάτι εμπεριέχει την πληροφορία του δυναμικού σχήματος. Εξειδικευμένες υποπεριοχές στη V1 περιοχή, προσλαμβάνουν τα ανεξάρτητα αυτά τα χαρακτηριστικά, τα οποία επεξεργάζονται και διανέμουν. Μάλιστα, θεωρείται ότι στην περιοχή αυτή, υπάρχει μια αρκετά ακριβής απεικόνιση της οπτικής σκηνής με τη μορφή παράλληλων χαρτών που επεξεργάζονται και αναπαριστούν ανεξάρτητα την πληροφορία των χαρακτηριστικών αυτών. Στο επόμενο στάδιο επεξεργασίας όπου συναντάται η περιοχή V2 όπου περιλαμβάνει ξανά το σύνολο της χρήσιμης πληροφορίας με σαφέστερο διαχωρισμό και επεξεργασία. Τα επόμενα στάδια της επεξεργασίας, από τις περιοχές V3-V5, είναι εξειδικευμένα σε πολύ μεγάλο βαθμό στην επεξεργασία ενός ανεξάρτητου πρωτογενούς οπτικού χαρακτηριστικού το κάθε ένα. Για παράδειγμα η περιοχή V4 δέχεται την πληροφορία του χρώματος, η περιοχή V3, V3A ενεργοποιούνται από ακμές με διαφορετικό προσανατολισμό και η περιοχή V5 είναι ευαίσθητη στην κίνηση.

Επομένως, ο άνθρωπος αποσυνθέτει την οπτική σκηνή στα επιμέρους πρωτογενή χαρακτηριστικά της, τα οποία επεξεργάζεται σε μεγάλο βαθμό ανεξάρτητα, αυτόνομα και παράλληλα. Στη συνέχεια, με κάποιο τρόπο ολοκληρώνει την επιμέρους πληροφορία, ώστε να αντιληφθεί το οπτικό περιεχόμενο και να δημιουργήσει τη δική του σημασιολογική εντύπωση της σκηνής. Το πώς συντίθεται η οπτική πληροφορία από τα επιμέρους πρωτογενή στοιχεία-μονάδες επεξεργασίας, παραμένει προς το παρόν αδιαφανές.

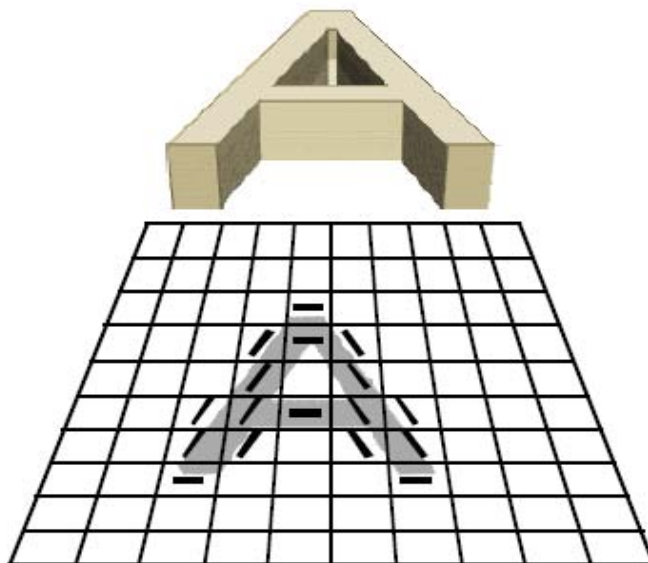
## **2.2 Πειραματική Ψυχολογία και Οπτική Αντίληψη**

Η βιολογική όραση αποτέλεσε και ακόμα αποτελεί ένα σημαντικό ερευνητικό κομμάτι της πειραματικής ψυχολογίας για την αντίληψη του ανθρώπου. Από τις αρχές του 20<sup>ου</sup> μόλις αιώνα θεμελιώθηκαν βασικές αρχές για την οπτική αντίληψη και την οπτική οργάνωση με ουσιαστικότερη και πλέον διαδεδομένη εκείνη της σχολής Gestalt. Καθώς οι μηχανισμοί κατανόησης του περιεχομένου του οπτικού σήματος παραμένουν ακόμα ασαφείς, συντελείται μέχρι σήμερα πολύ συστηματική έρευνα προς αυτή τη κατεύθυνση. Στην προσπάθειά μας να σχεδιάσουμε ένα μοντέλο ανάκτησης εικόνων που να προσομοιάζει τη βιολογική όραση δεν θα μπορούσαμε να παραλείψουμε από τη μελέτη μας τα πλέον σύγχρονα πειραματικά δεδομένα της ψυχολογίας της όρασης.

Όπως έχει ήδη ειπωθεί, ένα μοντέλο δύο ή περισσότερων επιπέδων επεξεργασίας του οπτικού περιεχομένου αρχίζει να γίνεται ευρέως αποδεκτό από την κοινότητα της πειραματικής ψυχολογίας (Scientific). Στο πρώτο επίπεδο που στην παρούσα εργασία θα το αποκαλούμε πρώιμο (pre-attentive) συναντάμε ένα σύνολο από τοπογραφικούς χάρτες, έναν για καθένα από τα πρωτογενή χαρακτηριστικά του οπτικού σήματος. Στο δεύτερο ή επιλεκτικό επίπεδο (attentive) συντελείται η σύνθεση αυτών των πρωτογενών στις αναπαραστάσεις των επιφανειών και των όγκων του περιβάλλοντος κόσμου. Πιθανώς σε ανώτερα επίπεδα γίνεται η αντιστοίχιση τους με εννοιολογικές αναπαραστάσεις του εγκεφάλου. Παρουσιάζει ιδιαίτερο ενδιαφέρον λοιπόν να αναφέρουμε ορισμένα από τα βασικότερα πειραματικά ευρήματα της ψυχολογίας της οπτικής αντίληψης και κυρίως να αιτιολογήσουμε και να ανακαλύψουμε τους μηχανισμούς που κρύβονται πίσω από αυτά.

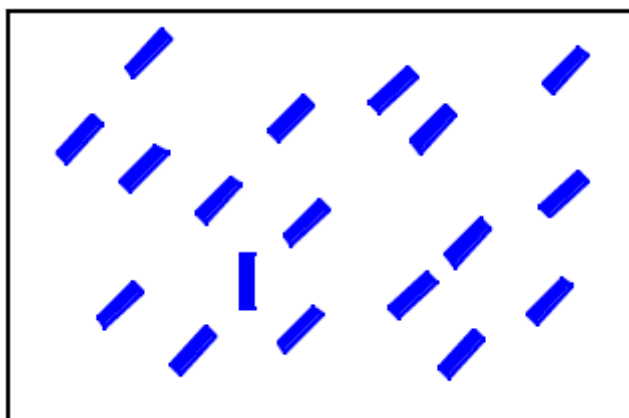
### **2.2.1 Πρώιμο στάδιο επεξεργασίας: preattentive level of cognition**

Σε αυτό το πρώτο στάδιο επεξεργασίας η οπτική πληροφορία είναι κωδικοποιημένη σε ένα σύνολο από παράλληλους ανεξάρτητους χάρτες κάθε ένας από τους οποίους κωδικοποιεί την εμφάνιση ενός πρωτογενούς χαρακτηριστικού της εικόνας σε κάθε θέση της οπτικής σκηνής [14],[15]. Τέτοια πρωτογενή χαρακτηριστικά είναι το χρώμα, ο προσανατολισμός των ακμών, η κίνηση, το βάθος κ.α. Για παράδειγμα, στην Εικόνα 8 παρουσιάζεται ο χάρτης του προσανατολισμού των ακμών σε δοσμένη οπτική είσοδο.



**Εικόνα 8: Πρωτογενής χάρτης ακμών σε οπτική είσοδο τον χαρακτήρα «Α»**

Πέρα από αυτούς τους ανεξάρτητους χάρτες είναι πολύ πιθανό να υπάρχει και ένας συνδυαστικός χάρτης ο οποίος κωδικοποιεί τη συνύπαρξη αυτών των χαρακτηριστικών σε κάθε θέση του οπτικού ερεθίσματος. Τα πειραματικά δεδομένα υποδεικνύουν πως η επεξεργασία διαφορετικών τοποθεσιών μέσα στον ίδιο χάρτη γίνεται παράλληλα και ταχύτατα. Οι πειραματικές διατάξεις για την κατανόηση των ιδιοτήτων του πρώιμου επιπέδου συνήθως περιλαμβάνουν την πολύ σύντομη παρατήρηση εικόνων σε χρόνο που δεν επιτρέπει την προσεκτική εξέτασή τους. Τότε οι πρώιμες ιδιότητες και χαρακτηριστικά που αναζητούνται φαίνονται να εξέρχονται της οπτικής σκηνής (pop out φαινόμενο). Για παράδειγμα κατά την αναζήτηση ακμών συγκεκριμένου προσανατολισμού στην Εικόνα 9 ο στόχος φαίνεται να πετάγεται μπροστά από την εικόνα.



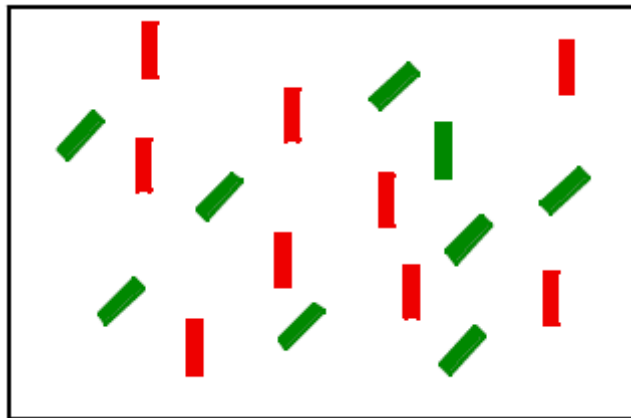
**Εικόνα 9: Διάταξη οπτικής αναζήτησης στο πρώιμο επίπεδο. Αναζήτηση ακμών με συγκεκριμένο προσανατολισμό**

Με άλλα λόγια σε αυτό το πρώτο επίπεδο, γίνεται ταυτόχρονη παρατήρηση όλων των πρωτογενών με τη συγκεκριμένη ιδιότητα ανεξάρτητα από την ύπαρξη άλλων πρωτογενών που θα μπορούσαν να αποσπάσουν την προσοχή (π.χ. ακμές με ίδια κλίση αλλά διαφορετικό χρώμα) και από την πολυπλοκότητα της εικόνας.

Επειδή αυτή η παρατήρηση γίνεται παράλληλα για όλο το χάρτη και πολύ γρήγορα, δεν απαιτείται προσεκτικός-σειριακός μηχανισμός για την εξαγωγή ενός στόχου ο οποίος μπορεί να περιγραφεί μονοσήμαντα από ένα τέτοιο χάρτη. Εξού και η αγγλική ονομασία αυτού του επιπέδου “preattentive”. Αντικείμενα που διαθέτουν μοναδική και σταθερή αναπαράσταση όσον αφορά ένα preattentive ή αλλιώς πρωτογενές χαρακτηριστικό, σε σχέση με την υπόλοιπη σκηνή, είναι τα μόνα που προσλαμβάνονται σε αυτό το στάδιο. Τέλος, πέρα από τα πρωτογενή χαρακτηριστικά που αναφέρθηκαν, υπάρχουν ενδείξεις πως ορισμένες επιπλέον ιδιότητες κωδικοποιούνται σε αυτό το πρώιμο στάδιο όπως το μέγεθος, η καμπυλότητα (curvature) και η κλειστότητα (closure). Στη βιβλιογραφία της πειραματικής ψυχολογίας για την οπτική αντίληψη μπορεί κανείς να συναντήσει πλήθος από πειράματα αναζήτησης τέτοιων χαρακτηριστικών.

### 2.2.2 Στάδιο επιλεκτικής επεξεργασίας: attentive level of cognition

Ο συνδυασμός των πρωτογενών χαρακτηριστικών κατά μήκος των διαφορετικών χαρτών [14],[15] δεν παρουσιάζει τις ιδιότητες που παρατηρήθηκαν στην προηγούμενη παράγραφο. Στην Εικόνα 10 για παράδειγμα, η αναζήτηση μίας πράσινης κάθετης ακμής δια μέσου τυχαία τοποθετημένων πράσινων-κεκλιμένων και κόκκινων κάθετων διασπαστών της προσοχής δεν συντελείται αυτόματα αλλά απαιτεί χρόνο.



Εικόνα 10: Διάταξη οπτικής αναζήτησης. Αναζήτηση ακμών με συγκεκριμένο προσανατολισμό και χρώμα δια μέσου διασπαστών με ίδιο προσανατολισμό αλλά διαφορετικό χρώμα

Ο χρόνος αυτός είναι ανάλογος του πλήθους των διασπαστών. Επιπρόσθετα, όταν ο στόχος αναζήτησης παρουσιάζεται στην εικόνα, ο εντοπισμός του γίνεται κατά μέσο όρο στο μισό χρόνο σε σχέση με εκείνον που απαιτείται στην περίπτωση όπου ο στόχος δε βρίσκεται στην εικόνα. Είναι προφανές πως στην πρώτη περίπτωση γίνεται σειριακή εξέταση των μισών κατά μέσο όρο αντικειμένων, ενώ όταν απαιτείται αρνητική απάντηση, είναι αναγκαία η εξέταση του συνόλου των αντικειμένων. Επομένως, η συνδυαστική αναζήτηση (conjunction search) πρωτογενών κατά μήκος δύο ή

περισσότερων πρωτογενών χαρτών απαιτεί σειριακή αναζήτηση του στόχου, αντικείμενο προς αντικείμενο. Καθώς απαιτείται προσεκτική εξέταση του κάθε αντικείμενου, το επίπεδο αυτό ονομάζεται «attentive» ή επιλεκτικό. Ο παραπάνω μηχανισμός συνήθως προσομοιάζεται με ένα προβολέα ο οποίος μετακινείται σειριακά από θέση σε θέση μέσα στην εικόνα φωτίζοντας και εξετάζοντας κάποια περιοχή και το μέγεθος του οποίου μπορεί να επεκταθεί για να εξετάσει μεγαλύτερη ή μικρότερη επιφάνεια ανάλογα με την περίπτωση.

Στην περίπτωση όπου ο στόχος αποτελεί συνδυασμό πολλαπλών πρωτογενών, η επιπλέον πληροφορία χρησιμοποιείται από τους μηχανισμούς της όρασης για επιλεκτική καθοδήγηση ανάμεσα στα εξεταζόμενα αντικείμενα. Για παράδειγμα, κατά την αναζήτηση τριπλών συνδυασμών, π.χ. μεγάλων- πράσινων- κατακόρυφων ακμών ανάμεσα σε μικρούς- κόκκινους- κατακόρυφους, μικρούς- πράσινους- κεκλιμένους και μεγάλους- κόκκινους- κεκλιμένους διασπαστές, η διαδικασία είναι συντομότερη από την αναζήτηση διπλού συνδυασμού. Επιπλέον, ο χρόνος του εντοπισμού αυξάνει με μικρότερο ρυθμό καθώς το πλήθος των διασπαστών αυξάνεται. Επομένως, η ολοκλήρωση των πρωτογενών, δεν είναι απλά αποτέλεσμα σειριακής (γραμμικής) αναζήτησης, αλλά καθοδηγείται από πιο σύνθετους μηχανισμούς που πιθανότατα μειώνουν επιλεκτικά τον αριθμό των αντικειμένων που εξετάζονται.

Τέλος σε πειράματα τριπλών συνδυασμών όπου οι διασπαστές εμφανίζουν τα δύο από τα τρία πρωτογενή του στόχου, οι μετρήσεις είναι αντίστοιχες με αυτές των διπλών συνδυασμών. Το προαναφερθέν αποτέλεσμα, και άλλα πρόσφατα πειραματικά ευρήματα οδηγούν στο συμπέρασμα ότι *όσο περισσότερο διαφέρει ο στόχος από τους διασπαστές τόσο ευκολότερη είναι η αναζήτηση*. Μολονότι αυτή η μελέτη δεν έχει ως στόχο να κάνει υποθέσεις πάνω στην ερμηνεία των παραπάνω πειραματικών παρατηρήσεων ένα πιθανό ενδεχόμενο είναι η ύπαρξη ενός χάρτη που κωδικοποιεί τη συνολική ενεργοποίηση των πρωτογενών από όλους τους υπόλοιπους χάρτες.

Μια διαφορετική προσέγγιση αιτιολογεί αυτά και άλλα φαινόμενα σε σχέση με τις χρονικές αποκρίσεις των νευρώνων. Για παράδειγμα, στο πρώτο επίπεδο, οι νευρώνες που κωδικοποιούν τα ανεξάρτητα πρωτογενή σε ολόκληρη την εικόνα συγχρονίζονται παράγοντας τα φαινόμενα που αναφέρθηκαν στο 2.2.1, όπως το pop out φαινόμενο. Στη συνέχεια για την ανάλυση μιας τοπικά εντοπισμένης περιοχής, οι νευρώνες από τα διαφορετικά πρωτογενή που κωδικοποιούν τη περιοχή συγχρονίζονται, ώστε να γίνει η προσεκτική μελέτη της με βάση τις συνδυαστικές ιδιότητές τους. Αυτή η προσέγγιση φυσικά απορρίπτει την ύπαρξη του αθροιστικού πρωτογενούς χάρτη. Ωστόσο, ο ενδιαφερόμενος αναγνώστης παροτρύνεται να μελετήσει και να εξάγει τα δικά του συμπεράσματα όσον αφορά την ερμηνεία αυτών των μηχανισμών, ανατρέχοντας σε δεδομένα που προκύπτουν από την ολοκλήρωση σύγχρονων πειραματικών δεδομένων της ψυχολογίας της όρασης.

### 2.2.3 Αρχές οπτικής οργάνωσης και μη τυχαίες ιδιότητες

Ένας από τους γενικότερους και αφαιρετικούς κανόνες πάνω στον οποίο θεμελιώνεται η επιστήμη της ψυχολογίας της οπτικής αντίληψης είναι η υπεραρχή της Prägnanz που διατυπώνεται ως εξής:

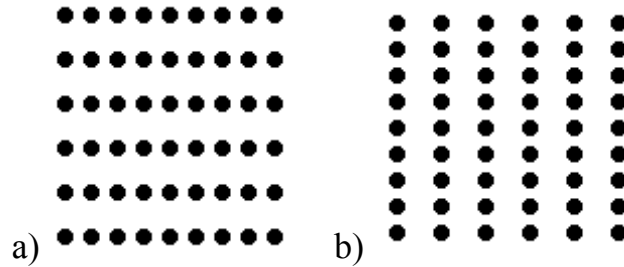
«οι απλές και σταθερές ερμηνείες ευνοούνται»

Επομένως, παρουσιάζει ιδιαίτερο ενδιαφέρον να εξεταστεί τι αποτελεί το απλό και το σταθερό στην όραση, τι είναι απλό και σταθερό στο οπτικό περιβάλλον που προσλαμβάνουμε. Είναι δυνατό να ειπωθεί πως οι ευνοούμενες διατάξεις οπτικών χαρακτηριστικών είναι εκείνες που αντιστοιχούν σε φυσικές οντότητες με νόημα. Κατά συνέπεια, η αντίληψη της οπτικής οργάνωσης πιθανότατα λειτουργεί χρησιμοποιώντας κανόνες που εκμεταλλεύονται την ύπαρξη μη τυχαίας οπτικής πληροφορίας που προκύπτει από την οργάνωση και τους κανόνες που διέπουν το φυσικό κόσμο. Ως μη τυχαίες οπτικές ιδιότητες μπορούν να θεωρηθούν εκείνες οι διατάξεις των χαρακτηριστικών για τις οποίες η πιθανότητα αναίτιας εμφάνισης (τυχαία, χωρίς νόημα εμφάνιση) προσεγγίζει το μηδέν. Διατυπώνοντας διαφορετικά είναι εκείνες οι οπτικές διατάξεις για τις οποίες ο αριθμός των εμφανίσεων ως σημασιολογικά αντιληπτές οντότητες είναι πολύ μεγαλύτερος από τον αριθμό των τυχαίων εμφανίσεων τους. *Επομένως, η βιολογική όραση προσαρμόζεται στο να εντοπίζει τις σημασιολογικές οντότητες της οπτικής σκηνής, ακολουθώντας αυτές τις μη τυχαίες οπτικές ιδιότητες που προσφέρει η οργάνωση και οι νόμοι του φυσικού κόσμου.* Στην επόμενη παράγραφο παρουσιάζονται στην πλειοψηφία τους οι μη τυχαίες ιδιότητες που καθοδηγούν την όραση κατά την ομαδοποίηση των πρωτογενών οπτικών χαρακτηριστικών και είναι ευρέως γνωστές ως αρχές ομαδοποίησης ή κανόνες Gestalt.

#### 2.2.3.1 Αρχές ομαδοποίησης

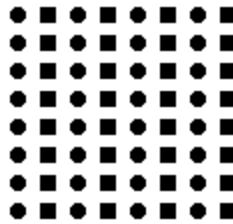
Οι αρχές ομαδοποίησης είναι αρχές που υπακούουν στην υπεραρχή της Prägnanz και βρίσκουν εφαρμογή σε πολλές πτυχές της ζωής, πρωτίστως όμως στην οπτική οργάνωση. Οι βασικότερες από αυτές διατυπώθηκαν από τη σχολή Gestalt. Οι κανόνες Gestalt είναι οι ακόλουθοι [8]:

- **Εγγύτητα** (proximity): Σύμφωνα με την αρχή της εγγύτητας σχήματα που είναι πιο κοντά μεταξύ τους τείνουν να ομαδοποιούνται. Κατά συνέπεια, στην Εικόνα 11a) αυτό, που είναι πιθανόν να παρατηρήσει κανείς σχετικά γρήγορα, είναι ότι δεν πρόκειται μόνο για ένα τετράγωνο σχήμα από τελείες αλλά μάλλον για μια σειρά από γραμμές με τελείες. Αντίθετα στην Εικόνα 11b) οι τελείες τείνουν να ομαδοποιούνται κατά στήλες.



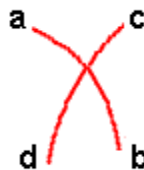
Εικόνα 11: Η αρχή της εγγύτητας

- **Ομοιότητα** (similarity). Σύμφωνα με την αρχή της ομοιότητας σχήματα που έχουν όμοιες ιδιότητες τείνουν να ομαδοποιούνται. Για παράδειγμα, στην Εικόνα 12 οι μικροί κύκλοι και τα τετράγωνα έχουν όμοιες αποστάσεις οριζόντια και κάθετα, ούτως ώστε η εγγύτητα να μην παίζει ρόλο. Παρατηρούμε ότι ευνοούνται οι στήλες από κύκλους ή τετράγωνα παρά οι μικτές γραμμές.



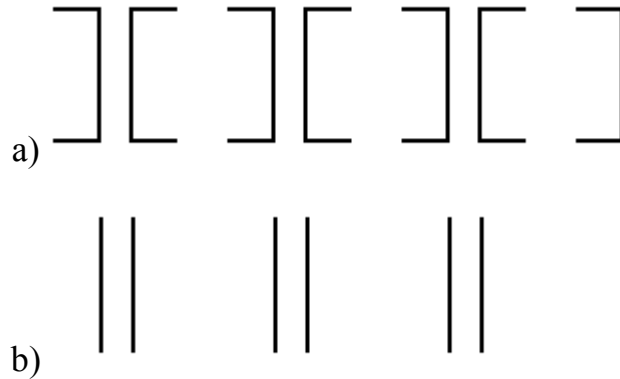
Εικόνα 12: Η αρχή της ομοιότητας

- **Ομαλή Συνέχεια** (good continuation). Σύμφωνα με αυτή την αρχή, τα περιγράμματα που εμφανίζουν σε ομαλή συνέχεια προτιμώνται από τις απότομες αλλαγές κατεύθυνσης. Για παράδειγμα, στην Εικόνα 13 είναι πιο πιθανό να δούμε τις γραμμές a-b και c-d να τέμνονται, παρά να δούμε τις a-d και c-b, ή a-c και d-b ως γραμμές.



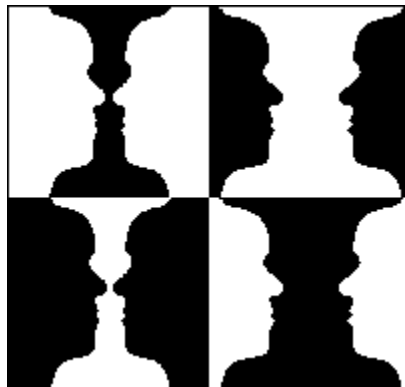
Εικόνα 13: Η αρχή της ομαλής συνέχειας

- **Κλειστότητα** (Closure). Σύμφωνα με την αρχή της κλειστότητας οι ερμηνείες που παράγουν ‘κλειστά’ παρά μάλλον ‘ανοικτά’ σχήματα ευνοούνται. Κατά συνέπεια, στην Εικόνα 14a) τείνουμε να βλέπουμε τρία σπασμένα ορθογώνια (και ένα μεμονωμένο σχήμα στην αριστερή άκρη), παρά τρία δοκάρια σε προφίλ (και ένα μεμονωμένο σχήμα δεξιά). Στην περίπτωση αυτή η αρχή της κλειστότητας διασταυρώνεται με την αρχή της εγγύτητας, αφού όταν αφαιρέσουμε τις αγκύλες, ομαδοποιούνται οι κοντινές μεταξύ τους γραμμές (Εικόνα 14b)).



Εικόνα 14: Η αρχή της κλειστότητας

- **Μικρό μέγεθος** (smallness): Σύμφωνα με την αρχή του μικρού μεγέθους, οι μικρότερες περιοχές τείνουν να θεωρούνται φιγούρες (αντικείμενα) παρουσία μεγαλύτερων περιοχών που τείνουν να προσλαμβάνονται ως φόντο. Για παράδειγμα στην εικόνα του Rubin (Εικόνα 15), το βάζο ενοείται ως φιγούρα στις δύο αριστερές εικόνες έναντι των προσώπων.



Εικόνα 15: Η αρχή του μικρού μεγέθους

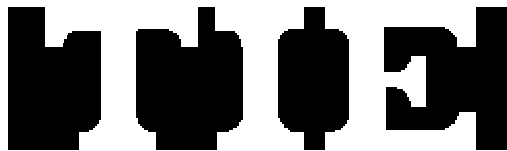
- **Συμμετρία** (symmetry): Σύμφωνα με την αρχή αυτή, οι συμμετρικές περιοχές τείνουν να θεωρούνται ως φιγούρες κι οι ασυμμετρικές ως υπόβαθρα. Για παράδειγμα, στην Εικόνα 16 η συμμετρική μαύρη περιοχή ενοείται ως φιγούρα σε σχέση με την άσπρη στα αριστερά της.



Εικόνα 16: Η αρχή της συμμετρίας



- **Περιβάλλοντος (surrounded ness):** Σύμφωνα με την αρχή αυτή, οι περιοχές που μπορεί να φανεί ότι περιβάλλονται από άλλες, τείνουν να γίνονται αντιληπτές ως φηγούρες (Εικόνα 17).



Εικόνα 17: Η αρχή του περιβάλλοντος

Πάνω σε αυτές τις βασικές αρχές είναι δυνατό να στηριχτούν και άλλες όπως για παράδειγμα η **παράλληλία** ως ομοιότητα στην γωνία κλίσης των ακμών, ιδιότητα που στη βιβλιογραφία αναφέρεται ως πρωτογενές (preattentive) χαρακτηριστικό της οπτική επεξεργασία, η συνδετικότητα (connectivity) κ.α. Επιπρόσθετα υπάρχουν άλλες αρχές οργάνωσης που σχετίζονται με την κίνηση των αντικειμένων, όπως η **αρχή της κοινής μοίρας** (law of common fate) που δηλώνει πως αντικείμενα που έχουν παρόμοια κίνηση (π.χ. ίδια κατεύθυνση) τείνουν να ομαδοποιούνται

*Σε αυτό το σημείο είναι σημαντικό να αναφερθεί η παρατήρηση μας πως πολλές από τις ιδιότητες ομαδοποίησης μπορούν να αναχθούν σε άλλες γενικότερες και απλούστερες. Για παράδειγμα, η αρχή της καλής συνέχειας μπορεί να θεωρηθεί πως ανάγεται στην αρχή της ομοιότητας αν η τελευταία εφαρμοστεί στον πρωτογενή χάρτη της κλίσης των ακμών. Ομοίως, η αρχή της κοινής μοίρας στην ομοιότητα που εφαρμόζεται στον πρωτογενή χάρτη της κίνησης (οπτικής ροής). Κατά συνέπεια, είναι πολύ πιθανό ένα μικρότερο σύνολο από αυτές τις ιδιότητες να αρκεί για να ομαδοποιήσει και να επανασυνθέσει την πληροφορία από τους πρωτογενείς χάρτες δημιουργώντας τις νοητικές αναπαραστάσεις της οπτικής σκηνής.*

## 2.3 Βιολογική Όραση και ευφυΐα

Η οπτική αντίληψη, η εξαγωγή σημασίας από το οπτικό ερέθισμα είναι μία ευφυής διαδικασία. Ωστόσο, είναι αυθαίρετο να γίνεται λόγος για ευφυείς διαδικασίες ενόσω δεν είναι ξεκάθαρη η ίδια η έννοια της ευφυΐας. Μια πολύ ενδιαφέρουσα και ολοκληρωμένη προσέγγιση την οποία υιοθετεί και η παρούσα εργασία, γίνεται στο [4].

### 2.3.1 Η έννοια της ευφυΐας

Μολονότι στην καθημερινή ζωή ο άνθρωπος τείνει να ποσοτικοποιεί και να διαβαθμίζει την ευφυΐα, τέτοιες προσπάθειες δεν είναι αξιόπιστες, απουσία κάποιου γενικού και αποδεκτού ορισμού της. Χαρακτηριστικό παράδειγμα τέτοιας προσπάθειας, αποτελεί η μέτρηση της ευφυΐας με ένα μοναδικό αριθμό, τον “factor g” των IQ tests.

Αντίθετα, αυτό που παρουσιάζει ενδιαφέρον είναι η φιλοσοφική αναζήτηση της έννοιας της ευφυΐας. Διεργασίες όπως η σκέψη και η επίλυση προβλημάτων, η μάθηση και η απομνημόνευση, ο χειρισμός της γλώσσας και η επικοινωνία, η διαίσθηση και η συνείδηση, η δημιουργικότητα και η φαντασία, η ικανότητα επιβίωσης στη φύση κ.α., συχνά χρησιμοποιούνται για να χαρακτηρίσουν την ευφυΐα ενός ατόμου υποκειμένου. Επομένως παρουσιάζει ιδιαίτερο ενδιαφέρον η αναζήτηση ενός κοινού αφαιρετικού παρανομαστή τους. Πράγματι οι προαναφερθείσες λειτουργίες παρουσιάζουν ένα βαθμό αλληλοσυσχέτισης, που σχετίζεται με την ικανότητα για καινοτομικής σκέψη, σύλληψης νέων ιδεών και κατάλληλης στρατηγικής στη συμπεριφορά, σε ένα περιβάλλον που συνεχώς μεταβάλλεται. Η επιλογή της εκάστοτε «κατάλληλης συμπεριφοράς» είναι μια σύνθετη διαδικασία που προϋποθέτει την αξιολόγηση όχι μόνο των καινοτομικών ενεργειών αλλά και των παλαιότερων κανόνων συμπεριφοράς μέσω της συνεχούς επιβεβαίωσης τους (diversity-compliance tradeoff). Η δυνατότητα της δημιουργίας ποικιλίας στη συμπεριφορά εν μέσω παρόμοιων καταστάσεων είναι ένας πολύ σημαντικός παράγοντας μάθησης και ευφυΐας. Η αξιολόγηση μίας ενέργειας είναι δυνατό να ποσοτικοποιηθεί ως ο βαθμός επίτευξης ενός συνόλου από σταθμισμένους αρχικούς στόχους. Κατά συνέπεια, ένας ευφυής οργανισμός πρέπει είναι σε θέση να αντιλαμβάνεται την αποτελεσματικότητα της εκάστοτε ενέργειας του και να τη συσχετίζει με τη συγκεκριμένη συγκυρία. Η πρόσληψη του περιβάλλοντος μέσω των αισθητήρων αλλά και των εσωτερικών μηχανισμών αντίληψης και κατανόησης είναι εξαιρετικά σύνθετη διαδικασία. Σε ένα δυναμικό περιβάλλον, απαιτείται ο διαχωρισμός και ανάλυση εκείνων των μεταβολών που προκύπτουν ως αποτέλεσμα της συγκεκριμένης ενέργειας και όχι λόγω εξωτερικών ανεξάρτητων παραγόντων. *Ολοκληρώνοντας την παραπάνω ανάλυση αναζητούμε την ευφυΐα στην ύπαρξη της προσαρμοστικότητας, που δεν αφορά μόνο τη παθητική συγκατάβαση και απόκριση (reactiveness) σε ανώτερα εξωτερικά γεγονότα, άλλα κυρίως την επινόηση νέων κανόνων συμπεριφοράς για την καλύτερη εκμετάλλευση των τυχαίων και μη συγκυριών μέσα σε ένα δυναμικό και απρόσμενο περιβάλλον.* Την παραπάνω προσέγγιση της έννοιας της ευφυΐας υιοθετεί σε μεγάλο βαθμό η θεωρία των αυτόνομων πρακτόρων που αποτελεί την πλέον σύγχρονη και αποδεκτή κατεύθυνση της τεχνητής νοημοσύνης.

### **2.3.2 Μοντελοποίηση της ευφυΐας - Η θεωρία των αυτόνομων πρακτόρων**

Η πλέον σύγχρονη και ελπιδοφόρα κατεύθυνση της τεχνητής νοημοσύνης, οι αυτόνομοι πράκτορες κάνουν λόγο για αυθύπαρκτες οντότητες που προσπαθούν να εκπληρώσουν ένα σύνολο από στόχους σε ένα δυναμικό περιβάλλον. Καθώς ο χρόνος περνά, ένας πράκτορας συλλέγει πληροφορίες, συνεχώς μαθαίνοντας για το περιβάλλον του και προσπαθώντας να βελτιώσει τις επιδόσεις του όσον αφορά τους στόχους που έχει αναλάβει. Οι ιδιότητες που χαρακτηρίζουν έναν πράκτορα έχουν προκύψει από τη μελέτη των ικανοτήτων που παρουσιάζουν οι βιολογικοί οργανισμοί. Το 1986 ο Brooks, στην αρχιτεκτονική του, subsumption architecture, πρότεινε ένα σύστημα παράλληλων, ασύγχρονων, ελαφρά συνδεδεμένων διεργασιών για τη μοντελοποίηση των πρώιμων σταδίων στις σύνθετες βιολογικές διεργασίες [4]. Τέτοια συστήματα είναι ευρέως γνωστά ως συστήματα πρακτόρων (multi-agent systems). Κάθε διεργασία σε αυτό το

μοντέλο αναλαμβάνει ένα μικρό ανεξάρτητο υπολογιστικό κομμάτι του συνολικού προβλήματος.

Τρεις είναι οι κύριες κατηγορίες αυτόνομων πρακτόρων

- **Βιολογικοί Πράκτορες**
- **Τεχνητοί**
  - **Ρομποτικοί Πράκτορες**
  - **Πράκτορες Λογισμικού**

Ενώ στο σύνολο τους οι βιολογικοί οργανισμοί πληρούν όλες τις προϋποθέσεις της θεωρίας πρακτόρων, δεν συμβαίνει το ίδιο και με τους τεχνητούς πράκτορες. Οι πράκτορες λογισμικού αποτελούν ένα κοινό πεδίο ενδιαφέροντος για πολλές περιοχές της επιστήμης υπολογιστών, καθώς μπορούν να εξεταστούν από διαφορετικές για κάθε πεδίο σκοπιές. Έτσι για την Τεχνητή Νοημοσύνη οι πράκτορες θεωρούνται ευφυείς οντότητες, για τον αντικειμενοστραφή προγραμματισμό θεωρούνται οντότητες (αντικείμενα) που επικοινωνούν μεταξύ τους, για τον παράλληλο προγραμματισμό είναι οντότητες που εκτελούνται παράλληλα και για το πεδίο διασύνδεσης ανθρώπου-μηχανής είναι οντότητες που παίρνουν πρωτοβουλία σε συνεργασία με το χρήστη για να επιτελέσουν τα καθήκοντά τους [5]. Αν και στον επιστημονικό χώρο πλανάται μια σχετική ασάφεια στο τι αποτελεί έναν πράκτορα, υπάρχει μια γενική αποδοχή σε ένα σύνολο από ιδιότητες ικανές να χαρακτηρίσουν έναν αυτόνομο πράκτορα λογισμικού [2]. Όταν τηρούνται αυτές οι ιδιότητες από μόνες τους υποχρεώνουν την ύπαρξη νοημοσύνης. Οι πιο σημαντικές είναι οι ακόλουθες:

- **Χρονική Συνέχεια - Αυτάρκεια:** ένας πράκτορας λογισμικού είναι μία διαρκής διαδικασία, η οποία σε αντίθεση με τα απλά προγράμματα επιτελεί τους στόχους της καθ' όσο διατηρεί την αυτάρκειά της ενώ δεν είναι προκαθορισμένο από το δημιουργό της πότε θα καταλήξει αλλά αυτό εξαρτάται κυρίως από τη βιωσιμότητα του σε ένα απρόσμενο περιβάλλον.
- **Αλληλεπίδραση με το Περιβάλλον - Στόχοι:** Ο πράκτορας αισθάνεται το περιβάλλον στο οποίο λειτουργεί και επιδρά πάνω σε αυτό επηρεάζοντας τις μελλοντικές του αισθήσεις, επιδιώκοντας την επίτευξη των προσωπικών του στόχων. Η αλληλεπίδραση με το περιβάλλον άλλοτε έχει τη μορφή της άμεσης αντίδρασης (reactive) και άλλοτε της προμελετημένης κίνησης (purposeful). Ωστόσο το περιβάλλον θα πρέπει να θεωρείται δυναμικό και απρόσμενο, διαφορετικά η αντιμετώπισή του δεν θα απαιτούσε ευφυή προσέγγιση.
- **Αυτονομία:** ο πράκτορας ασκεί προσωπικό έλεγχο στις ενέργειές του, απουσία εξωτερικού ή κεντρικού ελέγχου στην περίπτωση των συστημάτων πολλαπλών πρακτόρων.
- **Προσαρμοστικότητα - Μάθηση:** η δράση του πράκτορα δεν είναι προγραμματισμένη αλλά πρέπει ο ίδιος να τη διαφοροποιεί ανάλογα με την εμπειρία του, πιθανότατα με τη δημιουργία ποικιλίας στη συμπεριφορά και αξιολόγησης του αποτελέσματος από τη διαφοροποιημένη δράση του.
- **Κοινωνική Ικανότητα – Επικοινωνία:** ο πράκτορας επικοινωνεί με άλλους πράκτορες (πιθανότατα και ανθρώπους μέσω διεπαφών)

- **Κινητικότητα :** η ικανότητα του πράκτορα να μεταφέρεται από ένα μηχάνημα σε ένα απομακρυσμένο

Σε αυτό το σημείο πρέπει να σημειωθεί ότι ένας αυτόνομος πράκτορας λογισμικού δεν είναι αναγκαίο να παρουσιάζει στο σύνολό τους τις παραπάνω ιδιότητες. Ωστόσο, οι τέσσερις πρώτες είναι πολύ σημαντικές και παρουσιάζουν μεγάλη αλληλεξάρτηση. Η αδυναμία προσέγγισης κάποιας από αυτές πιθανότατα οφείλεται σε λανθασμένη μοντελοποίηση του προβλήματος που μπορεί εν τέλει να μην απαιτεί αυτόνομη νοημοσύνη για να . Γενικότερα, η χρήση των αυτόνομων πρακτόρων επιλέγεται ώστε να διευκολύνει το χρήστη σε χρονοβόρες και δυσεπίλυτες διαδικασίες αναλαμβάνοντας πλήρως την επίτευξη των στόχων μιας εργασίας του χρήστη με μη προκαθορισμένο τρόπο [5].

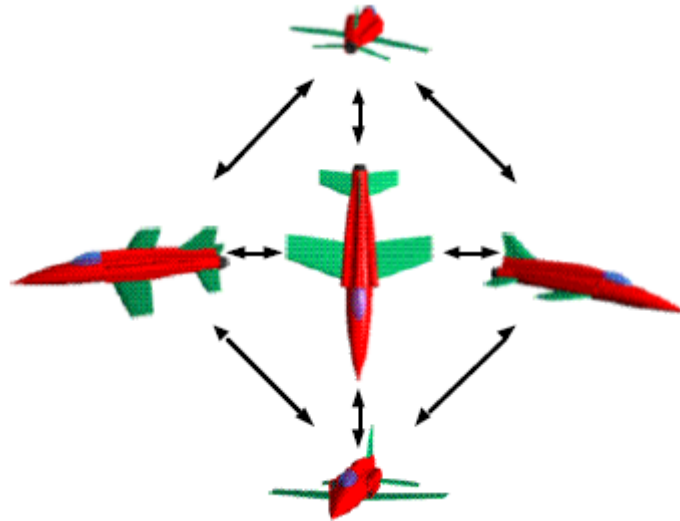
### 2.3.3 Όραση, μία ευφυής διαδικασία

Στις προηγούμενες παραγράφους έγινε η προσπάθεια ανάλυσης και προσδιορισμού της έννοιας της ευφυΐας. Είναι τελικά η όραση μία ευφυής διαδικασία; Όπως έγινε κατανοητό, λίγα γνωρίζουμε αναφορικά με το πως οι βιολογικοί οργανισμοί αντιλαμβάνονται το οπτικό περιεχόμενο και εξάγουν σημασία και νόημα από αυτό. Ωστόσο, είναι φανερό πως οι μηχανισμοί της όρασης περιλαμβάνουν και αλληλεπιδρούν με στοιχεία όπως η μνήμη και η μάθηση και εμφανίζουν ιδιότητες που χαρακτηρίζουν τις περισσότερες ευφυείς διαδικασίες, όπως η προσαρμοστικότητα. Ο άνθρωπος συνεχώς τείνει να δίνει νόημα σε αυτό που παρατηρεί και να παράγει κανόνες κατανόησης του. Οι κανόνες αυτοί του επιτρέπουν να επεξεργάζεται αποδοτικά την τεράστια οπτική πληροφορία που εισέρχεται στον αμφιβληστροειδή του. Ιδίως όσον αφορά τον εντοπισμό και την αναγνώριση σημασιολογικών αντικειμένων και οντοτήτων, ο μηχανισμός της όρασης καταφέρνει να αντεπεξέρχεται απέναντι σε εξαιρετικά δυσεπίλυτα προβλήματα όπως:

- Διαφορετικές συνθήκες φωτισμού των αντικειμένων
- Διαφορετικές όψεις του αντικειμένων συμπεριλαμβανομένης και της κλίμακας
- Παρεμβολή άλλων αντικειμένων και μερική απόκρυψη ενός αντικειμένου

Στην Εικόνα 18 παρατηρούμε 5 από τις άπειρες διαφορετικές όψεις του ίδιου αντικειμένου υπό τις σταθερές συνθήκες φωτισμού. Οι νοητικές αναπαραστάσεις των αντικειμένων αποτελούν τη βάση της οπτικής μας αντίληψης και χρησιμοποιούνται στην αναγνώριση προτύπων, στην οπτική αιτιολόγηση, σε εννοιολογικές-λεκτικές περιγραφές του περιβάλλοντος και στην κατηγοριοποίηση αντικειμένων. Οι μη τυχαίοι κανόνες που διέπουν το περιβάλλον οδηγούν σε επαναλαμβανόμενα πρότυπα που παρουσιάζουν μη τυχαίες οπτικές ιδιότητες. Αυτές οι ιδιότητες μέσω των μηχανισμών προσαρμογής, χρησιμοποιούνται από την όραση στη γρήγορη κατανόηση και αντίληψη της οπτικής σκηνής. Πειραματικά δεδομένα επιβεβαιώνουν πως σε ένα περιβάλλον που διέπεται από διαφορετικούς κανόνες και κατά συνέπεια διαφορετικές μη τυχαίες εμφανίσεις, η όραση

μπορεί να προσαρμόζεται και προοδευτικά να αντεπεξέρχεται σε μεγάλο βαθμό στις υποχρεώσεις της. Επομένως, το σύστημα της όρασης δεν θα πρέπει να αντιμετωπίζεται σαν μια ισχυρή στατική υπολογιστική μηχανή που δημιούργησε η εξέλιξη, παρά σαν ένα ευφύες σύστημα, ένα σύνολο από κανόνες διαρκούς προσαρμογής.



