

ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ ΔΥΤΙΚΗΣ ΕΛΛΑΔΑΣ

ΣΧΟΛΗ: Διοίκησης & Οικονομίας

ΤΜΗΜΑ: Διοίκησης Επιχειρήσεων

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

**«Εφαρμογή Τεχνητών Νευρωνικών Δικτύων στην  
Αναγνώριση Προσώπων»**

**FACE RECOGNITION  
TECHNOLOGY**



**Πτυχιακή Εργασία των  
Αρμπούτης Χρήστος  
Ντέντη Αικατερίνη**

**Επιβλέπων : Πιερρακέας Χρήστος**

**ΠΑΤΡΑ, Μάιος 2016**

## ΠΕΡΙΛΗΨΗ

Η Αναγνώριση Προσώπου (Face Recognition) έχει την τελευταία δεκαετία εξελιχθεί σε μία από τις σημαντικότερες εφαρμογές της Επεξεργασίας Εικόνας (Image Processing), εξαιτίας της ανάπτυξης και της διαθεσιμότητας των αντίστοιχων τεχνολογιών, αλλά και των απαιτήσεων για ασφάλεια και επιβολή των νόμων. Συνεπώς, σήμερα αναπτύσσονται και εξαπλώνονται γοργά πολλές εμπορικές εφαρμογές στο πεδίο αυτό. Το πρόβλημα της Αναγνώρισης Προσώπου δεν σχετίζεται μόνο με την ταυτοποίηση των χαρακτηριστικών ενός προσώπου, αλλά περιλαμβάνει την ανίχνευση ενός προσώπου μέσα σε μία ψηφιακή εικόνα ή σε ένα πλαίσιο ενός βίντεο, την εξαγωγή των χαρακτηριστικών του και την μετέπειτα αναγνώριση της ταυτότητας του προσώπου. Μολονότι πολλές αξιόπιστες μέθοδοι αναγνώρισης βιομετρικών ανθρώπινων χαρακτηριστικών υπάρχουν, όπως η ανάλυση αποτυπωμάτων, η σάρωση της ίριδας κ.α., αυτές οι μέθοδοι απαιτούν την συνεργασία των ανθρώπων. Αντίθετα, ένα σύστημα αναγνώρισης προσώπων βασίζεται στην ανάλυση εικόνων προσώπων εμπρόσθιων ή/και κατατομής χωρίς την συνεργασία ή την ενημέρωση των ανθρώπων των οποίων τα πρόσωπα απεικονίζονται στις εικόνες.

Η ανάπτυξη ενός συστήματος αναγνώρισης προσώπου είναι αρκετά δύσκολη και επίπονη διαδικασία, εξαιτίας του ότι τα ανθρώπινα πρόσωπα είναι αρκετά πολύπλοκα, πολυδιάστατα και υπόκεινται σε αλλαγές του περιβάλλοντος. Επιπρόσθετα, οι αλλαγές στα χαρακτηριστικά του προσώπου, και οι ποικίλες απεικονίσεις του ίδιου προσώπου μεταξύ των εικόνων εξαιτίας διαφορετικής γωνίας λήψης της φωτογραφίας, του φωτισμού και λοιπά, καταστούν το πρόβλημα του αυτόματου εντοπισμού ανθρώπων μέσω της αναγνώρισης ανθρώπινων προσώπων μία τεχνολογική πρόκληση. Ενδεικτικά, ορισμένα από τα κύρια ερευνητικά ζητήματα αφορούν τον τρόπο που τα χαρακτηριστικά υιοθετούνται για να περιγράψουν ένα πρόσωπο κάτω από περιβαλλοντικές αλλαγές και στο πως θα ταξινομηθούν με βάση μια νέα εικόνα ενός προσώπου στην επιλεγμένη αναπαράσταση. Στόχος της παρούσας διπλωματικής εργασίας είναι αφενός μεν η κατανόηση και η όσο το δυνατό ενδελεχής περιγραφή του προβλήματος της Αναγνώρισης Προσώπου, καθώς και η παρουσίαση των έως σήμερα προτεινόμενων μεθόδων, αφετέρου δε ο τρόπος εφαρμογής των Τεχνητών Νευρωνικών Δικτύων τονίζοντας τη συμβολή τους στην επίλυση του εν λόγω προβλήματος.

# Πίνακας Περιεχομένων

<b>ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ .....</b>	<b>2</b>
<b>ΛΙΣΤΑ ΕΙΚΟΝΩΝ.....</b>	<b>5</b>
<b>ΛΙΣΤΑ ΠΙΝΑΚΩΝ .....</b>	<b>7</b>
<b>ΚΕΦΑΛΑΙΟ 1<sup>ο</sup>.....</b>	<b>8</b>
1.1.Περιγραφή του προβλήματος .....	8
1.2 Υπόβαθρο .....	12
1.3 Δήλωση του Προβλήματος .....	12
1.4 Σκοπός της Διπλωματικής Εργασίας .....	12
1.5 Στόχοι της Διπλωματικής Εργασίας .....	13
1.6 Επισκόπηση της Διπλωματικής Εργασίας .....	13
<b>ΚΕΦΑΛΑΙΟ 2<sup>ο</sup>.....</b>	<b>14</b>
2.1Εισαγωγή .....	14
2.2Η Βιολογία της Αναγνώρισης Προσώπου .....	16
2.3Η Τεχνολογία της Αναγνώρισης Προσώπου .....	18
2.3.1 Ταξινόμηση .....	18
2.3.2Εφαρμογές της Αναγνώρισης Προσώπου .....	22
2.4 Εξαγωγή Χαρακτηριστικών και Συμπύεση Δεδομένων .....	23
2.4.1 Ανάλυση Κυρίων Συνιστωσών (Principal Components Analysis - PCA).....	23
2.5. Υφιστάμενες Τεχνικές.....	24
2.5.1.Ιδιοδιανύσματα πρόσωπων (Eigenfaces) .....	24
2.5.2.Απεικόνιση Εύρους (Range Imaging) .....	25
2.5.3. Χαρτογράφηση των άκρων της γραμμής (Line edge map) .....	26
2.5.4.Τεχνικές Αναγνώριση Προσώπου με βάση τα Νευρωνικά Δίκτυα (Neural Network based Face Recognition Techniques).....	26
2.5.5.Δίκτυα Κυμάτων Gabor (Gabor Wavelet Networks - GWN) .....	26
2.5.6.Αραιά Εκπροσώπηση (Sparse Representation) .....	27

2.5.7. Turk και Pentland από το Εργαστήριο Μέσων Πληροφόρησης στο MIT (Massachusetts Institute of Technology).....	28
2.5.8. Αναγνώριση Προσώπου στο Ινστιτούτο Τεχνολογίας της Πολεμικής Αεροπορίας στο AFIT (AirForce Institute of Technology).....	28
2.5.9. Ο Cottrell στο UCSD.....	29
2.5.10.Εθνικό Εργαστήριο Los Alamos.....	29
<b>2.6.Εξελικτική Αναζήτηση (Evolutionary Pursuit) .....</b>	<b>30</b>
<b>2.7.Αντιπαράβολή Γραφικής Παράστασης Ελαστικής Δέσμης (Elastic Bunch Graph Matching) .....</b>	<b>30</b>
<b>2.8.Μέθοδοι Πυρήνα(KernelMethods).....</b>	<b>31</b>
<b>2.9. Μετασχηματισμός Ίχνους (TraceTransform) .....</b>	<b>31</b>
<b>2.10.Πρόσθετες Έρευνες Δικτύου (Additional Research Network).....</b>	<b>32</b>
<b>2.11.Συμπεράσματα .....</b>	<b>32</b>
<b>ΚΕΦΑΛΑΙΟ 3<sup>ο</sup>.....</b>	<b>34</b>
<b>3.1 Εισαγωγή .....</b>	<b>34</b>
<b>3.2 Πολυστρωματικό Νευρωνικό Δίκτυο (Multilayer Perceptron - MLP).....</b>	<b>36</b>
<b>3.3 Προσδιορισμός του Αλγόριθμου Εσφαλμένης Ανάστροφης Διάδοσης .....</b>	<b>38</b>
<b>3.4 Συνάρτηση ενεργοποίησης (Activation function).....</b>	<b>40</b>
<b>3.5 Λειτουργίες της Εκπαίδευσης (Modes of Learning) .....</b>	<b>43</b>
<b>3.6 Τοπικό Ελάχιστο της Επιφάνειας Σφάλματος (Local Minima of Error Surface) .....</b>	<b>44</b>
<b>3.7 Μηχανισμοί Διανυσμάτων Υποστήριξης (Support Vector Machines) .....</b>	<b>45</b>
<b>ΚΕΦΑΛΑΙΟ 4<sup>ο</sup>.....</b>	<b>49</b>
<b>4.1Εισαγωγή.....</b>	<b>49</b>
<b>4.2 Ανάλυση Κύριων Συνιστωσών (Principal Components Analysis- PCA).....</b>	<b>49</b>
<b>4.3 Νευρωνικό Δίκτυο Cottrell (Cottrell Neural Network).....</b>	<b>59</b>
<b>4.4 Νευρωνικό Δίκτυο Επαλήθευσης Προσώπου με Ανάστροφη Διάδοση (Backpropagation Neural Network for Face Verification).....</b>	<b>61</b>
<b>4.5 Αναγνώριση Προσώπου βάση την Ανάλυση Ανεξαρτήτων Συνιστωσών (Face Recognition Based on Independent Component Analysis- ICA): .....</b>	<b>65</b>
<b>4.6 Οι Μηχανισμοί Διανόσματος Υποστήριξης στην Αναγνώριση Προσώπου (Support Vector Machines for Face Recognition).....</b>	<b>68</b>
4.6.1 Component-based προσέγγιση.....	72

4.6.2. Ανίχνευση.....	72
4.6.3. Αναγνώριση.....	74
4.6.4. Πειράματα.....	75
4.6.6. Συμπέρασμα.....	79
<b>4.7 Συμπεράσματα.....</b>	<b>79</b>
<b>ΚΕΦΑΛΑΙΟ 5<sup>ο</sup>.....</b>	<b>81</b>
5.1 Γενικά.....	81
5.2 Συμπεράσματα.....	82
5.3 Μελλοντική Εργασία.....	83
<b>ΒΙΒΛΙΟΓΡΑΦΙΑ.....</b>	<b>84</b>
<b>ΠΗΓΕΣ ΕΙΚΟΝΩΝ.....</b>	<b>87</b>

# Λίστα Εικόνων

Εικόνα 1 Παραδείγματα Εφαρμογών .....	10
Εικόνα 2 Παραδείγματα Εφαρμογών .....	10
Εικόνα 3 Αναγνώριση Προσώπου με τη Χρήση Χαρακτηριστικών (Προφίλ) .....	14
Εικόνα 4 Αναγνώριση Προσώπου με τη Χρήση Χαρακτηριστικών (Ολιστική Προσέγγιση) .....	15
Εικόνα 5 Εποπτευόμενη μάθηση .....	20
Εικόνα 6 Μη Εποπτευόμενη ή αυτο-οργανωμένη μάθηση .....	20
Εικόνα 7 Ορθή γενίκευση .....	21
Εικόνα 8 Μη ορθή γενίκευση .....	21
Εικόνα 9 Το διάγραμμα της ανάλυσης ICA .....	25
Εικόνα 10 Ο αισθητήρας εικόνας παρέχει υψηλού δυναμικού εικόνα εύρους .....	25
Εικόνα 11 Χαρτογράφηση των άκρων της γραμμής (Line edge map) .....	26
Εικόνα 12 Δίκτυο Κυμάτων Gabor .....	27
Εικόνα 13 Προοπτική Αραιής Εκπροσώπησης .....	27
Εικόνα 14 Παράθυρο Διαδρομής Gauss .....	29
Εικόνα 15 Προσέγγιση της Εξελικτικής Αναζήτησης .....	30
Εικόνα 16 Ο αληθινός χάρτης της Ρώμης και χάρτες ταξινόμησης που λαμβάνονται με διάφορες μεθόδους πυρήνα .....	31
Εικόνα 17 Πληροφορίες που προέρχονται από το Μετασχηματισμό Ίχνους .....	31
Εικόνα 18 Σχηματικό διάγραμμα του νευρικού συστήματος .....	34
Εικόνα 19 Νευρωνικά Δίκτυα Εμπρόσθιας Τροφοδότησης, Πολυστρωματικά Δίκτυα Προελέγχου και Επαναλαμβανόμενα Δίκτυα .....	35
Εικόνα 20 Μη γραμμικό μοντέλο ενός νευρώνα .....	35
Εικόνα 21 Πολυστρωματικό Νευρωνικό Δίκτυο με ένα κρυφό στρώμα .....	37
Εικόνα 22 -Συνήθεις μορφές συνάρτησης ενεργοποίησης νευρώνων. ....	40
Εικόνα 23 Συναρτήσεις ενεργοποίησης με τιμές στο διάστημα [-1,+1]. ....	42
Εικόνα 24 Επιφάνεια Σφάλματος .....	45
Εικόνα 25 Γενική δομή του μηχανισμού διανυσμάτων υποστήριξης SVM .....	47
Εικόνα 26 Εικόνες από την βάση εκπαίδευσης .....	53
Εικόνα 27 Ιδιοτιμές .....	54
Εικόνα 28 Τρία πρώτα ιδιοδιανύσματα .....	54
Εικόνα 29 Τελευταία τρία ιδιοδιανύσματα .....	54
Εικόνα 30 Ευκλείδια απόσταση ανάμεσα στην εικόνα δοκιμής και στις εικόνες από την βάση δεδομένων .....	56
Εικόνα 31 Εικόνα δοκιμής και αναγνωρισμένη εικόνα από την βάση εκπαίδευσης .....	56
Εικόνα 32 απόσταση ανάμεσα στην εικόνα δοκιμής και στην βάση δεδομένων .....	57
Εικόνα 33 Εικόνα δοκιμής και αναγνωρισμένη εικόνα από την βάση εκπαίδευσης .....	57
Εικόνα 34 Ευκλείδια απόσταση ανάμεσα σε εικόνα δοκιμής και στις εικόνες από την βάση δεδομένων .....	58
Εικόνα 35 Εικόνα από ένα άγνωστο πρόσωπο .....	58
Εικόνα 36 Αριστερά: Δίκτυο αυτόματης σύνδεσης του στρώματος εισόδου για το δίκτυο στα δεξιά με τη χρήση της αναστροφής διάδοσης. ....	60

Εικόνα 37 Αναγνώριση Προσώπου.....	61
Εικόνα 38 Δείγματα Εικόνων από την βάση δεδομένων του Yale .....	62
Εικόνα 39 (α) Εικόνα Εισόδου Εικόνα 39(β). Μέση Εικόνα Εικόνα 39(γ) Εικόνα εξόδου μέσω της PCA	62
Εικόνα 40(α) Σύνολο Εκπαίδευσης 40(β) Ιδιοδυναμύσματα προσώπων 40(γ) Αναγνωρισμένο πρόσωπο από την μέθοδο οπισθοδιάδοσης .....	62
Εικόνα 41 Γραμμική σύγκριση ποσοστού επίδοσης και χρόνου εκτέλεσης σε λεπτά σε σχέση με τον αριθμό των εικόνων.....	63
Εικόνα 42 Δίκτυο με Ανάστροφη Διάδοση.....	65
Εικόνα 43 Αριστερά, Αρχικές εικόνες, Κεντρικά PCA και δεξιά ICA εικόνες από την Yale βάση δεδομένων.....	66
Εικόνα 44 Παρουσιάζει το συσσωρευμένο βαθμό σφάλματος και την ταξινόμηση του σφάλματος ως συνάρτηση του αριθμού των συντελεστών για το σύνολο προσώπων από την AR. ....	68
Εικόνα 45 Αναγνώριση .....	70
Εικόνα 46 Δυαδικό δέντρο προσώπων που δημιουργείται από την διαιρετική ομαδοποίηση.....	72
Εικόνα 47 Στο σύστημα φαίνονται οι τέσσερις συνιστώσες όπου χρησιμοποιεί ο ανιχνευτής προσώπου. Στο πρώτο επίπεδο, τα παράθυρα του μεγέθους των συνιστωσών αλλάζουν την εικόνα του προσώπου και την ταξινομούν σύμφωνα με τους ταξινομητές των συνιστωσών. Στο δεύτερο επίπεδο, οι μέγιστοι εξοδοι των ταξινομητών των συνιστωσών, στην προκαθορισμένη περιοχή αναζήτησης και οι θέσεις των ανιχνευσιμων συνιστωσών τροφοδοτούνται από την γεωμετρική διάταξη του ταξινομητή. ....	73
Εικόνα 48 Η εικόνα (α) δείχνει τις 14 συνιστώσες του ανιχνευτή προσώπων. Τα κέντρα των συνιστωσών είναι μαρκαρισμένα με ένα λευκό σταυρό. Οι 10 συνιστώσες που χρησιμοποιήθηκαν για την αναγνώριση του προσώπου, φαίνονται στο (β) .....	74
Εικόνα 49 Παραδείγματα με βάση συνιστωσών ανίχνευση προσώπου. Παρουσιάζονται τα τμήματα του προσώπου καλυμμένα με τις 10 συνιστώσες που χρησιμοποιήθηκαν για την αναγνώριση προσώπου. ....	75
Εικόνα 50 Οι καμπύλες ROC όταν εκπαιδευτούν και δοκιμαστούν σε μετωπικά και περιστραμμένα πρόσωπα.....	76
Εικόνα 51 Οι καμπύλες ROC όταν εκπαιδευτούν σε μετωπικά πρόσωπα και δοκιμαστούν σε μετωπικά και περιστραμμένα πρόσωπα. ....	77
Εικόνα 52 Παραδείγματα των με βάση συνιστωσών αναγνώριση προσώπων. Οι τρεις πρώτες στήλες και η πρώτη εικόνα από την τελευταία στήλη, δείχνουν μια σωστή αναγνώριση. Οι δύο τελευταίες εικόνες στην τελευταία στήλη δείχνουν μια λάθος ταξινόμηση λόγω την μεγάλης περιστροφής και έκφρασης του προσώπου. ....	78

# Λίστα Πινάκων

Πίνακας 1 Το αποτέλεσμα της αναγνώρισης προσώπου, χρησιμοποιώντας ιδιοδιανύσματα .....	55
Πίνακας 2 Ποσοστό αναγνώρισης για διαφορετικό αριθμό εικόνων ανά άτομο .....	55
Πίνακας 3 Ποσοστό Επίδοσης & Χρόνος Εκτέλεσης .....	63
Πίνακας 4 Το ποσοστό αναγνώρισης για τις Yale και AR εικόνες χρησιμοποιούν το πλησιέστερο μέσο ταξινομητή(nearest mean classifier -NMC) και SVM .....	67



# ΚΕΦΑΛΑΙΟ 1<sup>ο</sup>

## ΕΙΣΑΓΩΓΗ

Στο κεφάλαιο αυτό παρουσιάζονται συνοπτικά οι βασικές έννοιες που θα αναλυθούν στα επόμενα κεφάλαια της εργασίας. Συγκεκριμένα, θα περιγράφεται εν συντομία το πρόβλημα της αναγνώρισης προσώπου και των τρόπων αντιμετώπισής του κυρίως μέσω των Τεχνητών Νευρωνικών Δικτύων (Artificial Neural Networks).

### 1.1.Περιγραφή του προβλήματος

Τα ανθρώπινα όντα έχουν δυνατότητες αναγνώρισης που είναι απαράμιλλη στη σύγχρονη εποχή των υπολογιστών, κυρίως λόγω του υψηλού βαθμού διασύνδεσης, της προσαρμοστικής φύσης, των δεξιοτήτων μάθησης και των ικανοτήτων γενίκευσης του νευρικού συστήματος. Ο ανθρώπινος εγκέφαλος έχει πολλούς εξαιρετικά διασυνδεδεμένους βιολογικούς νευρώνες που, σε ορισμένα ειδικά καθήκοντα, μπορούν να ξεπεράσουν τους υπερυπολογιστές. Ένα παιδί μπορεί να προσδιορίσει με ακρίβεια ένα πρόσωπο, αλλά για έναν υπολογιστή είναι μια πολύπλοκη διαδικασία. Ως εκ τούτου, η βασική ιδέα είναι να κατασκευαστεί ένα σύστημα που μπορεί να μιμηθεί αυτό που ένα παιδί μπορεί να κάνει. Εξελίξεις στην υπολογιστική ικανότητα κατά τη διάρκεια των τελευταίων δεκαετιών επέτρεψαν ανάλογες δυνατότητες αναγνώρισης από τέτοια συστήματα μηχανικής με μεγάλη επιτυχία. Αλγόριθμοι πρώιμης αναγνώρισης προσώπου χρησιμοποίησαν απλά γεωμετρικά μοντέλα, αλλά πρόσφατα η διαδικασία αναγνώρισης έχει πλέον ωριμάσει σε μια επιστήμη των εξελιγμένων μαθηματικών αναπαραστάσεων και συνδυαστικών διαδικασιών. Σημαντικές εξελίξεις και πρωτοβουλίες ώθησαν την τεχνολογία αναγνώρισης προσώπου στο προσκήνιο. (Samal & Prasana, 2012) (Meadows, 2004)

Η *Αναγνώριση Προσώπου* είναι η διαδικασία με την οποία εντοπίζονται και προσδιορίζονται πρόσωπα σε μια σκηνή με τη χρήση τεχνικών επεξεργασίας εικόνας. Παρόλο που οι άνθρωποι αναγνωρίζουν πρόσωπα πολλές φορές την ημέρα, με προφανή ευκολία, η αυτοματοποίηση της διαδικασίας αυτής έχει αμφισβητηθεί από τους

ερευνητές τουλάχιστον τις δύο τελευταίες δεκαετίες. (Meadows, 2004) Παρά ταύτα, ένα αυτοματοποιημένο σύστημα αναγνώρισης προσώπου έχει μελλοντική προσφορά στην ανθρωπότητα, ώστε να δικαιολογεί τα έτη της έρευνας που έχουν λάβει χώρα έως σήμερα πάνω στο πρόβλημα αυτό.

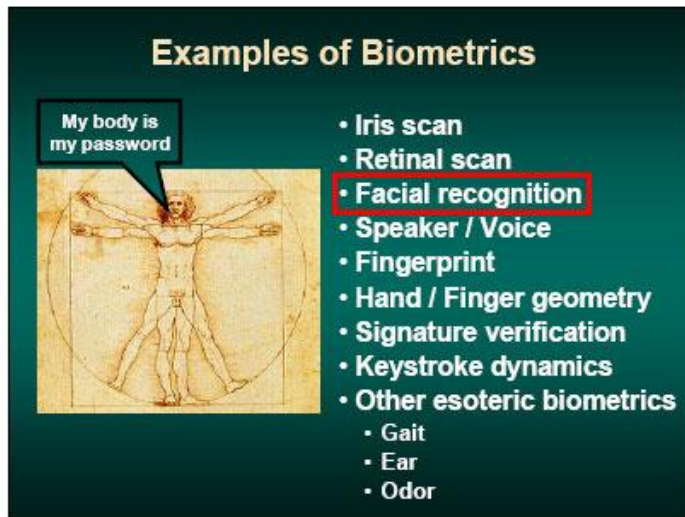
Ένα σύστημα που αναγνωρίζει πρόσωπα αυτόματα θα ήταν χρήσιμο για ποικίλους λόγους (μερικά παραδείγματα εφαρμογών παρουσιάζονται στην *Εικόνα 1 και 2*). Μια σημαντική πτυχή είναι ότι η εν λόγω τεχνολογία θα πρέπει να είναι σε θέση να ασχοληθεί με διάφορες αλλαγές στις εικόνες του προσώπου, όπως η περιστροφή και οι τροποποιήσεις στην έκφραση. Απροσδόκητα, οι μαθηματικές διαφορές μεταξύ των εικόνων του ίδιου προσώπου λόγω του φωτισμού και της κατεύθυνσης θέασης είναι σχεδόν πάντα μεγαλύτερες από τις μεταβολές της εικόνας λόγω των μεταβολών της ταυτότητας του προσώπου. Αυτό αποτελεί μια μεγάλη πρόκληση για την αναγνώριση προσώπου. Κυριότερα, δύο θέματα είναι κεντρικής σημασίας για την επιτυχή αναγνώριση προσώπου: Πρώτον, η επιλογή των χαρακτηριστικών που χρησιμοποιείται για να αντιπροσωπεύει ένα πρόσωπο. Δεδομένου ότι οι εικόνες υπόκεινται σε αλλαγές στη θέαση, στο φωτισμό και στην έκφραση, η αποτελεσματική εκπροσώπηση θα πρέπει να είναι σε θέση να ασχοληθεί με αυτές τις πιθανές αλλαγές. Δεύτερον, η κατάταξη μιας νέας εικόνας προσώπου που χρησιμοποιείται για την επιλεγμένη εκπροσώπηση. (Samal & Prasana, 2012)

Η Αναγνώριση προσώπου μπορεί να είναι δύο τύπων:

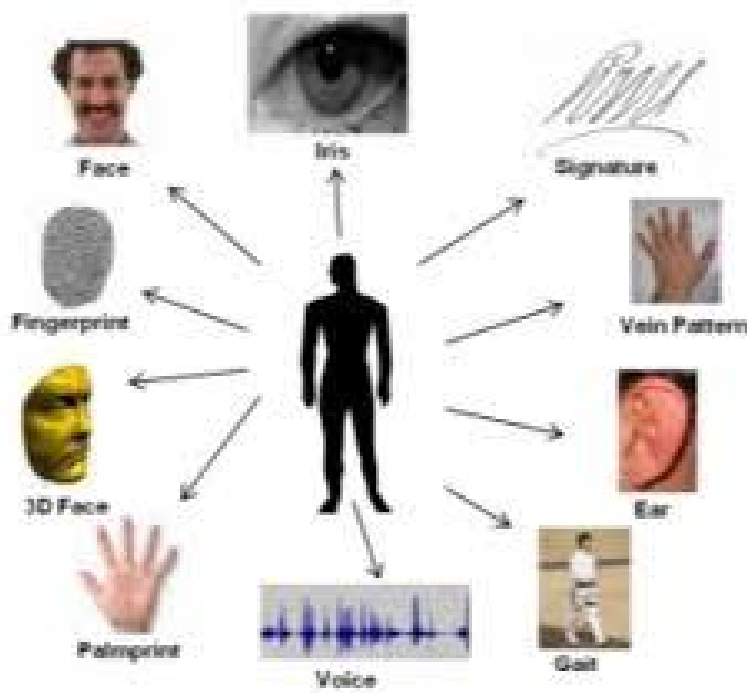
- Με βάση τα *χαρακτηριστικά (γεωμετρική - Geometric)*.
- Με βάση το *πρότυπο (φωτομετρική - Photometric)*.

Στις γεωμετρικές ή με βάση τα χαρακτηριστικά μεθόδους, ανιχνεύονται τα χαρακτηριστικά του προσώπου όπως τα μάτια, τη μύτη, το στόμα και το πηγούνι. Οι ιδιότητες και σχέσεις, όπως οι περιοχές, οι αποστάσεις και οι γωνίες μεταξύ των χαρακτηριστικών γνωρισμάτων χρησιμοποιούνται ως περιγραφείς των προσώπων. Αν και αυτή η κατηγορία μεθόδων είναι οικονομική και αποτελεσματική στην επίτευξη μείωσης δεδομένων και δεν είναι ευαίσθητη σε μεταβολές φωτισμού και οπτικής, στηρίζεται σε μεγάλο βαθμό σε σχετικά με την εξαγωγή και τη μέτρηση των χαρακτηριστικών του προσώπου. (Damasio, 2005) Δυστυχώς, η εξαγωγή χαρακτηριστικών και οι τεχνικές μέτρησης και οι αλγόριθμοι που αναπτύχθηκαν μέχρι σήμερα δεν ήταν αρκετά αξιόπιστα για να καλύψουν αυτή την ανάγκη.

Σε αντίθεση, το ταίριασμα προτύπων και οι νευρωνικές μέθοδοι γενικά λειτουργούν άμεσα σε μια εικόνα για την αναπαράσταση των προσώπων, όπως η συστοιχία της έντασης των εικονοστοιχείων (pixelintensityarray). Επειδή δεν απαιτείται η ανίχνευση και η μέτρηση των γεωμετρικών χαρακτηριστικών του προσώπου, αυτό το είδος μεθόδου είναι πιο πρακτικό και πιο εύκολο στην εφαρμογή, σε σύγκριση με γεωμετρικές ή με βάση τα χαρακτηριστικά μεθόδους. (Samal & Prasana, 2012)



Εικόνα 1 Παραδείγματα Εφαρμογών



Εικόνα 2 Παραδείγματα Εφαρμογών

Μια σημαντική πτυχή είναι ότι μια τέτοια τεχνολογία θα πρέπει να είναι σε θέση να αντιμετωπίσει διάφορες μεταβολές στις εικόνες προσώπου, όπως η περιστροφή και συχνές αλλαγές στην έκφραση. Απροσδόκητα, οι μαθηματικές παραλλαγές, μεταξύ των ειδώλων του ίδιου προσώπου που οφείλονται στην φωταψία και τον προσανατολισμό θέασης, είναι σχεδόν πάντα μεγαλύτερες από τις διαφοροποιήσεις της εικόνας λόγω των αλλαγών στην ταυτότητα του προσώπου. Αυτό ενέχει σημαντική πρόκληση για την αναγνώριση του προσώπου. Στο επίκεντρο, δύο θέματα είναι καίριας σημασίας για την επιτυχία των αλγορίθμων στην αναγνώριση του προσώπου: Πρώτον, η επιλογή των λειτουργιών που χρησιμοποιούνται για να αναπαραστήσουν ένα πρόσωπο. Δεδομένου ότι οι εικόνες υπόκεινται σε αλλαγές στην οπτική γωνία, στο φωτισμό και στην έκφραση, μια αποτελεσματική εκπροσώπηση πρέπει να είναι σε θέση να αντιμετωπίσει αυτές τις πιθανές αλλαγές. Δεύτερον, η κατάταξη μιας νέας εικόνας προσώπου χρησιμοποιώντας την επιλεγμένη αναπαράσταση. (Meadows, 2004) (Valentine & Ferrara, 2012)

Από την προοπτική της ασφάλειας, ένα αυτόματο σύστημα αναγνώρισης προσώπων θα μπορούσε να ενισχύσει τα τρέχοντα συστήματα ελέγχου πρόσβασης με την επικύρωση της ταυτότητας των χρηστών. Μερικά παραδείγματα τέτοιων συστημάτων ελέγχου πρόσβασης είναι συστήματα ασφάλειας του υπολογιστή, τα τραπεζικά μηχανήματα αυτόματης ανάληψης και οι αυτόματες κάρτες ανάγνωσης. Στην πράξη, οποιοσδήποτε οργανισμός ή σύστημα που επιτρέπει την πρόσβαση με βάση την ταυτότητα ενός ατόμου, θα επωφελούνταν από ένα σύστημα αναγνώρισης προσώπων. Άλλες χρήσεις ασφαλείας για ένα σύστημα αναγνώρισης προσώπων θα ήταν η αναγνώριση εγκληματιών και η σάρωση των αεροδρομίων για τρομοκράτες. Τέλος, το σύστημα αυτό θα μπορούσε να προσαρμοστεί για χρήση σε ένα σύστημα αναγνώριση ή ένα σύστημα οπτικής επικοινωνίας. (Samal & Prasana, 2012)

Το υπόλοιπο της παρούσας περιγραφής του προβλήματος αρχίζει με μία ανασκόπηση του υπόβαθρου της αναγνώρισης προσώπων, καθώς συνδέεται με την αναγνώριση μορφών και την παρούσα διπλωματική εργασία. Επίσης, συνοδεύεται από τη δήλωση του προβλήματος, το σκοπό της διπλωματικής εργασίας, τους στόχους της διπλωματικής εργασίας και τη μεθοδολογία. Τέλος, συζητείται η προσέγγιση του προβλήματος και το κεφάλαιο ολοκληρώνεται με μια επισκόπηση των υπολοίπων κεφαλαίων.