**Εικόνα 18: Διαφορετικές όψεις του ίδιου αντικειμένου**



### 3 Σύγχρονα Συστήματα Ανάκτησης Εικόνων με Βάση το Περιεχόμενο

Σε αυτό το κεφάλαιο παρουσιάζονται τα σημαντικότερα συστήματα ανάκτησης εικόνων με βάση το περιεχόμενο. Επίσης, γίνεται μία σύντομη αναφορά στα βασικότερα προβλήματα που παρουσιάζουν και στις βασικότερες μεθόδους αντιμετώπισης τους. Καθώς το ειδικό πεδίο εφαρμογής της παρούσας εργασίας είναι οι ιατρικές εικόνες, γίνεται ιδιαίτερος λόγος για τα συστήματα ανάκτησης ιατρικών εικόνων και για τα ιδιαίτερα χαρακτηριστικά αυτού του πεδίου.

#### 3.1 Ανάκτηση Εικόνων Γενικού Περιεχομένου

Πολλά είναι τα συστήματα CBIR γενικού περιεχομένου που συναντά κανείς στη βιβλιογραφία, ωστόσο τα περισσότερα διαχειρίζονται την οπτική πληροφορία με παρόμοιο τρόπο. Εν γένει, τα συστήματα ανάκτησης εικόνων βασίζονται σε ένα σύνολο από ανεξάρτητα κυρίαρχα οπτικά χαρακτηριστικά, όπως το χρώμα, η υφή (texture), το σχήμα ως ακμές ή περιγράμματα και σε μικρότερο βαθμό σε πιο εξειδικευμένα χαρακτηριστικά όπως μετασχηματισμοί εικόνων και ομοιόμορφες περιοχές [16]. Οι αντίστοιχες τιμές από αυτά τα χαρακτηριστικά συγκρίνονται ανάμεσα στις εικόνες συνήθως με κάποιο μέτρο απόστασης ή ομοιότητας. Στη συνέχεια τα επιμέρους αποτελέσματα συμπηφίζονται παράγοντας το τελικό αποτέλεσμα ανομοιότητας (ή ομοιότητας) των εικόνων. *Επομένως, η ανάκτηση παρόμοιων εικόνων συνήθως ανάγεται-υποβαθμίζεται σε ανάκτηση παρόμοιων οπτικών χαρακτηριστικών.*

Το κύριο πρόβλημα που αντιμετωπίζουν τέτοιες προσεγγίσεις είναι το **σημασιολογικό κενό** (semantic gap) ανάμεσα στο πραγματικό νόημα της εικόνας και την πληροφορία που αναπαριστούν αυτά τα ανεξάρτητα χαρακτηριστικά. Ωστόσο η ανεξάρτητη διαχείριση της πληροφορίας που προσφέρουν τα κυρίαρχα χαρακτηριστικά δεν εμπεριέχει το ουσιαστικό νόημα των εικόνων. Το ουσιαστικό νόημα των εικόνων εντοπίζεται στις σημασιολογικές οντότητες τους, οι οποίες αναπαριστούν αντικείμενα του πραγματικού ή ενός ιδεατού κόσμου. Όπως έγινε κατανοητό στο προηγούμενο κεφάλαιο, η βιολογική όραση μόνο στα πρώιμα στάδια της χρησιμοποιεί μία τέτοια ανεξάρτητη κωδικοποίηση. Ο εντοπισμός των σημασιολογικών οντοτήτων ή αλλιώς η τμηματοποίηση της εικόνας προκύπτει σε ανώτερα επίπεδα επεξεργασίας με το συνδυασμό αυτών των χαρακτηριστικών. Ωστόσο, η τμηματοποίηση των εικόνων δίχως τη γνώση του πεδίου εφαρμογής είναι ένα από τα άλυτα θέματα της υπολογιστικής όρασης. Ακόμα και αν ήταν δυνατός ο διαχωρισμός των αντικειμένων της εικόνας θα έπρεπε να υπάρχει η πληροφορία εντοπισμού εκείνων που περιέχουν το ουσιαστικό νόημά της. Κατά συνέπεια, τα περισσότερα από τα γνωστά CBIR συστήματα χρησιμοποιούν μεθόδους όπως διαδραστικές διεπαφές (π.χ. relevance feedback και incremental learning) ή/ και στοιχεία από την τεχνητή νοημοσύνη (π.χ. long-term learning) για να ξεπεράσουν αυτό το πρόβλημα. Στην πρώτη περίπτωση επιτυγχάνεται η προσαρμογή στις εκάστοτε επιθυμίες του χρήστη ενώ στη δεύτερη υπάρχει μία περισσότερο ευφυής διαχείριση και προσαρμογή στις ιδιότητες των εικόνων εισόδου. Μία άλλη ενδιαφέρουσα μέθοδος που δεν απαιτεί συνεχή αλληλεπίδραση με το χρήστη είναι η

καταγραφή αρχείων με την πληροφορία των παλαιότερων ανακτήσεων και προτιμήσεων του κάθε χρήστη (user log files). Στον αντίποδα τα εξειδικευμένα συστήματα μπορούν να χρησιμοποιούν την γνώση του πεδίου εφαρμογής (domain knowledge) για να επιλύσουν με αποδοτικούς μηχανισμούς τέτοια προβλήματα, χρησιμοποιώντας πολλές φορές και στοιχεία τεχνητής νοημοσύνης.

Ένα ακόμη πρόβλημα που παρουσιάζουν τα CBIR συστήματα γενικού περιεχομένου είναι η απουσία καλών μεθόδων αξιολόγησης τους [17]. Το γεγονός αυτό οφείλεται κυρίως στην υποκειμενικότητα του περιεχομένου των εικόνων που αναλύθηκε στην παράγραφο 1.1. Δύο είναι οι μετρικές που συνήθως χρησιμοποιούνται, η ακρίβεια της ανάκτησης (precision) και η ανάκληση (recall):

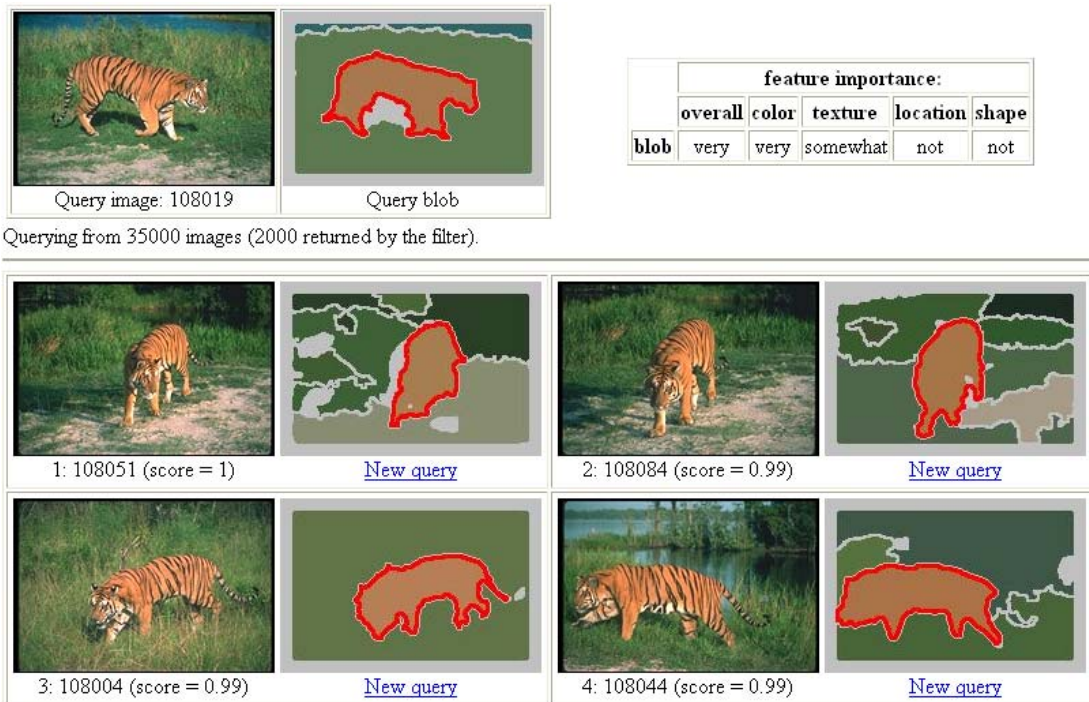
$$\text{ακρίβεια} = \frac{\text{πληθος σχετικών αντικειμένων που ανακτήθηκαν}}{\text{πλήθος αντικειμένων που ανακτήθηκαν}} \quad (1)$$

$$\text{ανάκληση} = \frac{\text{πληθος σχετικών αντικειμένων που ανακτήθηκαν}}{\text{πληθος σχετικών αντικειμένων}} \quad (2)$$

Πέρα από την αδυναμία αντικειμενικού προσδιορισμού της καλής ομοιότητας δύο εικόνων η αξιολόγηση των μεγεθών αυτών είναι σχετική ως προς τις παραμέτρους και την περιγραφή του εκάστοτε πειράματος. Όσο περισσότερη πληροφορία προσδίδεται στο σύστημα (π.χ. ανάκτηση από συγκεκριμένες κατηγορίες εικόνων, καθοδήγηση του συστήματος από το χρήστη) τόσο πιο εύκολη γίνεται η εργασία της ανάκτησης. Για παράδειγμα στην Εικόνα 19, όπου ο χρήστης υποδεικνύει στο σύστημα το αντικείμενο ενδιαφέροντος της εικόνας μόνο παρόμοια στη σημασιολογία αντικείμενα είναι δυνατό να θεωρηθούν σχετικά. Ομοίως, σε εξειδικευμένα συστήματα είναι δυνατό να καθοριστεί ο βαθμός συσχέτισης δύο εικόνων από το ίδιο πεδίο εφαρμογής με βάση τους αντικειμενικούς στόχους που έχουν καθοριστεί.

Το πιο γνωστό CBIR σύστημα γενικού περιεχομένου είναι το εμπορικό QBIC (Query By Image Content) της IBM. Ο χρήστης εκτελεί ανάκτηση ορίζοντας μία εικόνα δείγμα, ενώ έχει τη δυνατότητα να επιλέξει το επιθυμητό χρώμα ή texture από μία παλέτα. Επίσης του παρέχεται η δυνατότητα να σχεδιάσει ο ίδιος ένα σχήμα (human sketches) βάσει του οποίου θα γίνει η ανάκτηση. Στο Blobworld, του University of California, Berkeley, υπάρχει η δυνατότητα καθορισμού του θεματικού περιεχομένου της εικόνας επιλέγοντας μία από τις προσφερόμενες κατηγορίες εικόνων, όπως και του καθορισμού μιας περιοχής της εικόνας στην οποία θα επικεντρωθεί η ανάκτηση.



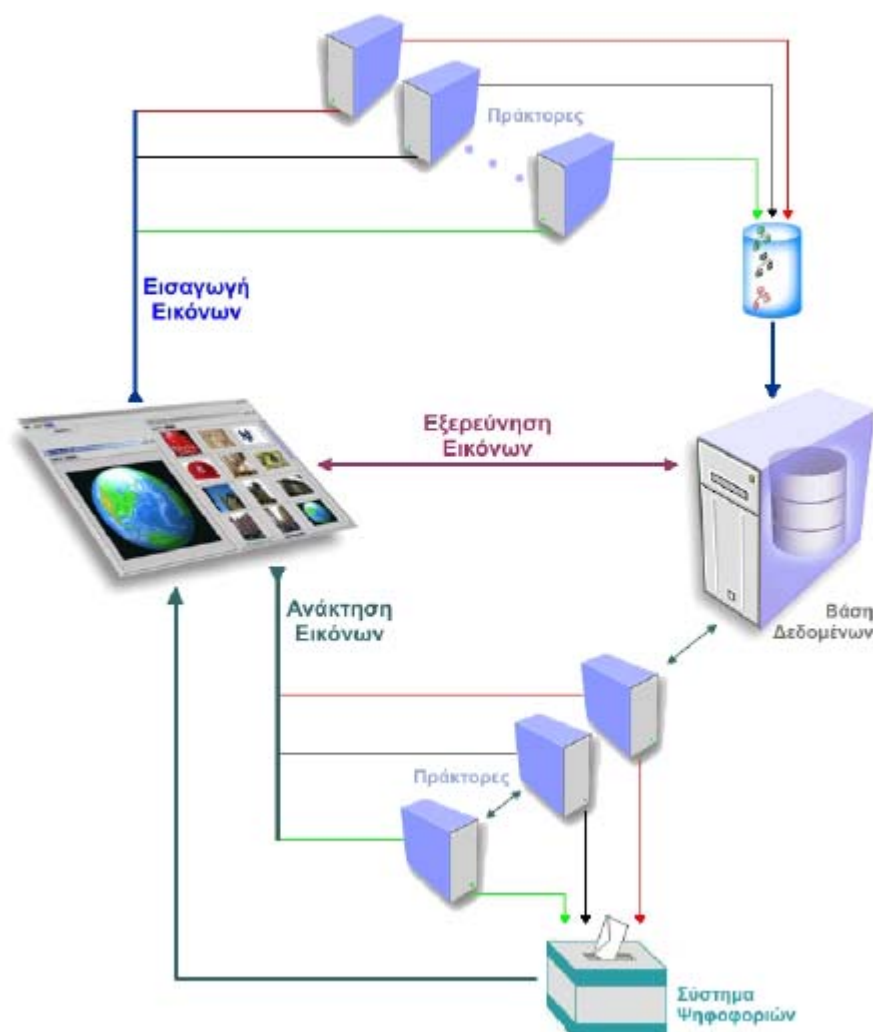


**Εικόνα 19: Τρόποι γεφύρωσης του σημασιολογικού κενού. Επιλογή περιοχών ενδιαφέροντος και άλλων χαρακτηριστικών από το χρήστη στο σύστημα Blobworld**

Το σύστημα ImageRover του Boston University χρησιμοποιεί μία τεχνική που ονομάζεται relevance feedback. Η διαδικασία της ανάκτησης επαναλαμβάνεται συνεχώς, δίνοντας στο χρήστη τη δυνατότητα να βελτιώνει σε κάθε επανάληψη τα αποτελέσματα που του προτείνονται, παρέχοντας έτσι μία μορφή ανάδρασης που επιτρέπει την προσαρμογή στις προτιμήσεις του. Τέλος, στο σύστημα Surfimage του INRIA, ο χρήστης επιλέγει τα χαρακτηριστικά που επιθυμεί και ορίζει τη συνεισφορά του καθ' ενός στην ανάκτηση, όπως επίσης και τις μετρικές που θα χρησιμοποιηθούν για την σύγκρισή τους ανάμεσα στις εικόνες. Μερικά άλλα σημαντικά συστήματα είναι τα Virage, το PicHunter, κ.α. Για μια εκτενέστερη περιγραφή και συγκριτική παρουσίαση των βασικότερων CBIR συστημάτων, ο αναγνώστης παραπέμπεται στο [16].

Τέλος, στα πλαίσια προηγούμενης εργασίας του Τμήματος Επιστήμης Υπολογιστών του Πανεπιστημίου Κρήτης [1],[18],[19], σχεδιάστηκε και υλοποιήθηκε μια πειραματική πλατφόρμα ανάκτησης εικόνων γενικού περιεχομένου. Η λειτουργικότητα της πλατφόρμας επιτυγχάνεται με ένα σύστημα πολλαπλών πρακτόρων. Οι πράκτορες στην προκειμένη περίπτωση αποτελούν παράλληλες ασύγχρονες διαδικασίες (χαλαρή θεώρηση της θεωρίας πρακτόρων) κάθε μία από τις οποίες είναι υπεύθυνη για τη διαχείριση ενός κυρίαρχου χαρακτηριστικού της εικόνας. Επομένως, η αρχιτεκτονική αυτή προσομοιάζει σε μεγάλο βαθμό το πρώιμο στάδιο επεξεργασίας της βιολογικής όρασης με τους παράλληλους χάρτες ανεξάρτητων πρωτογενών χαρακτηριστικών. Τα χαρακτηριστικά που εξετάστηκαν είναι διάφοροι τύποι χρώματος, υφής και σχήματος. Ο κάθε πράκτορας είναι υπεύθυνος τόσο για την εξαγωγή του χαρακτηριστικού που διαχειρίζεται κατά την εισαγωγή μίας εικόνας στη βάση όσο και για τη σύγκριση αυτού κατά τη διαδικασία της ανάκτησης. Το αποτέλεσμα ομοιότητας από κάθε πράκτορα οδηγείται μαζί με τα υπόλοιπα σε ένα σύστημα ψηφοφορίας το οποίο εξάγει ένα συμπέρασμα για τη

γενική ομοιότητα ανάμεσα σε δύο εικόνες. Η αρχιτεκτονική της πλατφόρμας απεικονίζεται στην Εικόνα 20.



Εικόνα 20: Η αρχιτεκτονική της πειραματικής πλατφόρμας ανάκτησης εικόνων [1]

### 3.2 Ανάκτηση Εικόνων στην Ιατρική

Οι ιατρικές εικόνες έχουν χρησιμοποιηθεί κατά κόρο για τη δοκιμή, αξιολόγηση και παρουσίαση των αποτελεσμάτων από τα διάφορα συστήματα ανάκτησης εικόνων. Περαιτέρω, ο τομέας της Ιατρικής συχνά αναφέρεται ως ένα από τα πρωταρχικά πεδία εφαρμογής τέτοιων συστημάτων, κυρίως λόγω των δυνητικών ωφελειών και της απήχησης τους στην κοινωνία. Για τους ίδιους λόγους, έχει μελετηθεί σε μεγάλο βαθμό η ενσωμάτωση των CBIR εργαλείων σε ολοκληρωμένα ιατρικά συστήματα. Για παράδειγμα στο [20] γίνεται λόγος για τη δυνατότητα ενσωμάτωσης των συστημάτων CBIR στο PACS (Picture Archiving and Communications Systems) Επίσης στο [17] γίνεται μία ολοκληρωμένη μελέτη των

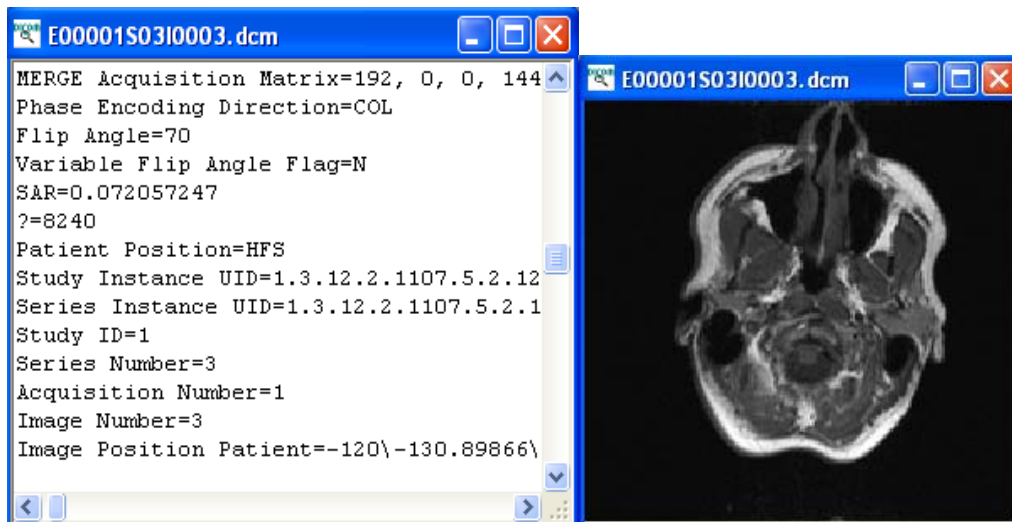
και των βασικότερων θεμάτων ανάκτησης εικόνων στην Ιατρική και των σημαντικότερων συστημάτων που έχουν υλοποιηθεί. Σε αυτή την παράγραφο εξετάζονται τα ιδιαίτερα χαρακτηριστικά της ανάκτησης εικόνων στο πεδίο της Ιατρικής και του ρόλου της σε ένα ολοκληρωμένο σύστημα υγείας.

### **3.2.1 Ιδιαίτερα χαρακτηριστικά ανάκτησης ιατρικών εικόνων**

Οι τεχνικές που χρησιμοποιούνται στην ανάκτηση ιατρικών εικόνων είναι εν γένη παραπλήσιες με εκείνες που αναφέρθηκαν στην παρ. 3.1 για τις γενικές εφαρμογές. Εντούτοις, το πεδίο της Ιατρικής παρουσιάζει ορισμένες ιδιαιτερότητες που εστιάζονται κυρίως σε τρεις παράγοντες: τα σημασιολογικά μεταδεδομένα που συχνά συνοδεύουν τις ιατρικές εικόνες, την δυνατότητα χρήσης εικόνων αναφοράς και τέλος τα ιδιαίτερα οπτικά χαρακτηριστικά που χρησιμοποιούνται για να εκφράσουν το περιεχόμενο των ιατρικών εικόνων.

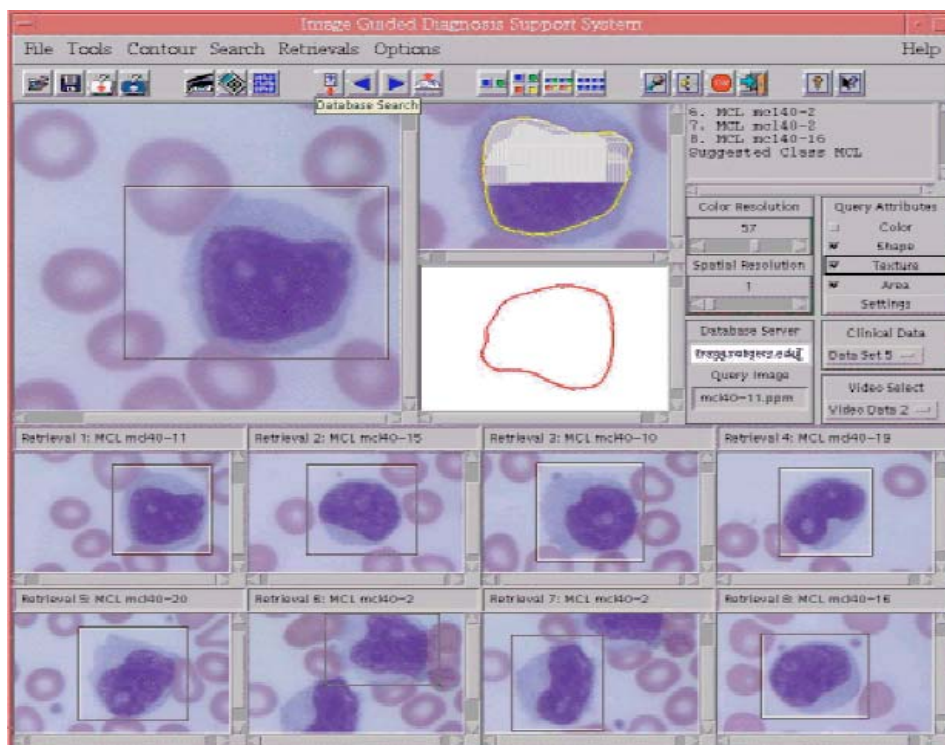
Η προσάρτηση μεταδεδομένων αν και «δαπανηρή» διαδικασία δεν είναι τόσο σπάνιο φαινόμενο στις ιατρικές εικόνες. Στο πρότυπο DICOM (Digital Imaging and Communications in Medicine) μία ιατρική εικόνα συνοδεύεται από ένα πολύ μεγάλο σύνολο από οπτικά χαρακτηριστικά της, στοιχεία του εξεταζόμενου και πιθανότατα περιγραφή ή διάγνωση του περιστατικού (Εικόνα 21). Μερικές από αυτές τις παραμέτρους καθορίζουν την τεχνική απεικόνισης (modality), τη σχετική θέση του ασθενούς ως προς το μηχάνημα, το εξεταζόμενο όργανο (Image Position Patient) κ.α. Είναι προφανές ότι αυτή η σημασιολογική συνοδευτική πληροφορία κειμένου μπορεί να χρησιμοποιηθεί ώστε να διευκολύνει και να εμπλουτίσει τα συστήματα ανάκτησης ιατρικών εικόνων.

Επιπρόσθετα, σημαντική δουλειά έχει γίνει τα τελευταία χρόνια για την ομογενοποίηση των ιατρικών όρων, με πρωτοπόρο την Εθνική Βιβλιοθήκη της Ιατρικής των Η.Π.Α., NLM (National Library of Medicine) που δημιούργησε την οντολογία ιατρικών όρων MeSH (medical subject headings) και τέλος τον εξυπηρετή γνώσης UMLS (Universal Medical Language System). Κατά συνέπεια, υπάρχει ένα κοινό σημείο αναφοράς στη χρησιμοποιούμενη ορολογία των μεταδεδομένων που διευκολύνει τη χρήση τους (για παράδειγμα την ανάγνωση ή την προσάρτησή τους στις εικόνες). Τα παραπάνω σημασιολογικά χαρακτηριστικά στις ιατρικές εικόνες καθιστούν αρκετά αποδοτική την ανάκτηση τους από μηχανές ανάκτησης με βάση το κείμενο. Καθώς όμως η ανάκτηση που βασίζεται στο οπτικό περιεχόμενο έχει να επιδείξει ιδιαίτερα πλεονεκτήματα (παρ 3.2.2), υβριδικά συστήματα που χρησιμοποιούν τόσο τη λεκτική όσο και την οπτική πληροφορία, σίγουρα θα αποτελέσουν στο μέλλον μία σημαντική σχεδιαστική επιλογή.



Εικόνα 21: Η αναπαράσταση μίας εικόνας και των μεταδεδομένων της από το πρότυπο DICOM

Πέρα από τη χρήση λεκτικής πληροφορίας, μία ακόμη δυνατότητα που διευκολύνει την ανάλυση και ανάκτηση ιατρικών εικόνων είναι η δυνατότητα χρήσης εικόνων αναφοράς. Οι εικόνες αναφοράς σε επίπεδο ανατομίας αφορούν ένα πρότυπο απεικόνισης ενός συγκεκριμένου οργάνου του υγιούς ανθρώπινου σώματος υπό διάφορες γωνίες και απεικονιστικές τεχνικές. Κάτι αντίστοιχο σε επίπεδο παθήσεων είναι δυνατό να υπάρχουν διάφορα χαρακτηριστικά οπτικά πρότυπα των παθήσεων. Η χρήση αυτής της γνώσης από το πεδίο εφαρμογής είναι δυνατό να διευκολύνει τη διαδικασία ανάκτησης, ωστόσο πρέπει να συσταθούν ραδιομετρικά πρότυπα εικόνων ώστε να υπάρχει ένα κοινό σημείο αναφοράς.



Εικόνα 22: Καθοδηγούμενη από οπτικό περιεχόμενο κλινική υποστήριξη για εικόνες παθολογίας στο σύστημα IGDS. Κυρίαρχο χαρακτηριστικό, το σχήμα του πυρήνα των κυττάρων.

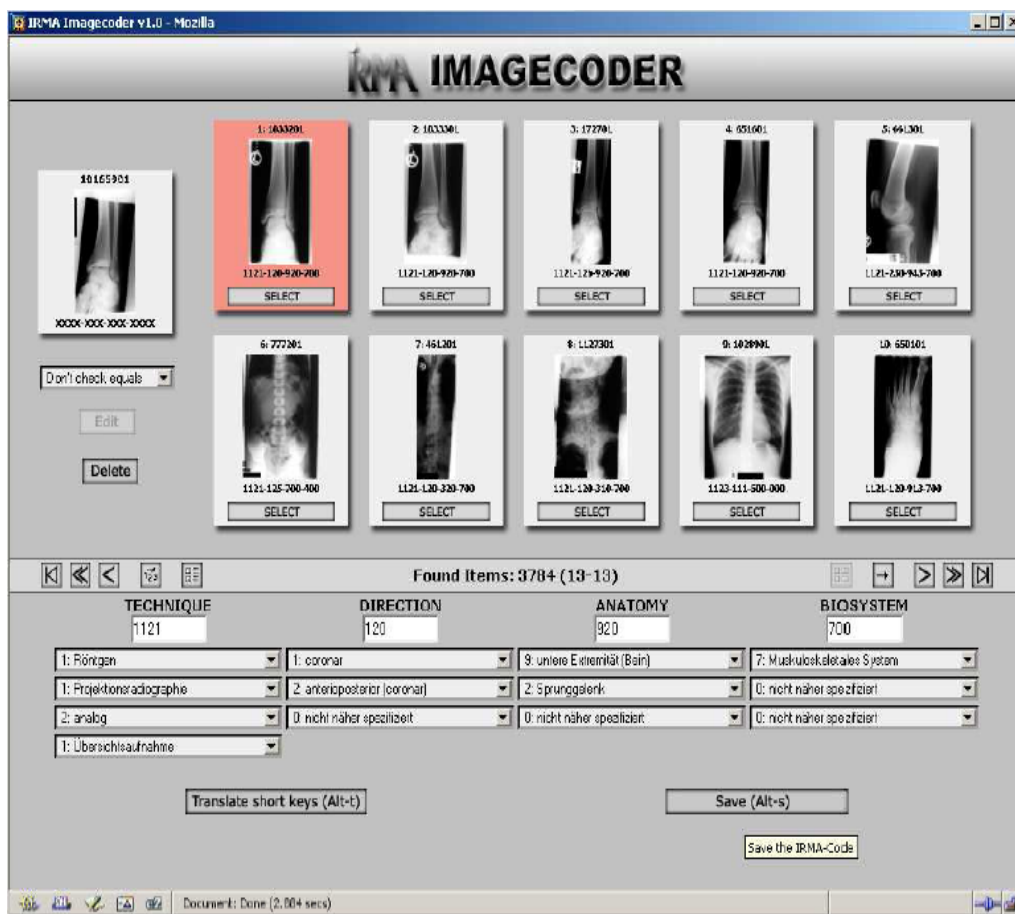
Τέλος, όσον αφορά τα οπτικά χαρακτηριστικά που προτιμούνται για την εξαγωγή εκφραστικής πληροφορίας από ιατρικές εικόνες, αυτά είναι εν γένη τα ίδια με εκείνα των γενικών εφαρμογών. Ωστόσο, σε κάθε κατηγορία ιατρικών εικόνων είναι διαφορετικά τα οπτικά χαρακτηριστικά που αποκτούν ιδιαίτερη σημασία. Για παράδειγμα, στις στεφανιαίες μαγνητικές τομογραφίες εγκεφάλου η συμμετρία είναι μία πολύ χαρακτηριστική ιδιότητα από την οποία μπορεί να εξαχθεί χρήσιμη πληροφορία. Επίσης, το σχήμα είναι συνήθως αρκετά σημαντικό στο χαρακτηρισμό και κατηγοριοποίηση περιοχών παθολογίας στις περισσότερες ασθένειες (σχήμα όγκου, σχήμα πυρήνα κυττάρου (Εικόνα 22) κ.α). Περαιτέρω, η τοποθεσία των περιοχών αυτών στο ανθρώπινο σώμα αποτελεί σημαντική κλινική πληροφορία. Συνεπώς, στις ιατρικές εφαρμογές είναι πολύ συνήθεις οι γεωμετρικές συσχετίσεις των διαφόρων περιοχών στην εικόνα. Εν τέλει, τα συστήματα ανάκτησης ιατρικών εικόνων συνήθως χρησιμοποιούν διαφορετικό τρόπο αντιμετώπισης ανάλογα με την εκάστοτε κατηγορία εικόνων, ενώ πολλά συστήματα έχουν υλοποιηθεί αποκλειστικά για συγκεκριμένες ιατρικές εικόνες και εφαρμογές.

### 3.2.2 Χρησιμότητα του CBIR σε κλινικές εφαρμογές

Στις μέρες μας, ένα σημαντικό μέρος της ιατρικής πληροφορίας συναντάται με τη μορφή ιατρικών εικόνων. Υπάρχουν διάφορα είδη ιατρικών απεικονιστικών μεθόδων, κάθε μία από τις οποίες προσπαθεί να αναδείξει την ανθρώπινη μορφολογία και παθολογία βασισμένη στις διαφορετικές ιδιότητες που παρουσιάζουν οι ιστοί και τα οστά. Κατά συνέπεια, η κλινική διάγνωση ενός υπό εξέταση περιστατικού βασίζεται σε ένα μεγάλο μέρος στη σωστή ανάγνωση αυτών των εικόνων. Μολονότι θα ήταν επιθυμητή μία αυτοματοποιημένη εξαγωγή της διάγνωσης από το ίδιο το σύστημα ανάκτησης, συνήθως την αποστολή αυτή επιφορτίζονται εξειδικευμένοι ιατροί-ακτινολόγοι. Πράγματι, η ανθρώπινη όραση και αντίληψη, είναι ένα πολύ ισχυρό εργαλείο, για την ανάλυση και εξαγωγή της χρήσιμης πληροφορίας από εικόνες. Ωστόσο, ορισμένες άλλες ιδιότητες συνυφασμένες με την ανθρώπινη φύση, όπως το ανθρώπινο σφάλμα και η παράληψη, η υποκειμενικότητα της κρίσης, και η περιορισμένη ταχύτητα επεξεργασίας μεγάλου όγκου δεδομένων, καθιστούν αναγκαία την αναζήτηση αυτοματοποιημένων αμερόληπτων στατιστικών διαγνωστικών συστημάτων. Αν και η αυτοματοποιημένη διάγνωση δεν είναι εν γένει ασφαλής, μία διαφορετική άποψη είναι πάντοτε χρήσιμη σε ζητήματα τόσο σοβαρά όπως η ιατρική διαγνωστική. Σε αυτόν τον τομέα, το CBIR, πέρα από τις αντικειμενικές δυσκολίες, παρουσιάζει σημαντικά πλεονεκτήματα έναντι των μεθόδων ανάκτησης με χρήση λεκτικών μεταδεδομένων. *Τα οπτικά χαρακτηριστικά δεν επιτρέπουν μόνο την ανάκτηση περιστατικών από ασθενείς με παρόμοια διάγνωση αλλά και περιστατικά που παρουσιάζουν στενή οπτική ομοιότητα παρά τη διαφορετική διάγνωση.* Αντίθετα, η χρήση λέξεων κλειδιών προφανώς δεν θα μπορούσε να αποδώσει παρόμοια αποτελέσματα.

Σε επίπεδο ανατομικής ομοιότητας η ανάκτηση παρόμοιων εικόνων επιτρέπει την αυτόματη κατηγοριοποίηση των εικόνων ως προς το απεικονιζόμενο όργανο και την τεχνική απεικόνισης. Σε αυτό μπορούν να φανεί χρήσιμη η χρήση χαρακτηριστικών εικόνων αναφοράς. Σε επίπεδο εντοπισμένων περιοχών, είναι δυνατό να οδηγήσει στην ταξινόμηση των ασθενειών και στην υποστήριξη της **κλινικής διάγνωσης** (clinical decision support), πιθανότατα σε συνδυασμό με άλλες εξετάσεις. Επιπρόσθετα, όπως ειπώθηκε στην παράγραφο 3.2.1, οι εικόνες στα σύγχρονα ιατρικά συστήματα, συχνά συνοδεύονται από μεταδεδομένα τα οποία είναι

δυνατό να περιέχουν, συν τοις άλλοις, και την ασφαλή διαγνωστική πληροφορία (π.χ. από διενέργεια βιοψίας). Η στατιστική επεξεργασία των μεταδεδομένων, από τις ανακτηθείσες εικόνες, προφανώς θα μπορούσε να αποτελέσει ένα διαγνωστικό κριτήριο (Image and Metadata Based Reasoning) για το υπό εξέταση περιστατικό. Επιπρόσθετα, η διάγνωση που συνοδεύει μία εικόνα είναι δυνατό στο μέλλον και πρέπει να συνοδεύεται από τον παράγοντα της αξιοπιστίας της χρησιμοποιούμενης μεθόδου (Evidence Quality), όπως ορίζει η σύγχρονη, βασισμένη στις ενδείξεις ιατρική (Evidence Based Medicine). Με αυτόν τον τρόπο, στο τελικό αποτέλεσμα είναι δυνατό να συνοπλογίζεται και ο βαθμός ποιότητας της κάθε διαγνωστικής μεθόδου. Το γεγονός αυτό οδηγεί ρητά στην ανάγκη αξιολόγησης των CBIR συστημάτων που όπως ειπώθηκε στη παράγραφο 3.1 αποτελεί ένα από τα ουσιαστικά προβλήματα των εφαρμογών αυτών.



Εικόνα 23: Ημιαυτόματος σχολιασμός ιατρικών εικόνων από το σύστημα IRMA

Πέρα από την ιατρική διαγνωστική, που αποτελεί τον πιο ουσιαστικό αλλά και δύσκολο στόχο, η ανάκτηση ιατρικών εικόνων είναι σημαντική όσον αφορά επιμορφωτικούς και εκπαιδευτικούς λόγους. Η ανάκτηση εικόνων επιτρέπει την εξερεύνηση σε μεγάλες βάσεις εικόνων με στόχο τη σύγκριση της διάγνωσης ανάμεσα σε παρόμοια οπτικά περιστατικά. Επίσης, είναι ιδανική για μελέτη του εύρους των πιθανών απεικονίσεων κάποιας πάθησης και ενδιαφέροντες ή και επικίνδυνους οπτικούς συσχετισμούς ανάμεσα σε παρόμοιες οπτικά παθήσεις. Στο άμεσο μέλλον λέκτορες θα έχουν τη δυνατότητα επιλεκτικής πρόσβασης σε μεγάλες βάσεις δεδομένων, όσον αφορά ενδιαφέροντα ή εκφραστικά παραδείγματα

περιστατικών. Επιπρόσθετα, περιστατικά χωρίς προσωπικά δεδομένα είναι δυνατό να διακινούνται ελεύθερα σε φοιτητές Ιατρικής ή στο διαδίκτυο για εκπαιδευτικούς λόγους.

Τέλος, η ανάκτηση εικόνων μπορεί να φανεί εξαιρετικά χρήσιμη σε ερευνητικές προσπάθειες, καθώς είναι πιθανό να προκύψουν νέες συσχετίσεις ανάμεσα σε παθήσεις και τα οπτικά χαρακτηριστικά που αυτές συνήθως εμφανίζουν. Επίσης, δεν πρέπει να παραλείπεται η χρονική παράμετρος των απεικονίσεων. Η παρακολούθηση των οπτικών ιδιοτήτων μίας πάθησης στα διάφορα στάδια εξέλιξης της αποτελεί ένα εξαιρετικά ενδιαφέρον ζήτημα, στο οποίο το CBIR μπορεί να φανεί εξαιρετικά χρήσιμο. Επιπρόσθετα, με την ανάκτηση εικόνων είναι δυνατό να προκύψουν πληροφορίες σχετικά με τη ηλικία των ασθενών και τις οπτικές ιδιότητες της εκάστοτε πάθησης που παρουσιάζουν. Επομένως η ανάκτηση περιστατικών με βάση το οπτικό τους περιεχόμενο μπορεί να οδηγήσει στην εξόρυξη σημαντικής γνώσης.

### 3.2.3 Σημαντικά συστήματα ανάκτησης ιατρικών εικόνων

Όπως ειπώθηκε, οι ιατρικές εικόνες έχουν χρησιμοποιηθεί κατά κόρο για τη δοκιμή, αξιολόγηση και παρουσίαση των αποτελεσμάτων από τα διάφορα συστήματα ανάκτησης εικόνων. Πολλά είναι τα συστήματα CBIR γενικού περιεχομένου που στη συνέχεια εξειδικεύθηκαν σε κάποιο πεδίο ιατρικών εικόνων. Περαιτέρω, αρκετά είναι και τα συστήματα που υλοποιήθηκαν με αποκλειστικό στόχο την ταξινόμηση και ανάκτηση εικόνων στο πεδίο της Ιατρικής. Στο [17] γίνεται μία σύνοψη όλων αυτών των συστημάτων και των εξειδικεύσεών τους σε συγκεκριμένες ανατομικές και παθολογικές κατηγορίες (Εικόνα 24).

Images Used	Names of The Systems
HRCTs of the lung	ASSERT
Functional PET	FICBDS
Spine X-rays	CBIR2, MIRS
Pathologic images	IDEM, I-Browse, PathFinder, PathMaster, IGDS
CTs of the head	MIMS
Mammographies	APKS
Images from biology	BioImage, BIRN
Dermatology	MELDOQ, MEDS
Breast cancer biopsies	BASS
Varied images	I <sup>2</sup> C, IRMA, KMed, COBRA, MedGIFT, ImageEngine

Εικόνα 24: Διάφοροι τύποι Ιατρικών εικόνων και τα συστήματα ανάκτησης που τις χρησιμοποιούν

Πέρα από αυτά τα ολοκληρωμένα συστήματα, σημαντική εργασία έχει γίνει και στις μαγνητικές τομογραφίες εγκεφάλου [20],[22],[23],[24]. Ωστόσο, αν και προτείνονται εξειδικευμένοι αλγόριθμοι για την ανάλυση και ανάκτηση των εικόνων αυτών, δεν γίνεται καμία προσπάθεια ενσωμάτωσής τους (medical integration) στα σύγχρονα ιατρικά συστήματα.