## 1.2 Υπόβαθρο

Η παρούσα ενότητα εξετάζει εν συντομία τη διαδικασία αναγνώρισης προτύπου και τονίζει στη συνέχεια ορισμένες από τις τρέχουσες έρευνες στον τομέα της αυτόνομης αναγνώρισης προσώπου.

Παραδοσιακά, η αναγνώριση προτύπων αναλύεται σε τρία πεδία: α) την *κατάτμηση (segmentation)*, β) την *εξαγωγή χαρακτηριστικών (feature extraction)* και γ) την *ταξινόμηση (classification)*. **Σφάλμα! Το αρχείο προέλευσης της αναφοράς δεν βρέθηκε.** Η κατάτμηση αποτελεί το πρώτο πεδίο, ανιχνεύοντας περιοχές ενδεχόμενων σημάτων. Το δεύτερο πεδίο είναι η εξαγωγή χαρακτηριστικών γνωρισμάτων, ανιχνεύοντας για τα πιο σημαντικά ή καίρια χαρακτηριστικά των περιοχών που πέρασαν από τον τεμαχιστή (*segmentor*) και τα οποία μπορούν να χρησιμοποιηθούν στο τελευταίο πεδίο, την ταξινόμηση. Η ταξινόμηση συγκρίνει τα εξαγόμενα χαρακτηριστικά με εκείνα του προηγουμένως ταυτοποιημένων αντικειμένων και εντοπίζει το αντικείμενο ως μία προηγουμένως διαπιστωμένη κατηγορία. Για την εργασία αυτή, οι περιοχές που επιχειρούνται να διατμηθούν και να προσδιοριστούν είναι τα πρόσωπα, αλλά η προσέγγιση είναι η ίδια ακόμη και για άλλα χαρακτηριστικά, όπως οι λέξεις στην ομιλία.

## 1.3 Δήλωση του Προβλήματος

Η παρούσα πτυχιακή εργασία επικεντρώνεται στη σημαντικότητα ενός αυτοματοποιημένου συστήματος αναγνώρισης προσώπου και στη μελλοντική προσφορά στην ανθρωπότητα, καθώς και στις τεχνολογικές προκλήσεις που θέτει στους ερευνητές η εφαρμογή του συστήματος αυτού.

## 1.4 Σκοπός της Διπλωματικής Εργασίας

Ο σκοπός αυτής της διπλωματικής εργασίας είναι αφενός η κατανόηση και η όσο το δυνατό ενδελεχής περιγραφή του προβλήματος της Αναγνώρισης Προσώπου, καθώς και η παρουσίαση των έως σήμερα προτεινόμενων μεθόδων και αφετέρου οι τρόποι εφαρμογής των Τεχνητών Νευρωνικών Δικτύων, τονίζοντας τη συμβολή τους στην επίλυση του εν λόγω προβλήματος. Για την πραγματοποίηση των σκοπών της πτυχιακής εργασίας αξιοποιήθηκε η διεθνής και ελληνική βιβλιογραφία αναφορικά με τις Τεχνικές Αναγνώρισης Προσώπου και τα Τεχνητά Νευρωνικά Δίκτυα

## 1.5 Στόχοι της Διπλωματικής Εργασίας

Οι στόχοι της παρούσας διπλωματικής εργασίας είναι ηκατανόηση των εννοιών των Τεχνητών Νευρωνικών Δικτύων, καθώς και του τρόπου εφαρμογής τους σε προβλήματα που άπτονται της Αναγνώρισης Προτύπων και συγκεκριμένα της Αναγνώρισης Προσώπου.

## 1.6 Επισκόπηση της Διπλωματικής Εργασίας

Το Δεύτερο κεφάλαιο αναφέρεται στην Αναγνώριση Προσώπου και παρουσιάζει διεξοδικά το πρόβλημα της αναγνώρισης προσώπου, τη σημαντικότητά του, καθώς και τις διάφορες προσεγγίσεις που έχουν προταθεί, μέσα από μια ανασκόπηση της υπάρχουσας βιβλιογραφίας σχετικά με τα συστήματα αναγνώρισης προσώπου, με πρωταρχική έμφαση στην εξαγωγή χαρακτηριστικών και την ταξινόμηση.

Το Τρίτο κεφάλαιο αναφέρεται στα Τεχνητά Νευρωνικά Δίκτυα και αναλύεται διεξοδικά η έννοια των Τεχνητών Νευρωνικών Δικτύων, οι διάφοροι τύποι τους, τα πλεονεκτήματα και τα μειονεκτήματά τους και οι εφαρμογές τους, παρέχοντας μια λεπτομερήπεριγραφήαπό τη μεθοδολογία που χρησιμοποιήθηκε σε αυτή τη διπλωματική εργασία.

Το Τέταρτο κεφάλαιο αναφέρεται στα Νευρωνικά Δίκτυα στην Αναγνώριση Προσώπων. Στο κεφάλαιο αυτό θα αναλυθούν οι μελέτες – προσεγγίσεις που έχουν προταθεί για την επίλυση του προβλήματος της Αναγνώρισης Προσώπων με τη χρήση Τεχνητών Νευρωνικών Δικτύων και θα σχολιαστούν τα αποτελέσματα της κάθε προσέγγισης.

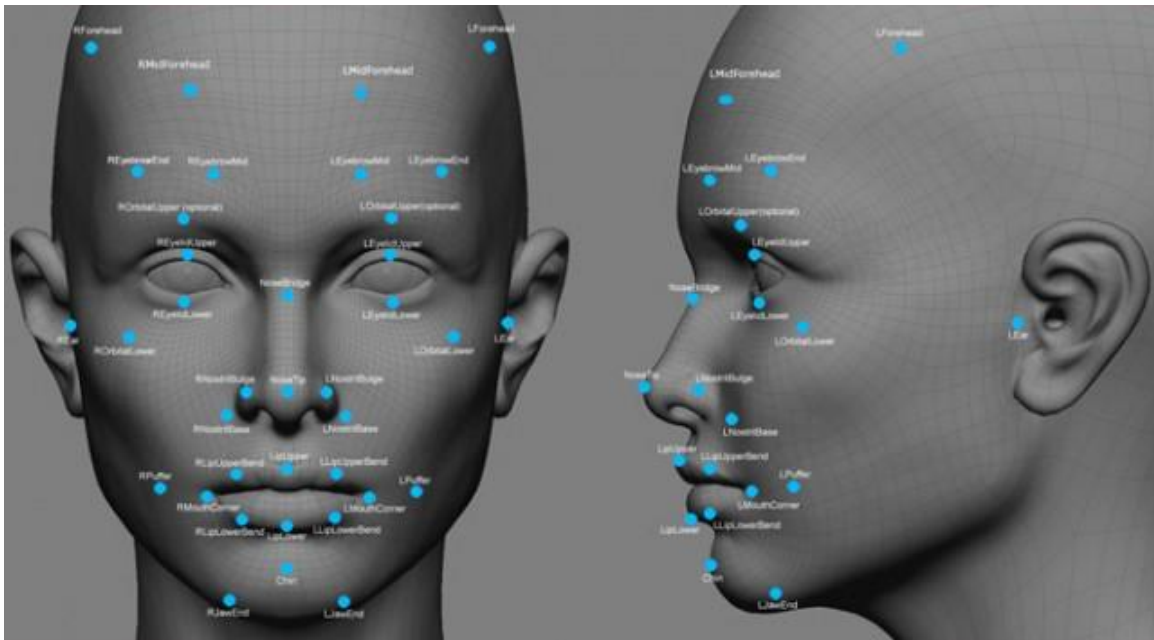
Το Πέμπτο κεφάλαιο αναφέρεται στα συμπεράσματα με βάση τη βιβλιογραφική ανασκόπηση και προβαίνει σε συστάσεις για τη μελλοντική μελέτη.

## ΚΕΦΑΛΑΙΟ 2<sup>Ο</sup>

### 2.1 Εισαγωγή

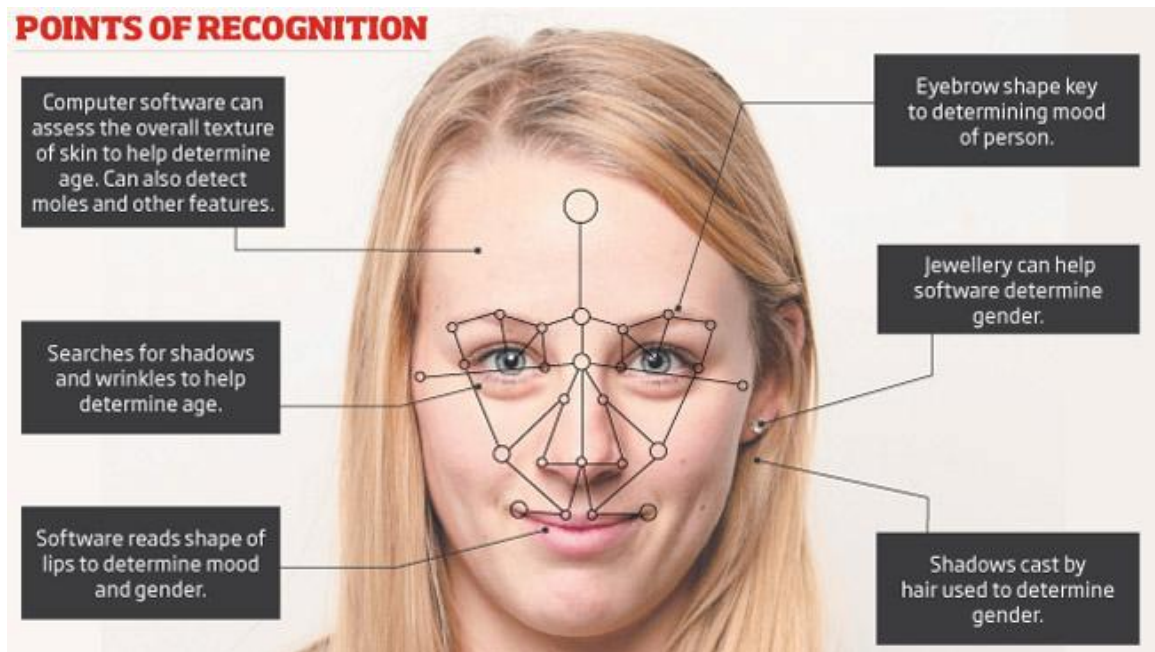
Η έρευνα για την αναγνώριση προσώπου κατά τη διάρκεια των τελευταίων ετών, κατατάσσεται σε δύο κατηγορίες: στην αναγνώριση με τη χρήση των χαρακτηριστικών και στην αναγνώριση με τη χρήση ολόκληρου του προσώπου ή ολιστική προσέγγιση (holistic approach).

Η πρώτη απόπειρα για την αναγνώριση προσώπου με τη χρήση χαρακτηριστικών πραγματοποιήθηκε από τον L.D Harmon στις αρχές του 1970, με την εξαγωγή χαρακτηριστικών από τα προφίλ για τον εντοπισμό προσώπων. Τα χαρακτηριστικά καθορίζονταν ως η απόσταση από την άκρη της μύτης προς το στόμα, η απόσταση από τη μύτη προς το πηγούνι, η απόσταση από τα μάτια στη μύτη και άλλες παρόμοιες μετρήσεις (Εικόνα 3). (Harmon, 2011) Εκτός από την μέθοδο του Harmon, άλλα είδη αναγνώρισης προσώπου με τη χρήση χαρακτηριστικών περιλαμβάνει την κατάτμηση του προσώπου και, στη συνέχεια, την εξαγωγή χαρακτηριστικών από τα τμήματα. Όποια και αν είναι η μέθοδος, η αναγνώριση προσώπου με τη χρήση χαρακτηριστικών συνεχίζεται έως σήμερα από τους ερευνητές σε όλο τον κόσμο. (Turk & Pentland, 2010)



Εικόνα 3 Αναγνώριση Προσώπου με τη Χρήση Χαρακτηριστικών (Προφίλ)

Η δεύτερη κατηγορία της αναγνώρισης προσώπου είναι η ολιστική προσέγγιση. Η έρευνα στην αναγνώριση προσώπου κινείται πλέον προς μια ολιστική άποψη μεερευνητές στο Ινστιτούτο Τεχνολογίας της Μασαχουσέτης (MIT) (Turk & Pentland, 2010),στο Πανεπιστήμιο της Καλιφόρνια στο Σαν Ντιέγκο (UCSD) (Cottrell, 1990) και στο (AFIT). (Wright, 2009) (Nan & W, 2009) Η ολιστική προσέγγιση εξακολουθεί να περιλαμβάνει την εξαγωγή χαρακτηριστικών, αλλά τα χαρακτηριστικά γνωρίσματα, τα οποία εξάγονται με τη χρήση κάποιου τύπου ανάλυσης των κυρίων συνιστωσών, λαμβάνονται πλέον από ολόκληρη την εικόνα του προσώπου, και όχι μόνο κατά τμήματα ή προφίλ (Εικόνα 4).



Εικόνα 4 Αναγνώριση Προσώπου με τη Χρήση Χαρακτηριστικών (Ολιστική Προσέγγιση)

Η διπλωματική αυτή εργασία βασίζεται στην ολιστική προσέγγιση και είναι το θέμα που θα συζητηθεί στα ακόλουθα τμήματα.

Το υπόλοιπο αυτής της βιβλιογραφικής ανασκόπησης ξεκινά με μερικές συνοπτικές βιολογικές σημειώσεις σχετικά με την ανθρώπινη αναγνώριση προσώπου. Ακολουθεί μια συζήτηση για την εξαγωγή χαρακτηριστικών γνωρισμάτων και τη συμπίεση των δεδομένων με τη χρήση της ανάλυσης των κυρίων συνιστωσών και τα νευρωνικά δίκτυα. Η βιβλιογραφική ανασκόπηση συνεχίζεται με αναφορές σε διάφορα δίκτυα που χρησιμοποιούνται για την αναγνώριση προσώπου, με έμφαση στο δίκτυο Cottrell, στο πρότυπο πολυστρωματικό νευρωνικό δίκτυο (standard multiple layer perceptron) με τη



χρήση ανάστροφης διάδοσης(backpropagation) και στο δίκτυο Tarr / Ruck για μετασχηματισμούς και ταξινόμηση Karhunen-Loeve.

## 2.2Η Βιολογία της Αναγνώρισης Προσώπου

Η αναγνώριση προσώπου από τον ανθρώπινο εγκέφαλο εξακολουθεί να αποτελεί ένα μυστήριο. Εμπειρογνώμονες πιστεύουν ότι η βιολογική διαδικασία της αναγνώρισης προσώπου παρουσιάζεται στην κάτω πλευρά και των δύο ημισφαιρίων του εγκεφάλου, στον κροταφικό και ινιακό λοβό. (Geschwind, 2009) Δηλαδή, η εν λόγω διαδικασία εντοπίζεται (βρίσκεται σε μια συγκεκριμένη περιοχή) στον ανθρώπινο εγκέφαλο μας. Αυτή η θεωρία υποστηρίζεται από τη μελέτη των Rolls, Baylis, Hasselmo και Nalwa για την αντιμετώπιση των συγκεκριμένων νευρώνων σε πρόσωπα. Μέσα από ένα πείραμα σε μια ομάδα νευρώνων του προσώπου που σε μακάκους πιθήκους, κατέγραψαν τις νευρωνικές αποκρίσεις (αιχμές δυναμικού δράσης ανά δευτερόλεπτο) σε ένα σύνολο προσωπικών και μη-προσωπικών ερεθισμάτων. Τα αποτελέσματα της μελέτης δείχνουν ότι υπάρχουν νευρώνες που αντιδρούν κυρίως σε πρόσωπα. (Rolls, 2009)

Ο εντοπισμός (localization) υποστηρίζεται, επίσης, από το έργο των Damasio (2005) και Meadows (2004) από την έρευνά τους σε άτομα που έχουν χάσει την ικανότητα να αναγνωρίζουν πρόσωπα, μια κατάσταση γνωστή ως *προσωπαγνωσία (prosopagnosia)*. Και οι δύο ερευνητές συμφωνούν ότι στα άτομα με προσωπαγνωσία, όταν πέρασαν από τη διαδικασία της αυτοψίας, εντοπίστηκαν αμφίπλευρες βλάβες στις κροταφο-ινιακές περιοχές του εγκεφάλου τους. (Damasio, 2005) (Meadows, 2004) Αντιστρόφως, τα άτομα με διμερείς αλλοιώσεις σε άλλα τμήματα του εγκεφάλου (δηλ., κροταφο-βρεγματική περιοχή) δεν έχουν προσωπαγνωσία. Αυτό το γεγονός υποστηρίζει επίσης, αν και σε περιορισμένο βαθμό, την έννοια ότι η διαδικασία είναι εντοπισμένη. Αποδέχοντας τον εντοπισμό ως γεγονός, ποιες πληροφορίες απαιτεί αυτή η περιοχή του εγκεφάλου για να αναγνωρίσει πρόσωπα;

Μια θεωρία, που βασίζεται στον ψυχολογικό πειραματισμό, δείχνει ότι οι πληροφορίες για τα πρόσωπα αποθηκεύονται από ένα παγκόσμιο προς ένα τοπικό σύστημα. (Wright, 2009) (Nan & W, 2009) Σε αυτό το πείραμα για πρώτη φορά ζητήθηκε από τους εθελοντές να *ταξινομήσουν (classify)* τις εικόνες με βάση τα πρόσωπα, καταγράφοντας και τον αντίστοιχο χρόνο. Στη συνέχεια, από κάθε εθελοντή ζητήθηκε να *εντοπίσει (identify)* γνωστά πρόσωπα που είτε ήταν διακριτά/ασυνήθιστα ή τυπικά πρόσωπα. Διαπιστώθηκε ότι η *προσωπο-γνωσία (faceness)* μπορούσε να αξιολογηθεί πολύ

γρήγορα, ενώ η αναγνώριση διήρκησε περισσότερο χρόνο. Επίσης, ένα διακριτό πρόσωπο μπορούσε να προσδιοριστεί πιο γρήγορα από ό,τι ένα τυπικό πρόσωπο. Ομοίως, προτάθηκε ότι οι πληροφορίες σχετικά με τη *διαφορά* (*difference*) (δηλ., το χρώμα των ματιών, το σχήμα της μύτης, το στυλ των μαλλιών, κλπ.) χρησιμοποιούνται στη διαδικασία αναγνώρισης και στη διαδρομή. Πληροφορίες που είναι ίδιες (δηλαδή, το γενικό σχήμα ενός προσώπου, δύο μάτια, δύο αυτιά, κλπ.) χρησιμοποιούνται σε υψηλότερο επίπεδο στη διαδικασία αναγνώρισης και δεν αποθηκεύονται για κάθε συγκεκριμένο πρόσωπο. Αυτό το στοιχείο θα μπορούσε να θεωρηθεί ως από την σφαιρική προς την τοπική διαδικασία τύπου στην οποία ταιριάζει η ανθρώπινη εμπειρία με την αναγνώριση προσώπου, διότι συχνά λέγεται ότι ένα πρόσωπο μοιάζει με κάποιο άλλο, εκτός από τα μάτια ή τη μύτη, κλπ. Αυτό το στοιχείο από την καθολική προς την τοπική ιδέα υποστηρίζεται, επίσης, στις μελέτες για την προσωπαγνωσία. Τα άτομα με προσωπαγνωσία μπορούν ακόμα να εντοπίσουν ένα πρόσωπο ως πρόσωπο, αλλά αδυνατούν να προσδιορίσουν σε ποιον ανήκει το πρόσωπο αυτό. Στην πραγματικότητα, πολλά άτομα με προσωπαγνωσία έχουν, επίσης, δυσκολίες με την αναγνώριση και άλλων αντικειμένων, όπως μια συγκεκριμένη μάρκα αυτοκινήτου. Όπως και με τα πρόσωπα, μπορούν να αναγνωρίσουν ένα αυτοκίνητο, ως αυτοκίνητο, αλλά αδυνατούν να αναγνωρίσουν το συγκεκριμένο τύπο του αυτοκινήτου. (Meadows, 2004) (Valentine & Ferrara, 2012) Αυτό συνεπάγεται ότι γίνεται επεξεργασία των γενικών πληροφοριών και στη συνέχεια αναλύονται οι λεπτομέρειες για την πραγματική διαδικασία αναγνώρισης, αλλά ποιες είναι οι λεπτομέρειες;

Οι ειδικοί συμφωνούν ότι κάποια στοιχεία αποθηκεύονται ή κωδικοποιούνται στους νευρώνες αναγνώρισης προσώπου. Ωστόσο, διαφωνούν ως προς το ποια είναι αυτά τα στοιχεία. Μία θεωρία είναι ότι οι πληροφορίες για ένα συγκεκριμένο πρόσωπο αποθηκεύονται σε ένα μητρικό κύτταρο. (Meadows, 2004) Με άλλα λόγια, όλες οι πληροφορίες για ένα συγκεκριμένο άτομο είναι κωδικοποιημένες σε ένα μονό κύτταρο (νευρώνας) και όταν βρίσκεται αυτό το κύτταρο γίνεται η αναγνώριση του προσώπου. Μια δεύτερη θεωρία, υποστηριζόμενη από το έργο του Rolls και των συναδέλφων του (2009), είναι ότι οι πληροφορίες για την αναγνώριση ενός προσώπου αποθηκεύονται σε ένα κωδικοποιημένο σύνολο νευρώνων και αυτοί οι κωδικοί επανυπολογίζονται κάθε φορά που ένα νέο πρόσωπο προστίθεται στο σύστημα. Η θεωρία αυτή ελέγχθηκε με καταγραφή της δραστηριότητας απόκρισης των νευρώνων προσώπου σε ένα σύνολο γνωστών προσώπων και, στη συνέχεια, με την πρόσθεση ενός νέου προσώπου στο σύνολο και την καταγραφή των απαντήσεων σε αυτό το νέο σύνολο. Οι δοκιμές έδειξαν

ότι η πρόσθεση ενός νέου προσώπου μεταβάλλει την απόκριση της *σταθερής κατάστασης* (*steady state*) σε ορισμένους από τους εν λόγω νευρώνες για σύντομο χρονικό διάστημα και, στη συνέχεια, φτάνουν σε μια νέα, στατιστικά σημαντική διαφορά, σε σταθεροποιημένη κατάσταση. (Rolls, 2009) Αυτό σημαίνει ότι οι πληροφορίες για την αναγνώριση προσώπου αποθηκεύονται σε νευρώνες του προσώπου, γνωρίζοντας το άτομο με αυτό τον τρόπο σε ποιο πρόσωπο βασίζει την πυροδότηση του μοτίβου αυτής της ομάδας νευρώνων. Αυτό δείχνει, επίσης, ότι το σύνολο του κώδικα έχει αλλάξει, αν και πολύ γρήγορα, κάθε φορά που ένα νέο πρόσωπο προστίθεται στο σύστημα.

Σε κάθε περίπτωση, υπάρχουν αρκετές και διαφορετικές βιολογικές θεωρίες γύρω από την αναγνώριση προσώπων. Η αποδεκτή θεωρία σήμερα είναι ότι υπάρχουν συγκεκριμένοι νευρώνες αναγνώρισης προσώπου στον ανθρώπινο εγκέφαλο και το στοιχείο από την καθολική (προσωπογνωσία) προς την τοπική (αναγνώριση) ιδέα κερδίζει έδαφος. Από άποψη αναγνώρισης προτύπων, αυτό το στοιχείο από την παγκόσμια στην τοπική ιδέα θα μπορούσε να θεωρηθεί ως συμπύεση δεδομένων, η οποία είναι σημαντική για το σχεδιασμό ενός συστήματος αναγνώρισης προσώπου.

## 2.3Η Τεχνολογία της Αναγνώρισης Προσώπου

### 2.3.1 Ταξινόμηση

Η Αναγνώριση προσώπου δεν είναι τίποτε άλλο παρά η ικανότητα μιας μηχανής να ταξινομήσει με επιτυχία μια σειρά από εικόνες που βασίζονται σε ορισμένα χαρακτηριστικά που εισάγουν διακρίσεις. Η ταξινόμηση ή η αναγνώριση προτύπων μπορεί να είναι ένα πολύ δύσκολο πρόβλημα και εξακολουθεί να είναι ένα πολύ ενεργό πεδίο της έρευνας. Όπως αναφέρθηκε προηγουμένως, το ανθρώπινο είδος έχει την ικανότητα να επεξεργάζεται ένα συγκεκριμένο σενάριο σε ένα γενικό πλαίσιο και να ταυτοποιεί τα στοιχεία του. Φυσικά δεν είναι εντελώς εμφανές πως μπορεί η μηχανή να διακρίνει μεταξύ των στοιχείων διαφόρων αντικειμένων από διαφορετικές ταξινομήσεις. Το έργο της ταξινόμησης μπορεί να παρουσιαστεί μαθηματικά με τον εξής καθορισμό: (Rolls, 2009)

$$f: A \times B = C$$

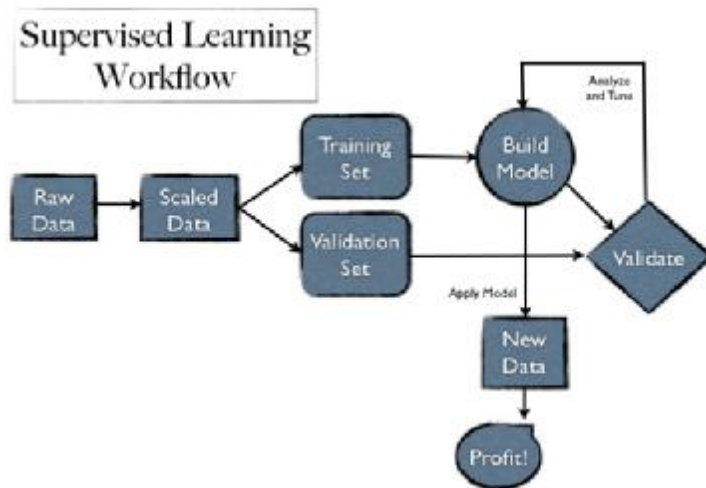
όπου το  $A$  αναπαριστά τα στοιχεία που έχουν ταξινομηθεί, π.χ. τα διανύσματα της έντασης των εικονοστοιχείων (pixel intensity vectors), χρησιμοποιώντας το  $B$  ως τις παραμέτρους λειτουργίας (function parameters). Το αποτέλεσμα  $C$  πρέπει να διαχωρίσει μεταξύ εικόνων από διαφορετικές ταξινομήσεις, έτσι ώστε η ταξινόμηση κάθε στοιχείου να καθορισθεί. Η ταξινόμηση διαμέσου μηχανημάτων απαιτεί δύο βήματα:

A) Πρώτον, θα πρέπει να αναγνωριστούν οι ιδιότητες οι οποίες διακρίνουν ένα στοιχείο μίας ταξινόμησης από μία άλλη ταξινόμηση.

B) Δεύτερον, το μηχάνημα θα πρέπει να καταρτιστεί ώστε να γνωρίζει πώς να διακρίνει μεταξύ των ταξινομήσεων, καθορίζοντας ένα μοντέλο εκμάθησης (learning model). Το μοντέλο εκμάθησης περιγράφει τη διαδικασία που θα πρέπει να χρησιμοποιηθεί για την πραγματική εκπαίδευση (actual learning).

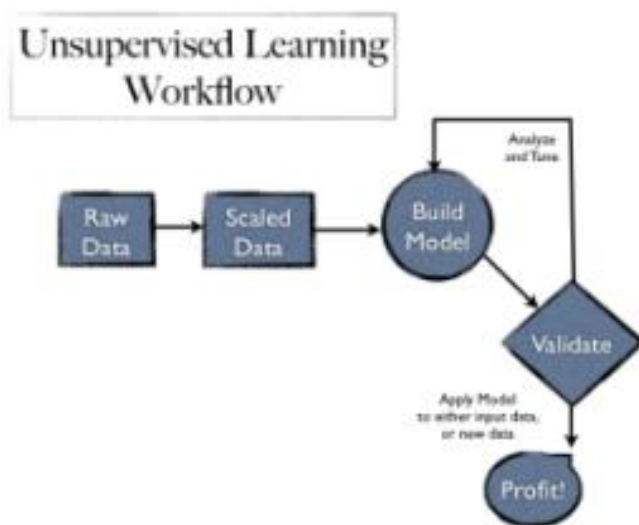
Βασικά, υπάρχουν δύο τύποι εκπαίδευσης που αναφέρονται συχνά ως εποπτευόμενη και μη εποπτευόμενη μάθηση.

- Εποπτευόμενη Μάθηση (supervised learning): Σύμφωνα με αυτό τον τύπο μάθησης, ο εκπαιδευτής (trainer) τροφοδοτεί το μηχάνημα με ένα δείγμα (sample) και το μηχάνημα το κατηγοριοποιεί. Ο εκπαιδευτής ενημερώνει έπειτα το μηχάνημα αν ταξινόμησε σωστά το δείγμα ή όχι. Σε περίπτωση που έχει ταξινομήσει εσφαλμένα το δείγμα, το μηχάνημα θα πρέπει να προσαρμόσει τις παραμέτρους ταξινόμησης για να ταξινομήσει καταλληλότερα το δεδομένο δείγμα. Η εποπτευόμενη μάθηση παρουσιάζεται στην *Εικόνα 5*. Παραδείγματα αυτής περιλαμβάνουν τον Ταξινομητή Bayes (BayesClassifier) και τα Νευρωνικά Δίκτυα (Neural Networks). Μία άλλη σημαντική άποψη που εμπλέκεται είναι η μέθοδος εκπαίδευσης που μεταδίδεται σε αυτά τα μηχανήματα. Είναι σημαντικό να υπάρχει ένα σύνολο εκπαίδευσης (learning set), το οποίο να είναι αντιπροσωπευτικό για τις δεδομένες ταξινομήσεις, ώστε η μελλοντική ταξινόμηση να είναι επιτυχής.