## 4 Η Προτεινόμενη Αρχιτεκτονική Ανάκτησης Εικόνων

Σε προηγούμενο κεφάλαιο εξεξηγήθηκε πως η βιολογική όραση και αντίληψη αποτελούν μια σύνθετη και ευφυή διαδικασία στους μηχανισμούς της οποίας δεν υπάρχει σαφής και πλήρης επιστημονική γνώση. Κατά συνέπεια, η μοντελοποίηση και η επιτυχημένη λειτουργία υπολογιστικών συστημάτων που προσομοιάζουν τη βιολογική όραση είναι εξαιρετικά δύσκολη. Τα περισσότερα συστήματα ανάκτησης εικόνων γενικού περιεχομένου προσομοιάζουν μόνο τα πρώιμα στάδια της βιολογικής όρασης διαχωρίζοντας την οπτική πληροφορία στα κυρίαρχα οπτικά χαρακτηριστικά της. Περαιτέρω, σε εξειδικευμένα πεδία όπως αυτό των ιατρικών εικόνων πολλά συστήματα προχωρούν στην τμηματοποίηση των περιοχών ενδιαφέροντος και την εξαγωγή εξειδικευμένης οπτικής πληροφορίας από αυτές. Εκεί εμπεριέχεται σε μεγάλο βαθμό η σημασιολογία των εικόνων. Ωστόσο, είναι λιγιστά εκείνα τα συστήματα που εκμεταλλεύονται τις σύγχρονες τάσεις της τεχνητής νοημοσύνης. Οι τεχνικές που χρησιμοποιούνται βασίζονται στην αλληλεπίδραση με το χρήστη για προσαρμογή στις προτιμήσεις της συγκεκριμένης ανάκτησης (π.χ. incremental learning) ή και στην χρήση αρχείων καταγραφής παλαιότερων ανακτήσεων του κάθε χρήστη (user log files). Λίγα είναι τα συστήματα που προσπαθούν να εξειδικευθούν στους συγκεκριμένους στόχους ενός συγκεκριμένου πεδίου εικόνων, μέσω μεθόδων μακροπρόθεσμης μάθησης. Σε αυτό το κεφάλαιο παρουσιάζεται μία πρωτότυπη βιολογικά εμπνευσμένη αρχιτεκτονική για εξειδικευμένα συστήματα ανάκτησης εικόνων, που θέτει τις κατάλληλες βάσεις για μια περισσότερο ευφυή προσέγγιση με βάση τη θεωρία των αυτόνομων πρακτόρων.

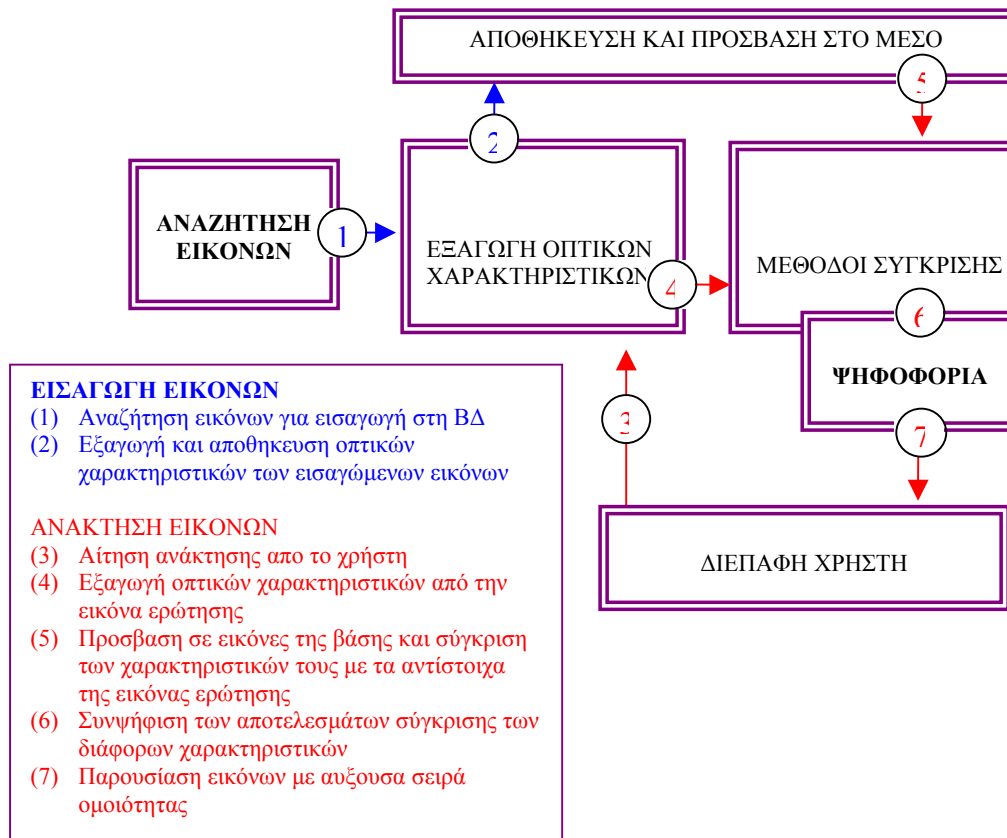
### 4.1 Παρουσίαση των Βασικών Λειτουργιών του Συστήματος

Σε αυτήν την παράγραφο παρουσιάζονται οι βασικές λειτουργίες του προτεινόμενου συστήματος που αποτελεί έναν πράκτορα ανάκτησης εξειδικευμένων εικόνων. Δύο είναι οι κύριοι στόχοι του προτεινόμενου πράκτορα: Πρώτον, η αναζήτηση εικόνων του εκάστοτε πεδίου εφαρμογής για επεξεργασία και εισαγωγή στον αποθηκευτικό του χώρο. Δεύτερον, η εξυπηρέτηση αιτήσεων ανάκτησης από τους απομακρυσμένους χρήστες. Οι βασικές λειτουργίες του συστήματος παρουσιάζονται στην Εικόνα 25.

Η αναζήτηση εικόνων γίνεται από καθορισμένους -προς το παρόν- πόρους και είναι μία συνεχής, δια-βίου λειτουργία. Για παράδειγμα, ένας πράκτορας εξειδικευμένος σε μαγνητικές τομογραφίες εγκεφάλου ελέγχει διαρκώς νέες εγγραφές εικόνων από συγκεκριμένες ιατρικές βάσεις. Με αυτόν τον τρόπο επιτυγχάνεται ο καθορισμός ενός δυναμικού περιβάλλοντος εικόνων. Όπως αναφέρθηκε στην παράγραφο 2.3.2 η ύπαρξη ενός δυναμικού περιβάλλοντος είναι απαραίτητη κατά τη μοντελοποίηση μίας αυτόνομης οντότητας σύμφωνα με τη θεωρία πρακτόρων. Καθώς νέες εικόνες καταγράφονται στους πόρους αυτούς με αδιαφανή για τον πράκτορα τρόπο εκείνος είναι επιφορτισμένος με την αναγνώριση των οπτικών ιδιοτήτων τους που τις εντάσσουν στο συγκεκριμένο πεδίο εφαρμογής. Η ύπαρξη μεταδεδομένων όπως θα γίνει κατανοητό στην παράγραφο μπορεί να διευκολύνει και να επιταχύνει τη διαδικασία αυτή. Από κάθε εικόνα του πεδίου εφαρμογής, εξάγεται

η απαραίτητη οπτική πληροφορία και αποθηκεύεται στον αποθηκευτικό χώρο του πράκτορα. Οι εικόνες αυτές αποτελούν τη συλλογή των εικόνων του. Στην παρούσα εργασία γίνεται η υπόθεση πως οι πηγές εικόνων περιέχουν μόνο εικόνες από το εξειδικευμένο πεδίο εφαρμογής, ωστόσο ο πράκτορας αναλαμβάνει δυναμικά και αυτόνομα τον εντοπισμό και την εισαγωγή τους στη βάση εικόνων του, δίχως να απαιτείται παρεμβολή του χρήστη. Ο πράκτορας αυτός αποτελεί τον υπερ-πράκτορα της προτεινόμενης πλατφόρμας, καθώς, τη διαχείριση των οπτικών χαρακτηριστικών αναλαμβάνουν σε χαμηλότερο επίπεδο ανεξάρτητοι οπτικοί υπο-πράκτορες.

Κατά τη διενέργεια της εξυπηρέτησης των αιτήσεων ανάκτησης, αρχικά εξάγεται η οπτική πληροφορία της εικόνας ερώτησης του χρήστη. Στη συνέχεια, ανακτώνται, μία προς μία, οι εικόνες από τη συλλογή του πράκτορα και συγκρίνονται με την εικόνα ερώτησης ως προς την οπτικό τους περιεχόμενο. Οι εικόνες της συλλογής εμφανίζονται στη δυναμική διεπαφή του χρήστη κατά τη χρονική σειρά ανάκτησης τους, αφού όμως τοποθετηθούν στη σωστή θέση κατά φθίνουσα σειρά ομοιότητας. Η διαδικασία τελειώνει οποιαδήποτε στιγμή ο χρήστης ικανοποιηθεί με τις ήδη ανακτημένες εικόνες ή όταν ο εξαντληθούν οι εικόνες της συλλογής του πράκτορα.



Εικόνα 25: Οι βασικές λειτουργίες του συστήματος ανάκτησης εικόνων

Όπως φαίνεται στην Εικόνα 25, το οπτικό περιεχόμενο των εικόνων αναλύεται υπό τη μορφή κυρίαρχων οπτικών χαρακτηριστικών. Κατά συνέπεια, κατά τη διαδικασία ανάκτησης, το αποτέλεσμα της οπτικής ομοιότητας ανάγεται σε σύγκριση των επιμέρους οπτικών χαρακτηριστικών. Για την εξαγωγή του συνολικού αποτελέσματος της σύγκρισης, τα αποτελέσματα ομοιότητας από τα διάφορα χαρακτηριστικά συμψηφίζονται. Όπως ειπώθηκε στην παράγραφο 3.1, τέτοιες

προσεγγίσεις παρουσιάζουν το πρόβλημα του σημασιολογικού κενού ανάμεσα στο πραγματικό νόημα της εικόνας και την εκφραστικότητα που προσφέρει η ανεξάρτητη διαχείριση των χαρακτηριστικών αυτών. Στην παράγραφο 4.2, επεξηγείται πως η προτεινόμενη αρχιτεκτονική CBIR, χρησιμοποιώντας αποκλειστικά εικόνες από συγκεκριμένα πεδία εφαρμογής, επιτυγχάνει μία περισσότερο σημασιολογική αναπαράστασή τους. Ωστόσο, πριν γίνει αυτό, κρίνεται απαραίτητο να αναλυθούν οι βασικές μονάδες εξαγωγής και σύγκρισης των οπτικών χαρακτηριστικών .

#### **4.1.1 Πράκτορες ανάλυσης και σύγκρισης οπτικών χαρακτηριστικών, η λειτουργική μονάδα του συστήματος**

Η λειτουργική μονάδα οπτικής επεξεργασίας στην προτεινόμενη αρχιτεκτονική είναι ο οπτικός πράκτορας. Ένας οπτικός πράκτορας είναι επιφορτισμένος με την διαχείριση συγκεκριμένου τύπου οπτικής πληροφορίας από μία εικόνα ή μία οροθετημένη περιοχή της εικόνας. Κατά την ανάλυση μίας εικόνας (ή υπο-περιοχής της), ο οπτικός πράκτορας εξάγει από αυτήν την απαραίτητη πληροφορία, που συνήθως αποτελεί ένα κυρίαρχο οπτικό χαρακτηριστικό όπως το χρώμα, η υφή, οι ακμές, το σχήμα, η συμμετρία κ.α. Στη συνέχεια, αποθηκεύει την πληροφορία αυτή στον προσωπικό του χώρο αποθήκευσης μέσα στη ΒΔ του συστήματος. Κατά τη σύγκριση δύο εικόνων (ή υπο-περιοχών τους) ο πράκτορας συγκρίνει τη συγκεκριμένη πληροφορία από τις δύο εικόνες και εξάγει ένα μέτρο ανομοιότητας ανάμεσά τους. Συνεπώς ένας οπτικός πράκτορας διαθέτει διαδικασίες εξαγωγής, αποθήκευσης και σύγκρισης συγκεκριμένου τύπου οπτικής πληροφορίας. Πέρα από τα παραπάνω, στη συγκεκριμένη αρχιτεκτονική οι πράκτορες διαθέτουν διαδικασίες διαρκούς απόκτησης στατιστικής εμπειρίας από τα αποτελέσματα των συγκρίσεων που διενεργούν, όπως η τυπική απόκλιση του συνόλου των αποτελεσμάτων σύγκρισης, τις μέγιστες και ελάχιστες ανομοιότητες που έχουν συναντήσει κ.α. Η χρησιμότητα αυτών των μετρήσεων σχετίζεται με την απόκτηση γνώσης από τον υπο-πράκτορα σχετικά με το συγκεκριμένο πεδίο εικόνων της εφαρμογής και θα αναλυθεί περαιτέρω στη συνέχεια.

#### **4.1.2 Ομάδες πρακτόρων για αποφάσεις οπτικής ομοιότητας**

Ενώ οι οπτικοί πράκτορες είναι επιφορτισμένοι με τη διαχείριση ενός συγκεκριμένου χαρακτηριστικού της εικόνας, ομάδες πρακτόρων χρησιμοποιούνται για τη συνολική απόφαση ομοιότητας δύο εικόνων ή δυο περιοχών. Σε αυτό το υψηλότερο επίπεδο της αρχιτεκτονικής οι αποφάσεις από τους διάφορους πράκτορες συμψηφίζονται. Τόσο κατά τη διαδικασία εξαγωγής πληροφορίας από μία εικόνα όσο και κατά την διαδικασία σύγκρισης των εικόνων οι επιμέρους πράκτορες λειτουργούν παράλληλα και ασύγχρονα. Μία τέτοια πλατφόρμα πρακτόρων προτάθηκε στο [1] (Εικόνα 20), και δοκιμάστηκε για διάφορους συνδυασμούς οπτικών χαρακτηριστικών και διάφορες μεθόδους ψηφοφορίας ανάμεσά τους. Τέτοιες ομάδες πρακτόρων όπως ειπώθηκε στην παράγραφο 2.3.2 αποτελούν στη «χαλαρή» τους θεώρηση τα συστήματα πρακτόρων του Brooks [4] .

Για το συμψηφισμό των αποφάσεων ομοιότητας από τους επιμέρους πρακτόρων μιας ομάδας επιλέχθηκε η μέθοδος της σταθμισμένης ψηφοφορίας. Κάθε ένας από τους πράκτορες σε μία ομάδα διαθέτει μία τιμή  $w_i$ , που αποτελεί την

ποσοστιαία συνεισφορά του στην τελική απόφαση ομοιότητας. Οι τιμές  $w_i$  των πρακτόρων επιλέγονται ούτως ώστε να εκφράζουν την σχετική αξία των οπτικών χαρακτηριστικών στην εξειδικευμένη οπτική εργασία και να επιτυγχάνεται καλύτερη σημασιολογική ανάκτηση. Για παράδειγμα, σε άλλες εφαρμογές έχει περισσότερη αξία η χρωματική ομοιότητα δύο αντικειμένων ενώ σε άλλες το σχήμα τους. Στην παράγραφο 4.4.1 γίνεται ιδιαίτερη αναφορά σε αυτήν την παράμετρο. Αν  $N$  το πλήθος των πρακτόρων  $i$  που συμμετέχουν στην ομάδα απόφασης και  $d_i$ ,  $i: [1..N]$ , το διάνυσμα των αντίστοιχων ανεξάρτητων αποτελεσμάτων ανομοιότητας, τότε η συνολική ανομοιότητα δύο εικόνων ή υπο-περιοχών δίνεται από τη σχέση:

$$S = \sum_{i=1}^N w_i \frac{d_i}{\sigma_i(D_i)} \quad (3)$$

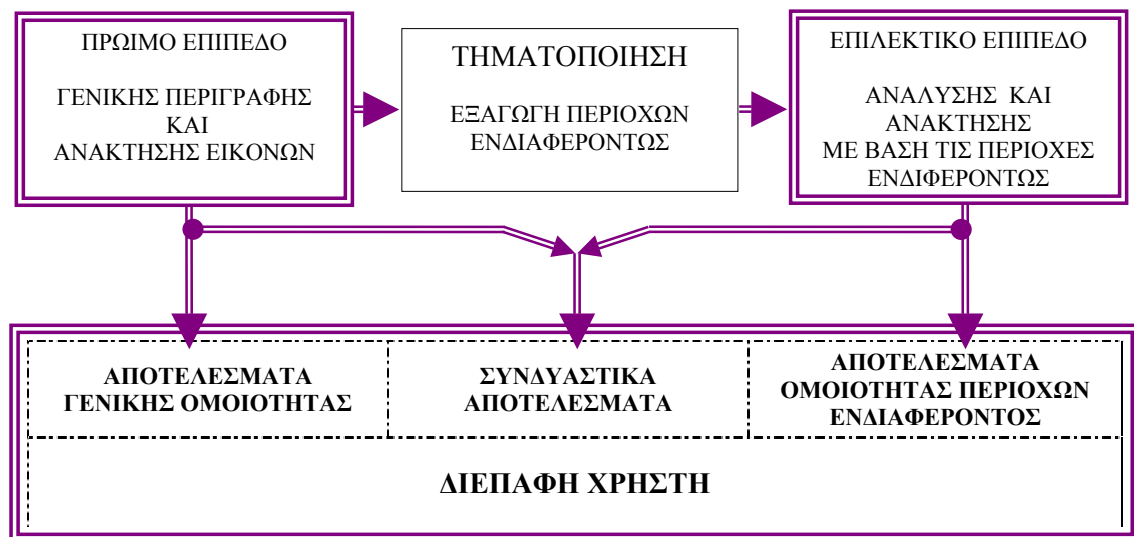
όπου  $D_i$  είναι το διάνυσμα των αποτελεσμάτων από όλα τα δείγματα σύγκρισης που έχει υπολογίσει ο οπτικός πράκτορας  $i$  όχι μόνο στη συγκεκριμένη ανάκτηση αλλά από τη γέννηση του συστήματος. Το  $\sigma(D_i)$  αντιπροσωπεύει τη τυπική απόκλιση όλων αυτών των αποστάσεων και χρησιμοποιείται στην εξίσωση (3) ώστε να κανονικοποιήσει τις διαφορετικές ποσότητες που ο κάθε πράκτορας εξάγει. Κάθε πράκτορας εκτελεί ανεξάρτητους και διαφορετικούς υπολογισμούς, συνεπώς οι αποφάσεις τους  $d_i$ , δεν ανήκουν σε μια προκαθορισμένη κλίμακα ούτε και έχουν συγκεκριμένη κατανομή. Αυτός είναι ο λόγος που απαιτείται η κανονικοποίηση των αποστάσεων πριν αυτές συμψηφιστούν. Καθώς το σύστημα αποκτά εμπειρία μέσω ανακτήσεων, τα μεγέθη αυτά σταθεροποιούνται και αντικατοπτρίζουν την κλίμακα των τιμών σύγκρισης του κάθε πράκτορα για εικόνες της συγκεκριμένη εφαρμογής

## **4.2 Βιολογικά Εμπνευσμένη Αρχιτεκτονική Δύο Επιπέδων Ανάλυσης και Ανάκτησης**

Η εντύπωση που αποκτά κανείς εξετάζοντας την οπτική σκηνή ή μία εικόνα, σχεδόν ποτέ δεν προκύπτει μόνο από τη γενική και ανεξάρτητη περιγραφή των κυρίαρχων οπτικών χαρακτηριστικών της. Αντίθετα, αυτό που συμβαίνει, τουλάχιστον στις περιπτώσεις όπου το οπτικό ερέθισμα διατηρείται για κάποιο ικανοποιητικό χρονικό διάστημα, είναι η προσεκτική εξέταση μίας ή περισσότερων περιοχών της εικόνας, που ονομάζονται περιοχές ενδιαφέροντος. Ωστόσο, όπως παρατηρήθηκε στην παράγραφο 2.2, ο ίδιος ο μηχανισμός αναζήτησης περιοχών ενδιαφέροντος, καθοδηγείται από τοπικές ιδιότητες και χαρακτηριστικά που παρουσιάζουν ενδιαφέρον τα οποία προσλαμβάνονται κατά τα πρώιμα στάδια της όρασης. Στη συνέχεια, οι περιοχές αυτές εξετάζονται σειριακά, σαν ένας προβολέας να τις φωτίζει μέχρι το υποκείμενο αποφασίζει να αποσύρει το ενδιαφέρον του, φωτίζοντας κάποια άλλη ενδιαφέρουσα περιοχή. Επομένως, το ουσιαστικό νόημα των εικόνων προκύπτει από την ανάλυση και αναγνώριση των οπτικών οντοτήτων της, ιδίως αυτών που προσελκύουν με τις ιδιότητες τους το ενδιαφέρον του παρατηρητή. Ομοίως, η ουσιαστική και σημασιολογική ανάκτηση εικόνων με βάση το περιεχόμενο τους περνά από την εξέταση, κατανόηση και σύγκριση αυτών των περιοχών ανάμεσα στις εικόνες. Κατά συνέπεια, προτείνουμε δύο επίπεδα επεξεργασίας και ανάκτησης των εικόνων, το πρώτο για τη γενική περιγραφή και ομοιότητα τους και το δεύτερο για την περιγραφή και σύγκριση των περιοχών ενδιαφέροντος. Η αρχιτεκτονική των δύο

επιπέδων είναι εμπνευσμένη από τα πειραματικά δεδομένα της ψυχολογίας της οπτικής αντίληψης που παρουσιάστηκαν στην παράγραφο 2.2. Επίσης τα δύο επίπεδα που προτείνονται παρουσιάζουν τις κύριες ιδιότητες των αντίστοιχων επιπέδων επεξεργασίας της βιολογικής όρασης, το πρώιμο (*preattentive*) και το επιλεκτικό (*attentive*). Η προτεινόμενη αρχιτεκτονική διαχείρισης της οπτικής πληροφορίας παρουσιάζεται στην Εικόνα 26.

Για την υλοποίηση της αρχιτεκτονικής απαιτείται η γενική περιγραφή των εικόνων, η μέθοδος εντοπισμού των περιοχών ενδιαφέροντος και στη συνέχεια η επιλεκτική περιγραφή τους. Κατά την εισαγωγή μίας εικόνας στη ΒΔ του συστήματος γίνεται μία γενική περιγραφή της με βάση τα κυρίαρχα χαρακτηριστικά της, στη συνέχεια χρησιμοποιείται μέρος αυτής της περιγραφής για την εξαγωγή των περιοχών όπου εντοπίζεται το ενδιαφέρον της συγκεκριμένης εφαρμογής και στη συνέχεια γίνεται η περιγραφή αυτών των περιοχών. Τα αποτελέσματα αποθηκεύονται στη ΒΔ του συστήματος.



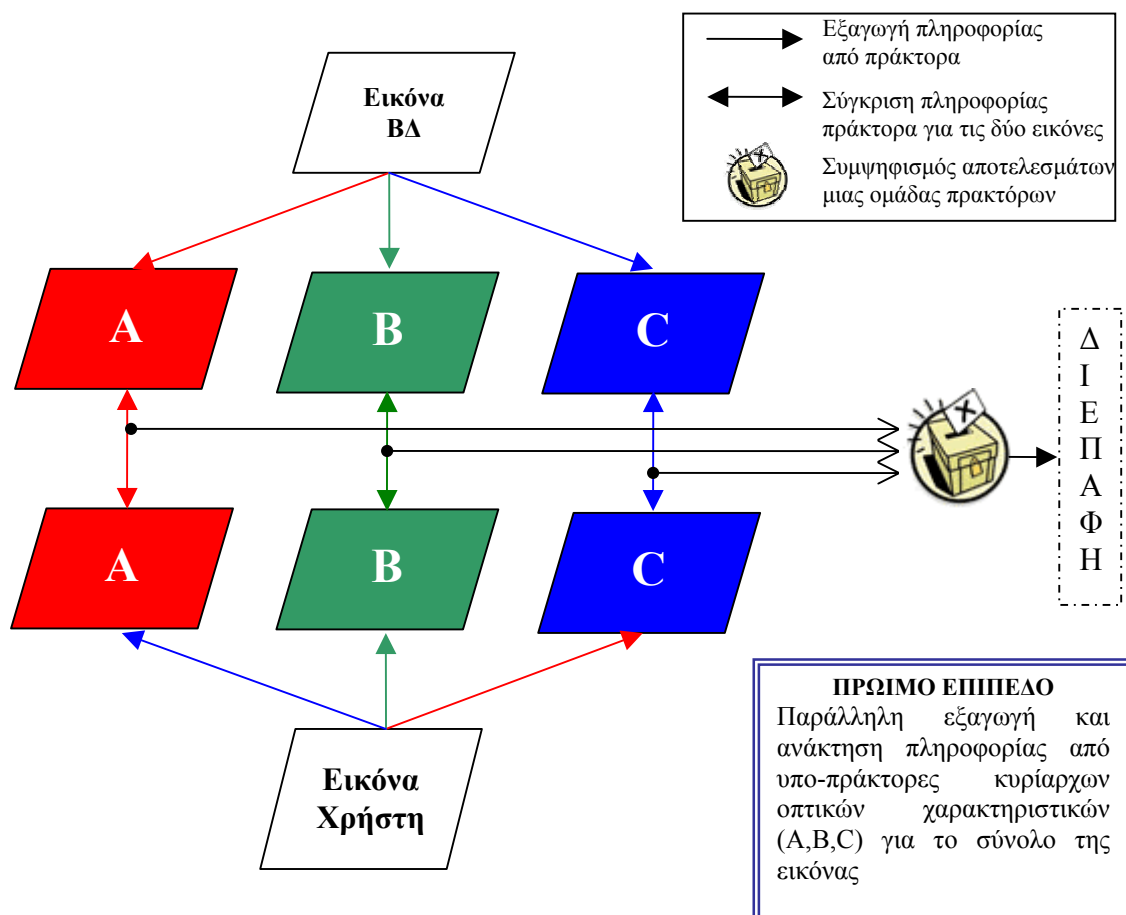
Εικόνα 26: Αρχιτεκτονική δύο επιπέδων ανάλυσης και ανάκτησης: Πρώιμο και Επιλεκτικό

Κατά την ανάκτηση η εικόνα ερώτησης του χρήστη περιγράφεται με τον ίδιο τρόπο. Στη συνέχεια, ανακτώνται μία προς μία οι εικόνες από τη συλλογή του πράκτορα και συγκρίνονται με την εικόνα ερώτησης ως προς την οπτική τους πληροφορία. Στο πρώιμο επίπεδο συγκρίνεται η γενική περιγραφή της εικόνας του χρήστη με τη γενική περιγραφή της ανακτημένης εικόνας ενώ στο επιλεκτικό συγκρίνονται περιγραφές των περιοχών ενδιαφέροντος. Τα αποτελέσματα ανάκτησης από τα δύο επίπεδα είναι ανεξάρτητα και παρουσιάζονται στο χρήστη σε ανεξάρτητες δυναμικές διεπαφές. Οι ανακτημένες εικόνες εμφανίζονται μία προς μία σε κάθε μία από τις διεπαφές του χρήστη κατά τη χρονική σειρά ανάκτησής τους πρώτα στο πρώιμο και αμέσως μετά στο επιλεκτικό επίπεδο της διεπαφής. Ωστόσο, τοποθετούνται με διαφορετική ταξινόμηση σε κάθε επίπεδο κατά τη φθίνουσα σειρά ομοιότητας με βάση τα αποτελέσματα του εκάστοτε επιπέδου. Επομένως, η αρχιτεκτονική επιτρέπει στο χρήστη να ανακτήσει εικόνες που παρουσιάζουν γενική ομοιότητα με την εικόνα ερώτησης και εικόνες που παρουσιάζουν ομοιότητα όσον αφορά τα σημασιολογικά αντικείμενα της εφαρμογής του. Επιπρόσθετα, παρέχεται ο σταθμισμένος συνδυασμός των αποτελεσμάτων από τα δύο επίπεδα. Η συνεισφορά του κάθε επιπέδου στο τελικό συνδυαστικό αποτέλεσμα καθορίζεται από το χρήστη.

Συνεπώς, στο συνδυαστικό επίπεδο ευνοούνται εικόνες που παρουσιάζουν τόσο γενική ομοιότητα όσο και ομοιότητα των περιοχών ενδιαφέροντος. Στα επόμενα, αναλύονται με μεγαλύτερη λεπτομέρεια οι αποφάσεις μας για κάθε λειτουργικό τμήμα της αρχιτεκτονικής.

#### 4.2.1 Το πρώιμο επίπεδο περιγραφής και ανάκτησης των εικόνων

Σύμφωνα με τα δεδομένα των νευροεπιστημών (παρ. 2.1) που επιβεβαιώνονται και από τα πειραματικά δεδομένα της ψυχολογίας της οπτικής αντίληψης (παρ. 2.2.1), η βιολογική όραση σε ένα πρώτο επίπεδο επεξεργασίας αποσυνθέτει την οπτική σκηνή στα επιμέρους πρωτογενή χαρακτηριστικά της, τα οποία επεξεργάζεται σε μεγάλο βαθμό ανεξάρτητα, αυτόνομα και παράλληλα. Κατά συνέπεια το πρώτο αυτό επίπεδο μπορεί να μοντελοποιηθεί από μία ομάδα πρακτόρων όπως αυτή περιγράφηκε στην παρ. 4.1.2. Τα πειραματικά αποτελέσματα που παρουσιάστηκαν στο [1] είναι ενθαρρυντικά στη χρήση τέτοιων συστημάτων για τη γενική περιγραφή των εικόνων, ιδίως όταν η ανάκτηση γίνεται μέσα από συλλογές συγκεκριμένης κατηγορίας εικόνων.



Εικόνα 27: Βαθμολόγηση Γενικής Ομοιότητας Εικόνων από μία Ομάδα Πρακτόρων για Χρήση στο Πρώιμο Επίπεδο Ανάκτησης.

Επομένως, η λειτουργικότητα του πρώιμου επιπέδου ανάλυσης και ανάκτησης επιτυγχάνεται με ένα σύνολο από παράλληλους ανεξάρτητους τοπογραφικούς χάρτες κάθε ένας από τους οποίους κωδικοποιεί την ύπαρξη του συγκεκριμένου οπτικού χαρακτηριστικού. Κατά τη διαδικασία ανάκτησης, κάθε πράκτορας αναλαμβάνει τη σύγκριση της πληροφορίας που ο ίδιος έχει εξάγει από τις δύο εικόνες (εικόνα ερώτησης και εικόνα προς ανάκτηση). Τα τελικά αποτελέσματα σύγκρισης από τους διάφορους πράκτορες οδηγούνται στο σύστημα ψηφοφορίας από όπου προκύπτει η τελική βαθμολογία της εικόνας ανάκτησης. Στη συνέχεια, η εικόνα τοποθετείται στη διεπαφή γενικής ομοιότητας μαζί με όλες τις προηγούμενες στη σωστή κατά φθίνουσα σειρά ομοιότητας. Στην Εικόνα 27 παρουσιάζεται η διαδικασία δημιουργίας των τοπογραφικών χαρτών από τους οπτικούς πράκτορες και η βαθμολόγησης μίας εικόνας ως προς την γενική ομοιότητα που προκύπτει από τη σύγκριση των ανεξάρτητων και παράλληλων αποφάσεων από τριών πρακτόρων

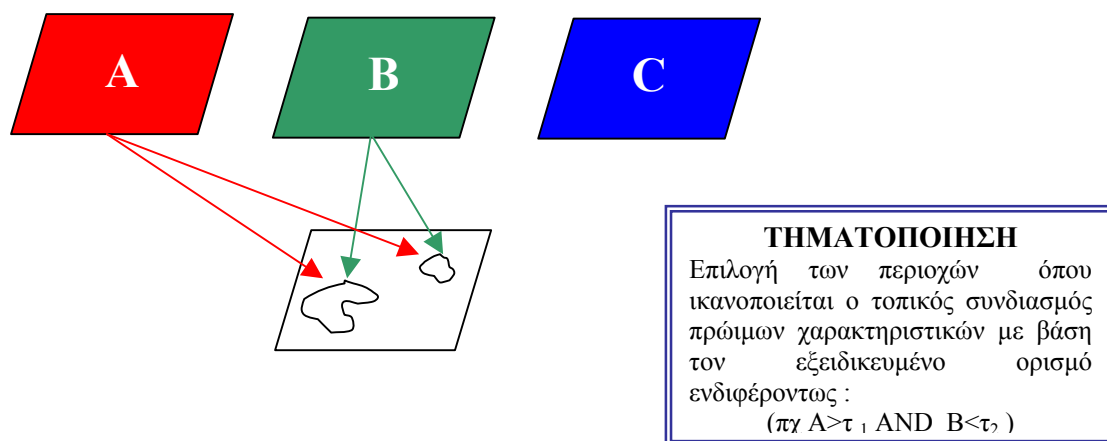
#### 4.2.2 Τμηματοποίηση των εικόνων σε εξειδικευμένες εφαρμογές

Η τμηματοποίηση εικόνων και γενικότερα ο εντοπισμός των οργανωμένων προτύπων σε μία εικόνα είναι ίσως το πιο ουσιαστικό πρόβλημα στην κατανόηση των εικόνων. Όπως αναφέρθηκε στο κεφάλαιο 2, η προσεκτική εξέταση για τον εντοπισμό και την αντίληψη οντοτήτων με νόημα, βασίζεται σε ένα σύνολο από μηχανισμούς, στους οποίους το υποκείμενο δεν έχει συνειδητή πρόσβαση. Στη διαδικασία εμπλέκονται στοιχεία της μνήμης και της υποκειμενικής επιθυμίας που μεταβάλλεται από άτομο σε άτομο και από στιγμή σε στιγμή. Κατά συνέπεια, είναι εξαιρετικά δύσκολο να μοντελοποιήσει κανείς ένα σύστημα όρασης ή σύγκρισης εικόνων, έξω από έναν ολοκληρωμένο μηχανισμό μάθησης, μνήμης και αντίληψης ίσως όχι έξω από έναν ολοκληρωμένο οργανισμό που αλληλεπιδρά μέσα σε ένα δυναμικό περιβάλλον. Η επιλογή των κατάλληλων πρωτογενών χαρακτηριστικών στο πρώιμο επίπεδο ανάλυσης της εικόνας και η χωρική ομαδοποίησή τους με βάση τις ιδιότητες Gestalt είναι ο κύριος διεκδικητής της λύσης του προβλήματος. Στο 2.2.3.1, παρατηρήθηκε ότι ορισμένες από τις αρχές Gestalt μπορούν να αναχθούν σε απλούστερες και γενικότερες ιδιότητες μειώνοντας την πολυπλοκότητα μίας τέτοιας αρχιτεκτονικής. Για παράδειγμα, η ιδιότητα της ομαλής συνέχειας μπορεί να αναχθεί στην ιδιότητα της ομοιότητας όταν αυτή εφαρμοστεί στον πρωτογενή τοπογραφικό χάρτη της κλίσης των ακμών. Ωστόσο, μία τέτοια ολοκληρωμένη αρχιτεκτονική δεν είναι διαθέσιμη στη βιβλιογραφία της οπτικής αντίληψης. Περαιτέρω, οι περισσότεροι αλγόριθμοι εξαγωγής οπτικής πληροφορίας από εικόνες, για παράδειγμα τα απλά ιστογράμματα, αφορούν στατιστική-γενική πληροφορία για το σύνολο της εικόνας που δεν περιέχει στις περισσότερες περιπτώσεις την πληροφορία της τοπογραφική εμφάνισης του υπό εξέταση χαρακτηριστικού. Προφανώς, αυτή η υποβαθμισμένη πληροφορία δεν είναι δυνατό να χρησιμοποιηθεί στη συνέχεια για τον εντοπισμό των αντικειμένων της εικόνας και κατά συνέπεια των περιοχών που εμπεριέχουν τη σημασιολογία και το νόημα της. Μέχρι σήμερα, δεν έχει υλοποιηθεί κάποια ασφαλής μέθοδος τμηματοποίησης που να πλησιάζει τις δυνατότητες της βιολογικής όρασης, για το σύνολο των εικόνων. Αντίθετα, στη βιβλιογραφία και τις επιστημονικές δημοσιεύσεις, πολύ συχνά, συναντά κανείς άρθρα για την αποδοτική τμηματοποίηση εικόνων ενός συγκεκριμένου πεδίου εφαρμογής, χρησιμοποιώντας την a priori γνώση για το πεδίο αυτό.

Τέλος, είναι φανερό, πως το νόημα μίας εικόνας είναι υποκειμενικό, δηλαδή υποκείμενα με διαφορετικό ενδιαφέρον και διαφορετική εμπειρία έχουν διαφορετική

άποψη για το περιεχόμενο μίας εικόνας (παρ 1.1) . Επομένως, ακόμη και αν υπήρχε μια ασφαλής μέθοδος τμηματοποίησης ολόκληρης της εικόνας, ο εντοπισμός εκείνων των τμημάτων που παρουσιάζουν ενδιαφέρον είναι μία υποκειμενική διαδικασία, μία εγκεφαλική εντολή. Πολλά από τα συστήματα ανάκτησης εικόνων χρησιμοποιούν την ανθρώπινη αλληλεπίδραση και καθοδήγηση για να κατανοήσουν το υποκειμενικό ενδιαφέρον του χρήστη (παρ. 3.1). Ωστόσο, κάτι τέτοιο αντιτίθεται στις αρχές της θεωρίας των ευφύων πρακτόρων και ιδιαίτερα στην ιδιότητα της αυτονομίας. Όπως θα γίνει φανερό, στην επόμενη παράγραφο, στόχος μας είναι να σχεδιάσουμε την αρχιτεκτονική για τα συστήματα CBIR, κατά τέτοιον τρόπο ώστε σταδιακά να την εμπλουτίσουμε με το σύνολο των ιδιοτήτων που διέπουν τους ευφείς πράκτορες. *Εν τέλη, θέλοντας να αποφύγουμε μια τέτοιου είδους αλληλεπίδραση προτείνουμε την εξειδίκευση του συστήματος σε εφαρμογές για τις οποίες είναι διαθέσιμη η περιγραφή των περιοχών ενδιαφέροντος.*

Κάθε πρώιμος τοπογραφικός χάρτης ποσοτικοποιεί την εμφάνιση μίας ιδιότητας ή ενός οπτικού χαρακτηριστικού σε κάθε σημείο της εικόνας. Επομένως, διαθέτοντας εξειδικευμένους τοπογραφικούς χάρτες και τον ορισμό των αντικειμένων ενδιαφέροντος καθίσταται δυνατό να εντοπίσουμε τις περιοχές εκείνες που εμφανίζουν συγκεκριμένες ιδιότητες και χαρακτηριστικά. Ουσιαστικά, αντί να διαχωρίσουμε οπτικά το σύνολο των αντικειμένων σε μία εικόνα, εντοπίζουμε απ' ευθείας μόνο εκείνα που παρουσιάζουν έναν συνδυασμό εξειδικευμένων ιδιοτήτων. Η επιλογή αυτών των ιδιοτήτων είναι καθοριστική και αποτελεί ένα δυσεπίλυτο πρόβλημα. Ωστόσο, η αρχιτεκτονική της προτεινόμενης πλατφόρμας, βασισμένη στη δια βίου λειτουργία πάνω σε ένα δυναμικό περιβάλλον εικόνων, επιτρέπει σε μελλοντικές επεκτάσεις της, την εκμάθηση των εξειδικευμένων περιορισμών που πρέπει να εφαρμοστούν πάνω σε ένα πλήρες σύνολο από πρώιμους χάρτες, βάσει της συνεχούς προσαρμογής σε εικόνες του πεδίου εφαρμογής, σύμφωνα με το βιολογικό μοντέλο. Στην Εικόνα 28, παρουσιάζεται ένα παράδειγμα τμηματοποίησης βασισμένο στην πληροφορία από ένα υποσύνολο των πρώιμων ανεξάρτητων χαρακτηριστικών.

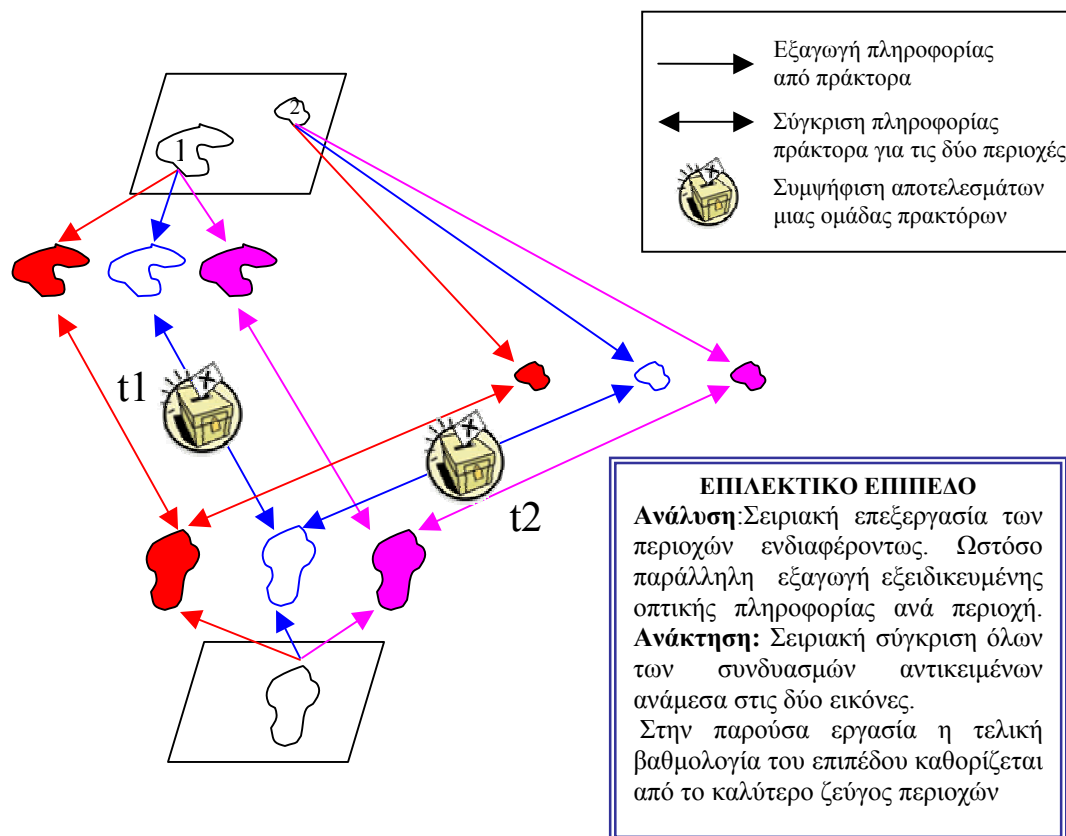


**Εικόνα 28:** Τμηματοποίηση με βάση τον ορισμό των περιοχών ενδιαφέροντος της εξειδικευμένης εφαρμογής με μερική χρήση της πληροφορίας των πρώιμων χαρακτηριστικών της εικόνας



### 4.2.3 Το Επίπεδο επιλεκτικής περιγραφής και ανάκτησης Εικόνων

Στο επιλεκτικό επίπεδο της βιολογικής όρασης (παρ. 2.2.2), ή αλλιώς το επίπεδο της προσεκτικής παρατήρησης (preattentive level of cognition), οι περιοχές ενδιαφέροντος εξετάζονται σειριακά, σαν ένας προβολέας να τις φωτίζει μέχρι το υποκείμενο αποφασίσει να αποσύρει το ενδιαφέρον του, φωτίζοντας κάποια άλλη ενδιαφέρουσα περιοχή. Ομοίως, στην προτεινόμενη αρχιτεκτονική, οι περιοχές ενδιαφέροντος, από την διαδικασία τμηματοποίησης, στέλνονται σειριακά για επεξεργασία σε μία εξειδικευμένη ομάδα πρακτόρων. Κατά συνέπεια, στην Εικόνα 29, η ομάδα πρακτόρων αναλύει πρώτα την περιοχή '1' και στη συνέχεια την περιοχή '2'. Κατά τη διαδικασία ανάκτησης η ομάδα των εξειδικευμένων πρακτόρων αναλαμβάνει να συγκρίνει έναν προς έναν όλους τους συνδυασμούς των περιοχών ενδιαφέροντος από τις δύο εικόνες. Ωστόσο, η σύγκριση κάθε ζεύγους περιοχών γίνεται παράλληλα από τους πράκτορες της ομάδας όπως ακριβώς και στο πρώιμο επίπεδο.



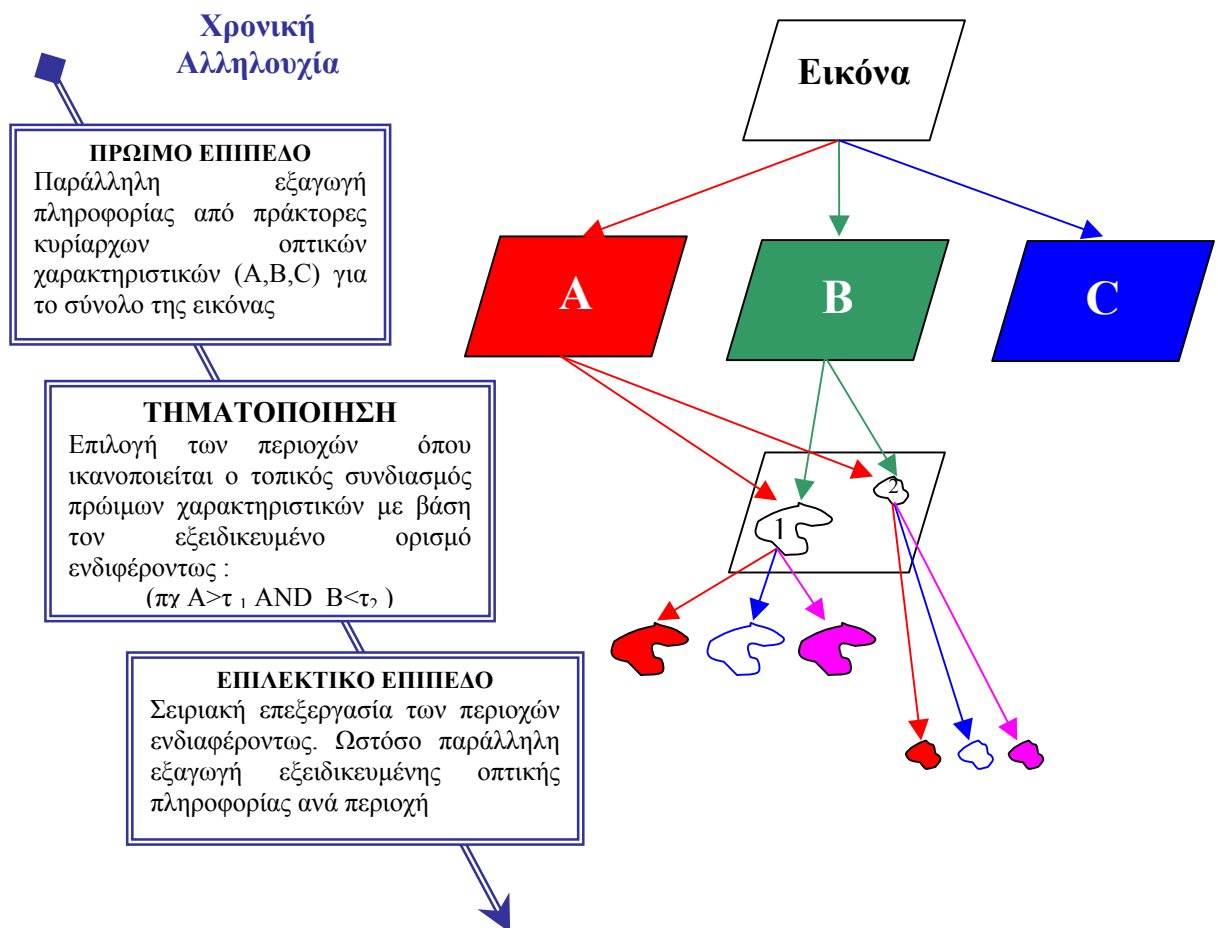
Εικόνα 29: Επιλεκτικό Επίπεδο Ανάλυσης και Ανάκτησης.

Στην Εικόνα 29 παρουσιάζεται ένα παράδειγμα λειτουργίας του επιλεκτικού επιπέδου. Μόνο αφότου τελειώσει η σύγκριση της περιοχής '1', ξεκινά η σύγκριση της περιοχής '2', με τη μοναδική περιοχή της δεύτερης εικόνας. Όσα ζεύγη περιοχών δημιουργούνται τόσα είναι και τα αποτελέσματα που λαμβάνονται. Για την εξαγωγή ενός τελικού αποτελέσματος ομοιότητας των δύο εικόνων ως προς τις περιοχές

ενδιαφέροντος, στην παρούσα εργασία επιλέγεται το καλύτερο αποτέλεσμα από όλους τους συνδυασμούς περιοχών. Με άλλα λόγια σε αυτό το επίπεδο, η ομοιότητα των δύο εικόνων υπολογίζεται ως η ομοιότητα του πιο όμοιου ζεύγους περιοχών ενδιαφέροντος.

#### 4.2.4 Ολοκλήρωση του σχήματος εξαγωγής πληροφορίας από τις εικόνες

Στην Εικόνα 30, παρουσιάζεται ολοκληρωμένη η διαδικασία εξαγωγής οπτικής πληροφορίας από μία εικόνα. Αρχικά η οπτική πληροφορία κωδικοποιείται απο ένα σύνολο από τοπογραφικούς χάρτες πρωιμων χαρακτηριστικών. Στη συνέχεια, η πληροφορία από τους χάρτες αυτούς συνδυάζεται για τον εντοπισμό των περιοχών ενδιαφέροντος με βάσει τον ορισμό τους στην εξειδικευμένη εφαρμογή. Τέλος, ένα σύνολο από τοπογραφικούς χάρτες μόνο των εντοπισμένων περιοχών αναλαμβάνει την εξειδικευμένη και ουσιαστική περιγραφή της εικόνας.



Εικόνα 30: Διαδικασία εξαγωγής πληροφορίας από τις εικόνες  
 Πρώιμη Παράλληλη Επεξεργασία → Τμηματοποίηση → Σειριακή Επιλεκτική Επεξεργασία.

Είναι σημαντικό σε αυτό το σημείο να ειπωθεί πως η πληροφορία από τους πρώιμους χάρτες είναι διαθέσιμη και στην επιλεκτική περιγραφή των περιοχών

ενδιαφέροντος. Για παράδειγμα στην Εικόνα 30, ο πράκτορας A, που κωδικοποιείται με κόκκινο χρώμα, μπορεί να χρησιμοποιηθεί τμηματικά στο επιλεκτικό επίπεδο για την περιγραφή μίας περιοχής. Περαιτέρω ο πράκτορας C που κωδικοποιείται με το μπλε χρώμα, αρχικά λειτουργεί πάνω σε ολόκληρη την εικόνα ενώ στη συνέχεια λειτουργεί μόνο πάνω στο περίγραμμα των περιοχών ενδιαφέροντος.

#### **4.2.5 Μορφή και αποθήκευση της οπτικής πληροφορίας από τους οπτικούς πράκτορες**

Όπως ειπώθηκε οι οπτικοί πράκτορες αποτελούν ανεξάρτητες διαδικασίες που διαχειρίζονται την εξαγωγή, αποθήκευση και σύγκριση ενός κυρίαρχου οπτικού χαρακτηριστικού ή ιδιότητας. Η πληροφορία που εξάγεται, αποθηκεύεται υπό τη μορφή ενός αρχείου κειμένου και ενός αρχείου εικόνας. Κάθε πράκτορας διαθέτει το δικό του χώρο αποθήκευσης με τα αρχεία κειμένου και εικόνες που δημιουργήσε για κάθε εικόνα στη συλλογή του συστήματος

Η εικόνα αποτελεί συνήθως ένα τοπογραφικό χάρτη κωδικοποίησης του συγκεκριμένου οπτικού χαρακτηριστικού και χρησιμοποιούνται από τα επόμενα επίπεδα της αρχιτεκτονικής και για την οπτική παρουσίαση της λειτουργίας του κάθε πράκτορα. Ωστόσο, υπάρχουν μερικοί πράκτορες που δεν παράγουν τοπογραφικούς χάρτες, για παράδειγμα ιστογράμματα, χρήσιμες μετρήσεις όπως σχετικές γωνίες κ.α. Αυτή η πληροφορία δεν μπορεί να χρησιμοποιηθεί στα επόμενα επίπεδα, αλλά είναι χρήσιμη μόνο για σύγκριση. Αυτό είναι αναμενόμενο καθώς η περιγραφή του βιολογικού μοντέλου στο κεφάλαιο 2, αφορά μόνο την κατανόηση του οπτικού περιεχομένου και όχι μηχανισμούς μνήμης και αναγνώρισης αντικειμένων που λαμβάνουν χώρα κατά την ανάκληση παρόμοιων εικόνων.

Τα αρχεία κειμένου περιέχουν συγκρίσιμη πληροφορία που χρησιμοποιείται από τον αντίστοιχο οπτικό πράκτορα κατά τη διαδικασία της ανάκτησης. Κατά τη σύγκριση δύο εικόνων ή περιοχών ως προς την ομοιότητα, κάθε οπτικός πράκτορας ανακτά τα αντίστοιχα αρχεία κειμένου από τον αποθηκευτικό του χώρο και στη συνέχεια συγκρίνει την πληροφορία που περιέχουν. Το αποτέλεσμα της σύγκρισης αποτελεί την ανεξάρτητη απόφαση ομοιότητας του συγκεκριμένου πράκτορα..

Καθώς υπάρχουν δύο επίπεδα λειτουργίας των πρακτόρων, για κάθε εικόνα υπάρχουν καταχωρίσεις της γενικής περιγραφής και πιθανώς καταχωρίσεις για όλες τις περιοχές ενδιαφέροντος που η εικόνα περιέχει. Για παράδειγμα, στην Εικόνα 30, ο πράκτορας C που κωδικοποιείται με το μπλε χρώμα αρχικά λειτουργεί πάνω σε ολόκληρη την εικόνα ενώ στη συνέχεια λειτουργεί πάνω στο περίγραμμα κάθε περιοχής ενδιαφέροντος. Επιπρόσθετα, ανάλογα με την εξειδίκευση του συστήματος ορισμένοι πράκτορες χρησιμοποιούνται αποκλειστικά σε κάποιο από τα επίπεδα, της εφαρμογής. Για παράδειγμα, ο πράκτορας B που κωδικοποιείται με το πράσινο χρώμα χρησιμοποιείται για τη γενική περιγραφή της εικόνας και στην τμηματοποίηση όμως δεν περιέχει χρήσιμη πληροφορία για την περιγραφή των περιοχών ενδιαφέροντος της συγκεκριμένης εφαρμογής.

### 4.3 Ιδιότητες του συστήματος

Όπως ειπώθηκε στην παρ. 2.3.2 μια από τις βασικότερες αρχές που αποτέλεσε και την αναγκαιότητα για τη θεωρία πρακτόρων είναι να επιτελούν μια συνεχή δυσεπίλυτη εργασία αποδεσμεύοντας το χρήστη από τη χρονοβόρα διαδικασία και επίβλεψη. Αυτό συντελείται στο υψηλότερο επίπεδο της προτεινόμενης αρχιτεκτονικής. Ένας αυτόνομος πράκτορας αναζητά συνεχώς εικόνες για επεξεργασία και εισαγωγή στη βάση του. Πέρα από την αυτοματοποιημένη αναζήτηση και ανάλυση εικόνων, ο πράκτορας αποτελεί έναν εξυπνέτη αιτήσεων ανάκτησης εικόνων με βάση το περιεχόμενό τους, με στόχο να βελτιώνει συνεχώς την επίδοση του όσον αφορά την διάγνωση ομοιότητας ανάμεσα σε εικόνες ενός πεδίου εφαρμογής. Κατά συνέπεια, έχει καθοριστεί ένα δυναμικό περιβάλλον, με το οποίο ο πράκτορας αλληλεπιδρά χωρίς καμία εξωτερική επίβλεψη.

Σε επίπεδο οπτικής επεξεργασίας, οι οπτικοί στόχοι της ανάλυσης και των αποφάσεων ομοιότητας διαιρέθηκαν σε ασύγχρονες παράλληλες διαδικασίες κάθε μία από τις οποίες διαχειρίζεται ανεξάρτητα την ύπαρξη κάποιου οπτικού χαρακτηριστικού. Τέτοιες ομάδες πρακτόρων όπως ειπώθηκε στην παράγραφο αποτελούν στη «χαλαρή» τους θεώρηση τα συστήματα πρακτόρων του Brooks. Περαιτέρω, σε αυτούς τους οπτικούς πράκτορες, ενσωματώθηκαν διαδικασίες «δια βίου» ανανέωσης διαφόρων στατιστικών μεγεθών όπως για παράδειγμα η διακύμανση των αποτελεσμάτων σύγκρισης. Καθώς το σύστημα αποκτά εμπειρία μέσω ανακτήσεων, τα μεγέθη αυτά σταθεροποιούνται και αντικατοπτρίζουν την κλίμακα των τιμών του κάθε πράκτορα για εικόνες της συγκεκριμένη εφαρμογής. Από αυτή την απλή και αναγκαία για τη λειτουργία του συστήματος τροποποίηση γίνεται φανερό *ότι θέτοντας το πρόβλημα στις σωστές βάσεις αυτονομίας, η ίδια η λειτουργικότητα της εφαρμογής καθιστά αναγκαία τα στοιχεία της ευφυΐας και της συνεχούς μάθησης.*

Τέλος τόσο η ανάλυση των εικόνων όσο και η σύγκριση τους γίνεται σε δύο επίπεδα. Το πρώτο ή αλλιώς πρώιμο επίπεδο αναλαμβάνει τη γενική περιγραφή των εικόνων και το δεύτερο ή αλλιώς επιλεκτικό την περιγραφή των περιοχών ενδιαφέροντος. Το πρώτο επίπεδο πέρα από τη γενική λειτουργία του είναι δυνατό να χρησιμοποιηθεί σε εξειδικευμένες εφαρμογές ως ένα φίλτρο αποκοπής των εικόνων που δεν αφορούν την εξειδικευμένη εφαρμογή. Επίσης, αν είναι διαθέσιμο ένα σύνολο από εικόνες αναφοράς του συγκεκριμένου πεδίου εφαρμογής, η γενική αυτή περιγραφή μπορεί να χρησιμοποιηθεί για μία γρήγορη κατηγοριοποίηση της εικόνας εισόδου (ΠΑΡ. 6.2.1). Το δεύτερο επίπεδο της αρχιτεκτονικής χρησιμοποιείται για μία περισσότερο σημασιολογική ανάκτηση με βάση τις περιοχές ενδιαφέροντος του εκάστοτε πεδίου εφαρμογής. Η αρχιτεκτονική είναι βασισμένη σε τμήμα της υπάρχουσας γνώσης σχετικά με τους μηχανισμούς λειτουργίας της βιολογικής όρασης και ειδικότερα της ψυχολογίας της οπτικής αντίληψης. Η προσαρμογή στα ιδιαίτερα χαρακτηριστικά της εκάστοτε εφαρμογής επιτυγχάνεται με την προαναφερθείσα, δια βίου απόκτηση στατιστικής πληροφορίας για το περιεχόμενο των εικόνων. Με αυτόν τον τρόπο η προτεινόμενη αρχιτεκτονική αντιμετωπίζει το πρόβλημα του σημασιολογικού κενού. Τέλος, η αρχιτεκτονική προσφέρει την υπηρεσία σταθμισμένης ανάκτησης των αποτελεσμάτων από τα δύο επίπεδα. Τα συνδυαστικά αποτελέσματα είναι εξαιρετικά χρήσιμα καθώς επιτρέπουν στο χρήστη να ανακτήσει εικόνες τόσο γενικής ομοιότητας όσο και επιλεκτικής ομοιότητας. Για παράδειγμα στο πεδίο των ιατρικών εικόνων, το πρώιμο επίπεδο διαχειρίζεται την ανάκτηση εικόνων από παρόμοιες ανατομικά περιοχές ενώ το επιλεκτικό την

ανάκτηση παρόμοιων ευρημάτων. Συνεπώς, στη συγκεκριμένη εφαρμογή το συνδυαστικό επίπεδο παρέχει τον εντοπισμό παρόμοιων οπτικά ευρημάτων σε παραπλήσια τοποθεσία μέσα στο ανθρώπινο σώμα.

Συμπερασματικά η προτεινόμενη αρχιτεκτονική είναι εξαιρετικά ευέλικτη και επιτυγχάνει μία περισσότερο σημασιολογική ανάκτηση καθώς επιτρέπει την ανάκτηση σε διάφορα σημασιολογικά επίπεδα τα οποία προσαρμόζονται στις οπτικές ιδιότητες του πεδίου εφαρμογής των εξεταζόμενων εικόνων. Κατά τη υλοποίηση της εφαρμογής συγκεντρωτικά αναφέρονται οι δυνατότητες προσθαφαίρεσης πρακτόρων, καθορισμός ομάδων πρακτόρων και ποσοστιαίας συνεισφοράς του κάθε ενός στο αποτέλεσμα της ψηφοφορίας. Επιπρόσθετα είναι επεκτάσιμη καθώς νέα επίπεδα περιγραφής των εικόνων είναι δυνατόν να προσαρτηθούν με βάση τα νέα δεδομένα που ενδέχεται να προκύψουν από τη σύγχρονη μελέτη των μηχανισμών της βιολογικής όρασης. Σε επίπεδο προσαρμοστικότητας κάθε πράκτορας προσαρμόζει την έκταση των αποτελεσμάτων του με βάση τα αποτελέσματα ομοιότητας των εικόνων της συγκεκριμένης εφαρμογής. Τέλος όσον αφορά τη συνεισφορά των οπτικών χαρακτηριστικών στο αποτέλεσμα ομοιότητας μίας ομάδας πρακτόρων, κατάλληλες παράμετροι έχουν καθοριστεί. Η προσάρτηση μεθόδων προσαρμογής αυτών των παραμέτρων στο εκάστοτε πεδίο εφαρμογής είναι δυνατό να βελτιώσει ακόμη περισσότερο τους στόχους της σημασιολογικής ανάκτησης και θεωρείται ένα από τα σημαντικότερα θέματα που αφήνονται για μελλοντική εργασία. Στην επόμενη παράγραφο γίνεται μία εκτενής αναφορά στις βασικότερες ελλείψεις της προτεινόμενης αρχιτεκτονικής

#### **4.4 Μελλοντική Εργασία και Συμπεράσματα**

Η προτεινόμενη αρχιτεκτονική θέτει τις σωστές βάσεις για μια περισσότερο ευφυή ανάκτηση εικόνων, ωστόσο σημαντικά θέματα έχουν αφηθεί για μελλοντική εργασία.

##### **4.4.1 Θέματα μάθησης**

Οι προτεινόμενοι οπτικοί πράκτορες διαθέτουν ορισμένα πολύ απλά στοιχεία μάθησης, ωστόσο κρίνεται σημαντικό να εμπλουτιστούν με μηχανισμούς εσωτερικής συλλογιστικής, βάση των οποίων θα βελτιώνουν τις μελλοντικές τους αποφάσεις ομοιότητας για το συγκεκριμένο πεδίο εφαρμογής. Επιπρόσθετα, στις ομάδες οπτικών πρακτόρων, η συνεισφορά (βάρος)  $w$  (εξ. (3)) του κάθε πράκτορα στο τελικό αποτέλεσμα ομοιότητας δύο περιοχών πρέπει να αποτελεί προϊόν μάθησης. Στην παρούσα εργασία τα βάρη των πρακτόρων καθορίζονται εφ' άπαξ κατά την εγκατάσταση της πλατφόρμας με κριτήριο τις απαιτήσεις της συγκεκριμένης εφαρμογής. Ωστόσο, ένα από τα σημαντικότερα θέματα που αφήνονται για μελλοντική εργασία είναι η επινόηση μίας ευφυούς διαδικασίας για την εκμάθηση αυτών των παραμέτρων.

Μία πολύ ελκυστική πρόταση στο πρόβλημα αυτό, αποτελεί η δημιουργία ποικιλίας στα βάρη των πρακτόρων και η γρήγορη ανάκτηση των πρώτων μόνο εικόνων της συλλογής με ταξινόμηση που προκύπτει από δύο ή περισσότερα στιγμιότυπα ποικιλίας. Στη συνέχεια, ο χρήστης καλείται να επιλέξει ανάμεσα από αυτές τις ομάδες γνωρίζοντας ότι τα υπόλοιπα αποτελέσματα που θα λάβει θα

ταξινομηθούν σύμφωνα με την επιλεγμένη ποικιλία βαρών. Ουσιαστικά, αυτή η διαδικασία αποτελεί μία έμμεση αξιολόγηση της επίδοσης των ομάδων πρακτόρων υπό διαφορετικές αποφάσεις συνεισφοράς των επιμέρους οπτικών χαρακτηριστικών. Είναι σαφές πως με κατάλληλη εσωτερική συλλογιστική τέτοιου είδους ανάδραση μπορεί να οδηγήσει στη βελτίωση των μελλοντικών αποφάσεων της συνολικής πλατφόρμας. Ωστόσο, είναι απαραίτητη η χρήση του συστήματος σε εξειδικευμένες εφαρμογές ώστε να υπάρχει συσχέτιση των επιθυμιών του εκάστοτε χρήστη με τους στόχους μακροπρόθεσμης προσαρμογής του προτεινόμενου υπερ-πράκτορα.

Είναι βασικό να κατανοήσουμε ότι οι πράκτορες πρέπει να λειτουργούν ως αυθύπαρκτες οντότητες με προσωπικούς στόχους. Οι στόχοι αυτοί είναι αδιαφανείς και μερικές φορές φαντάζουν παράλογοι και εγωιστικοί στον τελικό χρήστη. Η προτεινόμενη αλληλεπίδραση δεν συνιστά περιορισμό της αυτονομίας του πράκτορα διότι αφορά τη βελτίωση και την επίτευξη των προσωπικών του μακροπρόθεσμων στόχων και όχι την καθοδήγηση από τις επιθυμίες των διαφόρων χρηστών. Η αντικειμενικότητα των οπτικών στόχων του πράκτορα προκύπτει από το γεγονός της επιλογής ενός εξειδικευμένου πεδίου εφαρμογής. Ωστόσο αυτοί οι στόχοι είναι εξαιρετικά σύνθετοι ώστε να είναι προκαθορισμένοι είτε προγραμματισμένοι από το δημιουργό και επομένως αποτελούν προϊόν μάθησης που προκύπτει από την εκμετάλλευση της γνώσης των χρηστών για την αποκόμιση κοινού οφέλους.

Στο συγκεκριμένο πρόβλημα υπάρχουν επιπλέον δύο βασικές αρχές κατά την αξιολόγηση των επιμέρους οπτικών χαρακτηριστικών που πρέπει να λαμβάνει υπ' όψη η συλλογιστική διαχείρισης μίας ομάδας πρακτόρων

- Ένα οπτικό χαρακτηριστικό που συναντάται συχνά, ή έντονα σε μία εικόνα την περιγράφει καλά
- Ένα οπτικό χαρακτηριστικό που συναντάται συχνά στις εικόνες μίας εφαρμογής είναι κακό μέτρο διαχωρισμού των εικόνων της

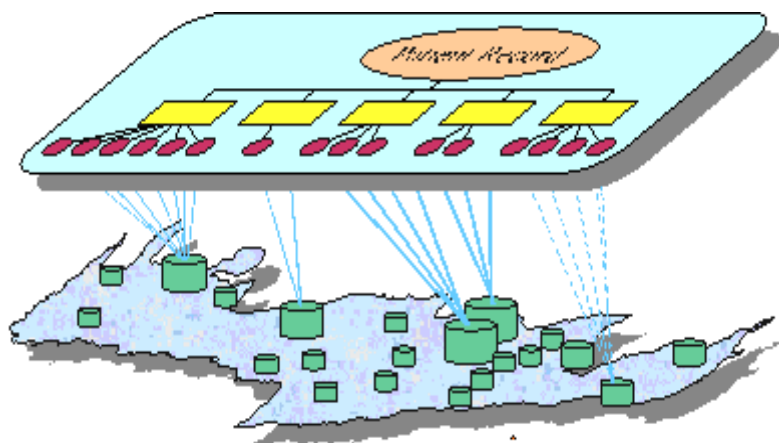
Με βάση τα παραπάνω, είναι δυνατή η εξειδίκευση των διαφόρων ομάδων οπτικών πρακτόρων, τόσο σε επίπεδο γενικής αλλά και επιλεκτικής περιγραφής των εικόνων ούτως ώστε να ανταποκρίνονται καλύτερα στη σημασιολογία του συγκεκριμένου στόχου που αναλαμβάνουν. Ιδανικά, θα ήταν επιθυμητή η δημιουργία μίας ομάδας από μεγάλο πλήθος οπτικών πρακτόρων που να χρησιμοποιείται αυτούσια για τις αποφάσεις ομοιότητας σε κάθε πεδίο εφαρμογής. Στη συνέχεια, η χρήση του συστήματος για ένα συγκεκριμένο πεδίο εικόνων θα οδηγούσε στην προσαρμογή των βαρών πρακτόρων και πιθανώς στην απενεργοποίηση των χαρακτηριστικών που δεν χρησιμεύουν στο σημασιολογικό διαχωρισμό των εικόνων του πεδίου αυτού.

Τέτοιες διαδικασίες μάθησης θα ήταν δυνατό να χρησιμοποιηθούν και για τον εντοπισμό των περιοχών ενδιαφέροντος του εκάστοτε πεδίου εικόνων με βάση το μοντέλο του τοπικού συνδυασμού των πρώιμων χαρακτηριστικών (Εικόνα 28). Ωστόσο, παραμένει πεποιθήση μας, ότι κάτι τέτοιο απαιτεί ένα πλήρες, δυναμικό και απρόσμενο οπτικό περιβάλλον. Για παράδειγμα, σε ένα περιβάλλον εικόνων απουσιάζει η πληροφορία της κίνησης των σημασιολογικών οντοτήτων που παρέχει μεγάλη πληροφορία αναφορικά με τις συμπαγείς και ενιαίες οντότητες του πραγματικού κόσμου. Επιπρόσθετα, σε έναν ενσώματο οργανισμό, πληροφορίες από τις άλλες αισθήσεις του επιδρούν με πολλούς και σύνθετους τρόπους στην κατανόηση του οπτικού περιεχομένου. Επομένως, στα συστήματα εικόνων χάνεται ένα σημαντικό κομμάτι της πληροφορίας που σχετίζεται με τις μη τυχαίες εμφανίσεις του περιβάλλοντος. Η έλλειψη αυτής της πληροφορίας καθιστά αδύνατη την προσομοίωση των μηχανισμών της βιολογικής όρασης δίχως την αλληλεπίδραση και καθοδήγηση από το χρήστη. Ωστόσο, η χρήση της προτεινόμενης αρχιτεκτονικής και

ο πειραματισμός σε ένα πλήρες, τουλάχιστον από οπτική πληροφορία, περιβάλλον αποτελεί μία πρόκληση, καθώς είναι δυνατό να προκύψουν βάσιμες υποθέσεις για τους ενδότερους μηχανισμούς της βιολογικής όρασης που παραμένουν ακόμα αδιαφανείς.

#### 4.4.2 Θέματα κινητικότητας και επικοινωνίας

Μέχρι στιγμής, δεν έγινε καθόλου λόγος για την ιδιότητα της κινητικότητας των ευφύων πρακτόρων. Αν και το σύστημα που υλοποιήθηκε δεν συμπεριλαμβάνει αυτήν την ιδιότητα, είναι πολύ εύκολο σε μελλοντικές εφαρμογές να ενσωματωθεί. Ουσιαστικά, ο κώδικας του συστήματος δεν είναι εξαρτημένος από τα δεδομένα του. Ένα στιγμιότυπο του πράκτορα μπορεί να λειτουργήσει πάνω στη βάση δεδομένων που δημιούργησε κάποιο άλλο στιγμιότυπο, καθώς η διαδικασία εισαγωγής και ανάκτησης εικόνων και από τους δύο ακολουθεί την ίδια λογική. Αυτό που διαφοροποιείται είναι η συνεισφορά του κάθε πράκτορα στο τελικό αποτέλεσμα οπτικής σύγκρισης από μία ομάδα πρακτόρων και τα δια βίου στατιστικά μεγέθη που έχει αποκομίσει με βάση την εμπειρία του. Μολονότι οι αποφάσεις του συστήματος εξαρτώνται από την εμπειρία του κάθε οπτικού πράκτορα, μία πλούσια βάση δεδομένων είναι εξίσου σημαντική στην πλευρά του χρήστη. Κατά συνέπεια, θα ήταν χρήσιμη η δυνατότητα μετακίνησης ενός πράκτορα σε μία περισσότερο ενημερωμένη βάση, όταν τα αποτελέσματα δεν ικανοποιούν το χρήστη. Είναι προφανές πως η μετακίνηση και ανταλλαγή δεδομένων και εικόνων ανάμεσα στις συλλογές των πρακτόρων δεν είναι καθόλου αποδοτική λύση όσον αφορά τους πόρους ενός δικτύου.



Εικόνα 31: Ιατρικό Δίκτυο Κατανεμημένης Ιατρικής Πληροφορίας στη νήσο Κρήτη

Όσον αφορά το πεδίο εφαρμογής της συγκεκριμένης εργασίας, που είναι οι ιατρικές εικόνες, υπάρχει σημαντική υποδομή για μια τέτοια επέκταση. Τα σύγχρονα ψηφιακά ιατρικά δίκτυα κάνουν λόγο για ένα περιβάλλον στο οποίο θα παρέχεται ολοκληρωμένη και συνεχής περίθαλψη στους ασθενείς όπου και αν βρίσκονται μέσω του ηλεκτρονικού και κατανεμημένου φακέλου υγείας (EHR) (Εικόνα 31). Σε ένα τέτοιο δίκτυο πληροφοριών οραματιζόμαστε τον πράκτορα αναζήτησης εικόνων, προσαρτημένο σε κάθε δικτυακό ιατρικό κέντρο (Εικόνα 31), δηλαδή κοντά στην πηγή εικόνων που διαχειρίζεται. Σε περίπτωση όπου ένας πράκτορα λόγω των

μειωμένων περιστατικών που διαθέτει αδυνατεί να ανταποκριθεί σε κάποια αίτηση ανάκτησης πρέπει να έχει τη δυνατότητα να μεταφέρεται σε ένα περισσότερο ενημερωμένο περιβάλλον. Περαιτέρω, οι νεαροί πράκτορες από τις διάφορες δικτυακές τοποθεσίες είναι δυνατό να επικοινωνούν με παλαιότερους πράκτορες για απόκτηση της στατιστικής εμπειρίας τους.

#### **4.4.3 Σημασιολογικά μεταδεδομένα εικόνων και δυνατότητα επέκτασης της αρχιτεκτονικής σε γενικότερες εφαρμογές**

Η προτεινόμενη πλατφόρμα προσλαμβάνει το περιβάλλον αναζητώντας διαρκώς εικόνες από προκαθορισμένες πηγές εικόνων. Ωστόσο, στην υλοποίηση που πραγματοποιήθηκε για τις μαγνητικές τομογραφίες εγκεφάλου (ΚΕΦ. 5), όλες οι εικόνες που εγγράφονται σε αυτές τις πηγές αποτελούν εικόνες της συγκεκριμένης εφαρμογής. Σε ένα δυναμικό περιβάλλον εικόνων, ο ίδιος ο πράκτορας θα έπρεπε να επιλέγει τις ενδιαφέρουσες εικόνες με βάση τα οπτικά χαρακτηριστικά τους. Κάτι τέτοιο είναι δυνατό να επιτευχθεί με χρήση ενός φίλτρου επιλογής εικόνων με βάση τη γενική οπτική περιγραφή τους, δηλαδή με χρήση του πρώιμου επιπέδου ανάκτησης. Κάθε οπτικός πράκτορας γι' αυτή τη λειτουργία διαθέτει στατιστικά μεγέθη όπως μέγιστες και ελάχιστες ανομοιότητες ανάμεσα στις εικόνες της εφαρμογής του με βάση την εμπειρία του. Κατά συνέπεια, η σύγκριση των εικόνων με ένα πλήρες προκαθορισμένο σύνολο από χαρακτηριστικές εικόνες αναφοράς είναι δυνατό να καθορίσει το βαθμό συσχέτισής τους με το πεδίο εφαρμογής. Αυτή η μέθοδος επίλυσης του προβλήματος είναι εξαιρετικά ενδιαφέρουσα σε επίπεδο υπολογιστικής όρασης (συνδέεται άμεσα με την αυτόματη κατηγοριοποίηση εικόνων), ωστόσο υπάρχει και μία πιο ασφαλής και αποδοτική λύση που κάνει χρήση των μεταδεδομένων που σε ορισμένες εφαρμογές συνοδεύουν τις εικόνες. Η λύση αυτή απαιτεί την ανάγνωση των λεκτικών περιγραφών που προσδιορίζουν την κατηγορία στην οποία ανήκει η εκάστοτε εικόνα. Σε μεγάλες, πολυποίκιλες βάσεις δεδομένων, αυτό το λεκτικό φίλτρο είναι πολύ αποδοτικότερο, ωστόσο προϋποθέτει ότι οι εικόνες συνοδεύονται από τα αντίστοιχα μεταδεδομένα και ότι υπάρχει κάποιο πρότυπο που να καθορίζει τις δυνατές τιμές τους. Επομένως, χρήση των μεταδεδομένων απαιτεί τον προσδιορισμό προτύπων που να εξασφαλίζουν την ορθή ανάγνωσή τους. Εν γένει, τέτοια πληροφορία σπάνια συνοδεύει τις εικόνες και σε πολλές περιπτώσεις δεν είναι ακριβής είτε παρουσιάζει σφάλματα.

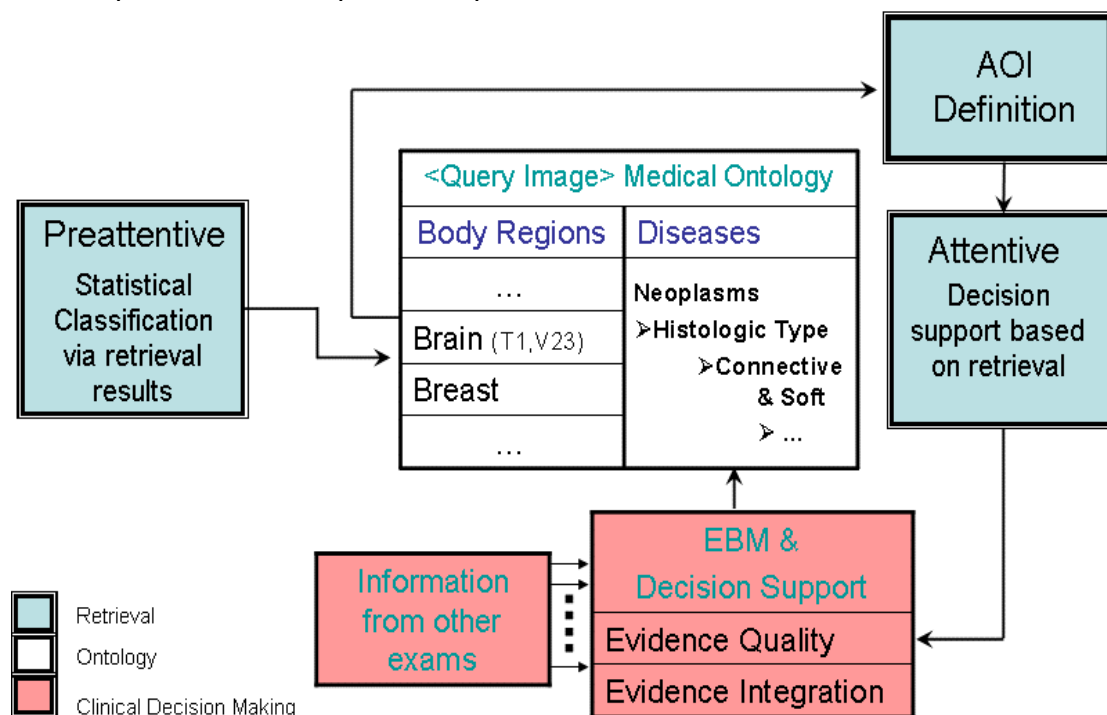
#### **4.4.4 Η προτεινόμενη αρχιτεκτονική στο πεδίο της Ιατρικής**

Η προσάρτηση μεταδεδομένων αν και «δαπανηρή» διαδικασία, δεν είναι τόσο σπάνιο φαινόμενο στις ιατρικές εικόνες καθώς είναι ιδιαίτερα χρήσιμη και συνηθίζεται λόγω της ανάγκης για την περιγραφή του εκάστοτε περιστατικού. Όπως ειπώθηκε στην παράγραφο 3.2.1, στο πρότυπο DICOM, μία ιατρική εικόνα συνοδεύεται από ένα πολύ μεγάλο σύνολο από περιγραφείς-ιδιότητες της εικόνας και του περιστατικού, από τις οποίες μπορεί να εξαχθεί η απαιτούμενη πληροφορία. Μερικές από αυτές τις παραμέτρους καθορίζουν την τεχνική απεικόνισης (modality), τη σχετική θέση του ασθενούς ως προς το μηχάνημα, το εξεταζόμενο όργανο (Image Position Patient), κ.α. Κατά συνέπεια με τη χρήση ενός φίλτρου διαχείρισης αυτής της πληροφορίας, είναι δυνατό να δρομολογούμε εικόνες στα διάφορα εξειδικευμένα



στιγμιότυπα του προτεινόμενου συστήματος, κάθε ένα από τα οποία κάνει χρήση της πρότερης γνώσης αλλά και προσαρμόζεται στο συγκεκριμένο πεδίο εφαρμογής. Η εξειδίκευση κυρίως αφορά τη διαδικασία τμηματοποίησης και τη συνεισφορά των οπτικών πρακτόρων στο αποτέλεσμα της ψηφοφορίας. Με αυτόν τον τρόπο το εξειδικευμένο μας σύστημα μπορεί να επεκταθεί σε ένα γενικότερο πεδίο εφαρμογής π.χ. για το σύνολο των ιατρικών εικόνων. Εναλλακτικά, όπως θα γίνει φανερό από τα αποτελέσματα που παρουσιάζονται στο κεφάλαιο 6 για τις μαγνητικές τομογραφίες εγκεφάλου, η προτεινόμενη πλατφόρμα είναι ικανή στο πρώιμο επίπεδο ανάκτησης να διαχωρίσει αυτόματα τις διάφορες απεικονιζόμενες ανατομικές περιοχές και στη συνέχεια να τις οδηγήσει στα επόμενα επίπεδα επεξεργασίας.

Η ολοκληρωμένη εφαρμογή του συστήματος ανάκτησης εικόνων στο πεδίο της Ιατρικής παρουσιάζεται στην Εικόνα 32. Σε αυτό το σχήμα συμπεριλαμβάνονται τα ιδιαίτερα χαρακτηριστικά και οι δυνατότητες της ανάκτησης εικόνων στο πεδίο της Ιατρικής που συζητήθηκαν στην παράγραφο 3.2.1 Συγκεκριμένα η στατιστική επεξεργασία των εικόνων ανάκτησης είναι δυνατό σε συνδυασμό με άλλες εξετάσεις του εκάστοτε περιστατικού να συντελέσει στην υποστήριξη της κλινικής του διάγνωσης (Image Based Reasoning and Decision Support). (Για παράδειγμα, το α% των ανακτημένων εικόνων να προέρχονται από την ίδια ασθένεια, b% από κάποια άλλη και c% από υγιή περιστατικά). Ο συνδυασμός των διαφόρων αποφάσεων από τις επιμέρους εξετάσεις προϋποθέτει την αξιολόγηση της κάθε μίας με βάση τη σύγχρονη βασισμένη στις ενδείξεις ιατρική (Evidence Based Medicine) ώστε να υπάρχει και ο συμψηφισμός της ποιότητας-βεβαιότητας τους (Evidence Quality). Κατά συνέπεια, είναι απαραίτητη η αξιολόγηση του συστήματος από ειδικούς, συναρτήσει της εμπειρίας του και του πλήθους των διαθέσιμων εικόνων στη βάση του. Είναι σημαντικό να αναφερθεί ότι η αξιολόγηση των συστημάτων ανάκτησης εικόνων είναι μία δύσκολη διαδικασία. Ωστόσο, σε εξειδικευμένα συστήματα όπου υπάρχει σημαντική πρότερη γνώση του πεδίου εφαρμογής, όπως στην Ιατρική μία τέτοια προσπάθεια είναι ρεαλιστική.



Εικόνα 32: Υποστήριξη κλινικής διάγνωσης από το σύστημα ανάκτησης ιατρικών εικόνων



## 5 Εξαγωγή Οπτικής πληροφορίας από Μαγνητικές Τομογραφίες Εγκεφάλου

Στο προηγούμενο κεφάλαιο υποστηρίχθηκε πως η σημασιολογική ανάκτηση εικόνων είναι δυνατό να επιτευχθεί αποδοτικά, όταν περιορίζεται σε εξειδικευμένες εφαρμογές, με τη χρήση της α-ργιοσι γνώσης από το συγκεκριμένο πεδίο εικόνων. Με γνώμονα την παραπάνω υπόθεση, καθορίστηκε η αρχιτεκτονική ενός συστήματος ανάκτησης εικόνων για εξειδικευμένες εφαρμογές. Περαιτέρω, έγινε φανερή η χρηστικότητα ενός τέτοιου συστήματος σε ένα ολοκληρωμένο ιατρικό περιβάλλον. Στην παρούσα εργασία επιλέγεται, ως εφαρμογή στην προτεινόμενη αρχιτεκτονική, το πεδίο των μαγνητικών τομογραφιών εγκεφάλου και συγκεκριμένα εκείνων των όψεων που απεικονίζουν τη συμμετρία του ανθρώπινου εγκεφάλου. Η ιδιότητα της συμμετρίας χρησιμοποιείται για την επινόηση και τον καθορισμό ενός ορισμού των περιοχών ενδιαφέροντος, που όπως θα γίνει φανερό συνήθως εμπεριέχουν την ουσιαστική πληροφορία των εικόνων αυτών. Επομένως, το κεφάλαιο αυτό είναι αφιερωμένο στη μελέτη των μαγνητικών τομογραφιών εγκεφάλου και συγκεκριμένα στην εξαγωγή οπτικής πληροφορίας με νόημα από αυτές.