Εικόνα 5 Εποπτευόμενη μάθηση

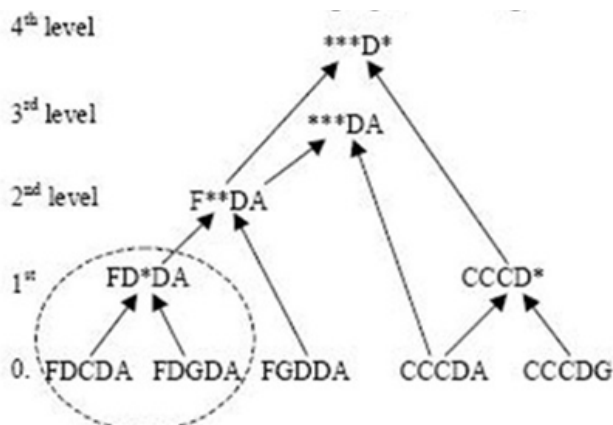
· Μη Εποπτευόμενη ή Αυτο-οργανωμένη μάθηση (unsupervised or self-organized learning): Σύμφωνα με αυτό τον τύπο μάθησης, δεν υπάρχει εξωτερικός εκπαιδευτής που να επιβλέπει την μαθησιακή διαδικασία. Εμφανώς υπολογίζεται μία ανεξάρτητη μέτρηση της ποιότητας της ταξινόμησης και οι ελεύθερες παράμετροι του δικτύου βελτιστοποιούνται με γνώμονα αυτή τη μέτρηση. Όταν το δίκτυο συντονιστεί με τις στατιστικές κανονικότητες των δεδομένων εισόδου, αναπτύσσει την ικανότητα να σχηματίζει εσωτερικές αναπαραστάσεις για χαρακτηριστικά κωδικοποίησης της εισόδου και, συνεπώς, να δημιουργεί αυτόματα καινούριες ταξινομήσεις. Μία τεχνική για την παραπάνω διαδικασία ονομάζεται *ομαδοποίηση (clustering)*. Αντίθετα, το μηχάνημα εκμάθησης καθορίζει ποια στοιχεία θα πρέπει να ανήκουν στην ίδια ομάδα και αναθέτει μία ανάλογη ταξινόμηση. Η Μη Εποπτευόμενη ή Αυτο-οργανωμένη μάθηση παρουσιάζεται στην *Εικόνα 6*.



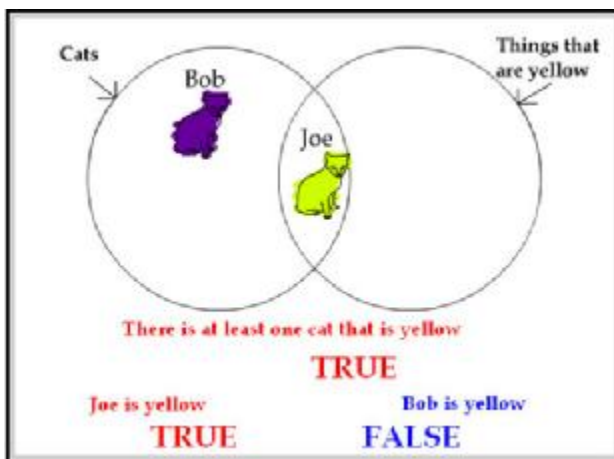
Εικόνα 6 Μη Εποπτευόμενη ή αυτο-οργανωμένη μάθηση

Αξίζει να σημειωθεί ότι η υπερ-μάθηση(over-learning) οδηγεί σε φτωχή γενίκευση (generalization). Για παράδειγμα, ένα υπερ-εκπαιδευμένο μηχάνημα θα αποδώσει διαφορετικές ταξινομήσεις για την ίδια εικόνα που υπόκειται σε διαφορετικές τροποποιήσεις. Όπως παρουσιάζεται στην *Εικόνα 7 και 8* η υπερ-εκπαίδευση απομνημονεύει απαραίτητως τις πληροφορίες και οδηγεί σε φτωχή γενίκευση.

Από την στιγμή που ένα δίκτυο βρίσκεται σε λειτουργική κατάσταση χαρακτηρίζεται από μία ικανότητα γενίκευσης όπου μας αποφέρει ορθές εξόδους για εισόδους και διαφορετικές τιμές από αυτές τι οποίες εκπαιδευτηκε.



Εικόνα 7 Ορθή γενίκευση



Εικόνα 8 Μη ορθή γενίκευση

Η αναγνώριση προσώπου είναι ένας πολύ πολύπλοκος τύπος αναγνώρισης προτύπου. Αποτελείται από την ταξινόμηση άκρως αμφιλεγόμενων σημάτων εισόδου, με πολλαπλές διαστάσεις και την αντιπαραβολή τους με οικεία σήματα. Τα ακόλουθα προβλήματα είναι αυτά που συνήθως συναντώνται:

- η 2D εγγενής δομή μιας μήτρας εικόνας (image matrix) είναι περισσότερο συχνά παρούσα από ό,τι μη αφαιρούμενη. Συνεπώς, οι χωρικές πληροφορίες που αποθηκεύονται εντός απορρίπτονται και δεν αξιοποιούνται αποτελεσματικά.
- Κατάρατης διάστασης (Curseofdimensionality): Κάθε δείγμα εικόνας μοντελοποιείται, αποτελώντας τυπικά ένα χωρικό σημείο υψηλών διαστάσεων. Συνεπώς, χρειάζεται συχνά ένας μεγάλος αριθμός δειγμάτων εκπαίδευσης για να ληφθεί αξιόπιστη και εύρωστη εκτίμηση των χαρακτηριστικών της διανομής πληροφοριών.
- Συνήθως πολύ περιορισμένες ποσότητες πληροφοριών είναι διαθέσιμες σε πραγματικές εφαρμογές, όπως η αναγνώριση προσώπου, η ανάκτηση εικόνας και η ταξινόμηση εικόνας.

Έχει προταθεί μία σειρά λύσεων για την επίλυση αυτών των προβλημάτων. Η εύρεση ενός αποτελεσματικού τρόπου για τη μείωση των διαστάσεων αποτελεί το πρώτο βήμα για την αναγνώριση προσώπου.

### 2.3.2 Εφαρμογές της Αναγνώρισης Προσώπου

Μερικές από τις εφαρμογές είναι οι εξής:

- Η προσπάθεια ανεύρεσης ενός προσώπου μέσα από μία τεράστια βάση δεδομένων από πρόσωπα. Σε αυτή την προσέγγιση, το σύστημα επιστρέφει μία πιθανή λίστα προσώπων από τη βάση δεδομένων. Οι πιο χρήσιμες εφαρμογές περιλαμβάνουν την επιτήρηση πλήθους, την ευρετηρίαση περιεχομένου βίντεο, την προσωπική ταυτοποίηση (π.χ. δίπλωμα οδήγησης), τη σύνδεση πλάνων σε κούπες κ.λ.π.
- Η αναγνώριση προσώπου σε πραγματικό χρόνο. Εδώ, η αναγνώριση προσώπου χρησιμοποιείται για να ταυτοποιήσει ένα άτομο επί τόπου και επιτρέπει την πρόσβαση σε ένα κτίριο ή σε μία χημική ένωση, αποφεύγοντας έτσι ταλαιπωρίες ασφάλειας. Σε αυτή την προσέγγιση, το πρόσωπο συγκρίνεται έναντι πολλαπλών δειγμάτων εκπαίδευσης ενός ατόμου.

## 2.4 Εξαγωγή Χαρακτηριστικών και Συμπίεση Δεδομένων

Η εξαγωγή χαρακτηριστικών και η συμπίεση δεδομένων αποτελούν σημαντικά προβλήματα στην αναγνώριση προτύπων και στην ανάλυση εικόνας. Αρκετές φορές ο στόχος είναι να παρατίθεται μία σειρά χαρακτηριστικών που αντιπροσωπεύουν τα δεδομένα όσο το δυνατόν εγγύτερα και να συμπιέζονται τα δεδομένα ταυτόχρονα. Υπάρχουν πολλές προσεγγίσεις για το πρόβλημα αυτό και μια από τις πιο γνωστές είναι η Ανάλυση Κύριων Συνιστωσών (PCA). Ομοίως, το πρόβλημα αυτό έχει αντιμετωπιστεί με γραμμικά και μη γραμμικά νευρωνικά δίκτυα PCA και σε δίκτυα αυτόματης σύνδεσης μη εποπτευόμενης μάθησης. (Oja, 2011) Κάθε ένα από αυτά θα συζητηθεί στα ακόλουθα τμήματα:

### 2.4.1 Ανάλυση Κυρίων Συνιστωσών (Principal Components Analysis - PCA).

Σε γενικές γραμμές, η PCA είναι μια στατιστική μέθοδος που χρησιμοποιείται για την εξαγωγή χαρακτηριστικών από ένα σύνολο δεδομένων με υψηλή διάσταση. Είναι μια λύση για την κατάρα του προβλήματος της διάστασης που βρέθηκε στην αναγνώριση προτύπων. (Engelbrecht, 2007) Η PCA είναι ένας γραμμικός, ορθογώνιος μετασχηματισμός ή μία προβολή των δεδομένων σε ένα νέο σύστημα συντεταγμένων όπου οι άξονες είναι ασυσχέτιστοι και η μέγιστη διακύμανση των αρχικών δεδομένων βρίσκεται σε έναν μικρό μόνο αριθμό των συντεταγμένων. (Rizon, 2006) Η μείωση της διάστασης επιτυγχάνεται στο χώρο αυτό με τη λήψη των συντεταγμένων που έχουν τη μέγιστη διακύμανση και αφήνοντας έξω τις συντεταγμένες με την ελάχιστη διακύμανση. Σε μαθηματικούς όρους, τα διανύσματα βάσης του νέου συστήματος συντεταγμένων είναι τα ιδιοδιανύσματα (eigenvectors) του πίνακα συνδιασποράς (covariance matrix) των δεδομένων και οι αποκλίσεις είναι οι αντίστοιχες ιδιοτιμές (eigenvalues). Έτσι από την άποψη της PCA, η καλύτερη προβολή, είναι ο καθορισμός ενός διανύσματος με πολύ χαμηλότερη διάσταση που προσεγγίζει καλύτερα κατά κάποιο τρόπο ένα δεδομένο διάνυσμα πληροφοριών. Επομένως, στην αναγνώριση προσώπου, λαμβάνει μία αναπαράσταση του διανύσματος διάστασης  $s$  από κάθε πρόσωπο ως δεδομένο και καθορίζει μία υπο-περιοχή διάστασης  $t$ , της οποίας η βάση διανύσματος ανταποκρίνεται μέγιστα στην αρχική εικόνα. Η διάσταση αυτής της νέας υπο-περιοχής είναι χαμηλότερη από την αρχική. ( $t \ll s$ ). Αν τα στοιχεία της αρχικής εικόνας θεωρούνται ως τυχαίες μεταβλητές, τότε οι κύριες συνιστώσες είναι κατά μήκος των ιδιοδιανυσμάτων που ανταποκρίνονται σε μεγαλύτερες ιδιοτιμές του πίνακα συσχέτισης και η ελαχιστοποίηση των λαθών πραγματοποιείται υπό μία ευρεία άποψη. (Turk & Pentland, 2010)



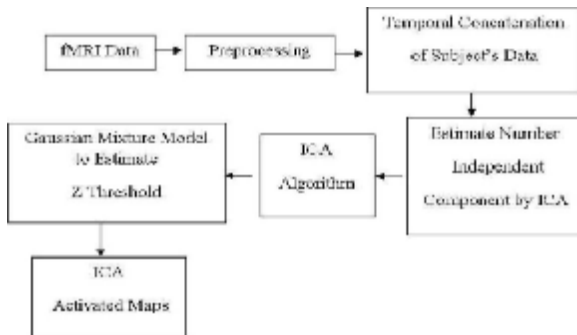
Οι ερευνητές στο Εργαστήριο Μέσων Πληροφόρησης του MIT και στο Ινστιτούτο Τεχνολογίας της Πολεμικής Αεροπορίας (AFIT) χρησιμοποιούν την ανάλυση κυρίων συνιστωσών στην ολιστική τους προσέγγιση για την αναγνώριση προσώπου, διότι πιστεύεται ότι ειδικά χαρακτηριστικά, όπως τα μάτια ή τη μύτη, μπορεί να μην είναι τόσο σημαντικά όσο το συνολικό σχήμα του προσώπου, όταν πρόκειται για την αναγνώριση. Μια λεπτομερής επισκόπηση του PCA χρησιμοποιώντας το μετασχηματισμό Karhunen-Loeve μπορεί να βρεθεί σε μια διπλωματική εργασία από τον Pedro Suarez. (Suarez, 2001) Αυτή η προσέγγιση υποστηρίζεται από την φυσιολογία και την ψυχολογία της διαδικασίας της αναγνώρισης προσώπου. (Turk & Pentland, 2010) (Valentine & Ferrara, 2012) Η παραπάνω διαδικασία έχει υπολογιστικά δαπανηρό υπολογισμό των ιδιοδιανυσμάτων και των συντελεστών προβολής. Ως εκ τούτου, το σύστημα αναγνώρισης προσώπου στο MIT λειτουργεί με τη βοήθεια τριών υπολογιστών: έναν επεξεργαστή εικόνας κύβου δεδομένων (Datacube), ένα Sun 3/160, και ένα Sun Sparcstation. Αυτό το σύστημα μπορεί να εκτελέσει το έργο αναγνώρισης σε ένα ρυθμό δύο ή τρεις φορές το δευτερόλεπτο, (Turk & Pentland, 2010) αλλά το μέγεθος του συστήματος το καθιστά μη πρακτικό για τις τρέχουσες εφαρμογές. Όπως και το σύστημα MIT, έτσι και το σύστημα AFIT φιλοξενείται σε πολλούς υπολογιστές. Το λογισμικό προεπεξεργασίας φιλοξενείται σε έναν υπολογιστή NEXT και σε έναν υπολογιστή Silicon Graphics και τουπόλοιπο λογισμικό τρέχει σε έναν δεύτερο υπολογιστή Silicon Graphics. Ωστόσο, ακόμη και με πολλαπλούς υπολογιστές, η διαδικασία εξακολουθεί να διαρκεί αρκετά λεπτά για να τρέξει.

## 2.5. Υφιστάμενες Τεχνικές

### 2.5.1. Ιδιοδιανύσματα προσώπων (Eigenfaces)

Πολλές προσεγγίσεις στο συνολικό πρόβλημα αναγνώρισης προσώπου έχουν ήδη διαμορφωθεί με την πάροδο των ετών, αλλά μία από τους πιο ακριβείς και γρήγορους τρόπους για να εντοπιστούν τα πρόσωπα είναι να χρησιμοποιήσουν την επονομαζόμενη τεχνική των ιδιοδιανυσμάτων προσώπου. Αυτή η τεχνική χρησιμοποιεί ένα συνδυασμό γραμμικής άλγεβρας και στατιστικής ανάλυσης για να δημιουργήσει μια σειρά από πρόσωπα βάσης - τα ιδιοδιανύσματα προσώπου - έναντι των οποίων δοκιμάζονται οι εισροές (Εικόνα 9). Τα ιδιοδιανύσματα προσώπων, ως επί το πλείστον, χρησιμοποιούν τα εργαλεία της PCA και της ICA για την αναγνώριση προσώπου. Χρησιμοποιώντας αυτά, η απόδοση είναι μεταβλητή. Αν και η ICA ξεπερνά σημαντικά το πρότυπο της PCA, έχει

υποστηριχθεί ότι η απόδοση της ICA εξαρτάται σε μεγάλο βαθμό από την εμπλεκόμενη διαδικασία της PCA. Η καθαρή προβολή της ICA έχει μικρή επίδραση στην απόδοση της αναγνώρισης προσώπου.



Εικόνα 9 Το διάγραμμα της ανάλυσης ICA

### 2.5.2.Απεικόνιση Εύρους (Range Imaging)

Πρόκειται για μια συλλογή τεχνικών, που χρησιμοποιούνται για την παραγωγή 2D εικόνας, που δείχνει την απόσταση από ένα σύνολο σημείων σε μια σκηνή από ένα συγκεκριμένο σημείο, συνήθως συνδεδεμένο με κάποιο τύπο συσκευής αισθητήρα. Η προκύπτουσα εικόνα, η *εικόνα εύρους*, έχει τιμές εικονοστοιχείων που αντιστοιχούν με την απόσταση, π.χ., φωτεινότερες τιμές σημαίνουν μικρότερη απόσταση, ή αντιστρόφως. Εάν ο αισθητήρας, ο οποίος χρησιμοποιείται για την παραγωγή της εικόνας εύρους βαθμονομηθεί σωστά, οι τιμές των εικονοστοιχείων μπορούν να αποδοθούν άμεσα σε φυσικές μονάδες, όπως τα εκατοστά (Εικόνα 10). (Chen, 2000)



Εικόνα 10 Ο αισθητήρας εικόνας παρέχει υψηλού δυναμικού εικόνα εύρους

### 2.5.3. Χαρτογράφηση των άκρων της γραμμής (Line edge map)

Πρόκειται για έναν αλγόριθμο αναγνώρισης προσώπου βασισμένος σε εικόνα που χρησιμοποιεί μια σειρά από τυχαία ευθύγραμμα τμήματα γραμμών του προσώπου σε 2D προβολές εικόνας ως την υποκείμενη αναπαράσταση της εικόνας, μαζί με τον πλησιέστερο γειτονιάζων ταξινομητή, όπως το συνδυασμό συστήματος γραμμής. Ο συνδυασμός των 1D τμημάτων γραμμής εκμεταλλεύεται την εγγενή συνοχή σε μία ή περισσότερες προβολές 2D εικόνας προσώπου στη σφαίρα προβολής(Εικόνα 11). (Tarr, 1991)



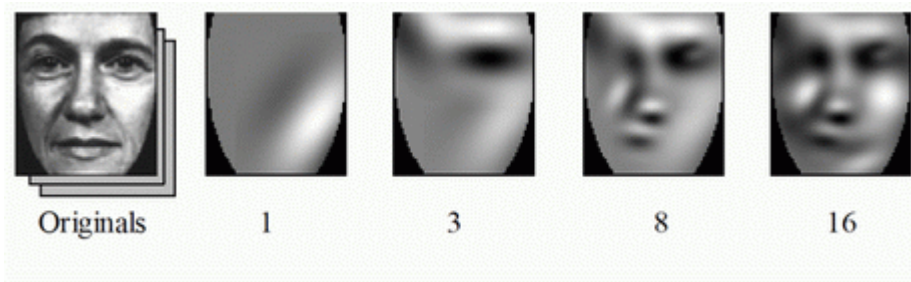
Εικόνα 11 Χαρτογράφηση των άκρων της γραμμής (Line edge map)

### 2.5.4. Τεχνικές Αναγνώριση Προσώπου με βάση τα Νευρωνικά Δίκτυα (Neural Network based Face Recognition Techniques)

Τα νευρωνικά δίκτυα χρησιμοποιούνται για τη δημιουργία βάσης δεδομένων προσώπου και να αναγνωρίσουν το πρόσωπο. Ένα ξεχωριστό δίκτυο για κάθε άτομο είναι κατασκευασμένο. Το πρόσωπο εισόδου προβάλλεται πρώτα στο χώρο των ιδιοδιανυσμάτων προσώπων για να λάβει ένα νέο περιγραφέα. Αυτός ο περιγραφέας χρησιμοποιείται ως είσοδος δικτύου και εφαρμόζεται στο δίκτυο του κάθε ατόμου. Επιλέγεται αυτός με τη μέγιστη έξοδο και αναφέρεται ως ο υποδοχέας εάν είναι μεγαλύτερος από ένα προκαθορισμένο όριο αναγνώρισης(Εικόνα.1,2 &15). (Turk & Pentland, 2010) (Samal & Prasana, 2012) (Damasio, 2005) (Meadows, 2004)

### 2.5.5. Δίκτυα Κυμάτων Gabor (Gabor Wavelet Networks - GWN)

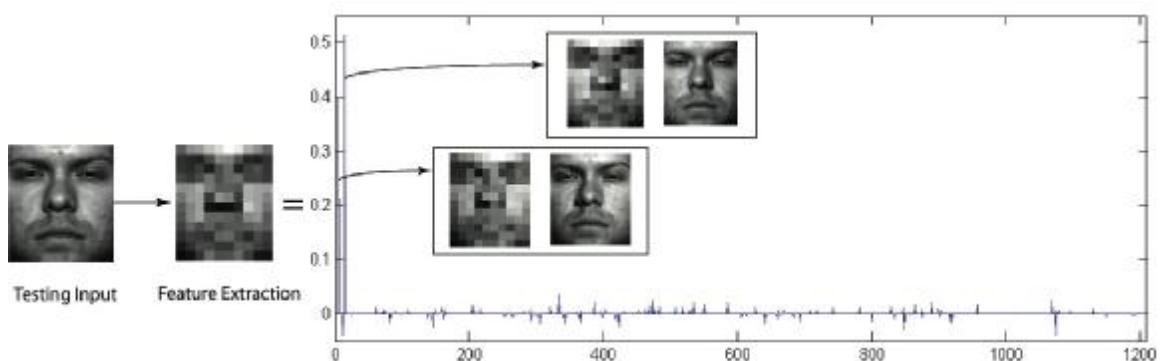
Το Δίκτυο Κυμάτων Gabor χρησιμοποιείται για μια αποτελεσματική αναπαράσταση του αντικειμένου (Εικόνα 12). Έχει αρκετά πλεονεκτήματα, όπως το αναλλοίωτο σε κάποιο βαθμό σε σχέση με την μετάφραση, την περιστροφή και τη διαστολή. Επιπλέον, έχει τη δυνατότητα να γενικευθεί και να απομακρυνθούν από τα δεδομένα εκπαίδευσης και να διαβεβαιώσει, για ένα δεδομένο δίκτυο. Πρόκειται για μία τεχνική γενίκευσης των Μεθόδων Πυρήνα.



Εικόνα 12 Δίκτυο Κυμάτων Gabor

### 2.5.6. Αραιά Εκπροσώπηση (Sparse Representation)

Η εικόνα εισόδου διαιρείται με ελαχιστοποίηση L-1. Η αραιώση συγκρίνεται με τη σποραδικότητα των δεδομένων εκπαίδευσης (Εικόνα 13). Υπάρχουν, επίσης, πολλές τεχνικές που έχουν προσπαθήσει να χρησιμοποιήσουν για τις τεχνικές αναγνώρισης προσώπου, αλλά μεταξύ όλων αυτών, η τεχνική των ιδιοδιανυσμάτων προσώπου δείχνει τα ταχύτερα και πιο ακριβή αποτελέσματα από ό,τι οι άλλες τεχνικές. Μέχρι στιγμής, η τεχνική των ιδιοδιανυσμάτων προσώπου έχει την ταχύτερη απόδοση από ό,τι άλλες τεχνικές. Τα τελευταία χρόνια οι ερευνητές και οι επιστήμονες προσπαθούν να επικεντρωθούν στα Ανθρώπινα Συστήματα Όρασης (Human Vision System HVS) για την αναγνώριση προσώπου. Ακόμη, δεν έχει εφαρμοστεί στην πράξη στον τομέα της αναγνώρισης προσώπου ακόμα, αλλά θέτει τα θεμέλια για μελλοντικές μελέτες. (Payne, 2002)



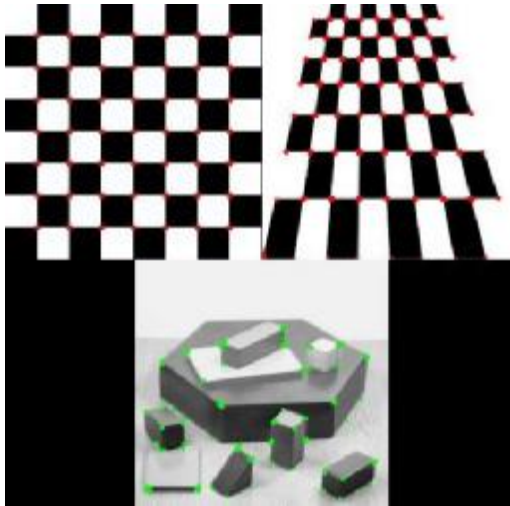
Εικόνα 13 Προοπτική Αραιής Εκπροσώπησης

### 2.5.7. Turk και Pentland από το Εργαστήριο Μέσων Πληροφόρησης στο MIT (Massachusetts Institute of Technology).

Οι Matthew Turk και ο Alex Pentland, από το Εργαστήριο Μέσων Ενημέρωσης Τεχνολογίας στο Ινστιτούτο της Μασαχουσέτης, έχουν θέσει σε εφαρμογή ένα σύστημα αυτόνομης αναγνώρισης προσώπου, που χρησιμοποιεί, επίσης, το Μετασχηματισμό της Karhunen-Loeve (Karhunen-Loeve Transform). (Turk & Pentland, 2010) Ενώ το σύστημα έχει καλή απόδοση, η εφαρμογή του είναι περιορισμένη λόγω της τεράστιας ποσότητας της υπολογιστικής ισχύος που είναι απαραίτητη για να λειτουργήσει το σύστημα, καθώς χρειάζεται τρεις εξειδικευμένους, υψηλής ταχύτητας επεξεργαστές.

### 2.5.8. Αναγνώριση Προσώπου στο Ινστιτούτο Τεχνολογίας της Πολεμικής Αεροπορίας στο AFIT (AirForce Institute of Technology).

Η Αναγνώριση προσώπου στο AFIT ξεκίνησε το 1985 με πρώιμα συστήματα αργής ταχύτητας, τα οποία απαιτούσαν σημαντικό μέρος της ανθρώπινης επέμβασης. (Wright, 2009) (Nan & W, 2009) Επί σειράς ετών, αρκετές βελτιώσεις προστέθηκαν και το σύστημα έχει εξελιχθεί σε αυτό που είναι σήμερα. (Harandi, 2004) (Tou & Rafei, 2004) Το υφιστάμενο σύστημα είναι παρόμοιο με το σύστημα του MIT που αναπτύχθηκε από τους Turk και Pentland. (Turk & Pentland, 2010) Ωστόσο, το σύστημα στο AFIT έχει μία σημαντική βελτίωση έναντι του συστήματος στο MIT. Δηλαδή, ο τρέχων προεπεξεργαστής του AFIT περιλαμβάνει μια διαδρομή κεντραρίσματος και ένα παράθυρο διαδρομής Gauss (Gaussian window routine) (Εικόνα 14). Αυτές οι διαδρομές επικεντρώνουν την εικόνα και σχεδιάζουν ένα παράθυρο Gauss γύρω από αυτή, με αποτέλεσμα τη μείωση της έμφασης το φόντο. Με αυτή τη βελτίωση, το σημερινό σύστημα του AFIT είναι 95 τοις εκατό επιτυχές στον εντοπισμό 55 εικόνων προσώπου. (Suarez, 2001) Το ισχύον σύστημα του AFIT χρησιμοποιείται ως σημείο αναφοράς για αυτή την εργασία.



Εικόνα 14 Παράθυρο Διαδρομής Gauss

#### 2.5.9. Ο Cottrell στο UCSD.

Ο Garrison Cottrell του Πανεπιστημίου της Καλιφόρνια στο Σαν Ντιέγκο έχει αναπτύξει ένα νευρωνικό μοντέλο προσέγγισης του δικτύου για την αναγνώριση. **Σφάλμα! Το αρχείο προέλευσης της αναφοράς δεν βρέθηκε.** Το μοντέλο του έχει εκπαιδευτεί σε μία προσέγγιση δύο σταδίων που χρησιμοποιεί ένα δίκτυο ταυτότητας για εξαγωγή χαρακτηριστικών γνωρισμάτων και ένα δίκτυο με μονό στρώμα ανάστροφης διάδοσης για την ταξινόμηση. Το δίκτυο αυτό έχει παρουσιάσει επιτυχία σε περιορισμένες βάσεις δεδομένων.

#### 2.5.10. Εθνικό Εργαστήριο Los Alamos

Μία άλλη νευρωνική δικτυακή προσέγγιση στην αναγνώριση προσώπου αναπτύχθηκε στο Εθνικό Εργαστήριο Los Alamos. **Σφάλμα! Το αρχείο προέλευσης της αναφοράς δεν βρέθηκε.** Το εργαστήριο χρησιμοποιεί ένα τυπικό δίκτυο με μονό στρώμα ανάστροφης διάδοσης, χρησιμοποιώντας τα πρώτα δεδομένα εικόνας ως εισροές (inputs). Επί του παρόντος, το δίκτυο περιορίζεται σε δύο προβλήματα κατηγορίας σε περιορισμένη βάση δεδομένων, αλλά τα προκαταρκτικά αποτελέσματα είναι πολύ αξιόσεβαστα.