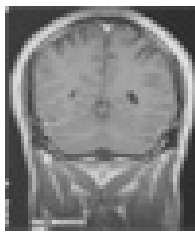
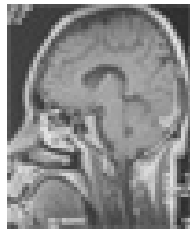
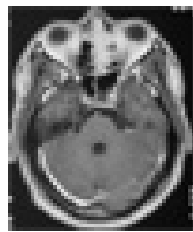
### 5.1 Μαγνητικές Τομογραφίες Εγκεφάλου

Σε αυτή τη παράγραφο παρουσιάζονται οι βασικότερες κατηγορίες μαγνητικών τομογραφιών εγκεφάλου και γίνεται μια σύντομη αναφορά σε μία πολύ χρήσιμη ιδιότητα που εμφανίζει ο ανθρώπινος εγκέφαλος, την ιδιότητα της συμμετρίας.

#### 5.1.1 Κατηγορίες μαγνητικών τομογραφιών εγκεφάλου

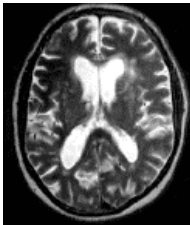
Οι μαγνητικές τομογραφίες του εγκεφάλου διακρίνονται κυρίως από την χρησιμοποιούμενη τεχνική απεικόνισης την διαφορετική γωνία λήψης των τομών και τέλος από τη συγκεκριμένη τομή που απεικονίζεται. Όπως φαίνεται στην Εικόνα 33.b, υπάρχουν διάφορες απεικονιστικές τεχνικές (modalities) με διαφορετικές ιδιότητες η κάθε μία. Οι βασικότερες είναι οι ακόλουθες:

- **PD:** απεικονίζει την πυκνότητα των πρωτονίων του ιστού δηλαδή ουσιαστικά την περιεκτικότητα σε νερό. Για τους περισσότερους μαλακούς ιστούς το μέγεθος αυτό δεν μεταβάλλεται αισθητά επομένως δεν συνεισφέρει στην καλή αντίθεση.
- **T1:** απεικονίζει το χρόνο επαναφοράς της διαμήκουσ συνιστώσας του διανύσματος μαγνήτισης (χρόνος διαμήκουσ χαλάρωσης). Παρουσιάζουν στα φωτεινά το λίπος και το γαδολίνιο(Gd) και στα σκοτεινά το νερό, τον αέρα, το εγκεφαλονωτιαίο υγρό, το ασβέστιο (Ca) και τα οιδήματα. Είναι εξαιρετικές για την απεικόνιση της ανατομίας του εγκεφάλου.
- **T2:** απεικονίζει το χρόνο επαναφοράς της εγκάρσιας συνιστώσας του διανύσματος μαγνήτισης (χρόνος εγκάρσιας χαλάρωσης). Παρουσιάζουν στα φωτεινά το εγκεφαλονωτιαίο υγρό, το νερό και τα οιδήματα και στα σκοτεινά το ασβέστιο τα κόκαλα και τον αέρα. Είναι εξαιρετικές στο διαχωρισμό των ανωμαλιών και βλαβών



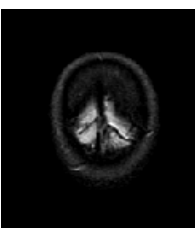
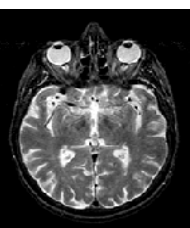
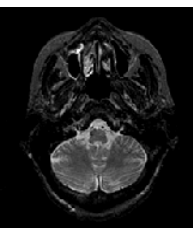
### a) Διαφορετικές Όψεις

Εγκάρσια – Οβελιαία - Στεφανιαία



### b) Τεχνικές Απεικόνισης

T1 - T2 - PD



### c) Διαφορετικές Τομές

10/53 - 25/53 - 50/53

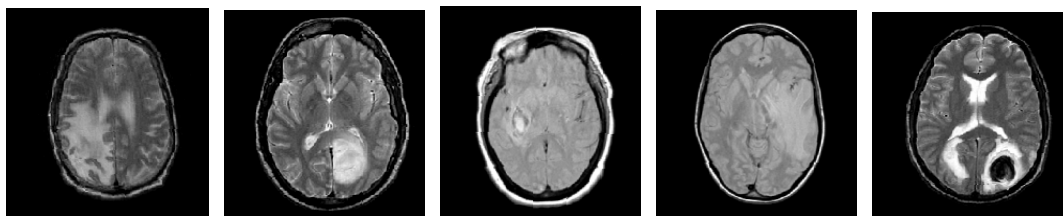
Εικόνα 33: Κατηγορίες μαγνητικών τομογραφιών

Επίσης, ανάλογα με τη γωνία λήψης των τομών υπάρχουν τρεις κύριες όψεις μαγνητικών τομογραφιών (Εικόνα 33.a), οι **εγκάρσιες** (axial), οι **στεφανιαίες** (coronal) και οι **οβελιαίες** (sagittal) τομές. Οι τομές (slices) από κάθε ομάδα εικόνων λαμβάνονται σε αποστάσεις της τάξης του 1 mm. Οπότε για κάθε τύπο μαγνητικής απεικόνισης λαμβάνονται ένας μεγάλος αριθμός από τομές (Εικόνα 33.c).

## 5.1.2 Η ιδιότητα της συμμετρίας στον ανθρώπινο εγκέφαλο

Ο υγιής ανθρώπινος εγκέφαλος παρουσιάζει εξαιρετική συμμετρία ως προς το μέσο οβελιαίο επίπεδο (mid-sagittal plane). Κατά συνέπεια οι εγκάρσιες (axial) και στεφανιαίες τομές (coronal) του εγκεφάλου εμφανίζουν αυτή τη συμμετρία σε αντίθεση με τις οβελιαίες τομές (sagittal). Το γεγονός αυτό είναι φανερό στην Εικόνα 33. Περαιτέρω, η αναγνώριση περιοχών ασυμμετρίας στον εγκέφαλο σε αρκετές περιπτώσεις υποδηλώνει την ύπαρξη κάποιας ασθένειας. Ασθένειες όπως η σχιζοφρένεια (Wang et al., 2000), η επιληψία (Hogan et al., 1999, 2000), και το Alzheimer (Csernansky et al., 2000) έχουν συνδεθεί με την ύπαρξη μορφολογικής-ανατομικής ασυμμετρίας ανάμεσα στο αριστερό και το δεξί ιππόκαμπο (hippocampi) του εγκεφάλου. Σαφώς, η ύπαρξη μορφωμάτων στον εγκέφαλο (π.χ. καρκίνωμα) όπως και εγκεφαλικών επεισοδίων (stroke) συνήθως απεικονίζονται με τη μορφή μεγάλων διαφορών ανάμεσα στο δεξί και αριστερό ημισφαίριο του εγκεφάλου. Στην περίπτωση των μορφωμάτων, η ασυμμετρία εμφανίζεται είτε υπό τη μορφή διαφορετικής ραδιομετρικής απόκρισης των ιστών (infiltrating tumors), είτε υπό τη μορφή γεωμετρική μετατόπισης και διατάραξης της φυσιολογικής δομής του εγκεφάλου (mass-effect tumors). Στην Εικόνα 34 εμφανίζονται ορισμένες

χαρακτηριστικές μαγνητικές τομογραφίες μη υγιούς εγκεφάλου σε διάφορες απεικονιστικές τεχνικές, όπου παρατηρείται μία σημαντική ποικιλία στη φωτεινότητα, το μέγεθος, το σχήμα, και τη σχετική θέση των περιοχών παθολογίας. Αντίθετα, η πρόκληση ασυμμετρίας από αυτές τις περιοχές αποτελεί τη πιο γενική ιδιότητα, ένα κοινό παρανομαστή για τις περισσότερες παθήσεις.



Εικόνα 34: Η εμφάνιση ασυμμετρίας στον ανθρώπινο εγκέφαλο υπό διάφορες παθήσεις

## 5.2 Εύρεση του Άξονα Συμμετρίας και Ευθυγράμμιση Εικόνας

Με τον όρο ευθυγράμμιση της εικόνας σε αυτή την εφαρμογή εννοούμε τη μετατόπιση του κέντρου βάρους της φιγούρας του εγκεφάλου στο κέντρο της εικόνας και τη στροφή της εικόνας γύρω από τον άξονα συμμετρίας του εγκεφάλου ούτως ώστε αυτός να συμπίπτει με τον κατακόρυφο άξονα,  $x=W/2$ , όπου  $W$  είναι το πλάτος της εικόνας. Αυτή η προεπεξεργασία διευκολύνει και απλουστεύει τους περαιτέρω υπολογισμούς στην εικόνα. Είναι φανερό, πως δουλεύοντας με 2Δ μαγνητικές τομογραφίες δεν είναι δυνατό να επιτευχθεί πλήρης ευθυγράμμιση της εικόνας. Γι' αυτό το λόγο, πολλές εφαρμογές κάνουν χρήση της 3Δ πληροφορίας κατά μήκος των τομών του εγκεφάλου (3D volumes) παρέχοντας μια πιο ακριβή ευθυγράμμιση των εικόνων.

### 5.2.1 Χρήση της θεωρίας των ροπών αδράνειας – moments

Η θεωρία των ροών αδράνειας (moments of inertia) χρησιμοποιήθηκε αρχικά στη μηχανική και αργότερα επεκτάθηκαν και σε άλλα πεδία. Ο Hu το 1962 απέδειξε ότι η χρήση τους διατηρεί την πληροφορία των σχημάτων. Ένα δισδιάστατο moment  $m_{pq}$  τάξης  $p+q$  μιας συνάρτησης  $f(x,y)$  ορίζεται ως

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x,y) dx dy \quad (4)$$

Αυτό ονομάζεται και κύριο moment. Η άπειρη σειρά των moments  $p,q$  προσδιορίζεται μοναδικά από ένα σχήμα αλλά και αντίστροφα η συνάρτηση  $f$  μπορεί να ανακατασκευαστεί με ακρίβεια από αυτά. Αν η συνάρτηση  $f$  αντικατασταθεί με το σήμα μιας δυαδικής εικόνας  $E$ , διαστάσεων  $M \times N$ , που λαμβάνει τιμή 1 στην επιφάνεια μίας περιοχής και τιμή 0 οπουδήποτε αλλού, τότε προκύπτει η σχέση:

$$m_{pq} = \sum_{x=1}^M \sum_{y=1}^N x^p y^q E(x,y) \quad (5)$$

Τα moments τάξης-0 προσδιορίζουν το εμβαδόν ενός σχήματος ενώ ο συνδυασμός τους με την τάξη-1 μας δίνει το κέντρο βάρους  $(\bar{x}, \bar{y})$  μίας περιοχής

$$(\bar{x}, \bar{y}) = \left( \frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad (6)$$

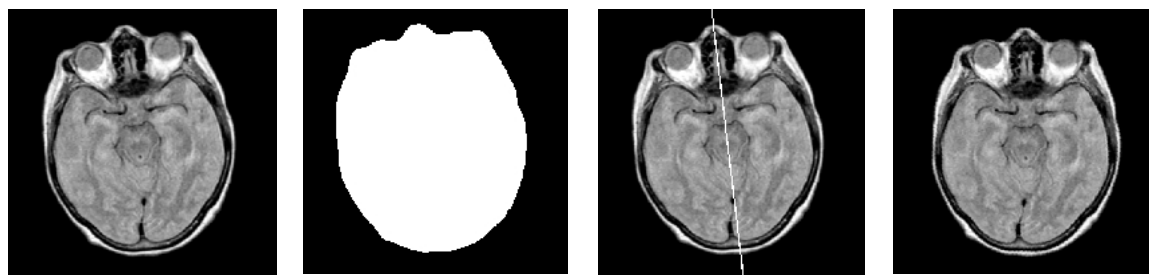
Τα moments τάξης-2 χρησιμοποιούνται για τον καθορισμό των κυρίαρχων αξόνων ενός σχήματος ενώ αντίστοιχες συσχετίσεις υπάρχουν και για κάποιες από τις μεγαλύτερες τάξεις. Η γωνία που σχηματίζουν αυτοί οι άξονες σε σχέση με την εικόνα.

$$\vartheta = \frac{1}{2} \tan^{-1} \left( \frac{2m_{11}}{m_{20} - m_{02}} \right) \quad (7)$$

Οι δύο κύριοι άξονες είναι ουσιαστικά οι κύριοι άξονες της έλλειψης που προσεγγίζει καλύτερα την περιοχή (best fitting ellipse). Επομένως η εκκεντρότητά τους  $e$  αποτελεί μία πολύ αφαιρετική και απλούστατη προσέγγιση του σχήματος

$$e = \frac{m_{20} + m_{02} + \sqrt{(m_{20} - m_{02})^2 + 4m_{11}^2}}{m_{20} + m_{02} - \sqrt{(m_{20} - m_{02})^2 + 4m_{11}^2}} \quad (8)$$

Σύμφωνα με τα παραπάνω, η διαδικασία ευθυγράμμισης της εικόνας είναι η ακόλουθη. Αρχικά δημιουργείται η δυαδική εικόνα με τη χρήση ενός κατωφλίου (thresholding) όπου η περιοχή του εγκεφάλου κωδικοποιείται με 1 και η υπόλοιπη εικόνα με 0. Στη συνέχεια υπολογίζεται το κέντρο βάρους της περιοχής αυτής από τη εξίσωση (6). Περαιτέρω, η γωνία του κύριου άξονα προκύπτει από την εξίσωση (7). Η διαδικασία που ακολουθείται παρουσιάζεται στην Εικόνα 35.



Εικόνα 35: Ευθυγράμμιση Μαγνητικών Τομογραφιών Εγκεφάλου με τη θεωρία των moments

Η μέθοδος των ροπών αδράνειας είναι ταχύτατη, ωστόσο, παρουσιάζει ένα σημαντικό μειονέκτημα. Βασίζεται μόνο στην συμμετρικότητα της φιγούρας του εγκεφάλου και καθόλου στην εσωτερική δομή του. Κατά συνέπεια, η ύπαρξη ασυμμετρίας στο εξωτερικό περίγραμμα της φιγούρας οδηγεί σε λανθασμένους υπολογισμούς, αρχικά του κέντρου βάρους του εγκεφάλου και στη συνέχεια της

γωνίας του κύριου άξονα συμμετρίας. Ένα χαρακτηριστικό παράδειγμα παρουσιάζεται στην Εικόνα 36. Επομένως, το κέντρο βάρους και ο άξονας συμμετρίας που προκύπτουν από αυτή τη μέθοδο είναι πιθανό να βρίσκονται πάνω στο γεωμετρικά αλλά όχι μορφολογικά μέσο οβελιαίο επίπεδο.



Εικόνα 36: Παράδειγμα Λανθασμένης Ευθυγράμμισης με τη Θεωρία των Moments

Καθώς ο αρχικός ορισμός των moments είναι εξαρτώμενος από τη θέση, την περιστροφή και την κλίμακα, έχουν προταθεί διάφορες παραλλαγές που είναι ανεξάρτητες αυτών των παραμέτρων και ονομάζονται moment invariants. Το πλεονέκτημά τους είναι η απλότητα και ο σαφής μαθηματικός ορισμός, ωστόσο μειονεκτούν όσον αφορά τη δυσκολία άμεσης συσχέτισης των υψηλότερων τάξεων με συγκεκριμένα χαρακτηριστικά και ιδιότητες των σχημάτων. Αυτός είναι και ο λόγος που συνήθως χρησιμοποιούνται μόνο μερικές από τις πρώτες τάξεις τους. Επιπλέον, όπως οι περισσότερες μέθοδοι που οδηγούν σε βαθμωτά αριθμητικά μεγέθη, έτσι και τα moments δεν ανιχνεύουν τοπικά χαρακτηριστικά των σχημάτων αλλά προσφέρουν μία συνολική πληροφορία τους.

### 5.2.2 Χρήση διαδοχικών δοκιμών εταιροσυσχέτισης

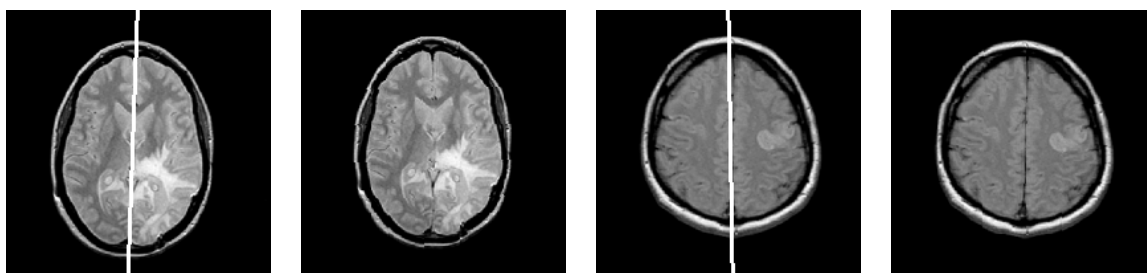
Σε αυτή τη δεύτερη μέθοδο το κριτήριο είναι η εταιροσυσχέτιση (cross-correlation) ανάμεσα στο αριστερό και το δεξί τμήμα της εικόνας εγκεφάλου μετά από διαδοχικές περιστροφές της γύρω από το κέντρο βάρους της. Η πράξη της εταιροσυσχέτισης  $C$  για έλεγχο της συμμετρίας ως προς τον άξονα  $W/2$ , για ένα σημείο  $(i, j)$  του αριστερού τμήματος της εικόνας  $I$ , δίνεται από τη σχέση:

$$C(i, j) = \sum_{k=-N}^N \sum_{l=-N}^N (I(i+k, j+l) - I(W-i+k, j+l))^2 \quad (9)$$

όπου  $W$  είναι το πλάτος της εικόνας και  $(2N+1) \times (2N+1)$  είναι το μέγεθος του παραθύρου υποστήριξης.

Αρχικά, λοιπόν, υπολογίζεται το κέντρο βάρους του εγκεφάλου με βάση τη θεωρία των moments, όπως και στην προηγούμενη μέθοδο. Όπως παρατηρήθηκε στην προηγούμενη παράγραφο, το κέντρο βάρους που υπολογίζεται πιθανώς να μη βρίσκεται πάνω στο μέσο οβελιαίο επίπεδο (mid-sagittal plane). Γι' αυτό το λόγο εκτελούνται δοκιμές για διάφορες εγγύς οριζόντιες μετατοπίσεις του. Για κάθε πιθανό κέντρο βάρους διενεργείται μία δυαδική αναζήτηση της μεγαλύτερης εταιροσυσχέτισης αριστερού-δεξιού τμήματος σε εικόνες που προκύπτουν από την περιστροφή της αρχικής εικόνας. Σε τούτη την εφαρμογή εξετάζονται περιστροφές

μεγέθους  $\pm 10$  μοιρών, σε σχέση με την αρχική εικόνα. Μετά από πέντε δοκιμές η σωστή γωνία του άξονα συμμετρίας υπολογίζεται με ακρίβεια  $0,625^\circ$ . Στην Εικόνα 37 παρουσιάζονται δύο παραδείγματα ευθυγράμμισης με βάση αυτή τη μέθοδο.



**Εικόνα 37: Ευθυγράμμιση μαγνητικών τομογραφιών εγκεφάλου με εταιροσυσχέτισης των δύο ημισφαιρίων**

Στο πρώτο παράδειγμα γίνεται φανερό πως η μέθοδος είναι εξαιρετική ακόμα και σε περιπτώσεις όπου η ύπαρξη μεγάλων και έντονων οιδημάτων έχει αλλοιώσει την εσωτερική-δομική συμμετρία του εγκεφάλου. Στο δεύτερο παράδειγμα, διαφαίνεται ότι η μέθοδος δεν επηρεάζεται ουσιαστικά από την ασυμμετρία του εξωτερικού περιγράμματος του εγκεφάλου σε αντιπαράθεση με τη μέθοδο των ροπών αδράνειας (Εικόνα 36).

Η μέθοδος των διαδοχικών δοκιμών εταιροσυσχέτισης είναι πολύ αποτελεσματική, ωστόσο, είναι εξαιρετικά απαιτητική σε υπολογιστικούς πόρους και χρόνο. Κατά συνέπεια, ανάλογα με τις προδιαγραφές της εκάστοτε εφαρμογής και φυσικά το μέγεθος των εικόνων είναι πολλές φορές επιθυμητό, να διενεργούνται λιγότερες δοκιμές μειώνοντας τη ακρίβεια της μεθόδου. Παραλλαγές αυτής της μεθόδου χρησιμοποιούν την εταιροσυσχέτιση μόνο για τα σημεία ακμών της εικόνας. Οι εικόνες ακμών άλλωστε διατηρούν σε μεγάλο βαθμό τη δομή του οπτικού περιεχομένου. Εναλλακτικά, είναι δυνατό να εξετάζεται η εταιροσυσχέτιση μόνο σε ένα τμήμα της εικόνας, για παράδειγμα σε μία κατακόρυφη ζώνη γύρω από τον εξεταζόμενο άξονα συμμετρίας.

### **5.3 Εξαγωγή πληροφορίας για τη γενική μορφολογία**

Σε αυτήν την παράγραφο γίνεται η προσπάθεια εξαγωγής πληροφορίας για τη γενική περιγραφή του εγκεφάλου από τις μαγνητικές απεικονίσεις του. Τα οπτικά χαρακτηριστικά που επιλέγονται θα αποτελέσουν το πρώτο επίπεδο ανάκτησης. Συνεπώς, πρέπει να είναι κατάλληλα για το διαχωρισμό των εικόνων ως προς τις βασικές κατηγορίες μαγνητικών τομογραφιών που αναφέρθηκαν στην παράγραφο 5.1.1, δηλαδή την απεικονιστική τεχνική (modality), την γωνία λήψης (view) και το ύψος της συγκεκριμένης τομής (slice). Επιπλέον, λαμβάνεται υπ' όψιν ότι μέρος αυτής της πληροφορίας θα χρησιμοποιηθεί για τον εντοπισμό των πιθανών περιοχών παθολογίας, που προφανώς αποτελούν τις περιοχές ενδιαφέροντος της εφαρμογής. Όπως ειπώθηκε στην παράγραφο 5.1.2, το γενικότερο κοινό χαρακτηριστικό των περιοχών αυτών, είναι η διαταραχή της καλής συμμετρίας που εν γένει παρουσιάζει ο εγκέφαλος. Γι' αυτό το λόγο υλοποιείται ένας πρωτότυπος τοπογραφικός χάρτης ασυμμετρίας ο οποίος στη συνέχεια μπορεί να χρησιμοποιηθεί για την τμηματοποίηση των περιοχών αυτών.



### 5.3.1 Χρωματική Πληροφορία – ιστόγραμμα

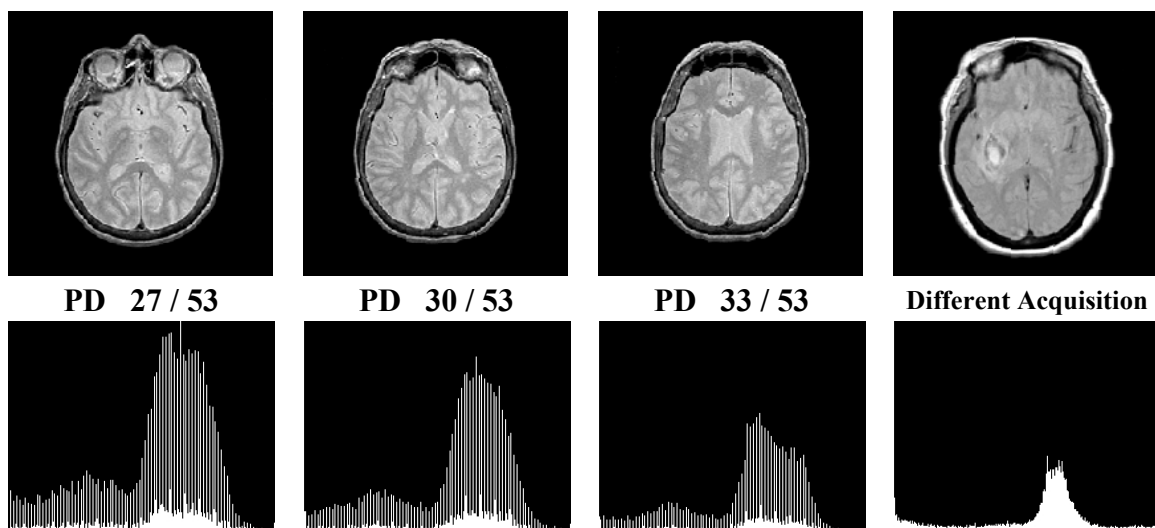
Η χρωματική πληροφορία μίας μαγνητικής τομογραφίας παρέχει την πληροφορία της σύστασης των υποκείμενων ιστών. Περαιτέρω, στην παράγραφο 5.1.1, παρατηρήθηκε πως διαφορετικές τεχνικές μαγνητικής απεικόνισης χρωματίζουν διαφορετικά τους ιστούς. Το πλέον διαδεδομένο μέτρο έκφρασης της γενικής πληροφορίας του χρώματος μίας εικόνας είναι το ιστόγραμμά της. Το ιστόγραμμα μίας εικόνας αποχρώσεων του γκρι είναι ως γνωστό η συνάρτηση πυκνότητας πιθανότητας (probability density function) της φωτεινότητας των εικονοστοιχείων της. Αν  $\delta$  είναι η συνάρτηση του Cronecker,

$$\delta(x) = \begin{cases} 1, & x = 0 \\ 0, & x \neq 0 \end{cases} \quad (10)$$

Τότε το ιστόγραμμα  $H(k)$ , μίας εικόνας  $I(i,j)$ , για κάθε διακριτή διαβάθμιση  $k$  των φωτεινότητων που περιέχονται σε αυτή (έστω 256), δίνεται από τη σχέση:

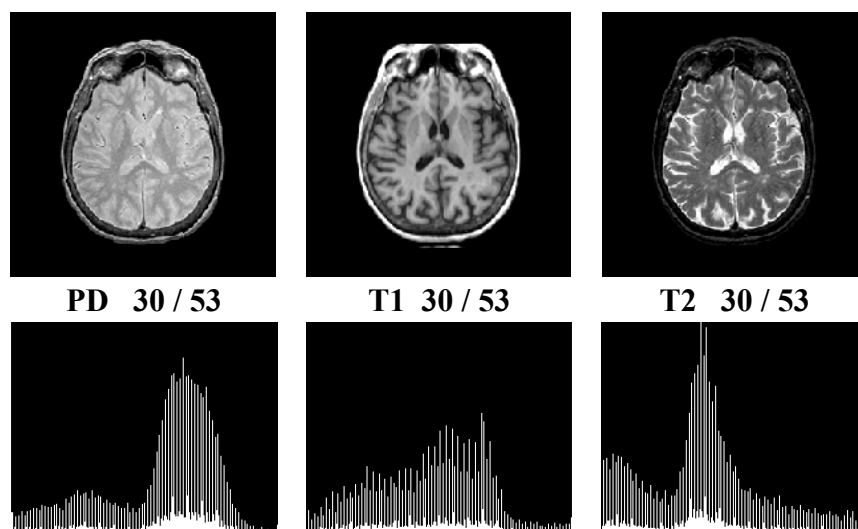
$$H(k) = \sum_{i,j} \delta(I(i,j) - k), \quad k = [1 \dots 256] \quad (11)$$

Σύμφωνα με τα παραπάνω, το ιστόγραμμα μίας μαγνητικής τομογραφίας εγκεφάλου παρέχει διαφοροποιημένη την πληροφορία των ιστών που απεικονίζονται ανάλογα με την τεχνική απεικόνισης. Στην Εικόνα 38, επιδεικνύεται η σαφής ομοιότητα του ιστογράμματος ανάμεσα σε παραπλήσιες τομές pd του ίδιου υποκειμένου. Περαιτέρω, η σύγκριση των ιστογραμμάτων αυτών με εκείνο μίας pd απεικόνισης παραπλήσιας τομής από διαφορετικό περιστατικό (που πιθανότατα προσλήφθηκε υπό διαφορετικές συνθήκες) εξάγει παρόμοια συμπεράσματα.



Εικόνα 38: Το ιστόγραμμα παραπλήσιων τομών PD του ίδιου και διαφορετικού υποκειμένου

Στον αντίποδα, στην Εικόνα 39 παρουσιάζεται η ασθενής συσχέτιση του ιστογράμματος ίδιων τομών, του ίδιου υποκειμένου υπό διαφορετικές απεικονιστικές τεχνικές (pd, t<sub>1</sub>, t<sub>2</sub>).



Εικόνα 39: Το ιστόγραμμα ίδιων τομών του εγκεφάλου υπό απεικονιστικές τεχνικές PD, T1, T2

Κατά τη διαδικασία ανάκτησης, η ανομοιότητα δύο ιστογραμμάτων υπολογίζεται με την σημείο προς σημείο διαφορά των δύο συναρτήσεων. Ωστόσο, υπάρχουν και άλλα παράπλευρα στατιστικά μεγέθη που εμπεριέχουν σημαντική πληροφορία για την εικόνα όπως η μέση τιμή και η διασπορά των φωτεινοτήτων της.

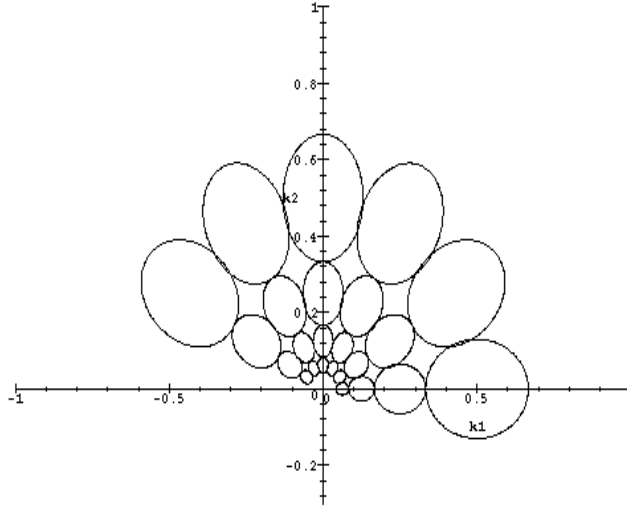
Σε αυτό το σημείο πρέπει να τονισθεί ότι παρότι τα ιστογράμματα είναι πολύ χρήσιμα για την απεικόνιση και την σύγκριση της γενικής χρωματικής πληροφορίας των εικόνων, δεν περιέχουν καμία πληροφορία σχετικά με την τοπογραφική κατανομή του χρώματος. Κατά συνέπεια δεν είναι δυνατό στη συνέχεια να χρησιμοποιηθούν για την οργάνωση της οπτικής πληροφορίας σε σημασιολογικές οντότητες.

### 5.3.2 Πληροφορία υφής – texture Gabor

Ένα ακόμη χαρακτηριστικό της εικόνας που εξετάστηκε είναι αυτό της υφής. Η μέθοδος που επιλέχθηκε για την εξαγωγή της υφής είναι αυτή των φίλτρων Gabor τα οποία φημίζονται για τη βιολογική τους προέλευση και τις τεχνικές τους ιδιότητες που τα καθιστούν ιδανικά για την ανάλυση της υφής [25],[26]. Οι κυματομορφές των φίλτρων Gabor έχουν σχήμα παρόμοιο με την αισθητήρια αντίδραση των κυττάρων του πρωτεύοντος οπτικού φλοιού V1. Διάφοροι ερευνητές στο παρελθόν έχουν χαρακτηρίσει τα κύτταρα αυτά ως ανιχνευτές ακμών, ενώ πειράματα στον αντίστοιχο οπτικό φλοιό ζώων έχουν δείξει ότι οι αντιδράσεις των κυττάρων στα ερεθίσματα αντιστοιχούν σε τοπικές μετρήσεις συχνοτήτων. Επιπλέον, διαπιστώθηκε ότι γειτονικές περιοχές του φλοιού παρουσίαζαν συμμετρία μεταξύ τους. Τα πειράματα αυτά αποτέλεσαν ουσιαστικά το κίνητρο και την έμπνευση των φίλτρων Gabor.

Τα φίλτρα Gabor είναι παραμετροποιημένα τόσο στο πεδίο του χώρου όσο και των συχνοτήτων. Η βάση των φίλτρων Gabor δημιουργείται από ένα σύνολο συναρτήσεων μετατοπισμένων μεταξύ τους στο χώρο και το πεδίο συχνοτήτων, τα όρια των οποίων προσδιορίζονται από γκαουσιανές συναρτήσεις (gaussian envelope function). Κάθε φίλτρο της βάσης αποτελεί μία δισδιάστατη μιγαδική συνάρτηση τα

δύο μέρα της οποίας μοντελοποιούν την προαναφερόμενη βιολογική συμμετρία. Όλα τα φίλτρα, τόσο το πραγματικό όσο και το φανταστικό τους μέρος, δημιουργούνται βάσει ενός μητρικού με κατάλληλη μεταφορά, κλιμάκωση και περιστροφή (Εικόνα 40). Η βάση επομένως, αποτελείται από όλους τους συνδυασμούς συναρτήσεων που προκύπτουν για τις παραπάνω παραμέτρους.



**Εικόνα 40: Μέρος της κατανομής μη επικαλυπτόμενων φίλτρων Gabor στο πεδίο των συχνοτήτων**

Η εξαγωγή της πληροφορίας επιτελείται σε τρία στάδια. Στο πρώτο, δημιουργείται η βάση των φίλτρων Gabor. Στο δεύτερο, γίνεται η συνέλιξη της εικόνας με τη βάση των φίλτρων και εξάγονται χαρακτηριστικά σύμφωνα με τα οποία επιτυγχάνεται η ομαδοποίηση σε κλάσεις με παρόμοια υφή. Αυτή η ομαδοποίηση σε πολλές εφαρμογές χρησιμοποιείται είτε αυτούσια είτε σε συνδυασμό με άλλες μεθόδους για την τμηματοποίηση των εικόνων. Στο τρίτο δημιουργείται και αποθηκεύεται ένα αντιπροσωπευτικό χαρακτηριστικό το οποίο είναι κατάλληλο για ανάκτηση.

Η δημιουργία της βάσης στηρίζεται σε μία γεννήτρια δισδιάστατη μιγαδική συνάρτηση  $g$  όπου:

$$g(x, y) = \left( \frac{1}{2\pi\sigma_x\sigma_y} \right) \exp \left[ -\frac{1}{2} \left( \frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) + 2\pi j W_x \right] \quad (12)$$

από την οποία με κατάλληλες περιστροφές και μεταθέσεις προκύπτουν οι υπόλοιπες:

$$g_{mn}(x, y) = a^{-m} g_{mn}(x', y'), \quad a > 1, \quad m, n \in \mathbb{Z} \quad (13)$$

$$x' = a^{-m} (x \cos \theta + y \sin \theta) \quad (14)$$

$$y' = a^{-m} (-x \sin \theta + y \cos \theta) \quad (15)$$

όπου  $\theta = \frac{n\pi}{K}$  και  $K$  το συνολικό πλήθος των κατευθύνσεων.

Προκειμένου κάθε ένα από τα φίλτρα της βάσης να περιοριστεί σε ένα πεδίο κάλυψης, όπως παρουσιάζεται στην Εικόνα 40, εξαλείφοντας έτσι την εισαγωγή πλεονασμού στη φιλτραρισμένη εικόνα, πρέπει να ορισθούν κατάλληλα οι διακυμάνσεις στην εξίσωση (12). Από το μετασχηματισμό Fourier έχουμε ότι:

$$\sigma_x = \frac{1}{2\pi\sigma_u} \quad \sigma_y = \frac{1}{2\pi\sigma_v} \quad , \text{όπου} \quad (16)$$

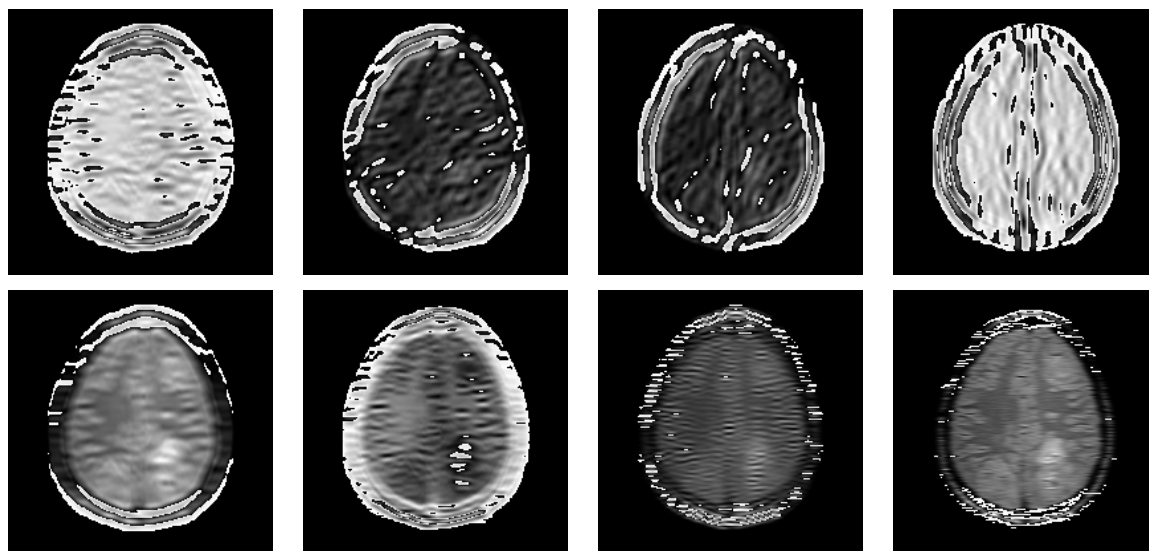
$$\sigma_u = \frac{(a-1)U_h}{(a+1)\sqrt{2\ln 2}} \quad \text{και} \quad (17)$$

$$\sigma_v = \tan\left(\frac{\pi}{K}\right) \left[ U_h - 2\ln 2 \left( \frac{\sigma_u^2}{U_h} \right) \right] \left[ 2\ln 2 - \frac{(2\ln 2)^2 \sigma_u^2}{U_h^2} \right]^{-\frac{1}{2}} \quad (18)$$

Οι παράμετροι  $U_1$  και  $U_h$  ορίζουν τα ακρότατα των συχνοτήτων της περιοχής ενδιαφέροντος,  $m$  είναι ο αύξων αριθμός των κλιμακώσεων ( $0, 1, \dots, S-1$ ),  $n$  ο αύξων αριθμός των περιστροφών ( $1, 2, \dots, K$ ) και

$$a = \left( \frac{U_h}{U_1} \right)^{\frac{1}{S-1}} \quad (19)$$

Στην Εικόνα 41, παρουσιάζεται το αποτέλεσμα της συνέλιξης μαγνητικής τομογραφίας εγκεφάλου με φίλτρα Gabor διαφορετικού προσανατολισμού (1<sup>η</sup> σειρά) και διαφορετικής κλίμακας (2<sup>η</sup> σειρά).



**Εικόνα 41: Αποτέλεσμα συνέλιξης εικόνας MR εγκεφάλου με φίλτρα Gabor διαφορετικής κλίσης και διαφορετικής κλίμακας**

Στην εφαρμογή που υλοποιήθηκε χρησιμοποιήθηκαν 4 διαφορετικές κλιμακώσεις και 6 διαφορετικοί προσανατολισμοί δηλαδή δημιουργούνται 24

διαφορετικές εικόνες. Σκοπός μας στη συνέχεια είναι η τμηματοποίηση της εικόνας σε περιοχές με παρόμοια υφή.

Η εικόνα χωρίζεται σε μη επικαλυπτόμενα τμήματα διαστάσεων 20x20 και για κάθε ένα από αυτά, σε κάθε φιλτραρισμένη εικόνα, υπολογίζουμε τη μέση τιμή  $\mu$  και τη διακύμανση  $\sigma$  της ενέργειας. Για κάθε τμήμα, υπολογίζεται η μέση τιμή και διακύμανση για κάθε κλίμακα και κατεύθυνση παράγοντας ένα διάνυσμα 2x24. Στη συνέχεια τα διανύσματα αυτά ταξινομούνται σε ένα μικρό αριθμό από κλάσεις (3-5) με τον αλγόριθμο k-means. Έτσι επιτυγχάνεται μία απλή τμηματοποίηση με βάση την υφή όπως προαναφέρθηκε. Καθώς, κάθε τμήμα της εικόνας έχει χαρακτηριστεί με την εμφάνιση (ή μη) κάποιου είδους υφής η κωδικοποίηση αυτή αποτελεί έναν τοπογραφικό χάρτη κατάλληλο για την τμηματοποίηση της εικόνας. Ωστόσο, σε αυτήν την εφαρμογή δεν χρησιμοποιήθηκε η πληροφορία της υφής για τον εντοπισμό των περιοχών παθολογίας.

Τέλος για τον υπολογισμό της μετρικής που χρησιμοποιείται κατά την ανάκτηση λαμβάνουν χώρα τα παρακάτω. Για κάθε κλάση δημιουργείται ένα αντιπροσωπευτικό διάνυσμα (2x24) των παραμέτρων  $\mu$  και  $\sigma$ , λαμβάνοντας υπόψη τα κομμάτια που ανήκουν στην εν λόγω κλάση, πλην αυτών που συνορεύουν με άλλες. Το αντιπροσωπευτικό αυτό διάνυσμα αποτελεί τη μέση τιμή και τη διακύμανση για το σύνολο των τμημάτων που ανήκουν σε μια κλάση σύμφωνα με το προαναφερόμενο κριτήριο. Στη διαδικασία ανάκτησης κριτήριο της ανομοιότητας της υφής ανάμεσα σε δύο εικόνες είναι η διαφορά του αντιπροσωπευτικού διανύσματος της κάθε κλάσης της πρώτης εικόνας από την αντίστοιχη κλάση της δεύτερης. Οι συνδυασμοί για κάθε ζεύγος εικόνων είναι το πολύ 9 (για 3 κλάσεις) και η μικρότερη απόσταση αυτών καθορίζει το βαθμό ομοιότητας του ζεύγους. Για μία εκτενέστερη και αναλυτικότερη περιγραφή της διαδικασίας εισαγωγής και ανάκτησης της πληροφορίας υφής που χρησιμοποιήθηκε μπορεί κανείς να ανατρέξει στο [1] από όπου ενσωματώθηκε η μέθοδος.