## 2.6.Εξελικτική Αναζήτηση (Evolutionary Pursuit)

Πρόκειται για μια προσέγγιση που βασίζεται στο χώρο, ψάχνοντας για το καλύτερο σύνολο των αξόνων προβολής προκειμένου να μεγιστοποιηθεί μια λειτουργία καταλληλότητας, μετρώντας κατά την ίδια στιγμή την ακρίβεια ταξινόμησης και την ικανότητα γενίκευσης του συστήματος (Εικόνα 15). Επειδή η διάσταση του χώρου των λύσεων είναι πολύ μεγάλη, αυτό λύνεται χρησιμοποιώντας ένα συγκεκριμένο είδος γενετικού αλγορίθμου που ονομάζεται Εξελικτική Αναζήτηση. (Foldiak, 2009) (Kung & Diamantaras, 1990)










Εικόνα 15 Προσέγγιση της Εξελικτικής Αναζήτησης

## 2.7.Αντιπαραβολή Γραφικής Παράστασης Ελαστικής Δέσμης (Elastic Bunch Graph Matching)

Όλα τα ανθρώπινα πρόσωπα μοιράζονται μια παρόμοια τοπολογία. Τα πρόσωπα παρουσιάζονται ως γραφικές παραστάσεις, με κόμβους που τοποθετούνται σε χαρακτηριστικά σημεία (μάτια, μύτη, κλπ) και ακμές που επισημαίνονται με διανύσματα αποστάσεως 2-D (Εικόνα 4). Κάθε κόμβος περιέχει ένα σύνολο από 40 σύνθετους κυματοειδείς συντελεστές Gabor σε διαφορετικές κλίμακες και κατευθύνσεις (φάση, πλάτος), οι οποίοι ονομάζονται «πίδακες». Η αναγνώριση βασίζεται σε ένα επισημασμένο γράφημα, δηλαδή ένα σύνολο κόμβων που συνδέεται με τις ακμές. Κάθε κόμβος χαρακτηρίζεται ως πίδακας και οι ακμές ως αποστάσεις. Η προσέγγιση αυτή αποτελεί μία γενίκευση του *Νευρωνικού Δικτύου Cottrell*. (Sanger, 2009)

## 2.8. Μέθοδοι Πυρήνα (Kernel Methods)

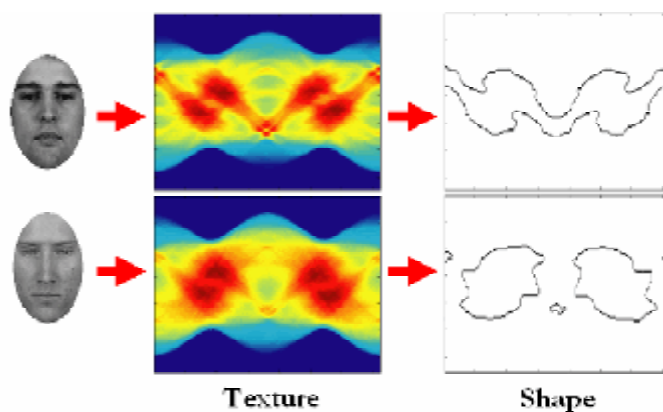
Η πολλαπλότητα του προσώπου στον υπο-χώρο δεν χρειάζεται να είναι γραμμική. Οι μέθοδοι του πυρήνα είναι μία γενίκευση των γραμμικών μεθόδων. Τα άμεσα μη-γραμμικά πολλαπλά σχεδιαγράμματα διερευνώνται για την εκμάθηση αυτής της μη γραμμικής πολλαπλότητας. Η προσέγγιση αυτή αποτελεί μία γενίκευση της *Εξαγωγής Χαρακτηριστικών και Συμπίεσης Δεδομένων σε Μη Γραμμικά Δίκτυα* (Εικόνα 16). (Wright, 2009) (Nan & W, 2009) (Geschwind, 2009) (Rolls, 2009)

True map	GC	MoG	k-NN	One-Class SVM	SVM	LapSVM
						
$\kappa$ stat. (accuracy)	0.58 (66.2)	0.61 (69.7)	0.48 (59.1)	0.63 (77.3)	0.77 (93.2)	0.78 (94.3)

Εικόνα 16 Ο αληθινός χάρτης της Ρώμης και χάρτες ταξινόμησης που λαμβάνονται με διάφορες μεθόδους πυρήνα

## 2.9. Μετασχηματισμός Ίχνους (Trace Transform)

Ο μετασχηματισμός Ίχνους αποτελεί ένα νέο εργαλείο για την επεξεργασία εικόνας που μπορεί να χρησιμοποιηθεί για την αναγνώριση αντικειμένων κάτω από μετασχηματισμούς, π.χ. την περιστροφή, τη μετάφραση και την κλιμάκωση (Εικόνα 17). Για να παραχθεί ο μετασχηματισμός ίχνους, θα πρέπει να υπολογιστεί μια λειτουργία της εικόνας κατά τον εντοπισμό των γραμμών της. Διαφορετικοί μετασχηματισμοί ίχνών μπορούν να παραχθούν από μια εικόνα χρησιμοποιώντας διαφορετικές λειτουργίες ίχνους. (Baldi & Hornik, 2009) (Eberhart & Yuhui, 2007)



Εικόνα 17 Πληροφορίες που προέρχονται από το Μετασχηματισμό Ίχνους



## 2.10.Πρόσθετες Έρευνες Δικτύου (Additional Research Network)

Ένα σύστημα αναγνώρισης εικόνας που χρησιμοποιεί ένα υπέρυθρο σύστημα φωτισμού για να ξεπεραστεί ο φωτισμός του περιβάλλοντος αναπτύχθηκε στην Αγγλία. (Farahati, 2002) Το σύστημα βασίζεται σε ένα νευρωνικό δίκτυο γνωστό ως WISARD και φέρεται στην ικανότητά του να αναγνωρίζει ένα πρόσωπο από σκηνές φόντου. Το μειονέκτημα αυτού του συστήματος είναι η πολυπλοκότητα: απαιτεί δύο φωτογραφικές μηχανές, υλικό ελέγχου, ένα σύστημα φωτισμού IR, εκπαίδευση για το λογισμικό και το υλικό WISARD. Επίσης, απαιτεί ένα σημαντικό ποσό μνήμης, επειδή οι δύο εικόνες 800x540 λαμβάνονται για κάθε σκηνή.

Επιπλέον προς τη βάση του συστήματος WISARD, ένα νευρωνικό δίκτυο για την αναγνώριση προσώπου με βάση ένα πολυστρωματικό νευρωνικό δίκτυο και τα κοινά βάρη έχει εισαχθεί στη Γαλλία. (Bouattour, 2002) Πιστεύεται ότι αυτό το σύστημα χειρίζεται τις αλλαγές στο φωτισμό και τις περιστροφές και τις μεταφράσεις στο πρόσωπο με την προσθήκη επιπλέον πρωτοτύπων. Προκαταρκτικά αποτελέσματα των δοκιμών έχουν δώσει αποδεκτά αποτελέσματα, αλλά η βάση δεδομένων αποτελείται από μόνο 10 άτομα. Το σύστημα αυτό θα απαιτήσει, επίσης, μεγάλες ποσότητες μνήμης, επειδή οι εικόνες που χρησιμοποιήθηκαν ήταν 256x256 εικονοστοιχεία, με 70 πρωτότυπα που λαμβάνονται ανά άτομο. Αυτό ισοδυναμεί με περίπου 640 MB δεδομένων.

## 2.11.Συμπεράσματα

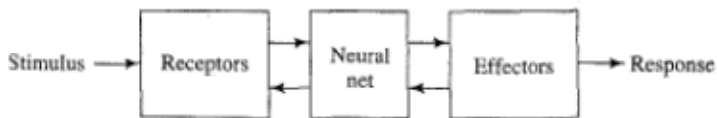
Στο κεφάλαιο αυτό παρουσιάζονται διεξοδικά το πρόβλημα της αναγνώρισης προσώπου, η σημαντικότητά του, καθώς και οι διάφορες προσεγγίσεις που έχουν προταθεί, μέσα από μια ανασκόπηση της υπάρχουσας βιβλιογραφίας σχετικά με τα συστήματα αναγνώρισης προσώπου, με πρωταρχική έμφαση στην εξαγωγή χαρακτηριστικών, τη συμπίεση δεδομένων και την ταξινόμηση. Η καθεμία από αυτές αναφέρθηκε στην ορθή εκτέλεση για την αναγνώριση προσώπου, ωστόσο, οι βάσεις δεδομένων που χρησιμοποιούνται ήταν περιορισμένες σε μέγεθος και / ή στον αριθμό των ταξινομήσεων. Αυτοί οι περιορισμοί θα πρέπει να αντιμετωπιστούν με την αύξηση του μεγέθους των βάσεων δεδομένων και την αύξηση του αριθμού των ταξινομήσεων στις βάσεις δεδομένων. Επιπλέον, οι δυνατότητες των δικτύων αυτών για τα δεδομένα που λαμβάνονται κατά τη διάρκεια πολλών ημερών δεν έχει ακόμη ερευνηθεί.

Το Κεφάλαιο 3 παρουσιάζει τη μεθοδολογία που χρησιμοποιήθηκε για την αντιμετώπιση αυτών των ζητημάτων.

## ΚΕΦΑΛΑΙΟ 3<sup>ο</sup>

### 3.1 Εισαγωγή

Όπως έχει ήδη αναλυθεί στο 2<sup>ο</sup> Κεφάλαιο, τα *Νευρωνικά Δίκτυα* είναι μη γραμμικά μοντέλα των νευρικών οδών (pathways) στο νευρικό σύστημα. Η ροή των σημάτων στο νευρικό σύστημα φαίνεται στην *Εικόνα 18*.

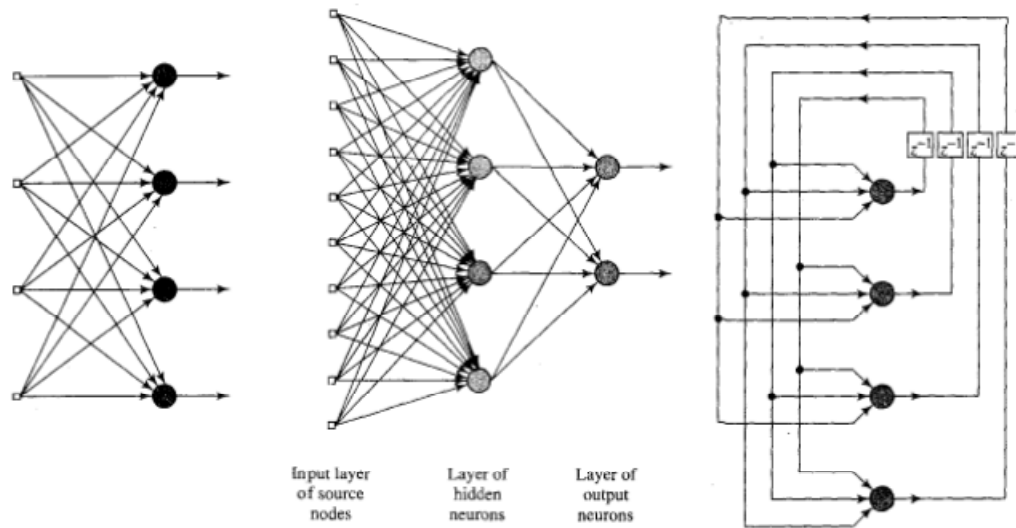


Εικόνα 18 Σχηματικό διάγραμμα του νευρικού συστήματος

Η διάδοση του ερεθίσματος προς οποιαδήποτε κατεύθυνση σε όλο το σύστημα των αισθητηρίων οργάνων, οφείλεται στην ενεργοποίηση των απλών στοιχείων που λειτουργούν παράλληλα. Αυτά τα στοιχεία μπορεί να είναι δομημένα ως εξής:

- Νευρωνικά Δίκτυα Εμπρόσθιας Τροφοδότησης ενός επιπέδου (Single-Layer Feedforward Networks)
- Πολυστρωματικά Νευρωνικά Δίκτυα Εμπρόσθιας Τροφοδότησης (Multilayer Feedforward Networks)
- Επαναλαμβανόμενα Δίκτυα (Recurrent Networks)

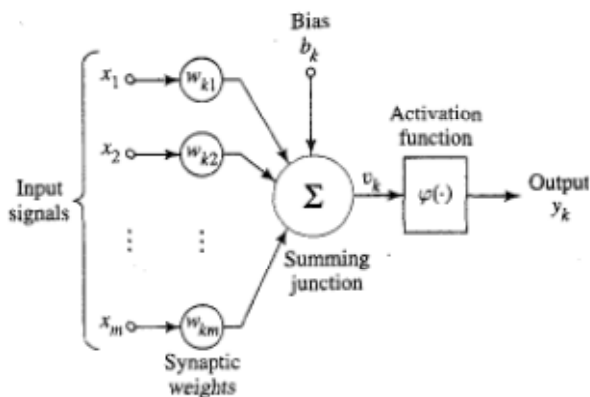
Αυτές οι δομές φαίνονται στην *Εικόνα 19*, στην οποία κάθε κόμβος αντιπροσωπεύει ένα μαθηματικό μοντέλο ενός νευρώνα.



Εικόνα 19 Νευρωνικά Δίκτυα Εμπρόσθιας Τροφοδότησης, Πολυστρωματικά Δίκτυα Προελέγχου και Επαναλαμβανόμενα Δίκτυα

Ο τρόπος με τον οποίο η εν λόγω διαβίβαση εξελίσσεται, καθορίζεται από τις ικανότητες μάθησης αυτών των νευρώνων. Στο 2<sup>ο</sup> Κεφάλαιο πραγματοποιήθηκε εκτενής αναφορά στα είδη εκπαίδευσης. Η εκπαίδευση μπορεί να είναι εκπαίδευση διόρθωσης σφάλματος, που βασίζεται σε μνημονική μάθηση, *Χεββιανή* μάθηση ή να βασίζεται σε πρότυπα όπως η επιβλεπόμενη μάθηση, η χωρίς επίβλεψη μάθηση, η *ενίσχυση της μάθησης* μέσω της συνεχούς αλληλεπίδρασης με το περιβάλλον.

Το πολυστρωματικό νευρωνικό δίκτυο του Rosenblatt είναι μια βασική μορφή ενός νευρωνικού δικτύου, κατασκευασμένο γύρω από ένα μη γραμμικό μοντέλο ενός νευρώνα, δηλαδή το *μοντέλο McCulloch-Pitts* ενός νευρώνα (Εικόνα 20).



Εικόνα 20 Μη γραμμικό μοντέλο ενός νευρώνα

Ο αθροιστής κόμβος του μοντέλου νευρώνα υπολογίζει ένα γραμμικό συνδυασμό των δεδομένων εισόδου που εφαρμόζεται στις συνάψεις και την τάση να είναι μεροληπτική, αποδιδόμενη από την τιμή  $b$ , κατά τη διάρκεια μιας εργασίας ταξινόμησης. Αυτό το επαγόμενο τοπικό πεδίο μετατρέπεται σε ένα δυαδικό σήμα,  $+1$  ή  $-1$ , το οποίο καθορίζει τη σταθμισμένη ικανότητα να ενεργοποιεί το γειτονικό νευρώνα.

Στην πιο γενική μορφή του, ένα νευρωνικό δίκτυο σχεδιάζεται για να διαμορφώσει τον τρόπο με τον οποίο ο εγκέφαλος εκτελεί μια συγκεκριμένη εργασία ή λειτουργία ενδιαφέροντος. Όπως αναφέρθηκε στο 2<sup>ο</sup> Κεφάλαιο, μοιάζει με τον εγκέφαλο στα ακόλουθα δύο σημεία:

1. Η γνώση αποκτάται από το δίκτυο από το περιβάλλον του μέσω μίας διαδικασίας εκπαίδευσης.
2. Ενδονευρικές δυνάμεις σύνδεσης, που είναι γνωστές ως συναπτικά βάρη, χρησιμοποιούνται για να αποθηκεύουν τις γνώσεις που αποκτήθηκαν.

Όπως αναλύθηκε στο 2<sup>ο</sup> Κεφάλαιο, η διαδικασία που χρησιμοποιείται για να εκτελέσει τη διαδικασία μάθησης ονομάζεται αλγόριθμος εκπαίδευσης, τροποποιώντας τα συναπτικά βάρη του δικτύου σε ένα μεθοδικό τρόπο προς την αναγνώριση προσώπου, με τη χρήση ενός τεχνητού νευρωνικού δικτύου ως προς την επίτευξη ενός επιθυμητού στόχου σχεδιασμού. Η τροποποίηση των συνοπτικών βαρών παρέχει την παραδοσιακή βάση για το σχεδιασμό των νευρωνικών δικτύων. Περαιτέρω, είναι, επίσης, δυνατόν για ένα νευρωνικό δίκτυο να τροποποιήσει τη δική του τοπολογία, η οποία δικαιολογείται από το γεγονός ότι οι νευρώνες στον ανθρώπινο εγκέφαλο μπορεί να πεθάνουν και ότι νέες συναπτικές συνδέσεις μπορούν να αναπτυχθούν.