### 5.3.3 Αμοιβαία πληροφορία – mutual information

Η αμοιβαία πληροφορία είναι ένα μέτρο που εκφράζει το κατά πόσο καλά μία εικόνα μπορεί να εξηγήσει την άλλη και παρουσιάζει ισχυρή εξάρτηση από τη γεωμετρική ευθυγράμμιση και την δομική μορφολογία μίας εικόνας [27]. Κατά συνέπεια είναι εξαιρετικά χρήσιμη για τη σύγκριση της γενικής μορφολογίας δύο ευθυγραμμισμένων μαγνητικών τομογραφιών εγκεφάλου. Η αμοιβαία πληροφορία δύο εικόνων  $A, B$  περιγράφεται από την ακόλουθη εξίσωση:

$$I(A, B) = H(A) + H(B) - H(A, B) \quad (20)$$

Όπου  $H(A)$  (και παρομοίως  $H(B)$ ) εκφράζουν την εντροπία της εικόνας  $A$ .

$$H(A) = - \sum_{a \in A} p_A(a) \cdot \log p_A(a) \quad (21)$$

και  $H(A, B)$  είναι η από κοινού εντροπία των δύο εικόνων και  $p_A$  (ομοίως  $p_B$ ) η συνάρτηση πυκνότητας πιθανότητας της φωτεινότητας των σημείων της εικόνας  $A$ , δηλαδή το ιστόγραμμα της.

$$H(A, B) = - \sum_{a \in A} \sum_{b \in B} p_{AB}(a, b) \log p_{AB}(a, b) \quad (22)$$

με  $p_{AB}$  την από κοινού συνάρτηση πυκνότητας πιθανότητας του ζεύγους εικόνων  $A, B$ . Διάφοροι ερευνητές πρότειναν παρόμοιες μετρικές. Οι περισσότερες διαδεδομένες είναι οι ακόλουθες.

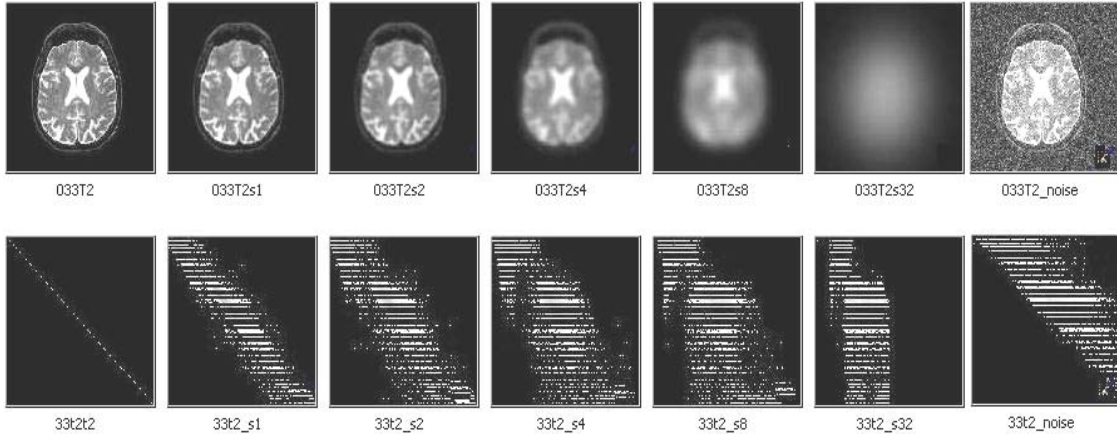
$$I_2(A, B) = \frac{2I(A, B)}{H(A, B)} \quad (23)$$

$$I_3(A, B) = H(A, B) - I(A, B) \quad (24)$$

$$I_4(A, B) = \frac{H(A) + H(B)}{H(A, B)} \quad (25)$$

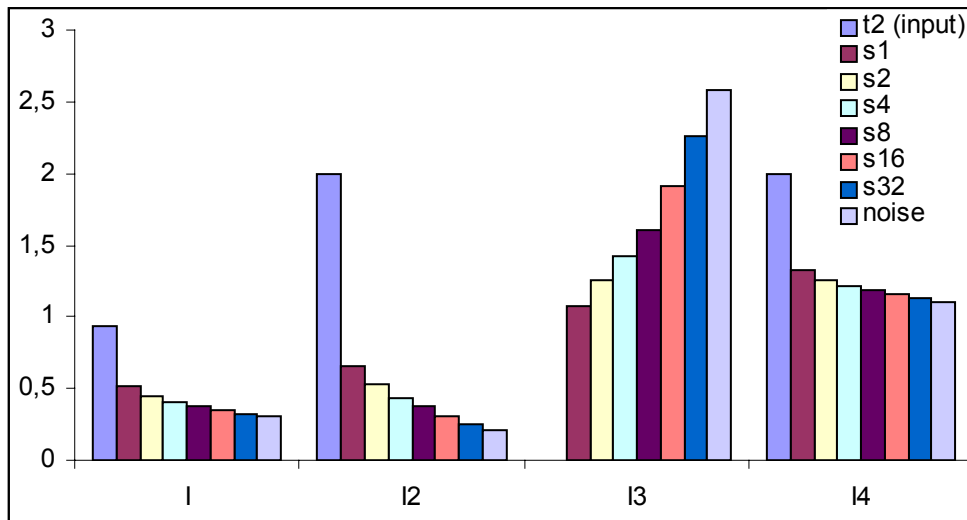
Τα παραπάνω μέτρα πιστεύεται ότι είναι περισσότερο ασφαλή (robust) και ανεξάρτητα από φωτομετρικούς παράγοντες σε σχέση με το MI.

Η Εικόνα 42, επιδεικνύει την ισχυρή εξάρτηση της αμοιβαίας πληροφορίας από τη γεωμετρική ευθυγράμμιση και τη μορφολογία. Στην πρώτη σειρά, η εικόνα εισόδου εξομαλύνεται σταδιακά με ένα γκαουσιανό φίλτρο τυπικής απόκλισης  $\sigma = \{1, 2, 4, 8, 16, 32\}$ . Τα αντίστοιχα  $2\Delta$  ιστογράμματα αμοιβαίας πληροφορίας με την αρχική εικόνα παρουσιάζονται στη δεύτερη σειρά, όπου διαφάνεται η σταδιακή διασπορά από την ιδανική διαγώνιο.



**Εικόνα 42: Εξάρτηση της αμοιβαίας πληροφορίας από τη μορφολογία της εικόνας. Πειράματα ανάμεσα στην αρχική εικόνα και εξομαλυσμένα και ενθόρυβα στιγμιότυπα της**

Στην Εικόνα 43, επιδεικνύεται διαγραμματικά η (αναμενόμενη) σταδιακή μείωση της αμοιβαίας πληροφορίας, παρόλα αυτά οι τιμές παραμένουν σημαντικά σχετικές με την ιδανική περίπτωση. Η μεγαλύτερη ανομοιότητα εμφανίζεται στην υποβαθμισμένη με θόρυβο εικόνα. Το μέτρο  $I_3$  παρέχει την ομοιότητα με αντίστροφο πρόσημο.



Εικόνα 43: Διάγραμμα εξάρτησης της αμοιβαίας πληροφορίας για τη διάταξη της Εικόνα 42

Η αμοιβαία πληροφορία είναι ένα στατιστικό μέτρο σύγκρισης των εικόνων. Κατά συνέπεια (όπως και στην περίπτωση του χρωματικού ιστογράμματος) δεν είναι δυνατό στη συνέχεια να χρησιμοποιηθεί για την οργάνωση της οπτικής πληροφορίας από τα ανεξάρτητα πρωτογενή σε σημασιολογικές οντότητες. Επιπρόσθετα, το ουσιαστικό μέρος των υπολογισμών γίνεται κατά τη διαδικασία ανάκτησης των εικόνων γεγονός που στο παρόν σύστημα καθυστερεί ανάλογα με το μέγεθος των εικόνων την εφαρμογή του χρήστη.

### 5.3.4 Τοπογραφικός χάρτης ασυμμετρίας

Στην παράγραφο 5.1.2 τονίστηκε η σημασία της συμμετρίας του ανθρώπινου εγκεφάλου. Περιοχές που προκαλούν ασυμμετρία στην μορφολογία του εγκεφάλου, πιθανότατα υποδηλώνουν την ύπαρξη κάποια παθολογίας. Περαιτέρω, στην παράγραφο 5.2.2, χρησιμοποιήθηκε η εταιροσυσχέτιση ανάμεσα στο αριστερό και το δεξί ημισφαίριο του εγκεφάλου (εξ: (9) ), ως ένα καλό μέτρο της γενικής συμμετρίας της εικόνας. Χρησιμοποιώντας τώρα την εξίσωση ως μετασχηματισμό της εικόνας εισόδου σε μία τοπογραφική εικόνα της συνεισφοράς των περιοχών στη συνολική ασυμμετρία, παρατηρούμε πως οι έντονες (ελαφρά μετατοπισμένες ως προς την καλή συμμετρία) ακμές ευνοούνται και επιπλέον διαπλατώνονται. Επομένως, ο μετασχηματισμός αυτός δεν αποτελεί ένα καλό αρχικό μέτρο εντοπισμού των επιθυμητών περιοχών. Ένας ελαφρώς καλύτερος μετασχηματισμός που μειώνει το φαινόμενο της διαπλάτυνσης των ακμών, είναι ο παρακάτω.

$$C_1(i, j) = \frac{1}{N^2} (I(i, j) - I(w - i + k, j + l)) \quad , k, l \in [-N \dots N] \quad (26)$$

Οι τοπογραφικοί χάρτες που παράγονται με αυτό το μετασχηματισμό παρουσιάζονται στη δεύτερη σειρά της Εικόνα 44, για τρεις τομογραφίες με εμφανείς παθήσεις και για μία υγιούς εγκεφάλου με έντονες ακμές. Παρατηρείται πως οι περιοχές παθολογίας διατηρούνται πολύ ικανοποιητικά παρουσία όμως έντονου θορύβου που παράγεται κυρίως από τις ακμές.

Στην παρούσα μελέτη προτείνουμε ένα νέο παραπλήσιο, περισσότερο εκφραστικό μέτρο της ασυμμετρίας. Ουσιαστικά, προτείνουμε την αντιστοίχιση των εικονοστοιχείων που βασίζεται σε κριτήρια συμμετρίας. Σημαντική μελέτη του προβλήματος της αντιστοίχισης διεξήγαγαν πρώτοι ο Ullman [28] και οι Marr και Poggio [29]. Συγκεκριμένα ο Ullman στη θεωρία του περί ελάχιστης αντιστοίχισης (minimal mapping), πρότεινε τρία διαισθητικά τοπικά κριτήρια καλής καθολικής (global) αντιστοίχισης τα οποία είναι: η ομοιότητα, η εγγύτητα και η αποκλειστικότητα ή 1-1 αντιστοίχιση. Στην παρούσα εργασία αρκούμαστε στην χρήση των δύο πρώτων κριτηρίων αν και η ενσωμάτωση του τρίτου κριτηρίου (της αποκλειστικότητας) πιθανότατα θα βελτιώνει την αποτελεσματικότητα της μεθόδου. Το σκεπτικό που ακολουθήθηκε είναι το ακόλουθο:

Όπως έχει αναφερθεί ο ανθρώπινος υγιής εγκέφαλος παρουσιάζει καλή συμμετρία. Παρ' όλα αυτά ακόμα και στον υγιή εγκέφαλο παρουσιάζονται μικρές χωρικές και ραδιομετρικές μετατοπίσεις των αντισυμμετρικών ιστών. Αυτές οι μικρές αποκλίσεις από τη συμμετρία οφείλονται απλά στην φυσιολογία του εγκεφάλου αλλά και στην απόκλιση της εικονιζόμενης τομής από το εγκάρσιο επίπεδο. Ωστόσο, ένα εικονοστοιχείο του υγιούς ιστού αναμένεται να «συναντά» στοιχείο παρόμοιου ιστού (κριτήριο ραδιομετρικής ομοιότητας), στη γειτονιά (κριτήριο εγγύτητας) γύρω από το αντισυμμετρικό του σημείο. Αντίθετα, ένα εικονοστοιχείο μη υγιούς ιστού αναμένεται να μην συναντά στην αντισυμμετρική περιοχή κάποιο στοιχείο παρόμοιας φωτεινότητας είτε λόγω διαφορετικής σύστασης (ραδιομετρική μετατόπιση) είτε λόγω γεωμετρικής μετατόπισης των ιστών (mass effect). Το μέτρο της αντιστοίχισης που προτείνουμε για ένα σημείο  $(i, j)$  είναι το ακόλουθο:

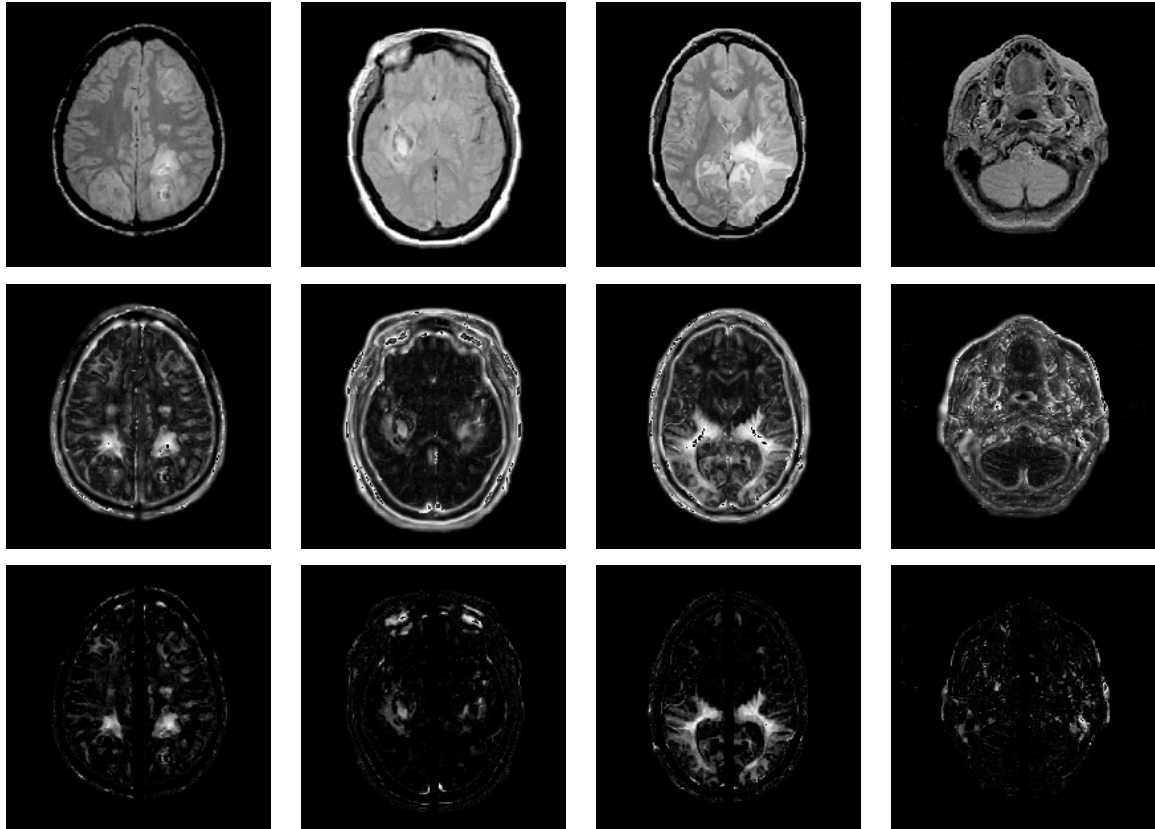
$$C_2(i, j) = \left| \min_{k,l} (I(i, j) - I(w - i + k, j + l)) \right|, k, l \in [-N \dots N] \quad (27)$$

για όλα τα  $k, l$  του παραθύρου υποστήριξης, μεγέθους  $(2N+1 \times 2N+1)$ .

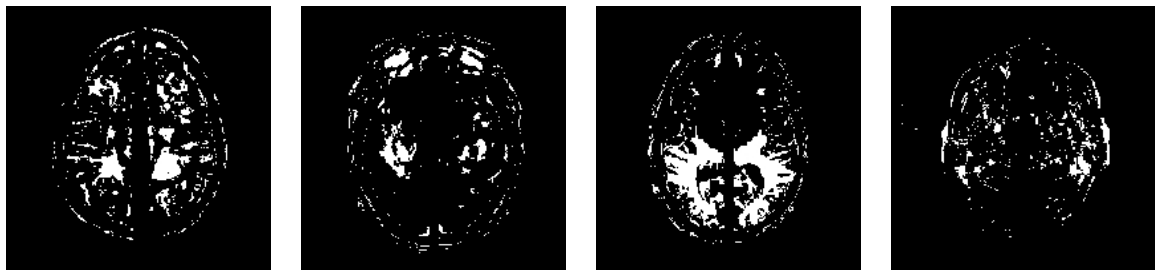
Με βάση αυτό το κριτήριο δημιουργείται ένας τοπογραφικός χάρτης της ασυμμετρίας των σημείων της εικόνας. Στην τρίτη σειρά της Εικόνας 44 παρουσιάζονται οι τοπογραφικοί χάρτες με βάση το δικό μας κριτήριο ασυμμετρίας σε διάφορες μαγνητικές τομογραφίες εγκεφάλου. Σε αυτές τις συγκριτικές εικόνες παρατηρείται ότι ο νέος μετασχηματισμός είναι πολύ πιο ανθεκτικός στις ακμές ενώ διατηρεί πολύ ικανοποιητικά τις περιοχές παθολογίας. Το γεγονός αυτό γίνεται ξεκάθαρο στην (τελευταία) εικόνα του υγιούς εγκεφάλου.

Στη συνέχεια αναζητείται ένα κατώφλι  $t$ , καλής αντιστοίχισης. Σε εικόνες  $200 \times 200 \times 256$ , επιλέχθηκε η παράμετρος του παραθύρου υποστήριξης  $N=2$  (μέγεθος παραθύρου  $5 \times 5$ ), και το κατώφλι  $t=15$ . Τα ζεύγη των αντιστοιχισμένων σημείων ταξινομούνται κατά φθίνουσα σειρά ομοιότητας, ενώ από τα σημεία που προκαλούν ασυμμετρία δημιουργείται ένας δυαδικός τοπογραφικός χάρτης ( $C_2(i, j) > 15$ ). Οι τοπογραφικοί δυαδικοί χάρτες ασυμμετρίας για τις παραπάνω τομογραφίες παρουσιάζονται στην Εικόνα 45.





Εικόνα 44: Τοπογραφικοί χάρτες ασυμμετρίας. (Απεικονίζεται μέγεθος: 2\*C)



Εικόνα 45: Δυαδικοί τοπογραφικοί χάρτες ασυμμετρίας

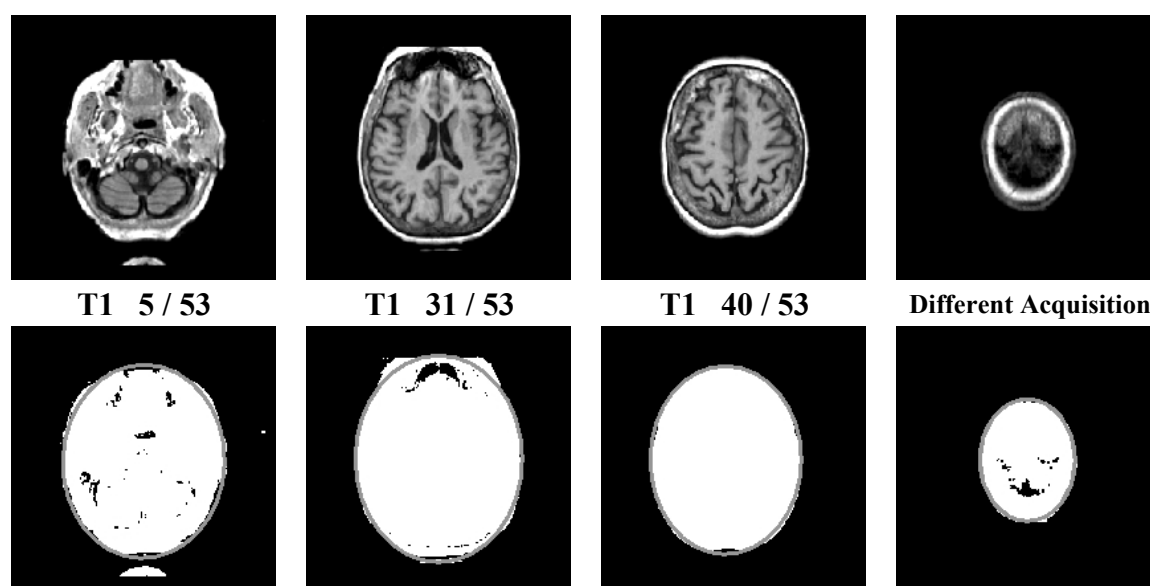
Παρατηρείται ότι αυτοί οι τελικοί χάρτες εντοπίζουν ικανοποιητικά τις περιοχές που πιθανότατα αποτελούν σημεία παθολογίας. Βέβαια συμπεριλαμβάνουν και τις αντισυμμετρικές τους περιοχές που όπως αναμενόταν δεν βρίσκουν αντίστροφα την αντιστοίχισή τους. Επίσης θόρυβος παρουσιάζεται και σε φυσιολογικές περιοχές λόγω φυσιολογικής δομικής ασυμμετρίας, ο οποίος ωστόσο είναι εύκολο να απαλειφθεί. Αυτό που είναι σημαντικό να τονισθεί, είναι ότι είναι διαθέσιμη μία πρώτη τμηματοποίηση των περιοχών ενδιαφέροντος από τη γενική ιδιότητα της συμμετρίας δίχως να χρησιμοποιηθεί καμία άλλη υπόθεση για τις ιδιότητες των περιοχών αυτών.

Κατά τη διαδικασία ανάκτησης, οι δυαδικοί τοπογραφικοί χάρτες ασυμμετρίας από δύο εικόνες συγκρίνονται σημείο προς σημείο με κριτήριο το πλήθος των επικαλυπτόμενων ασύμμετρων εικονοστοιχείων. Ένα πιο απλό μέτρο γενικής ασυμμετρίας θα μπορούσε να συγκρίνει το συνολικό πλήθος των σημείων που χαρακτηρίζονται ασύμμετρα, το αθροιστικό C για όλη την εικόνα κ.α. Ωστόσο το τοπογραφικό μέτρο είναι περισσότερο εκφραστικό, όπως και ο τοπογραφικός χάρτης

κατάλληλος για χρήση κατά τον συνδυασμό της πρωτογενούς πληροφορίας από τους διάφορους παράλληλους χάρτες.

### 5.3.5 Σχήμα ως έλλειψη– best fitting ellipse

Στην παράγραφο 5.2.1 παρατηρήθηκε πως η θεωρία των ροπών αδράνειας μας παρέχει τους κύριους άξονες μίας περιοχής που ουσιαστικά αποτελούν τους άξονες της έλλειψης που προσεγγίζει καλύτερα την περιοχή αυτή. Καθώς η φιγούρα μίας στεφανιαίας τομής του ανθρώπινου εγκεφάλου αφαιρετικά μπορεί να θεωρηθεί ωοειδής, είναι δυνατό να προσεγγίσουμε ικανοποιητικά το σχήμα της από την έλλειψη αυτή (Εικόνα 46).



Εικόνα 46: Προσέγγιση του σχήματος του εγκεφάλου με βάση την εκκεντρότητα του

Ουσιαστικά, το μέγεθος που χρησιμοποιείται για τη σύγκριση δύο σχημάτων εγκεφάλου είναι η εκκεντρότητα της έλλειψης που ορίζεται ως ο λόγος του μέτρου του κύριου δια τον δευτερεύοντα άξονά της. Αυτό το μέτρο είναι ανεξάρτητο της κλίμακας και της στροφής της εικόνας και υπολογίζεται με βάση τις ροπές αδράνειας της περιοχής σύμφωνα την εξίσωση (8).

## 5.4 Τμηματοποίηση Μαγνητικών Τομογραφιών Εγκεφάλου

Οι περιοχές ενδιαφέροντος στο πεδίο των μαγνητικών τομογραφιών του εγκεφάλου είναι προφανώς οι περιοχές όπου ο εγκέφαλος είναι προσβεβλημένος από κάποια πάθηση. Ωστόσο, είναι γνωστό ότι η τμηματοποίηση των περιοχών παθολογίας αποτελεί ένα εξαιρετικά δυσεπίλυτο πρόβλημα. Ουσιαστικά, μία εικόνα είναι δυνατό να περιέχει περισσότερες από μία τέτοιες περιοχές, κάθε μία από τις οποίες δεν παρουσιάζει προκαθορισμένα χαρακτηριστικά όσον αφορά το σχήμα την πυκνότητα ή τη γεωμετρική και ανατομική τοποθεσία. Ωστόσο όπως ειπώθηκε η διαταραχή της καλής συμμετρίας του εγκεφάλου συνήθως υποδηλώνει την ύπαρξη κάποιας ασθένειας. Σε αυτή την παράγραφο προτείνουμε μία μέθοδο εντοπισμού των

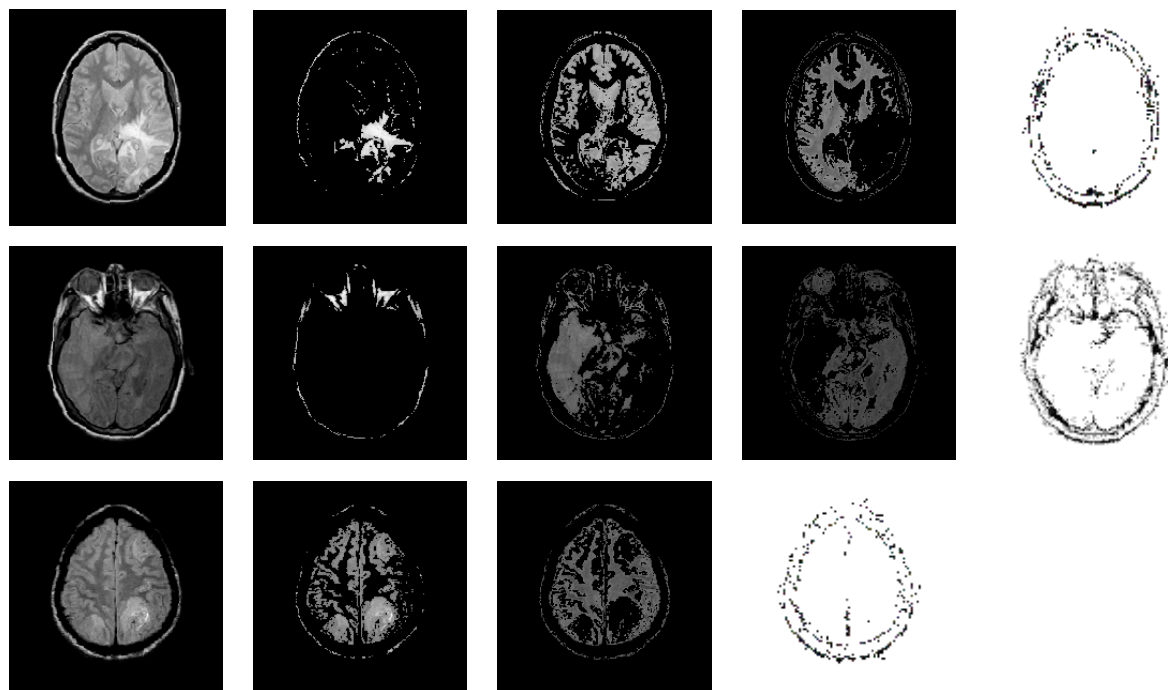
περιοχών που προκαλούν ασυμμετρία η οποία βασίζεται σε μία παραλλαγή του αλγόριθμου ομαδοποίησης kmeans. Η ομαδοποίηση των στοιχείων της εικόνας λαμβάνει υπ' όψιν της την φωτεινότητα των στοιχείων και τη συμμετρία τους σύμφωνα με την πληροφορία που παρέχει ο πρώτος τοπογραφικός χάρτης ασυμμετρίας. Όπως θα γίνει φανερό η μέθοδος είναι εξαιρετικά αποτελεσματική στον εντοπισμό των περιοχών παθολογίας στις μαγνητικές τομογραφίες του εγκεφάλου.

### 5.4.1 Χρωματικός διαχωρισμός με τον αλγόριθμο kmeans

Ο αλγόριθμος k-means είναι ένας αλγόριθμος για την ομαδοποίηση (ή το διαχωρισμό)  $N$  αντικειμένων σε  $k$  ξένα υποσύνολα ώστε να ελαχιστοποιείται κάποιο κριτήριο (συνήθως το sum-of-squares error) [30]. Ο αλγόριθμος είναι ιδιαίτερα απλός:

- Αρχικά τα αντικείμενα τοποθετούνται με τυχαίο τρόπο στα υποσύνολα
- Στη συνέχεια υπολογίζεται το κέντρο του κάθε υποσυνόλου (π.χ. η μέση τιμή)
- Τα αντικείμενα επανατοποθετούνται στο κοντινότερο (σχετικά με το κέντρο) υποσύνολο
- Η διαδικασία συνεχίζεται μέχρι να μην υπάρχουν γίνουν μετακινήσεις στο προηγούμενο βήμα (ή αν εκπληρωθεί κάποια άλλη συνθήκη)

Ο αλγόριθμος k-means χρησιμοποιείται στο χρωματικό διαχωρισμό των περιοχών μίας εικόνας. Οι μαγνητικές τομογραφίες εγκεφάλου απεικονίζουν χρωματικά τους ιστούς ανάλογα με τις φυσικές τους ιδιότητες. Επομένως ο αλγόριθμος kmeans θα μπορούσε να χρησιμοποιηθεί για το διαχωρισμό των διαφόρων ιστών. Τα αποτελέσματα σε μαγνητικές τομογραφίες εγκεφάλου για τον διαχωρισμό των διαφόρων ιστών σε 4 διαφορετικά υποσύνολα παρουσιάζονται στην Εικόνα 47.



Εικόνα 47: Αλγόριθμος k-means σε μαγνητικές τομογραφίες εγκεφάλου με παθολογία

Παρατηρείται πως η απλή στατιστική χρωματική ομαδοποίηση δεν διαχωρίζει ικανοποιητικά τους διάφορους ιστούς. Η εμφάνιση παθήσεων σε πολύ κοντινά χρωματικά επίπεδα με τους υποκείμενους ιστούς καθιστά το πρόβλημα ακόμη πιο

περίπλοκο. Είναι σαφές πως για την τμηματοποίηση ή διαχωρισμό των ανεξάρτητων αντικειμένων μίας εικόνας απαιτείται η οργάνωση της πληροφορίας του χρώματος με βάση την χωρική κατανομή της.

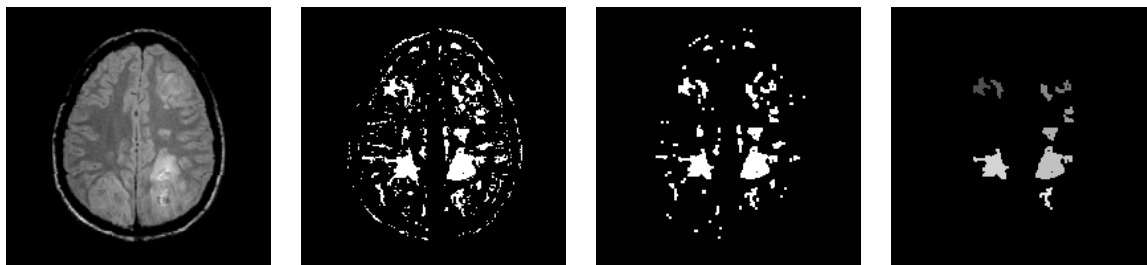
## 5.4.2 Τμηματοποίηση με βάση την ασυμμετρία

Η ομοιογένεια είναι μία ιδιότητα που διέπει σε μεγάλο βαθμό τις ανεξάρτητες φυσικές οντότητες. Περαιτέρω, το έντονο περίγραμμα είναι η ιδιότητα που συνήθως τις διαχωρίζει από το περιβάλλον και τα γύρω αντικείμενα. Ωστόσο, η τμηματοποίηση των εικόνων που βασίζεται σε αυτές τις ιδιότητες θεωρείται ένα άλυτο πρόβλημα της υπολογιστικής όρασης. Στη συγκεκριμένη εφαρμογή, δηλαδή στην τμηματοποίηση των μαγνητικών τομογραφιών εγκεφάλου, χρησιμοποιούμε τη γνώση μας από το πεδίο εφαρμογής για να διευκολύνουμε τους στόχους της τμηματοποίησης. Γνωρίζουμε πως οι περιοχές ενδιαφέροντος είναι οι περιοχές που προκαλούν την ενδεχόμενη ασυμμετρία της εικόνας. Περαιτέρω, αναμένουμε πως οι υπόλοιπες περιοχές της εικόνας, εμφανίζονται σε συμμετρικά ζεύγη. Στην προτεινόμενη μέθοδο η τμηματοποίηση λαμβάνει χώρα σε τρία στάδια και αντιμετωπίζει διαφορετικά τα συμμετρικά στοιχεία από εκείνα που διαταράσσουν τη συμμετρία. Η πληροφορία αυτή παρέχεται στην πλατφόρμα ανάκτησης από τον πρώιμο οπτικό πράκτορα που διαχειρίζεται τον τοπογραφικό χάρτη ασυμμετρίας.

Στο πρώτο βήμα της επεξεργασίας γίνεται ένας χρωματικός διαχωρισμός μόνο των εικονοστοιχείων της εικόνας που κρίθηκαν ως συμμετρικά με βάση το κριτήριο της εξίσωσης (27). Τα συμμετρικά στοιχεία ομαδοποιήθηκαν σε ζεύγη κάθε ένα από τα οποία χαρακτηρίστηκε από το βαθμό συμμετρικότητας του. Ως συμπλήρωμα της πληροφορίας του χάρτη ασυμμετρίας, ο πρώιμος πράκτορας ασυμμετρίας, παρέχει στο επίπεδο της τμηματοποίησης και μία ταξινομημένη λίστα με τα ζεύγη συμμετρικών σημείων. Η ταξινόμηση γίνεται κατά τη φθίνουσα σειρά καλής αντιστοίχισης των συμμετρικών σημείων. Εισάγοντας αυτή την πληροφορία στον αλγόριθμο kmeans, ξεκινούμε να τοποθετούμε στα χρωματικά υποσύνολα τα ζεύγη της λίστας, δηλαδή ξεκινώντας από εκείνα που παρουσιάζουν καλύτερη συμμετρία. Με αυτόν τον τρόπο αποφεύγουμε στο πρώτο βήμα να εντάξουμε τα στοιχεία του τοπογραφικού χάρτη ασυμμετρίας από τα οποία θα προκύψουν οι περιοχές ενδιαφέροντος. Ωστόσο, καταφέρνουμε μία ικανοποιητική ομαδοποίηση των γύρω περιοχών η οποία όπως θα γίνει φανερό είναι πολύ μεγάλης σημασίας. Η προτεινόμενη παραλλαγή του kmeans παρουσιάζει τρία πλεονεκτήματα. Πρώτον, λόγω της παραπλήσιας φωτεινότητας των συμμετρικών ζευγών, ο επαναληπτικός αλγόριθμος συγκλίνει πιο γρήγορα. Δεύτερον, προσαρμόζεται στις ιδιότητες και τη μορφολογία της εικόνας ευνοώντας τις ομαδοποιήσεις συμμετρικών περιοχών, δηλαδή παρέχει μία περισσότερο σημασιολογική παρά στατιστική τμηματοποίηση της εικόνας. Τέλος με τη εισαγωγή των εικονοστοιχείων ανά δύο στις χρωματικές ομάδες, κάνει χρήση της είδη υπολογισμένης πληροφορίας του πρώιμου πρακτορα, μειώνοντας το χρόνο εκτέλεσης της κάθε επανάληψης στο μισό.

Στο δεύτερο βήμα του προτεινόμενου αλγορίθμου εξετάζονται τα εικονοστοιχεία που προκαλούν ασυμμετρία, δηλαδή εκείνα που περιέχονται στον τοπογραφικό χάρτη ασυμμετρίας. Όπως φαίνεται από την Εικόνα 45 οι χάρτες αυτοί περιέχουν σημαντικό ωστόσο εύκολα απαλοιίσιμο θόρυβο. Για την απαλοιφή του θορύβου χρησιμοποιήθηκαν διαδοχικά οι μορφολογικοί μετασχηματισμοί ανοίγματος και κλεισίματος των περιοχών. Στη συνέχεια, έγινε η επιλογή και ο διαχωρισμός των μεγαλύτερων ανεξάρτητων τμημάτων (labeling). Η διαδικασία παρουσιάζεται στην

Εικόνα 48. Ανάμεσα σε αυτές τις ανεξάρτητες εντοπισμένες περιοχές βρίσκονται οι περιοχές ενδιαφέροντος της εφαρμογής. Στα αποτελέσματα που ακολουθούν επιλέχθηκε ως ελάχιστο μέγεθος των περιοχών ενδιαφέροντος τα 50px σε εικόνες 200x200. Όπως φαίνεται οι περιοχές εμφανίζονται ως επί το πλείστο σε αντισυμμετρικά ζεύγη. Ωστόσο μία μόνο περιοχή από κάθε ζεύγος προκάλεσε την ασυμμετρία και αυτή αποτελεί το στόχο της τμηματοποίησης.



Εικόνα 48: Διαχωρισμός των σημαντικών αντικειμένων του δυαδικού τοπογραφικού χάρτη.

Στο τρίτο βήμα λαμβάνει χώρα η επιλογή ανάμεσα στα ζεύγη των αντισυμμετρικών περιοχών. Από απλή παρατήρηση προκύπτει πως η διαφορά ανάμεσα στις δύο ανταγωνιστικές περιοχές ενός τέτοιου ζεύγους είναι ότι η περιοχή που προκάλεσε την ασυμμετρία παρουσιάζει όλες τις ιδιότητες των οντοτήτων του φυσικού κόσμου. Δηλαδή καλή ομοιογένεια και πιθανόν έντονο περίγραμμα. Αντίθετα, ο ανταγωνιστής της αποτελεί μία τυχαία επιφάνεια που ενδέχεται να περιέχει τμήματα από περισσότερες φυσικές περιοχές είτε να αποτελεί τμήμα ενός μεγαλύτερου αντικειμένου. Επομένως, απαιτείται η σύγκριση των δύο περιοχών ως προς την εσωτερική ομοιογένεια και την εξωτερική ανομοιογένεια. Σε αυτό το εγχείρημα είναι εξαιρετικά χρήσιμη η πληροφορία από το πρώτο βήμα του αλγορίθμου όπου συντελέστηκε μία αρχική τμηματοποίηση των ιστών από εικονοστοιχεία που χαρακτηρίστηκαν ως συμμετρικά. Ο αλγόριθμος που ακολουθήθηκε είναι επαναληπτικός. Σε κάθε βήμα, εξετάζονται τα αντίστοιχα εικονοστοιχεία του περιγράμματος των δύο περιοχών ως προς την ομοιότητα με τα γειτονικά σημεία του πλησιέστερου χρωματικά ιστού και τα γειτονικά εσωτερικά σημεία του εξεταζόμενου ιστού. Το κριτήριο  $T$  για κάθε τέτοιο σημείο του 35 περιγράμματος  $i, j$  είναι το ακόλουθο:

$$T(i, j) = \underset{k,l}{mean} [I(i, j) - Int(I(i+k, j+l))] - \underset{k,l}{mean} [I(i, j) - Ext(I(i+k, j+l))] - \underset{k,l}{mean} [Hext(i+k, j+l)] \quad , k, l \in [-N \dots N] \quad (28)$$

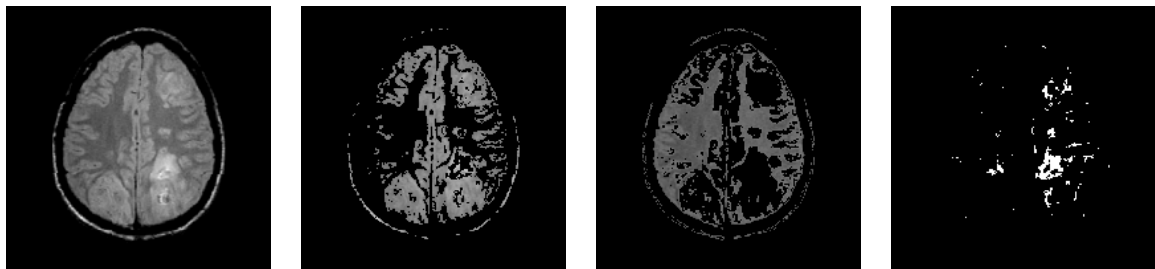
για όλα τα  $k, l$  του παραθύρου υποστήριξης, μεγέθους  $(2N+1 \times 2N+1)$ . Επεξηγηματικά, για κάθε στοιχείο στο περίγραμμα της περιοχής ενδιαφέροντος ορίζεται μία γειτονιά μεγέθους  $(2N+1 \times 2N+1)$ . Σε αυτή τη γειτονιά υπολογίζονται τα ακόλουθα:

- η μέση διαφορά της φωτεινότητάς του από όλα τα σημεία του πλησιέστερου σε φωτεινότητα εξωτερικού ιστού που περιέχονται στη γειτονιά του (σημεία  $Ext$ )
- η μέση διαφορά της φωτεινότητάς του από τα εσωτερικά σημεία της περιοχής που περιέχονται στη γειτονιά του (σημεία  $Int$ )

Επομένως, οι πίνακες  $Int(\cdot)$  και  $Ext(\cdot)$  περιορίζουν διαφορετικά ο κάθε ένας τις πιθανές τιμές των μεταβλητών  $k, l$  ανάλογα με την ομάδα του υποκείμενου ιστού (στοιχείο  $(i+k, j+l)$ ). Η διαφορά αυτών των δύο παρέχει την πληροφορία του κατά πόσο το εξεταζόμενο στοιχείο ταιριάζει καλύτερα με τον εσωτερικό παρά με τον πιο όμοιο περιβάλλοντα ιστό. Στην πρώτη επανάληψη το  $Hext$  είναι μηδέν για όλα τα στοιχεία της εικόνας. Αν το κριτήριο ομοιογένειας  $T(i, j)$  είναι μικρότερο από το αντίστοιχο συμμετρικό του  $T(w-i, j)$  όπου  $w$  το πλάτος της εικόνας και μικρότερο από ένα κατώφλι ομοιογένειας της τάξης του 10, το στοιχείο αυτό προσχωρεί στο περιβάλλοντα ιστό και αντίστροφα. Για εικονοστοιχεία που δεν συναντούν στο συμμετρικό ιστό κάποιο σημείο ασυμμετρίας εξετάζεται μόνο μία ελαφρώς πιο χαλαρή συνθήκη του κατωφλίου (πχ 15). Για κάθε εικονοστοιχείο  $i, j$  που εξέρχεται από την περιοχή του, υπολογίζεται η τιμή  $Hext(i, j)$  ως εξής,

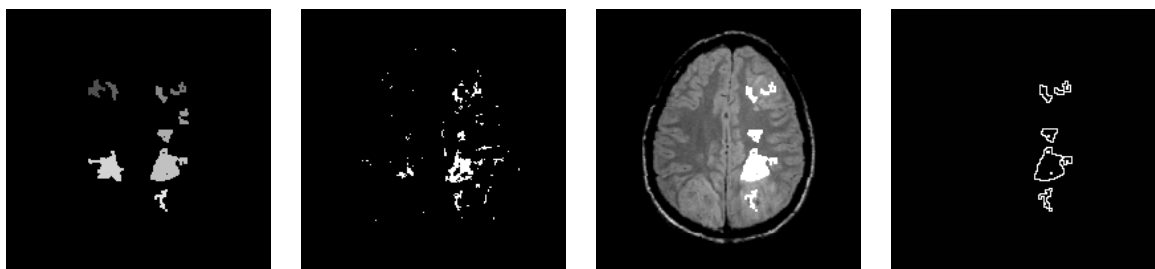
$$Hext(i, j) = T(i, j) \tag{29}$$

Η λογική της μεταβλητής  $Hext$  είναι να περιορίσει σε κάποιο βαθμό την επίδραση του εξερχόμενου στοιχείου στους επόμενους υπολογισμούς. Καθώς συνεχίζονται οι επαναλήψεις της προτεινόμενης μεθόδου, η μία περιοχή συνεχώς μειώνεται, ενώ η ανταγωνιστική της παραμένει σε μεγάλο βαθμό σταθερή. Τα αποτελέσματα μετά από την ολοκλήρωση της μεθόδου για την rd μαγνητική τομογραφία της Εικόνα 48 παρουσιάζονται στην Εικόνα 49 όπου απεικονίζονται και τα αποτελέσματα του συμμετρικού kmeans.



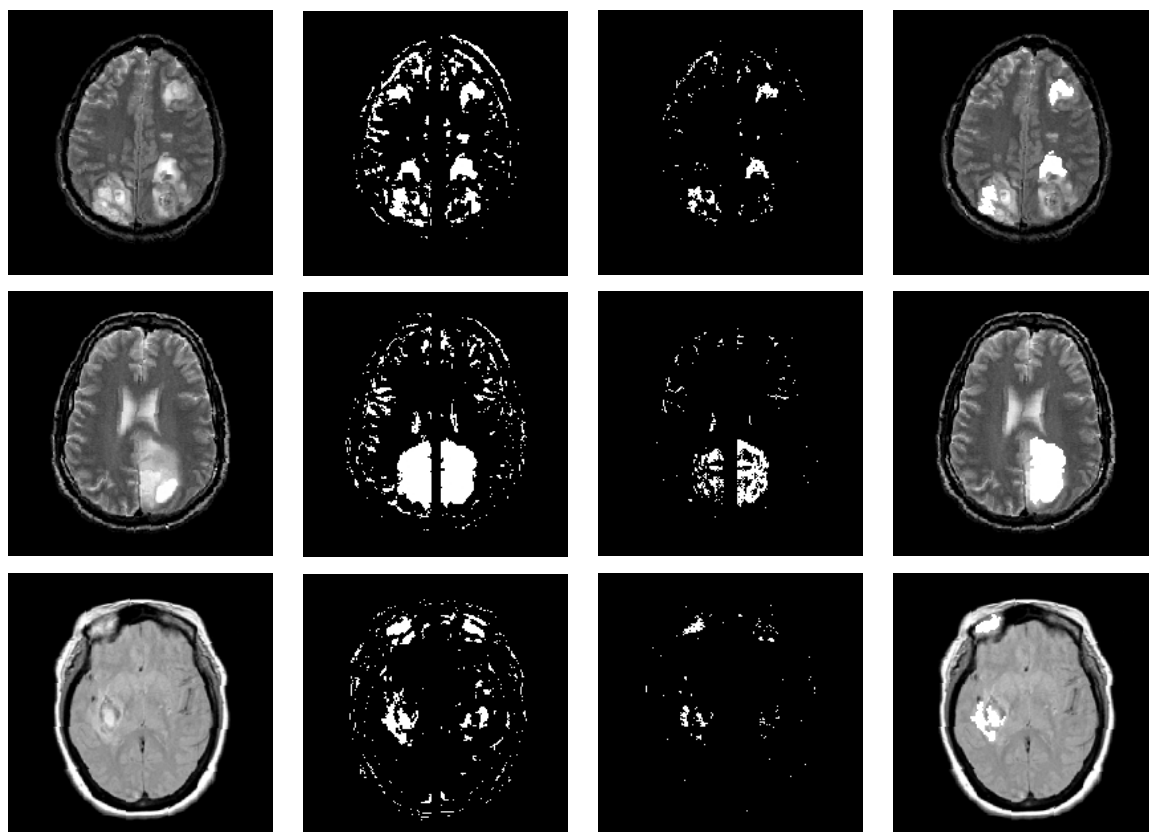
Εικόνα 49: Η επαναληπτική μέθοδος προσχώρησης των στοιχείων δύο ανταγωνιστικών περιοχών στους γύρω ιστούς

Εν τέλει, όποια περιοχή διατηρεί το μέγεθος της σε ένα ικανοποιητικό ποσοστό καθορίζεται ως περιοχή ενδιαφέροντος. Με αυτόν τον τρόπο σε σπάνιες περιπτώσεις είναι δυνατό να επιλεγούν και οι δύο ανταγωνιστικές περιοχές. Τα τελικά αποτελέσματα τμηματοποίησης παρουσιάζονται στην Εικόνα 50. Τα αποτελέσματα της μεθόδου κρίνονται εξαιρετικά δεδομένης και της δυσκολίας της επιλεγμένης εικόνας.



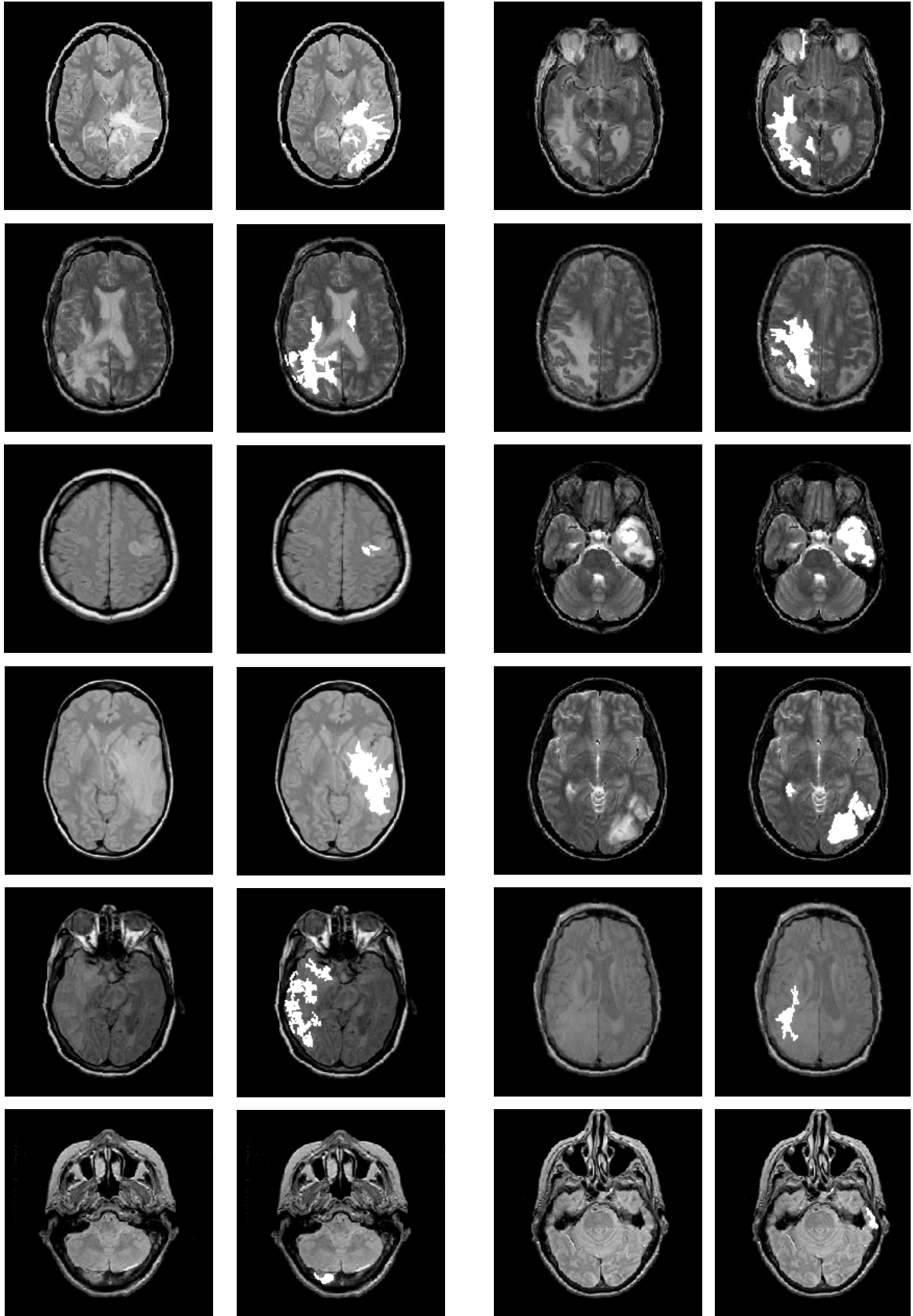
Εικόνα 50: Επιλογή των περιοχών ενδιαφέροντος με βάση το ποσοστό διατήρησής τους μετά τη λήξη του επαναληπτικού αλγόριθμου

Η ίδια διαδικασία παρουσιάζεται στην Εικόνα 51 για μία τρεις επιπλέον μαγνητικές τομογραφίες εγκεφάλου που παρουσιάζουν η κάθε μία ξεχωριστό ενδιαφέρον. Στην τελευταία από αυτές, παρατηρείται μία ψευδής θετική εμφάνιση(false positive) ως αποτέλεσμα της απόκλισης της τομής λήψης από το εγκάρσιο επίπεδο. Λόγω αυτής της απόκλισης απεικονίστηκε η πληροφορία μόνο του αριστερού ματιού η οποία στη συνέχεια επιλέχθηκε ως ενδιαφέρουσα ως προς την πρόκληση ασυμμετρίας.



**Εικόνα 51: Επιλογή των περιοχών ενδιαφέροντος με βάση το ποσοστό διατήρησής τους μετά τη λήξη του επαναληπτικού αλγόριθμου για τρεις χαρακτηριστικές εικόνες**

Η μέθοδος κρίνεται ικανοποιητική ως προς την τμηματοποίηση περιοχών που διαταράσσουν τη συμμετρία σε έντονα συμμετρικές εικόνες. Το κύριο μειονέκτημα της είναι η σύνθετη περιγραφή και ο διαισθητικός τρόπος που υπολογίζονται κρίσιμες παράμετροι. Γι' αυτό το λόγο έγιναν διάφορες προσπάθειες απλούστερης υλοποίησης. Η βασικότερη από αυτές αντιστοιχεί τις αντισυμμετρικές περιοχές του χάρτη ασυμμετρίας και στη συνέχεια επιλέγει εκείνη με το συνδυασμό σαφέστερου περιγράμματος και καλύτερης εσωτερικής ομοιογενείας. Χρησιμοποιήθηκαν διάφορες μετρικές και συνδυασμοί συνεισφοράς των δύο ιδιοτήτων ωστόσο δεν επιτεύχθηκαν καλύτερα αποτελέσματα. Παρ' όλα αυτά το θέμα παραμένει ανοικτό για περαιτέρω μελέτη. Αυτός είναι και ο κύριος λόγος που δεν γίνεται κάποια στατιστική αξιολόγηση των αποτελεσμάτων της τμηματοποίησης βάση των χαρακτηριστικών μετρήσεων για την εμφάνιση πραγματικών/ψευδών εντοπισμών (true/false positives) και ψευδών μη εντοπισμών (false negatives). Ωστόσο για την ικανοποίηση του αναγνώστη παρουσιάζονται επιπλέον αποτελέσματα σε μία μεγάλη γκάμα από μαγνητικές τομογραφίες εγκεφάλου.(Εικόνα 52).

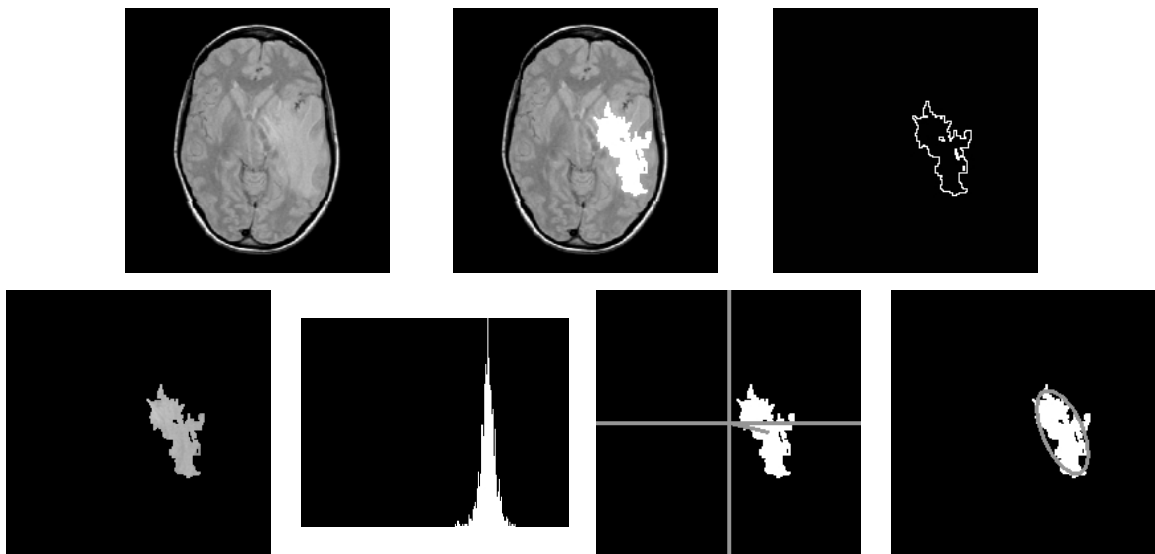


**Εικόνα 52:** Αποτελέσματα τμηματοποίησης μαγνητικών τομογραφιών εγκεφάλου. Απεικόνιση ανά δυάδες (αριστερα εικόνα, δεξιά εντοπισμός περιοχών ενδιαφέροντος). Τα δύο τελευταία παραδείγματα αποτελούν υγιείς εικόνες



## 5.5 Εξαγωγή Πληροφορίας για τις Περιοχές Ενδιαφέροντος

Έχοντας επινοήσει μία ικανοποιητική μέθοδο τμηματοποίησης των περιοχών ενδιαφέροντος, η προσπάθεια επικεντρώνεται στην αναζήτηση κατάλληλων ιδιοτήτων για το σημασιολογικό χαρακτηρισμό των περιοχών αυτών. Στη βιβλιογραφία είναι διαθέσιμη μία ποικιλία οπτικών χαρακτηριστικών που μπορούν να χρησιμοποιηθούν αυτό το σκοπό [20],[22],[23]. Στην παρούσα μελέτη επιλέγηκαν τέσσερα πολύ απλά χαρακτηριστικά όπως το χρώμα, το μέγεθος, το απλό σχήμα και η θέση της περιοχής στην εικόνα. Η πληροφορία που παρέχουν τα χαρακτηριστικά αυτά, παρουσιάζεται για τις περιοχές δύο MRI εγκεφάλου στις παρακάτω εικόνες (Εικόνα 53 και Εικόνα 54).



Εικόνα 53: Εξαγωγή χαρακτηριστικών για περιοχή ενδιαφέροντος από PD axial MRI

### 5.5.1 Χρωματική πληροφορία – ιστόγραμμα

Όπως και στην παράγραφο 5.3.1, χρησιμοποιείται το ιστόγραμμα για το χαρακτηρισμό της χρωματικής πληροφορία μίας περιοχής ενδιαφέροντος.

### 5.5.2 Σχετική θέση

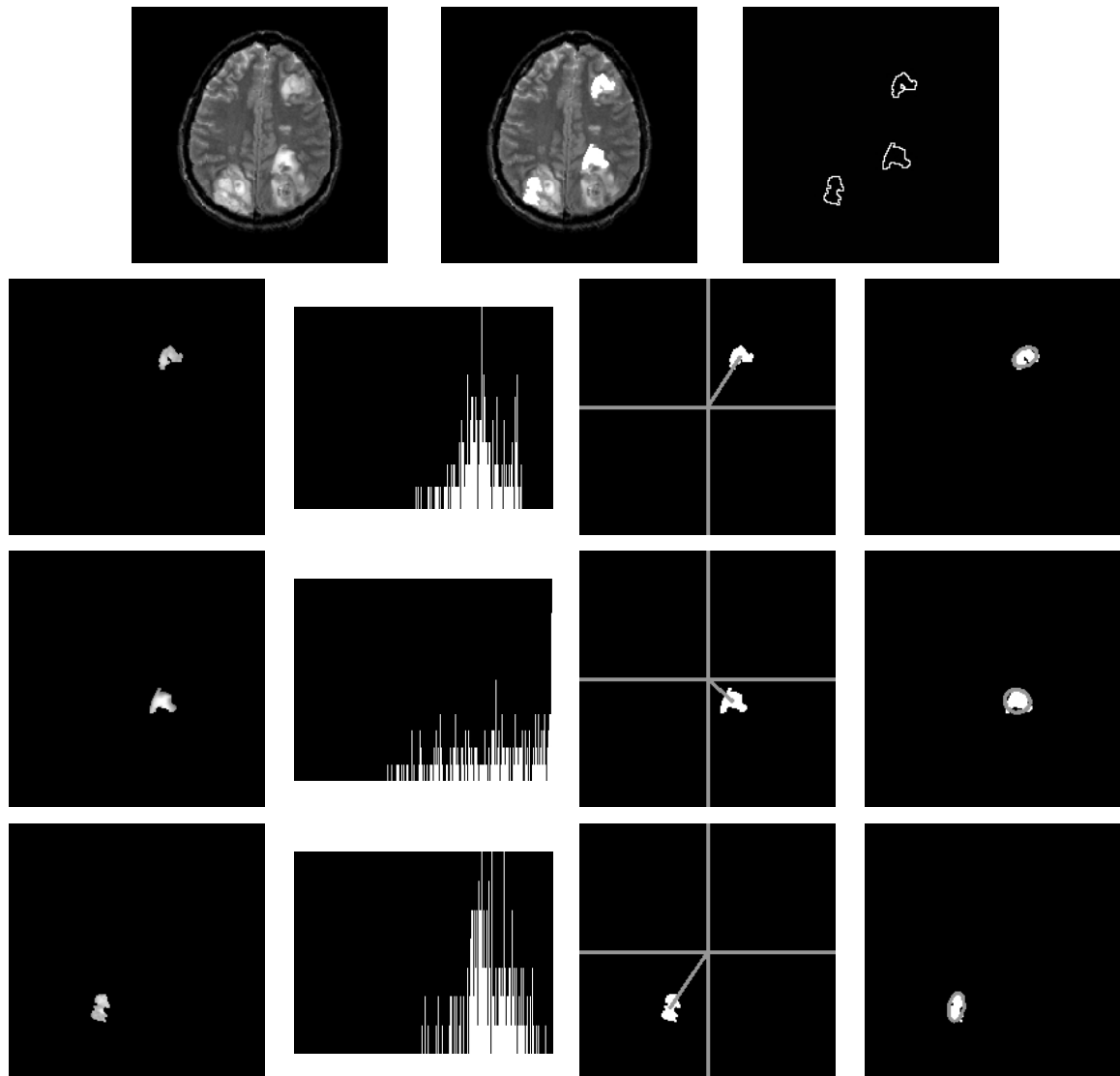
Η θέση κάθε περιοχής ενδιαφέροντος κωδικοποιείται από τη γωνία και την απόσταση του (γεωμετρικού) κέντρου βάρους της από το κέντρο της εικόνας. Κατά τη διαδικασία ανάκτησης ως κριτήριο ανομοιότητας επιλέχθηκε η απόσταση ανάμεσα στα κέντρα των δύο περιοχών.

### 5.5.3 Μέγεθος

Εξάγεται το μέγεθος της περιοχής σε εικονοστοιχεία. Κατά την ανάκτηση κριτήριο ανομοιότητας δύο εικόνων είναι η διαφορά ανάμεσα στο μέγεθός τους.

### 5.5.4 Μορφή

Όπως και στην παράγραφο 5.3.5, χρησιμοποιείται η προσέγγιση του σχήματος από μία έλλειψη. Όπως φαίνεται στην η εκκεντρότητα της έλλειψης είναι εξαιρετικά απλή για να εκφράσει τα διάφορα σχήματα των παθολογικών ιστών, τα οποία είναι ποικίλα και πολύπλοκα. Παρ' όλα αυτά το προαναφερόμενο μέτρο χρησιμοποιείται λόγω της απλότητάς του. Περαιτέρω, εξάγεται το περίγραμμα της περιοχής το οποίο εν γένη είναι χρήσιμο στις ιατρικές εφαρμογές καθώς ορισμένες ιδιότητες του μπορεί να δώσουν πληροφορία για την ύπαρξη οιδήματος γύρω από την περιοχή ενδιαφέροντος. Ωστόσο, στην παρούσα εφαρμογή δεν έγινε χρήση της πληροφορίας του περιγράμματος.



Εικόνα 54: Εξαγωγή χαρακτηριστικών για τις 3 περιοχές ενδιαφέροντος από T2 axial MRI

## 5.6 Συμπεράσματα

Σε αυτό το κεφάλαιο παρουσιάστηκαν αλγόριθμοι για την εξαγωγή σημασιολογικής πληροφορίας από τις μαγνητικές τομογραφίες εγκεφάλου. Οι περισσότεροι περιγράφουν γενικά οπτικά χαρακτηριστικά των εικόνων και κατά συνέπεια μπορούν να χρησιμοποιηθούν σε διάφορες εφαρμογές για τον ίδιο σκοπό. Επιπρόσθετα, χρησιμοποιήθηκε εξειδικευμένη γνώση από το πεδίο των εικόνων της εφαρμογής και συγκεκριμένα της ιδιότητας της συμμετρίας που χαρακτηρίζει τις εικόνες εγκεφάλου. Η πληροφορία της συμμετρίας χρησιμοποιήθηκε με δύο τρόπους. Ως ένα τοπογραφικό χαρακτηριστικό της γενικής μορφολογίας του εγκεφάλου και ως η κατάλληλη ιδιότητα για τον εντοπισμό των περιοχών ενδιαφέροντος της εφαρμογής. Η αποτελεσματική, πρωτότυπη μέθοδος τμηματοποίησης των εικόνων του εγκεφάλου που παρουσιάστηκε (παρ.5.4), επιτρέπει την επιλεκτική περιγραφή των περιοχών που εμπεριέχουν την ουσιαστική πληροφορία. Οι αλγόριθμοι οπτικής περιγραφής υλοποιήθηκαν στο πρότυπο των οπτικών πρακτόρων που περιγράφηκε στην παράγραφο 4.1.1. Αποτελούν παράλληλες ανεξάρτητες διαδικασίες με μηχανισμούς εξαγωγής, σύγκρισης και αποθήκευσης της οπτικής πληροφορίας των προς εξέταση εικόνων ή περιοχών. Περαιτέρω, διαθέτουν μηχανισμούς για τη δια βίου απόκτηση στατιστικής εμπειρίας που τους επιτρέπει να κανονικοποιούν τα αποτελέσματα ομοιότητας μέσα σε ομάδες πρακτόρων (παρ. 4.1.2). Επομένως οι αλγόριθμοι που υλοποιήθηκαν είναι κατάλληλοι για εφαρμογή στην προτεινόμενη αρχιτεκτονική του κεφαλαίου 4. Τα αποτελέσματα από τον πειραματισμό στη συνεισφορά των επιμέρους μεθόδων σε ένα περιβάλλον σύγκρισης εικόνων είναι δυνατό να παρέχει χρήσιμες πληροφορίες για την ικανότητά τους στην εξαγωγή σημασιολογικής πληροφορίας από το εξειδικευμένο πεδίο εφαρμογής.

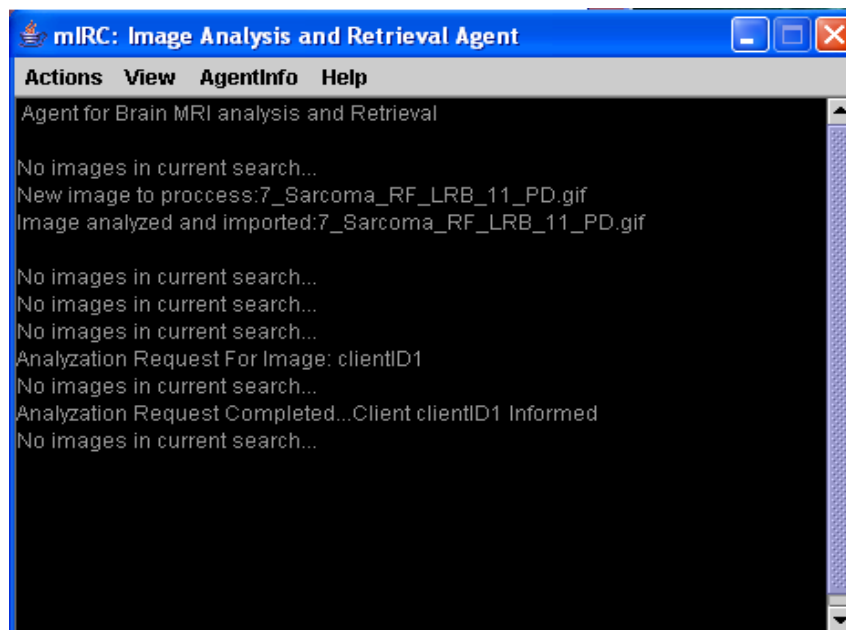


## 6 Εφαρμογή της Αρχιτεκτονικής Ανάκτησης σε Μαγνητικές Τομογραφίες Εγκεφάλου

Σε τούτο το κεφάλαιο παρουσιάζεται η πλατφόρμα ανάκτησης εικόνων που υλοποιήθηκε στα πλαίσια της παρούσας εργασίας. Το πεδίο εφαρμογής που χρησιμοποιείται για την επίδειξη της προτεινόμενης αρχιτεκτονικής είναι οι μαγνητικές τομογραφίες εγκεφάλου (ΚΕΦ 4). Η πλατφόρμα χρησιμοποιεί τους αλγορίθμους περιγραφής των μαγνητικών τομογραφιών εγκεφάλου, οι οποίοι περιγράφηκαν στο κεφάλαιο 5 και όπως ειπώθηκε είναι κατάλληλοι για χρήση στην αρχιτεκτονική δύο επιπέδων. Αρχικά, παρουσιάζονται ορισμένα βασικά χαρακτηριστικά του συστήματος και στη συνέχεια αποτελέσματα ανάκτησης εικόνων σε πρώιμο, επιλεκτικό και συνδυαστικό επίπεδο οπτικής πληροφορίας.

### 6.1 Το Σύστημα που Υλοποιήθηκε

Η εφαρμογή που υλοποιήθηκε ακολουθεί το μοντέλο πελάτη-εξυπρέτη (client-server). Στην πλευρά του εξυπρέτη ένας πράκτορας αναζητά και αναλύει στεφανιαίες μαγνητικές τομογραφίες εγκεφάλου. Η ανάλυση περιλαμβάνει τη γενική περιγραφή, την τμηματοποίηση και την περιγραφή των περιοχών ενδιαφέροντος της εικόνας. Ο εξυπρέτης ταυτόχρονα διαχειρίζεται αιτήσεις ανάλυσης και ανάκτησης μαγνητικών τομογραφιών εγκεφάλου. Στην Εικόνα 55 παρουσιάζεται ένα στιγμιότυπο της διεπαφής του πράκτορα-εξυπρέτη όπου φαίνεται η παράλληλη διαχείριση των λειτουργιών αναζήτησης/ ανάλυσης εικόνων και εξυπηρέτησης αιτήσεων ανάλυσης/ ανάκτησης.



Εικόνα 55: Η διεπαφή του πράκτορα ανάλυσης και ανάκτησης εικόνων

Οι διαδικασίες ανάλυσης αποθήκευσης και σύγκρισης των οπτικών χαρακτηριστικών σε κάθε επίπεδο της αρχιτεκτονικής επιτυγχάνονται από τους οπτικούς πράκτορες που περιγράφηκαν στο κεφάλαιο 5. Όπως ειπώθηκε η συνεισφορά του κάθε οπτικού-

πράκτορα στο αποτέλεσμα της ψηφοφορίας μίας ομάδας πρακτόρων καθορίζεται εφ' άπαξ κατά την εκκίνηση του υπερ-πράκτορα με στόχο την καλύτερη προσαρμογή των απαιτήσεων της εκάστοτε εφαρμογής. Στην παρούσα εφαρμογή για τις μαγνητικές τομογραφίες εγκεφάλου επιλέχθηκαν οι ακόλουθοι οπτικοί πράκτορες με την αντίστοιχη συνεισφορά ανά επίπεδο:

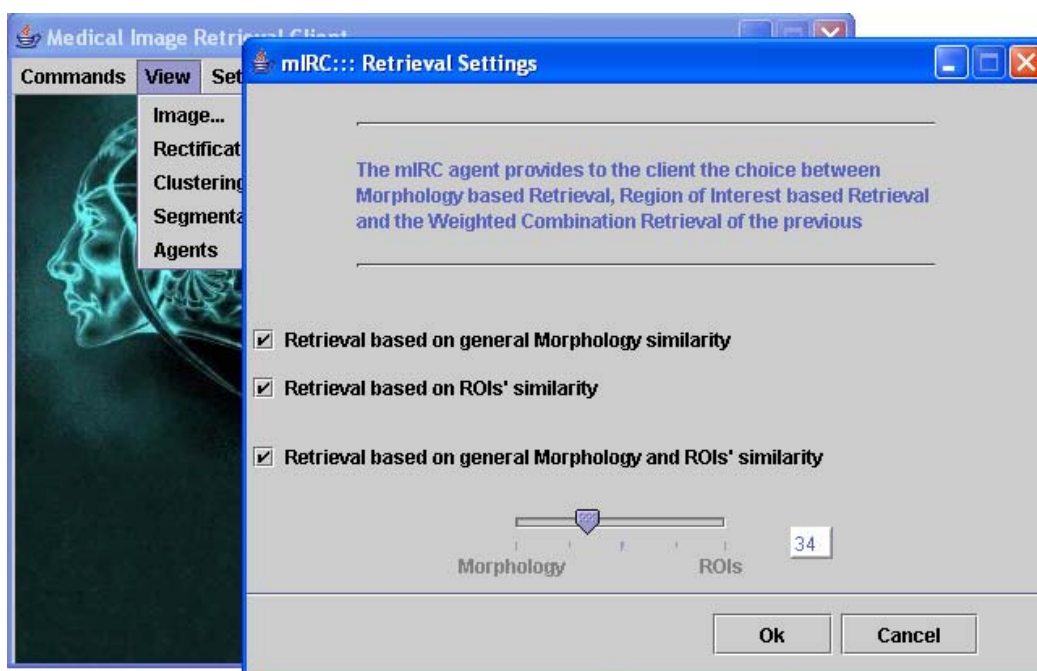
### Πρώιμο Επίπεδο

Οπτικοί Πράκτορες	Συνεισφορά (%)
Αμοιβαία Πληροφορία	45
Ιστόγραμμα Φωτεινοτήτων	20
Σχήμα ως Έλλειψη	20
Τοπογραφικός Χάρτης Ασυμμετρίας	15

### Επιλεκτικό Επίπεδο

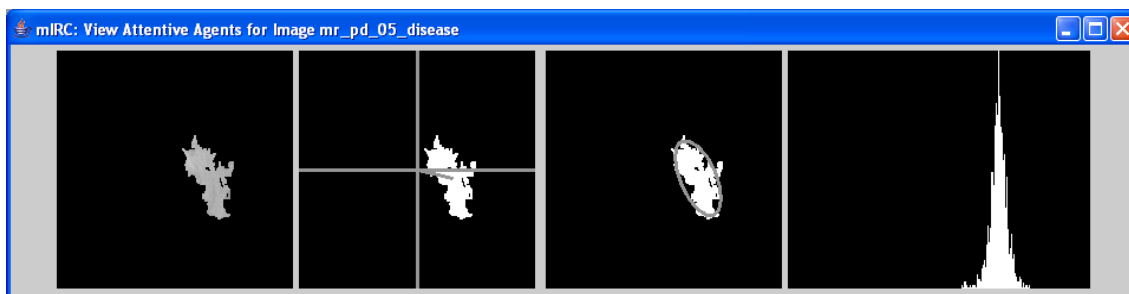
Οπτικοί Πράκτορες	Συνεισφορά (%)
Τοποθεσία	35
Μέγεθος	30
Ιστόγραμμα Φωτεινοτήτων	25
Σχήμα ως Έλλειψη	10

Όπως έχει ειπωθεί, πέρα από τα αποτελέσματα ανάκτησης γενικής και επιλεκτικής περιγραφής, η πλατφόρμα παρέχει υπηρεσίες σταθμισμένου συνδυασμού των αποτελεσμάτων από τα δύο αυτά επίπεδα. Η συνεισφορά του κάθε επιπέδου στο τελικό αποτέλεσμα καθορίζεται από τον πελάτη με βάση τις εκάστοτε προτιμήσεις του (.Εικόνα 56 )



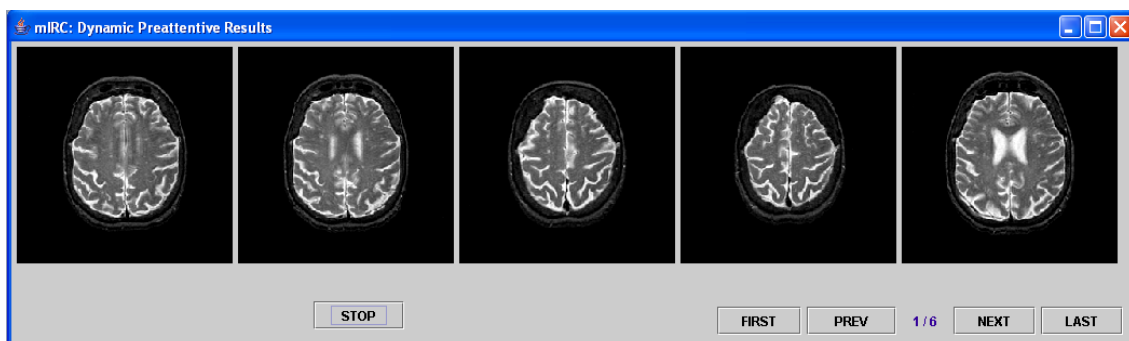
Εικόνα 56: Η διεπαφή του χρήστη της εφαρμογής ανάλυσης και ανάκτησης εικόνων

Η διεπαφή του πελάτη παρέχει την οπτική παρουσίαση της διαδικασίας ανάλυσης της εικόνας ερώτησης, δηλαδή την ευθυγράμμιση της εικόνας, την τμηματοποίηση της και την οπτική αναπαράσταση της πληροφορίας από κάθε οπτικό πράκτορα στα δύο επίπεδα επεξεργασίας. Για παράδειγμα, στην Εικόνα 57 παρουσιάζεται η πληροφορία από τους οπτικούς πράκτορες του επιλεκτικού επιπέδου για τη μοναδική περιοχή ενδιαφέροντος της εικόνας ερώτησης.



Εικόνα 57: Η πληροφορία των πρακτόρων στο επιλεκτικό επίπεδο ανάκτησης για μία περιοχή ενδιαφέροντος

Τα αποτελέσματα ανάκτησης παρουσιάζονται κατά φθίνουσα σειρά ομοιότητας σε δυναμικές διεπαφές, που ανανεώνονται όσο νέες εικόνες από τη βάση του εξυπηρέτη εξετάζονται. Κάθε επίπεδο παρουσιάζει τα αποτελέσματα ανάκτησης του σε ανεξάρτητες διεπαφές. Καθώς ο εξυπηρέτης τροφοδοτεί συνεχώς τις διεπαφές με εικόνες, ο χρήστης, οποιαδήποτε στιγμή το επιθυμεί, μπορεί να διακόψει τη διαδικασία. Ένα στιγμιότυπο της δυναμικής διεπαφής παρουσίασης των αποτελεσμάτων γενικής περιγραφής παρουσιάζεται στην Εικόνα 58.



Εικόνα 58: Η δυναμική διεπαφή παρουσίασης των αποτελεσμάτων ανάκτησης

Παρατηρείται πως σε αυτό το επίπεδο ανακτώνται εικόνες με παρόμοια ανατομική μορφολογία. Στην επόμενη παράγραφο παρουσιάζονται και σχολιάζονται περισσότερα αποτελέσματα ανακτήσεων για κάθε επίπεδο της προτεινόμενης αρχιτεκτονικής.

## 6.2 Αποτελέσματα Ανάκτησης

Σε αυτήν την παράγραφο παρουσιάζονται χαρακτηριστικά αποτελέσματα ανάκτησης από την προτεινόμενη πλατφόρμα για την εφαρμογή στις εγκάρσιες μαγνητικές τομογραφίες εγκεφάλου. Στο πρώιμο επίπεδο ανάκτησης στόχος είναι η ανάκτηση παρόμοιων ανατομικά-μορφολογικά τομών (παραπλήσιο slice) και παρόμοιων τεχνικών απεικόνισης (παραπλήσιο modality) του εγκεφάλου. Στο

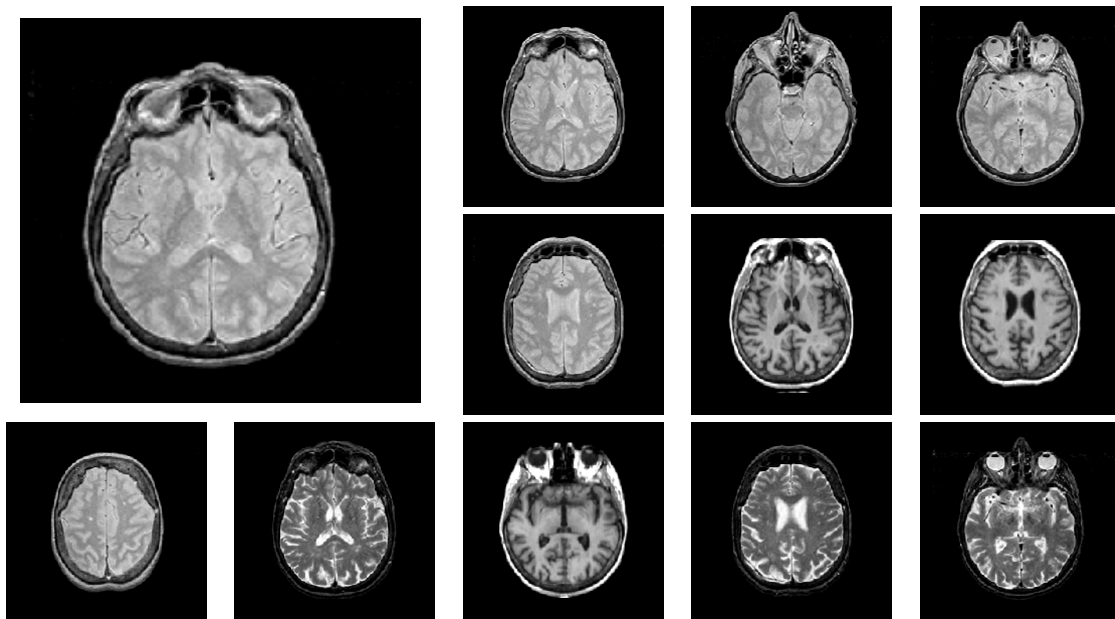
επιλεκτικό επίπεδο στόχος είναι η ανάκτηση εικόνων με παρόμοια ευρήματα παθολογίας. Καθώς, δεν ήταν διαθέσιμη εξειδικευμένη γνώση πάνω στις διάφορες ασθένειες, όπως επίσης και οι αλγόριθμοι που υλοποιήθηκαν (ΠΑΡ 5.5) δεν έλαβαν υπ' όψιν ιδιαίτερα χαρακτηριστικά τους, δεν είναι δυνατό να γίνει μία αξιολόγηση βασισμένη στην ιατρική γνώση. Κατά συνέπεια, η προσπάθεια αξιολόγησης της πλατφόρμας στο επιλεκτικό επίπεδο αφορά μόνο την οπτική ομοιότητα των περιοχών ενδιαφέροντος όπως την αντιλαμβάνεται ένας άνθρωπος της επεξεργασίας εικόνων. Η προσέγγιση αυτή, όπως ειπώθηκε στο παρουσιάζει σημαντικά πλεονεκτήματα καθώς τα γενικευμένα οπτικά χαρακτηριστικά επιτρέπουν την ανάκτηση περιστατικών που παρουσιάζουν γενικά στενή οπτική ομοιότητα, παρά τους εξειδικευμένες επιμέρους οπτικές ιδιότητες που χρησιμοποιούνται από τους ακτινολόγους για τον εντοπισμό και τη διάγνωσή τους.

Στα πειράματα που ακολουθούν, χρησιμοποιήθηκε η ακόλουθη συλλογή εικόνων ανάκτησης. Χρησιμοποιήθηκαν συνολικά 100 εικόνες, όλες δανεισμένες από το [31], από τις οποίες 67 παρουσιάζουν μορφώματα και 33 αποτελούν υγιείς απεικονίσεις. Όσον αφορά τις ασθενείς απεικονίσεις, αυτές πάρθηκαν από 8 διαφορετικά περιστατικά που παρουσιάζουν διάφορα είδη όγκων, όπως γλιώματα, γλιοβλαστώματα, σαρκώματα και μεταστατικό καρκίνο. Εικόνες από το ίδιο περιστατικό επιλέχθηκαν σε απομακρυσμένες τομές, ούτως ώστε να μην περιέχουν παραπλήσιες απεικονίσεις του ίδιου μορφώματος. Οι υγιείς απεικονίσεις πάρθηκαν από μία συγκεκριμένη εξέταση όπου ήταν διαθέσιμες 55 εικόνες ανά κύρια τεχνική απεικόνισης (pd, t<sub>1</sub>, t<sub>2</sub>). Αυτές οι εικόνες αποτελούν μία εγκάρσια δειγματοληψία του ανθρώπινου εγκεφάλου και όπως θα γίνει φανερό στη συνέχεια, είναι δυνατόν να χρησιμοποιηθούν για την κατηγοριοποίηση των εικόνων που εξετάζονται. Από αυτές, επιλέχθηκαν, για τους στόχους της εφαρμογής, 11 ισαπέχουσες, απομακρυσμένες τομές, ανά τεχνική, ούτως ώστε να μην υπάρχει πλήθος εικόνων με παραπλήσια μορφολογία του εγκεφάλου. Γενικότερα, όταν χρησιμοποιούνται τέτοια σύνολα εικόνων αναφοράς, πέρα από την απεικόνιση της γενικής μορφολογίας πρέπει να λαμβάνεται υπ' όψη ότι οι διάφορες εξετάσεις λαμβάνουν χώρα υπό διαφορετικές ραδιομετρικές συνθήκες. Κατά συνέπεια, είναι απαραίτητος ο καθορισμός ραδιομετρικών προτύπων για την επίτευξη καλύτερων αποτελεσμάτων. Αν αυτό δεν είναι εφικτό, τότε είναι απαραίτητη η χρήση πολλαπλών τέτοιων συνόλων ώστε να είναι διάφοροι διαθέσιμοι συνδυασμοί μορφολογίας και φωτεινότητας.

### **6.2.1 Αποτελέσματα με κριτήριο τη γενική μορφολογία**

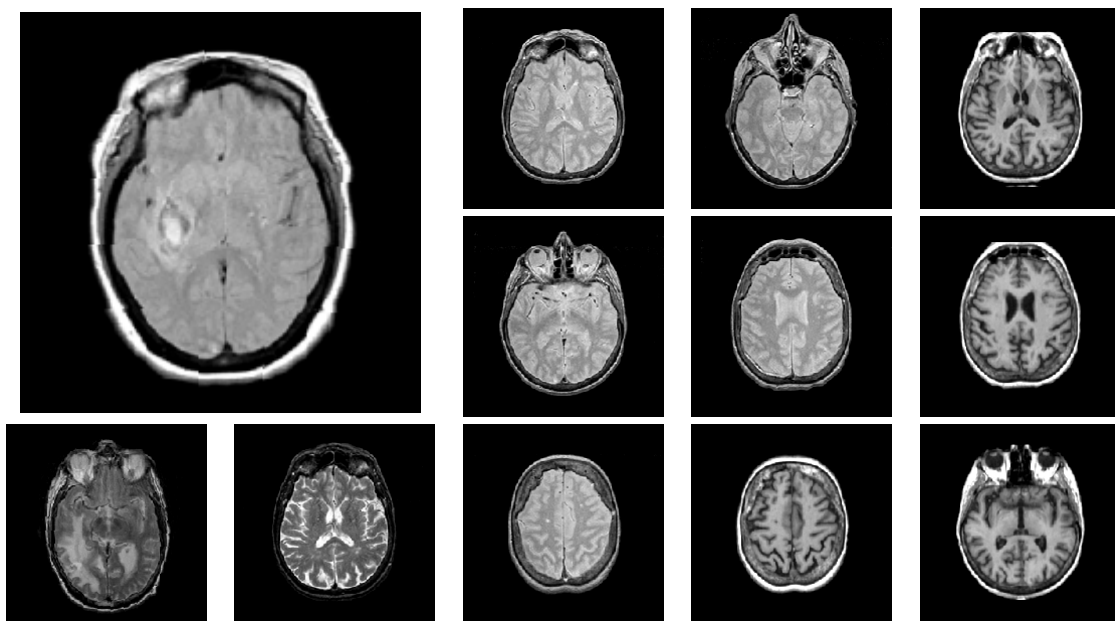
Στα πρώτα πειράματα για την αξιολόγηση του πρώιμου επιπέδου ανάκτησης εξετάστηκε η δυνατότητα χρήσης εικόνων αναφοράς για την κατηγοριοποίηση των μαγνητικών τομογραφιών του εγκεφάλου. Χρησιμοποιήθηκε ένα αραιό σύνολο εικόνων αναφοράς από 11 ισαπέχουσες τομές ανά τεχνική απεικόνισης. Στα αποτελέσματα που παρουσιάζονται στην Εικόνα 59, η εικόνα ερώτησης έχει επιλεγεί ανάμεσα από τις 55 pd διαθέσιμες απεικονίσεις του υγιούς περιστατικού ώστε να παρουσιάζει τα ίδια ραδιομετρικά κριτήρια με τις εικόνες αναφοράς. Ωστόσο δεν περιέχεται η ίδια (ως τομή) στο σύνολο των 33 εικόνων προς ανάκτηση. Παρατηρείται πως αν και χρησιμοποιήθηκε ένα αραιό σύνολο εικόνων η πλατφόρμα είναι κατάλληλη για την κατηγοριοποίηση της όσον αφορά τη μορφολογία και την τεχνική απεικόνισης.





**Εικόνα 59:** Αποτελέσματα πρώιμου επίπεδου ανάκτησης για υγιή rd MRI εγκεφάλου. Η εικόνα ερώτησης και το σύνολο των εικόνων ανάκτησης πληρούν τα ίδια ραδιομετρικά κριτήρια.

Περαιτέρω, σε ένα πιο σύνθετο εγχείρημα ανάκτησης στο πρώιμο επίπεδο, χρησιμοποιήθηκε μία εικόνα ερώτησης από διαφορετικό περιστατικό και υπό διαφορετικές συνθήκες λήψης (Εικόνα 60).

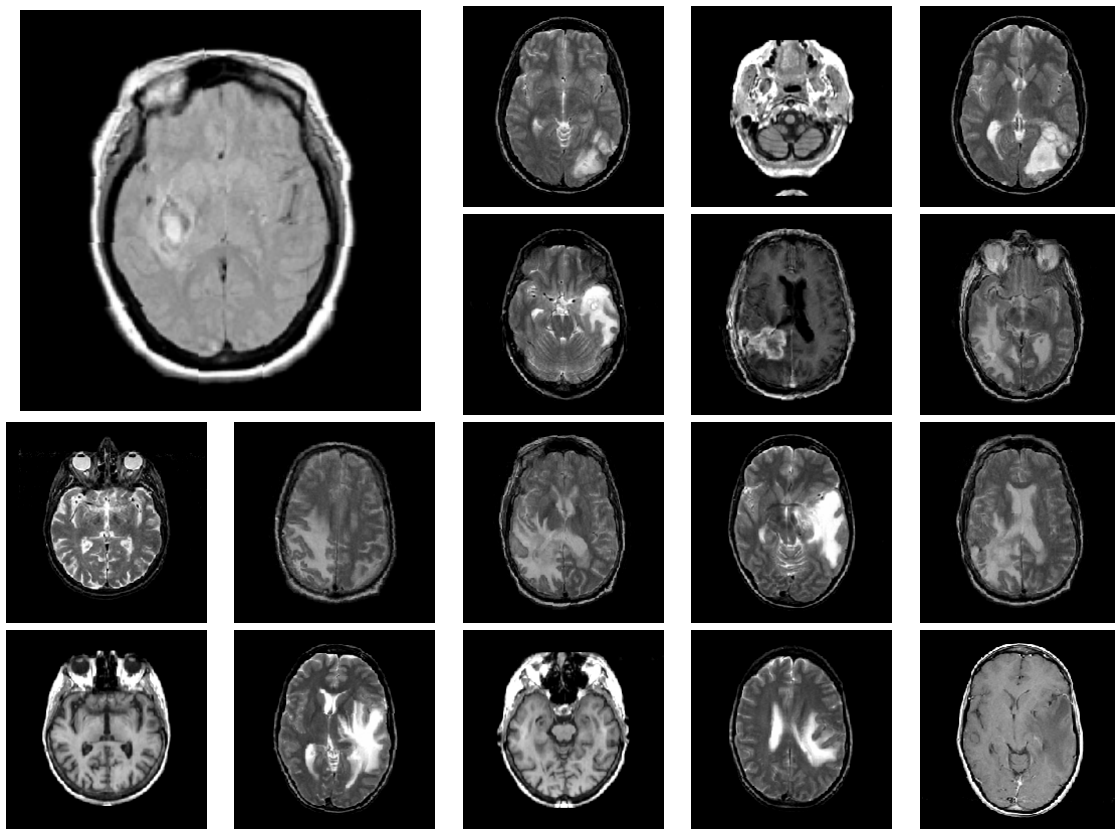


**Εικόνα 60:** Αποτελέσματα πρώιμου επίπεδου ανάκτησης για rd MRI εγκεφάλου με ευρήματα. Η εικόνα ερώτησης και το σύνολο των εικόνων ανάκτησης είναι ραδιομετρικά ανεξάρτητες.

Επιπλέον, σε σχέση με το προηγούμενο πείραμα, στο σύνολο των εικόνων αναφοράς έχουν προστεθεί και εικόνες με ευρήματα από διάφορες εξετάσεις (67 εικόνες με μορφώματα). Παρατηρείται, πως και σε αυτήν την περίπτωση, η πλατφόρμα ανάκτησης εικόνων είναι κατάλληλη για την κατηγοριοποίηση των εικόνων εισόδου.

## 6.2.2 Αποτελέσματα με κριτήριο τις περιοχές ενδιαφέροντος

Όπως ειπώθηκε, η περιγραφή των εικόνων στο επιλεκτικό επίπεδο περιέχει σε μεγάλο βαθμό τη σημασιολογία τους. Στη συγκεκριμένη εφαρμογή των MRIs εγκεφάλου οι περιοχές ενδιαφέροντος είναι πιθανές περιοχές πάθησης του απεικονιζόμενου περιστατικού. Η αποτελεσματική ανάκτηση παρόμοιων ασθενειών είναι εξαιρετικής σημασίας και θα μπορούσε να αποτελέσει ένα αυτόματο στατιστικό εργαλείο διάγνωσης. Ωστόσο στα αποτελέσματα που ακολουθούν δεν γίνεται προσπάθεια αξιολόγησης του αποτελέσματος για δύο κύριους λόγους. Αφ' ενός, δεν είναι διαθέσιμη η εξειδικευμένη ιατρική γνώση που απαιτείται, αφ' εταίρου, ένας τέτοιος στόχος κρίνεται μεγαλεπήβολος αναλογικά με την απλότητα των οπτικών πρακτόρων που χρησιμοποιήθηκαν για την περιγραφή των περιοχών ενδιαφέροντος. Επομένως, πέρα από κάποιες υποκειμενικές και διαισθητικές στατιστικές μετρήσεις που παρουσιάζονται, τα συμπεράσματα αφήνονται στη κρίση του αναγνώστη. Πρέπει επιπλέον να σημειωθεί πως τα αποτελέσματα της ανάκτησης παρόμοιων περιοχών ενδιαφέροντος εξαρτώνται σε μεγάλο βαθμό από το σύνολο των εικόνων της βάσης της πλατφόρμας καθώς είναι δυνατό να μην υπάρχουν παρόμοιες οπτικά απεικονίσεις.

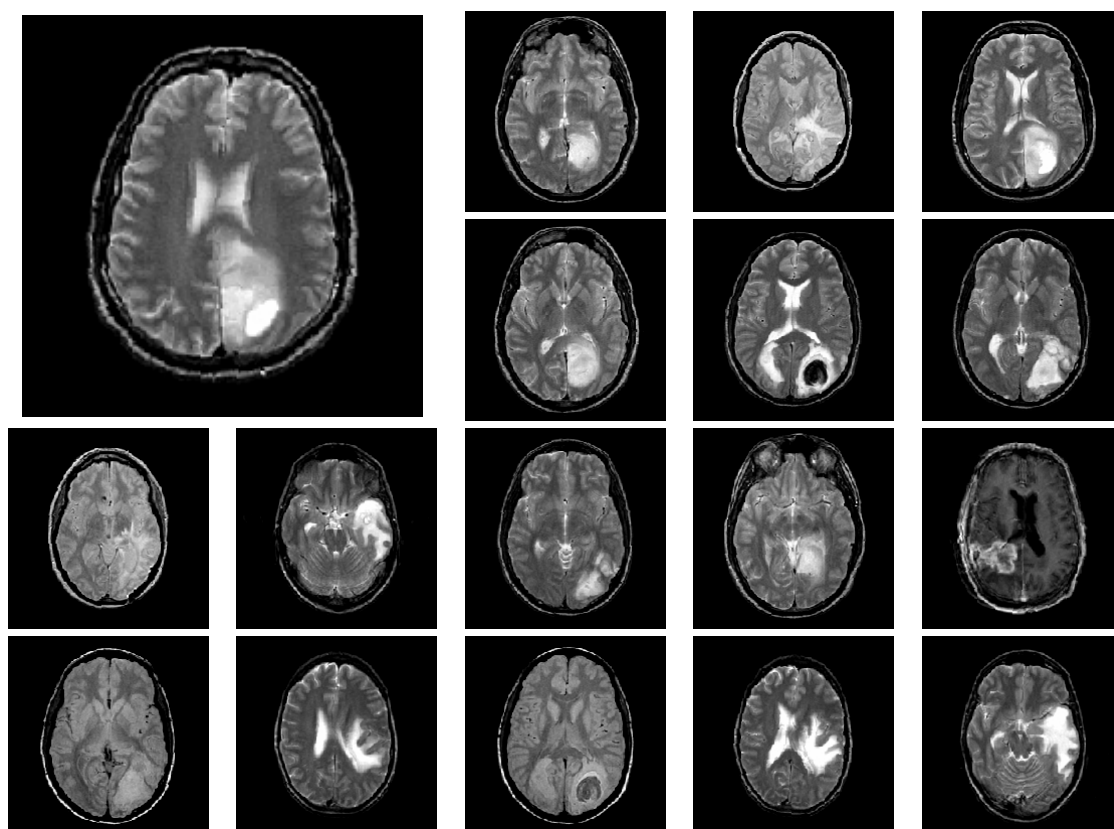


**Εικόνα 61:** Αποτελέσματα επιλεκτικού επίπεδου ανάκτησης για rd MRI εγκεφάλου με ευρήματα. Η εικόνα ερώτησης κατά την τμηματοποίηση εμφάνισε ένα ψευδές αρνητικό στην περιοχή του αρ. ματιού. Επιπρόσθετα, απουσίαζαν παρόμοιες εμφανίσεις περιοχών στη ΒΔ της πλατφόρμας με αποτέλεσμα να μην επιτευχθούν οι στόχοι της ανάκτησης

Στο πρώτο παράδειγμα ανάκτησης (Εικόνα 61), επιλέχθηκε μία εικόνα για την οποία δεν υπάρχουν αντίστοιχες απεικονίσεις στη ΒΔ σε επίπεδο περιοχών ενδιαφέροντος. Επιπρόσθετα, η εικόνα εμφάνισε ένα ψευδές θετικό εύρημα (false positive) στη περιοχή του αριστερού ματιού λόγω μικρής απόκλισης της γωνίας

λήψης από το εγκάρσιο επίπεδο (Εικόνα από 5<sup>ο</sup> ΚΕΦ false positive). Σύμφωνα με τα υποκειμενικά μας κριτήρια ομοιότητας δεν εμφανίστηκαν παρόμοιες εικόνες ανάμεσα στις 16 καλύτερες ανακτήσεις. Οι περιοχές ενδιαφέροντος που ανακτήθηκαν παρουσιάζουν ομοιότητα σε ανεξάρτητες και μεμονωμένες οπτικές ιδιότητες όπως μόνο στη φωτεινότητα, μόνο στο μέγεθος, ή μόνο στη τοποθεσία. Επιπρόσθετα λόγω του μικρού μεγέθους των περιοχών ενδιαφέροντος και της ασυμβατότητας τους με τις εμφανίσεις ασθενειών στη ΒΔ, ευνοήθηκαν τέσσερις εικόνες με ψευδή θετικά ( 2<sup>η</sup>, 7<sup>η</sup>, 12<sup>η</sup> και 14<sup>η</sup> εικόνα). Όπως θα γίνει φανερό στη συνέχεια, τα αποτελέσματα στο συνδυαστικό επίπεδο ανάκτησης (Εικόνα 63) συμπεριλαμβάνουν μία ενδιαφέρουσα εικόνα. Στην επόμενη παράγραφο γίνεται αιτιολόγηση αυτής της μη αναμενόμενης εμφάνισης.

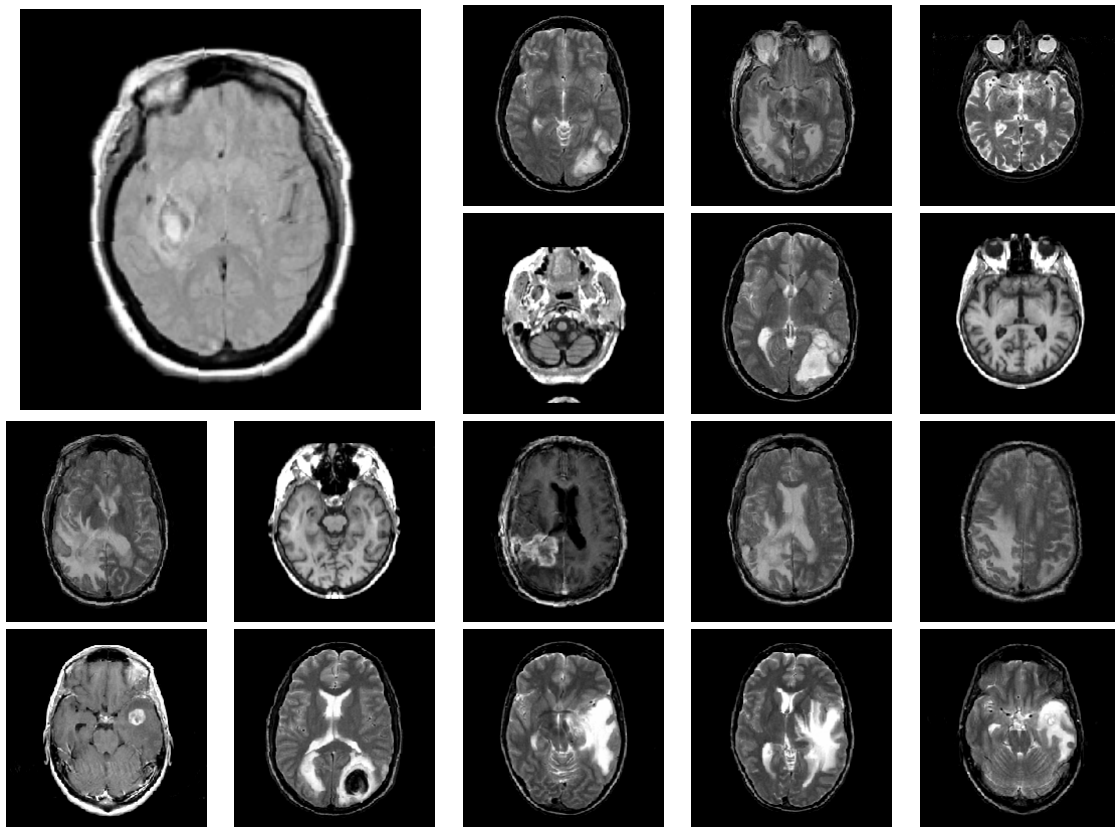
Στο δεύτερο παράδειγμα επιλέχθηκε μία εικόνα που παρουσιάζει ένα ευμεγέθες μόρφωμα για το οποίο υπάρχουν αρκετές παραπλήσιες εμφανίσεις στη ΒΔ. Λόγω του μεγέθους του μορφώματος είναι εξαιρετικά απίθανο να ανακληθούν εικόνες ψευδούς θετικού εντοπισμού. Όπως ειπώθηκε στην παράγραφο 5.4.2, αυτές με βάση τη χρησιμοποιούμενη μέθοδο είναι εν γένει πολύ μικρές. Τα αποτελέσματα ανάκτησης, παρουσιάζονται στην Εικόνα 62, και κρίνονται με βάση τα υποκειμενικά μας κριτήρια πολύ ικανοποιητικά. Αρκετές από τις περιοχές που ανακτήθηκαν παρουσιάζουν σημαντική ομοιότητα, όχι μόνο σε ένα, αλλά στο σύνολο των οπτικών χαρακτηριστικών του μορφώματος της εικόνας ερώτησης. Συνεπώς οι στόχοι της επιλεκτικής ανάκτησης έχουν επιτευχθεί.



**Εικόνα 62:** Αποτελέσματα επιλεκτικής ανάκτησης για  $t_2$  MRI εγκεφάλου με ευρήματα. Οι στόχοι της ανάκτησης έχουν επιτευχθεί.

### 6.2.3 Συνδυαστικά αποτελέσματα

Στα παρακάτω, συνδυαστικά, αποτελέσματα ανάκτησης η συνεισφορά των δύο επιπέδων περιγραφής είναι ισότιμη. Γίνεται κατά συνέπεια η υπόθεση ότι το ενδιαφέρον του χρήστη εντοπίζεται στην ανάκτηση παρόμοιων παθολογικά ευρημάτων σε παραπλήσιες περιοχές του εγκεφάλου. Η συνεισφορά του επιπέδου γενικής περιγραφής ευνοεί τόσο τις γειτονικές τομές όσο και την ίδια τεχνική απεικόνισης. Ο εντοπισμός ευρημάτων σε παραπλήσια περιοχή του εγκεφάλου επιτυγχάνεται και με τη συνεισφορά του επιλεκτικού επιπέδου όπου συνυπολογίζεται η τοποθεσία του ευρήματος στο επίπεδο της συγκεκριμένης τομής. Η πληροφορία αυτή είναι ουσιαστική, καθώς τα ευρήματα αποκτούν ειδική σημασία ανάλογα με την ανατομική περιοχή στην οποία εμφανίζονται. Για παράδειγμα, η εμφάνιση καρκινώματος κοντά στο οπτικό νεύρο μπορεί να συνδυαστεί με την πληροφορία από άλλες εξετάσεις που αφορούν προβλήματα όρασης του ασθενούς.

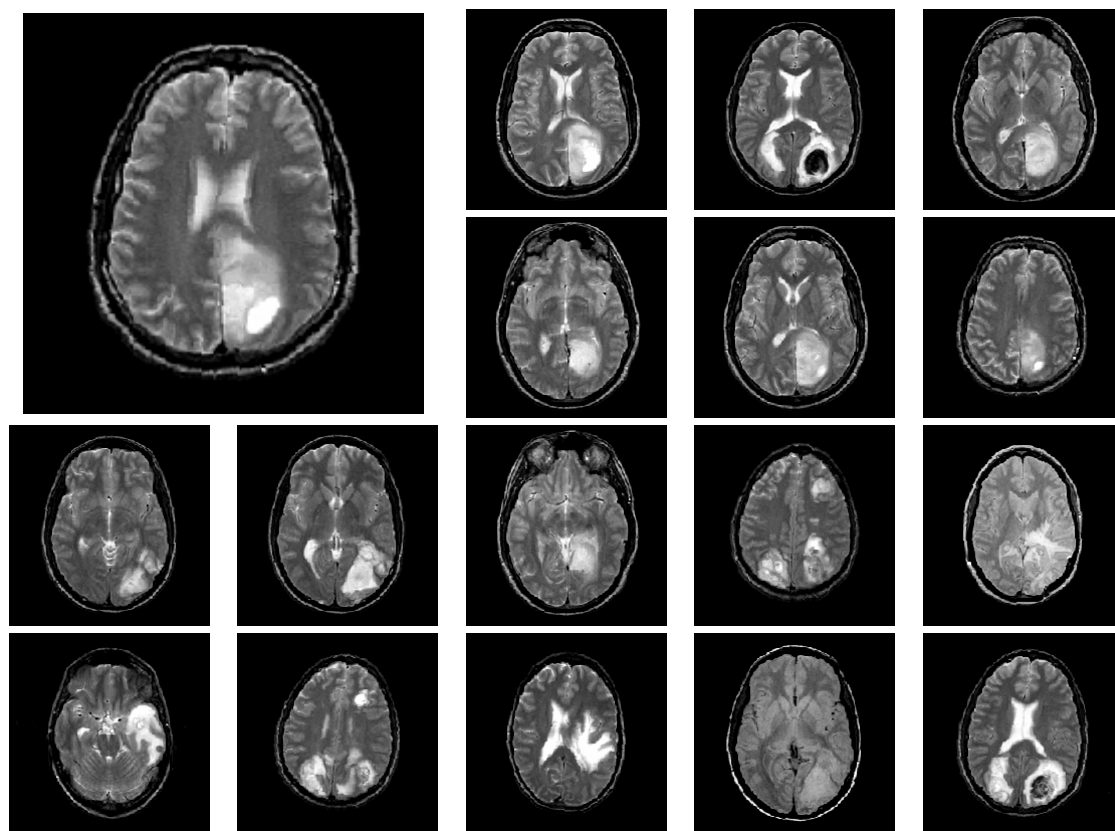


Εικόνα 63: Αποτελέσματα συνδυαστικής ανάκτησης για rd MRI εγκεφάλου με ευρήματα. Σε αντίθεση με τα αποτελέσματα του επιλεκτικού επιπέδου, η εικόνα στη θέση 12 κρίνεται ενδιαφέρουσα.

Σαν πρώτο πείραμα επιλέγεται η εικόνα ερώτησης που χρησιμοποιήθηκε για την επίδειξη της καλής λειτουργίας του πρώιμου επιπέδου. Η ίδια εικόνα στο επιλεκτικό επίπεδο δεν συνάντησε παρόμοιες περιοχές ενδιαφέροντος. Ωστόσο στο συνδυαστικό επίπεδο (Εικόνα 63) ανακλήθηκε μία ενδιαφέρουσα εικόνα στη 12<sup>η</sup> καλύτερη θέση. Η εικόνα αυτή περιέχει μία περιοχή σε παραπλήσιο μέγεθος σχήμα και φωτεινότητα με την αντίστοιχη περιοχή της εικόνας ερώτησης. Επιπλέον, η περιοχή εμφανίζεται σε συμμετρικό σημείο ως προς την εικόνα ερώτησης γεγονός που μπορεί να εμπεριέχει ουσιαστική πληροφορία λόγω της συμμετρικότητας του ανθρώπινου εγκεφάλου. Παρουσιάζει ωστόσο ενδιαφέρον να εξεταστούν οι λόγοι για

τους οποίους η εικόνα αυτή δεν τοποθετήθηκε ομοίως σε καλή σειρά κατάταξης από το επιλεκτικό επίπεδο. Ένας σαφής λόγος είναι η γενική ομοιότητα των δύο εικόνων σε επίπεδο μορφολογίας και απεικονιστικής τεχνικής. Επιπρόσθετα, η υλοποίηση του πρώιμου τοπογραφικού χάρτη ασυμμετρίας ευνοεί κατά την ανάκτηση τέτοιες περιοχές. Επεξηγηματικά, κατά τη πρώιμη ανάκτηση των εικόνων, ο οπτικός πράκτορας που διαχειρίζεται αυτή τη πληροφορία, μετρά το πλήθος των επικαλυπτόμενων εικονοστοιχείων των τοπογραφικών χαρτών από τις επιμέρους εικόνες (παρ. 5.3.4). Αυτοί οι χάρτες περιέχουν ζεύγη αντισυμμετρικών περιοχών που διαταράσσουν τη συμμετρία της εικόνας. Κατά συνέπεια, δεν ευνοούν μόνο τις περιοχές που προκάλεσαν με την μη αναμενόμενη ύπαρξή τους την ασυμμετρία, αλλά και τις φυσιολογικές αντισυμμετρικές τους. Στον αντίποδα, στο επιλεκτικό επίπεδο, ο οπτικός πράκτορας που διαχειρίζεται την πληροφορία της σχετικής θέσης των περιοχών ενδιαφέροντος μετρά την απόσταση των κέντρων των δύο περιοχών. Στην περίπτωση δύο αντισυμμετρικών περιοχών δεν υπάρχει καμία εγγύηση για την εγγύτητά τους. Από αυτό το χαρακτηριστικό παράδειγμα αποδεικνύεται η σημασία της επιλογής των κατάλληλων οπτικών πρακτόρων για την εξαγωγή και σύγκριση της πληροφορίας στις εξειδικευμένες εφαρμογές. Στην παρούσα εφαρμογή, αν τονιζόταν (πιθανότατα από ιατρούς) η σημασιολογική αξία της εμφάνισης ασθενειών σε αντισυμμετρικές περιοχές, τότε ίσως θα ήταν χρήσιμο να προσαρτηθούν κατάλληλοι πράκτορες στο επίπεδο επιλεκτικής ανάκτησης.

Στην Εικόνα 64 παρουσιάζονται τα αποτελέσματα της συνδυαστικής ανάκτησης για την  $t_2$  μαγνητική τομογραφία που χρησιμοποιήθηκε ως παράδειγμα επιτυχημένης ανάκτησης στο επιλεκτικό επίπεδο.

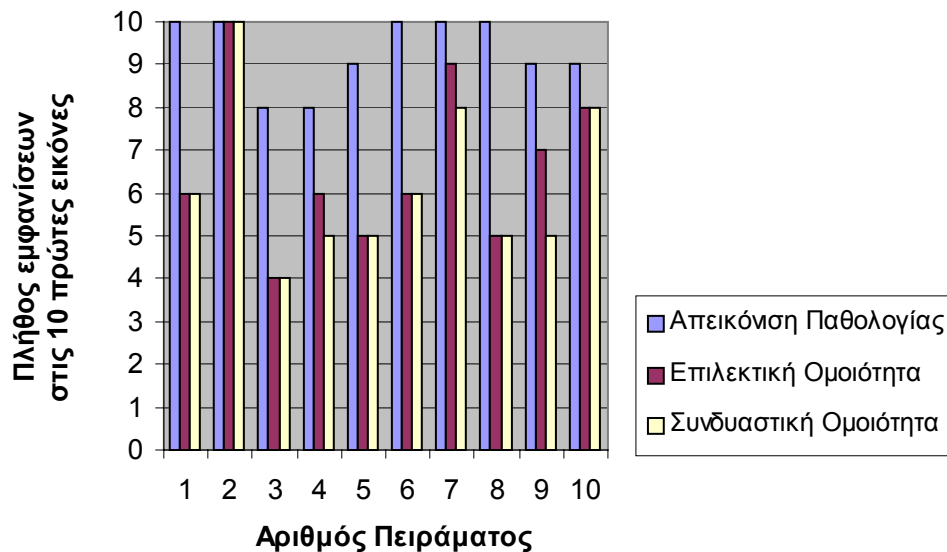


**Εικόνα 64:** Αποτελέσματα συνδυαστικής ανάκτησης για  $t_2$  MRI εγκεφάλου με ευρήματα. Οι στόχοι της ανάκτησης έχουν επιτευχθεί καθώς οι περισσότερες εικόνες παρουσιάζουν πολύ καλή ομοιότητα τόσο σε επίπεδο μορφολογίας όσο και σε επίπεδο περιοχών ενδιαφέροντος.

Τα αποτελέσματα υποδεικνύουν πως το συνδυαστικό επίπεδο παρέχει καλύτερη σημασιολογική περιγραφή των εικόνων συγκρινόμενο μετα δύο ανεξάρτητα επίπεδα.

Σε αυτό το σημείο παρουσιάζεται ένα διάγραμμα που απεικονίζει ορισμένα σημαντικά μεγέθη ανάκτησης από δέκα τυχαία πειράματα συνδυαστικής ανάκτησης (Εικόνα 65). Η συνεισφορά των δύο επιπέδων περιγραφής στο τελικό αποτέλεσμα καθορίστηκε ως ισότιμη. Οι εικόνες ερώτησης που χρησιμοποιήθηκαν εμφάνιζαν όλες παθολογικά ευρήματα. Για την αξιοπιστία των μετρήσεων δεν χρησιμοποιήθηκαν εικόνες από την εξέταση του ίδιου περιστατικού. Τα μεγέθη υπολογίστηκαν σε κάθε πείραμα είναι τα ακόλουθα:

- το πλήθος των εικόνων που παρουσιάζουν παθολογικά ευρήματα στις 10 καλύτερες ανακτήσεις. Δηλαδή εξετάστηκε η ικανότητα του συστήματα να επιστρέφει ασθενή περιστατικά σε ασθενή εικόνα εισόδου
- το πλήθος των εικόνων ανάμεσα στις 10 καλύτερες εμφανίσεις που παρουσιάζουν οπτική ομοιότητα στις περιοχές ενδιαφέροντος. Αυτό το μέτρο ουσιαστικά αντικατοπτρίζει την «ακρίβεια» του επιλεκτικού επιπέδου όπως αυτή εκφράστηκε στην εξίσωση (1).
- το πλήθος των εικόνων ανάμεσα στις 10 καλύτερες εμφανίσεις που παρουσιάζουν συνδυαστική ομοιότητα των περιοχών ενδιαφέροντος και της απεικονιζόμενης μορφολογίας. Αυτό το μέτρο ουσιαστικά αντικατοπτρίζει την «ακρίβεια» του συνδυαστικού επιπέδου όπως αυτή εκφράστηκε στην εξίσωση (1).



**Εικόνα 65:** Διάγραμμα απεικόνισης ορισμένων ουσιαστικών μετρήσεων επίδοσης της προτεινόμενης πλατφόρμας ανάκτησης στο συνδυαστικό επίπεδο. Οι μετρήσεις ομοιότητας ωστόσο βασίστηκαν σε οπτικά και όχι σημασιολογικά κριτήρια.

Από τις παραπάνω μετρήσεις εξάγονται τα ακόλουθα συμπεράσματα όσον αφορά την αποτελεσματικότητα του συνδυαστικού επιπέδου ανάκτησης:

- όταν οι εικόνες ερώτησης απεικονίζουν περιστατικά με ευρήματα, ευνοείται σε εξαιρετικό βαθμό η ανάκτηση ασθενών περιστατικών
- οι περισσότερες από τις ανακτώμενες εικόνες εμφανίζουν παρόμοιες οπτικά παθολογικές περιοχές με την εικόνα ερώτησης

- στο σύνολο σχεδόν των αποτελεσμάτων, η επιλεκτική ομοιότητα των ευνοημένων εικόνων συνδυάζεται και με τη γενική- μορφολογική ομοιότητα. Το γεγονός αυτό επιδεικνύει την αποτελεσματικότητα του πρώιμου επίπεδο επιπέδου στην ανάκτηση παρόμοιων ανατομικά απεικονίσεων του εγκεφάλου.





## 7 Συμπεράσματα και Επίλογος

Σε τούτη την εργασία έγινε μία προσπάθεια περιγραφής και ανάκτησης του σημασιολογικού περιεχομένου των εικόνων σε εξειδικευμένες εφαρμογές. Η αρχιτεκτονική που παρουσιάστηκε προτείνει τη χρήση δύο επιπέδων περιγραφής και ανάκτησης. Το πρώτο επίπεδο επιτελεί τη γενική περιγραφή των εικόνων ενώ το δεύτερο περιγράφει επιλεκτικά ένα σύνολο από περιοχές ενδιαφέροντος της εξειδικευμένης εφαρμογής. Η ανάκτηση μπορεί να διεξαχθεί στα δύο επίπεδα ανεξάρτητα, ενώ παρέχεται και η δυνατότητα σταθμισμένου συνδυασμού των δύο αποτελεσμάτων. Η πρόταση αυτή βασίσθηκε στη γνώση από τα δεδομένα της ψυχολογίας για την οπτική αντίληψη. Η αδυναμία της αρχιτεκτονικής εντοπίζεται στην υπόθεση της καλής τμηματοποίησης. Ωστόσο, η χρήση ευφών μεθόδων σε ένα δυναμικό περιβάλλον ανάκτησης εικόνων και αλληλεπίδρασης με τους χρήστες διευκολύνει τους στόχους της τμηματοποίησης. Επιπρόσθετα, η γνώση από το εξειδικευμένο πεδίο εφαρμογής είναι δυνατό να περιορίσει το χώρο της αναζήτησης των περιοχών ενδιαφέροντος. Στην παρούσα εργασία έγινε μία προσπάθεια εκμετάλλευσης αυτών των δυνατοτήτων με τον καθορισμό ενός δυναμικού περιβάλλοντος εικόνων και τη χρήση μέρους της πληροφορίας από τη θεωρία των αυτόνομων πρακτόρων. Περαιτέρω έγινε αναφορά στις διάφορες συμπληρωματικές τεχνικές που μπορούν να ενσωματωθούν στην αρχιτεκτονική σε μελλοντικές επεκτάσεις της.

Τα αποτελέσματα της εφαρμογής στις μαγνητικές τομογραφίες εγκεφάλου είναι ενθαρρυντικά καθώς επιτεύχθηκε μία καλή σημασιολογική περιγραφή τους. Το πρώτο επίπεδο της γενικής περιγραφής (πρώιμο επίπεδο της αρχιτεκτονικής) διαχειρίζεται θέματα ομοιότητας των τομογραφιών σε επίπεδο μορφολογίας του εγκεφάλου. Κατά συνέπεια είναι δυνατό να διακρίνει τις εικόνες ανάλογα με την απεικονιζόμενη τομή αλλά και την τεχνική απεικόνισης που χρησιμοποιήθηκε για τη λήψη της. Το δεύτερο επίπεδο (επιλεκτικό επίπεδο της αρχιτεκτονικής) διαχειρίζεται την πληροφορία των εντοπισμένων περιοχών, που στη συγκεκριμένη εφαρμογή αποτελούν τα πιθανά παθολογικά ευρήματα. Περαιτέρω, η δυνατότητα του συνδυαστικού επιπέδου ανάκτησης επιτρέπει την ανάκτηση παρόμοιων ευρημάτων σε παρόμοια τοποθεσία στον εγκέφαλο. Σημαντική προσπάθεια έγινε για τον αυτόματο εντοπισμό των περιοχών παθολογίας στις μαγνητικές τομογραφίες εγκεφάλου. Γι' αυτό το σκοπό χρησιμοποιήθηκε η πρότερη γνώση της καλής συμμετρίας του υγιούς εγκεφάλου. Ο τοπογραφικός χάρτης ασυμμετρίας, που προτάθηκε αποτελεί ένα εξαιρετικό μέτρο της ασυμμετρίας και μπορεί να χρησιμοποιηθεί με ποικίλους τρόπους στον εντοπισμό των περιοχών παθολογίας.

Η επιτυχημένη εφαρμογή στις μαγνητικές τομογραφίες εγκεφάλου είναι ενθαρρυντική, ωστόσο απαιτείται η εφαρμογή της προτεινόμενης αρχιτεκτονικής σε περαιτέρω πεδία εικόνων. Με αυτόν τον τρόπο θα γίνει σαφές πως μέσα σε ένα δυναμικό περιβάλλον εισαγωγής και ανάκτησης εικόνων ο αποτελεσματικός εντοπισμός των περιοχών ενδιαφέροντος δεν είναι ένα ουτοπικό ζήτημα. Άλλωστε ο απώτερος στόχος της παρούσας εργασίας δεν είναι μόνο η μελέτη των ευφών μηχανισμών της βιολογικής όρασης, αλλά η δημιουργία μίας πλατφόρμας ανάκτησης εικόνων που θα είναι ικανή να ενταχθεί στο ολοκληρωμένο σύγχρονο ιατρικό περιβάλλον. Τα οφέλη από μία τέτοια προσέγγιση είναι εξαιρετικής σημασίας στον τομέα της ερευνητικής, εκπαιδευτικής και διαγνωστικής ιατρικής.



## 8 Βιβλιογραφία

- [1] Σωκράτης Δημητριάδης, «Περιβάλλον Ανάκτησης Εικόνων Βασισμένο σε Βιολογικά Εμπνευσμένη Αρχιτεκτονική», Μεταπτυχιακή Εργασία, Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης, 2002
- [2] Stan Franklin, “*Is it an agent, or just a program. A taxonomy for Autonomous Agents*” Proceedings of the Third International Workshop on Agent Theories, Architectures and Languages, Springer-Verlag, 1996
- [3] Pattie Maes, “*Modeling Adaptive Autonomous Agents*”, Artificial Life Journal, edited by C. Langton, Vol. 1, No. 1 & 2, pp. 135-162, MIT Press, 1994
- [4] R. Pfeifer and C. Scheier, “*Understanding Intelligence*”, MIT Press 1999
- [5] Βλαχαβάς, Κεφαλάς, Βασιλειάδης, Ρεφανίδης, Κόκκορας, Σακελαρίου, «Τεχνητή Νοημοσύνη», Εκδόσεις Γαρταγάνη 2002
- [6] Nicholas R. Jennings, “On agent-based software engineering” Artificial Intelligence 117 (2000) pp. 277-296
- [7] Πλαϊνής Βεμμο “*Η νευροφυσιολογία του αμφιβληστροειδή και της οπτικής οδού*”
- [8] Daniel Chandler, “*Visual Perception*”, UWA, 1997, [www.aber.ac.uk/media/Modules/MC10220/visindex.html](http://www.aber.ac.uk/media/Modules/MC10220/visindex.html)
- [9] Chang, Beitaο Li, Chen Li, “*Towards Perception Based Image Retrieval*”, 2000
- [10] Irving Biederman, “*Recognition By Components: A Theory of Human Image Understanding*”, Psychological Review 1987, Vol 94, No 2, pp. 115-147
- [11] Steven M. Boker, “*The Represent of Color. Metrics and Mapings in perceptual color Space*” Jan 1997
- [12] Tagare, Toyama, Wang, “*A maximum Likelihood Strategy for Directing Attention during Visual Search*”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, NO 5, May 2001
- [13] Sarkar, Boyer, “*Perceptual Organization in Computer Vision: A Review and a Proposal for a Classificatory Structure*” IEEE Transactions on Man and Cybernetics, Vol. 23, NO 2, March/April 1993
- [14] Anne Treisman, “*Features and Objects in Visual Processing*” Scientific American Nov. 1986
- [15] Michael J. Tarr, “*Visual Representation: From Features To Objects*”, Encyclopedia of Human Behavior, V.S Ramachandran, (ed), San Diego, CA: Academic Press, Vol 4, pp 503-512
- [16] Remco C. Veltkamp, Mirela Tanase and Danielle Sent, “*Features in Content Based Image retrieval Systems: A Survey*”, Kluwer Academic Publishers, 2001
- [17] Henning Muller, Nicolas Michoux, David Bandon and Antoine Geissbuhler, “*A Review of Content-Based Image Retrieval Systems in Medical Applications - Clinical Benefits and Future Directions*”, Division for Medical Informatics, University Hospital of Geneva
- [18] S. Dimitriadis, K. Marias and S.C. Orphanoudakis, "" In press

- [19] S. Dimitriadis, K. Marias and S.C. Orphanoudakis, “*A Versatile Image Retrieval Platform based on a Multiagent Architecture*”, 6th International Conference on Visual Information Systems, Miami, 2003
- [20] Thomas M Lehmann, Berthold B Wein, Hayit GreenSpan, “*Integration of Content-based Image Retrieval Systems to Picture Archiving and Communication Systems*”, Dep. of Medical Informatics, Dep. of Diagnostic Radiology, RWTH and Dep. of Biomedical Engineering, TAU, 2003
- [21] Joshi, Lorenzen, Gerig, Bullit, “*Structural and Radiometric Asymmetry in Brain Images*”, *Medical Image Analysis* 7 (2003) 155-170
- [22] Yanxi Liu and Frank Dellaert, “*Classification Driven Medical Image Retrieval*”, The Robotics Institute, CMU, 1998
- [23] Yanxi Liu, Frank Dellaert and William E. Rothfusa, “*Classification Driven Semantic Based Medical Image Indexing and Retrieval*”, The Robotics Institute, CMU, 1998
- [24] Punam K.Saha, Jayarem K. Udupa, “*Optimum Image Thresholding via Class Uncertainty and Region Homogeneity*”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, VOL. 23, NO. 7, July 2001
- [25] B. S. Manjunath and W. Y. Ma, “*Texture Features for Browsing and Retrieval of Image Data*”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.18, No.8, August 1996
- [26] J.D. Daugman, “*Complete discrete 2-d Gabor transforms by neural networks for image analysis and compression*”, in *IEEE Trans. Acoustics, Speech, and Signal Processing*, 36:1169-1179, 1988
- [27] Wells, W., Viola, P., and Kikinis, R., “*Multi-modal volume registration by maximization of mutual information*”, in *Proceedings of the Second International Symposium on Medical Robotics and Computer Assisted Surgery*, pages 55-62. Wiley, 1995
- [28] S. Ullman. “*The interpretation of Visual Motion*”. MIT Press, Cambridge, MA, 1979.
- [29] D. Marr and T. Poggio. “*A computational theory of human stereo vision.*” *Proc.Royal Society London*, B 204:301 & 328, 1979.
- [30] Jain, A.K., and Dubes, R.C., “*Algorithms for Clustering Data*”, Prentice Hall: New Jersey, 1988
- [31] “The Whole Brain Atlas” [www.harvard.edu](http://www.harvard.edu)