### **3.2 Πολυστρωματικό Νευρωνικό Δίκτυο (Multilayer Perceptron - MLP)**

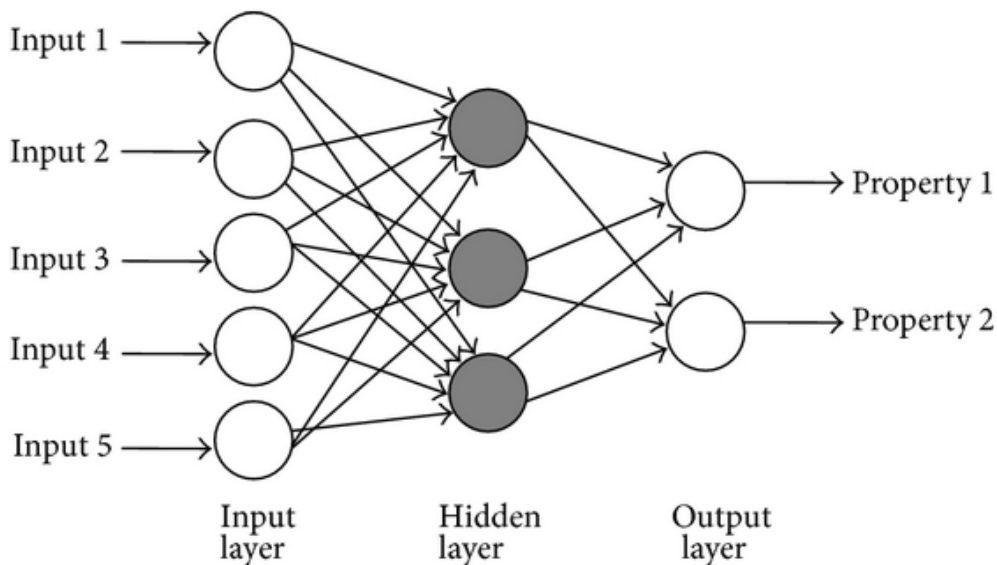
Τα πολυστρωματικά νευρωνικά δίκτυα είναι πολυστρωματικά δίκτυα εμπρόσθιας τροφοδότησης που αποτελούνται από τουλάχιστον τρία στρώματα: το στρώμα εισόδου, το κρυφό στρώμα (-τα) και το στρώμα εξόδου:

1. Το στρώμα εισόδου αποτελείται από είσοδο αισθητηρίων κόμβων στους οποίους παρουσιάζονται τα δεδομένα, εκπροσωπούμενα από τα διανύσματα της έντασης των εικονοκυττάρων της προ-επεξεργασίας εικόνων. Αυτά διαδίδονται μέσω του νευρωνικού

δικτύου σε μια κατεύθυνση προς τα εμπρός, σε μία βάση στρώμα προς στρώμα και, ως εκ τούτου, είναι γνωστά ως δίκτυα εμπροσθόδρομης τροφοδοσίας πολλαπλών στρώσεων.

2. Το κρυφό στρώμα εκτελεί ένα μη-γραμμικό μετασχηματισμό με το σήμα εισόδου σε ένα νέο χώρο που ονομάζεται ως χώρος "χαρακτηριστικού". Περισσότερα από ένα κρυφά στρώματα μπορεί να χρησιμοποιηθούν για να επιτευχθεί αυτός ο σκοπός. Δεδομένου ότι η μάθηση προχωρά, οι κρυμμένοι νευρώνες σταδιακά "ανακαλύπτουν" τα "χαρακτηριστικά" που αντιπροσωπεύουν τις εικόνες. Ο αριθμός των κρυμμένων νευρώνων σε κάθε κρυμμένο στρώμα και ο αριθμός των κρυφών στρωμάτων είναι κρίσιμοι για τη διαδικασία μάθησης και είναι ποικίλη για μεγαλύτερη ακρίβεια.

3. Στο στρώμα εξόδου, λαμβάνεται το σήμα (-τα) εξόδου / λειτουργίας, που εκφράζονται σαν μια συνεχής μη γραμμική συνάρτηση του σήματος εισόδου και τα συναπτικά βάρη που συνδέονται με αυτό το νευρώνα. Η *Εικόνα 21* δείχνει την αρχιτεκτονική γράφημα πολυστρωματικό νευρωνικό δίκτυο.



Εικόνα 21 Πολυστρωματικό Νευρωνικό Δίκτυο με ένα κρυφό στρώμα

Το δίκτυο φαίνεται να είναι πλήρως συνδεδεμένο, δηλαδή, ένας νευρώνας σε κάθε επίπεδο του δικτύου είναι συνδεδεμένο με όλους τους νευρώνες στο προηγούμενο στρώμα. (Bouattour, 2002)

Στην περίπτωση της εποπτευόμενης μάθησης, αυτό το σήμα εξόδου αφαιρείται από την επιθυμητή ανταπόκριση για να δώσει το σήμα σφάλματος. Στο πρόβλημα ταξινόμησης, η επιθυμητή ανταπόκριση αντιστοιχεί στις κατηγοριών των εικόνων. Για αυτό το συγκεκριμένο πρόβλημα, ο αριθμός των νευρώνων εξόδου ισούται με τον αριθμό των κατηγοριών των εικόνων. Η ανάστροφη διάδοση της εκτίμησης του διανύσματος κλίσης λάθους, δηλαδή, οι κλίσεις της επιφάνειας σφάλματος σε σχέση με τα βάρη που συνδέονται με τις εισόδους ενός νευρώνα, αποτελεί τον πυρήνα της εποπτευόμενης διαδικασίας της μάθησης. Στη φάση της εκπαίδευσης, λαμβάνονται τα βέλτιστα βάρη των κρυφών στρωμάτων και των στρωμάτων εξόδου, τα οποία ελαχιστοποιούν την εκτίμηση σφάλματος για τα επιθυμητά αποτελέσματα. Η εκμάθηση των βέλτιστων βαρών λαμβάνει χώρα με τη διάδοση του σήματος σφάλματος ανάστροφα και ενάντια στις συναπτικές συνδέσεις. Αυτός ο εποπτευόμενος αλγόριθμος μάθησης είναι γνωστός ως αλγόριθμος εσφαλμένης αναστροφής διάδοσης, ο οποίος περιγράφεται λεπτομερώς στην Ενότητα 3.3.

### 3.3 Προσδιορισμός του Αλγόριθμου Εσφαλμένης Ανάστροφης Διάδοσης

Ο εποπτευόμενος αλγόριθμος μάθησης απαιτεί την προσαρμογή των βαρών για να μειώσει το σφάλμα.

Έστω ότι:

$d_k(t)$ : η επιθυμητή έξοδος του  $k$  νευρώνα στη χρονική στιγμή  $t$

$x(t)$ : το διάνυσμα εισόδου που εφαρμόζεται στο στρώμα εισόδου και

$y_k(t)$ : η πραγματική έξοδος του  $k$  νευρώνα

Η διαφορά μεταξύ της επιθυμητής εξόδου και της πραγματικής εξόδου είναι:

$$e_k(t) = d_k(t) - y_k(t) \quad \Gamma.1$$

Με την συγκεκριμένη διαφορά ορίζουμε την συνάρτηση κόστους

$$I = E[0.5 \sum_k e_k^2(t)] \quad \Gamma.2$$

Το αποτέλεσμα αυτό είναι η συνάρτηση μέσου τετραγωνικού σφάλματος (Mean Square Error criterion, MSE) και εκφράζει την μέση τιμή του αθροίσματος των τετραγώνων των σφαλμάτων.

Το αποτέλεσμα που δίνει στο I είναι ότι η συνάρτηση αυτή εκτείνεται σε όλους του νευρώνες εξόδου, δηλαδή σε όλα τα k εξόδου. Το E[ ] είναι η αναμενόμενη μέση τιμή και θεωρείται ότι η εκπαίδευση γίνεται σε ένα περιβάλλον με άγνωστες κατανομές πιθανότητας.

Το πρόβλημα όμως είναι πως πρέπει να ελαχιστοποιηθεί το μέσο τετραγωνικό σφάλμα αφού πρώτα επιλεγθούν τα βάρη των νευρώνων.

Εμείς θα περιοριστούμε σε μια προσεγγιστική λύση για την επίλυση του προβλήματος αυτού γιατί η ακριβής λύση του προβλήματος απαιτεί την γνώση των στατικών ιδιοτήτων και στοχαστικών ανελίξεων που υπάρχουν σε κάθε περίπτωση. Οπότε έχουμε την συνάρτηση:

$$J = 0.5 \sum_k e_k^2(t) \quad \Gamma.3$$

Έτσι τα συναπτικά βάρη  $w_{kj}$  του ΤΝΔ, όπου  $w_{kj}$  είναι το βάρος της σύναψης j του νευρώνα k. Ο κανόνας μάθησης ο οποίος βασίζεται στη διόρθωση του σφάλματος είναι :

$$\Delta w_{kj}(t) = \gamma e_k(t) x_j(t) \quad \Gamma.4$$

όπου  $\Delta w_{kj}(t)$  είναι ο ρυθμός ανανέωσης στο χρόνο t του συναπτικού βάρους  $w_{kj}$  και  $\gamma$  είναι μια θετική σταθερά η οποία καθορίζει το ρυθμό μάθησης (learningrate). Ο κανόνας μάθησης ο οποίος στη βιβλιογραφία είναι γνωστός ως «κανόνας δέλτα» λέει ότι η διόρθωση του συναπτικού βάρους είναι του σφάλματος  $e_k(t)$  και του σήματος εισόδου  $x_j(t)$  της υπό εξέταση σύναψης j.

Η ανανεωμένη (νέα) τιμή  $w_{kj}(t+1)$  του θεωρούμενου συναπτικού βάρους δίνεται από τη σχέση:

$$w_{kj}(t + 1) = w_{kj}(t) + \Delta w_{kj}(t) \quad \Gamma.5$$

Για να μπορέσει να ξεκινήσει ο κανόνας μάθησης (Γ.5), θα πρέπει να γνωρίζουμε τις αρχικές τιμές των βαρών για την χρονική στιγμή t=0. Σε περίπτωση που το ΤΝΔ περιέχει μόνο γραμμικά στοιχεία προς επεξεργασία, αυτό σημαίνει πως το j είναι ακριβώς

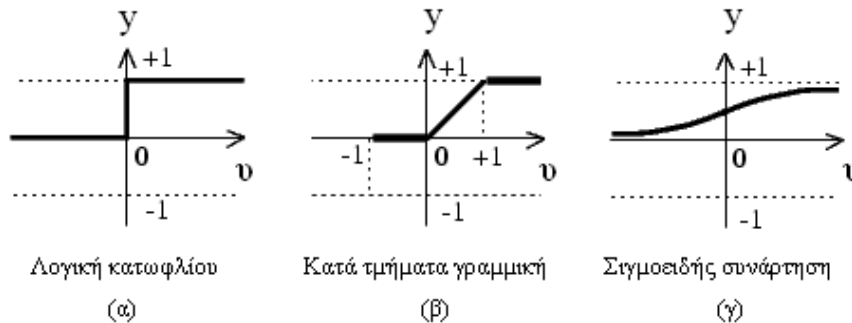


τετραγωνικό και αυτό έχει σαν αποτέλεσμα ότι ο αλγόριθμος οδηγεί βήμα προς βήμα στο ολικό ελάχιστο. Τώρα, σε περίπτωση που το ΤΝΔ περιέχει μη γραμμικά στοιχεία τότε δεν μπορούμε να έχουμε το ολικό ελάχιστο γιατί υπάρχει κίνδυνος ο αλγόριθμος να “εγκλωβισθεί” σε κάποιο τοπικό ελάχιστο.

Επιπρόσθετα, ο ρυθμός μάθησης έχει μεγάλη επίδραση στην απόδοση της μεθόδου, στην ταχύτητα της μάθησης αλλά και αποτέλεσμα. Για το λόγο αυτό η μάθηση διόρθωσης σφάλματος συμπεριφέρεται σαν να είναι σε ένα σύστημα κλειστής αναδρομής και πρέπει να επιλέγει με απόλυτη προσοχή η τιμή του ρυθμού μάθησης  $\gamma$ , ώστε να εξασφαλίσει την σταθερότητα της διαδικασίας. Σε περίπτωση που η τιμή  $\gamma$  έχει μεγάλη τιμή, η μάθηση επιταχύνεται αλλά υπάρχει ο κίνδυνος η διαδικασία να αποκλίνει και το σύστημα να γίνει ασταθές. Αντιθέτως, δηλαδή στην περίπτωση που το  $\gamma$  έχει μικρή τιμή, η μάθηση προχωράει κανονικά, αλλά υπάρχει μεγάλη πιθανότητα να πάρει πολύ χρόνο στο σύστημα μέχρι να φτάσει σε σταθερή λύση.

### 3.4 Συνάρτηση ενεργοποίησης (Activation function)

Μία συνάρτηση ενεργοποίησης μπορεί να έχει κάποια από τις παρακάτω μορφές:



Εικόνα 22 -Συνήθεις μορφές συνάρτησης ενεργοποίησης νευρώνων.

Οι συναρτήσεις της εικόνας 22 ορίζονται ως :

- Λογική κατωφλίου (Εικόνα 22 α)

$$f(u) = \begin{cases} 1, & u \geq 0 \\ 0, & u < 0 \end{cases}$$

Ο αντίστοιχος νευρώνας έχει τις εξισώσεις:

$$v = \sum_{j=1}^n w_j x_j - \theta$$

$$y = \begin{cases} 1, & u \geq 0 \\ 0, & u < 0 \end{cases}$$

- Κατά τμήματα γραμμική συνάρτηση (Εικόνα 22 β)

$$f(u) = \begin{cases} 1, & u \geq 1 \\ u, & 0 < u < 1 \\ 0, & u \leq 0 \end{cases} \quad (\Gamma.6)$$

Η συνάρτηση αυτή παριστά ένα γραμμικό ενισχυτή με κορεσμό και μεταπίπτει στη συνάρτηση κατωφλίου αν η απολαβή (ενίσχυση) του ενισχυτή γίνει πάρα πολύ μεγάλη (θεωρητικά άπειρη).

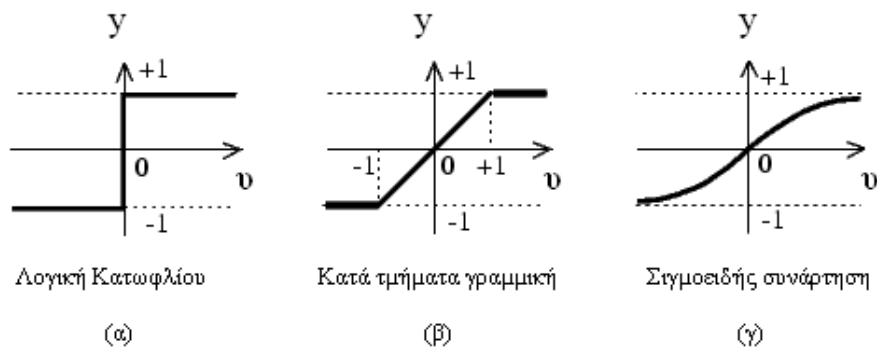
- Σιγμοειδής συνάρτηση (Εικόνα 22 γ)

Η σιγμοειδής συνάρτηση είναι η ευρύτερα χρησιμοποιούμενη συνάρτηση ενεργοποίησης και μπορεί να οριστεί με πολλούς τρόπους. Ένας από αυτούς είναι η λογιστική (logistic) συνάρτηση:

$$f(u) = \frac{1}{1 + \exp(-u)} \quad (\Gamma.7)$$

όπου  $\lambda$  είναι η «παράμετρος κλίσης» της σιγμοειδούς συνάρτησης. Για διάφορα  $\lambda$  παίρνουμε σιγμοειδείς συναρτήσεις με διαφορετική κλίση. Η κλίση στην αρχή ( $u=0$ ) είναι ίση με  $\lambda/4$ .

Οι συναρτήσεις ενεργοποίησης της Εικόνας 22 παίρνουν τιμές μεταξύ 0 και 1, δηλαδή  $0 \leq f(u) \leq 1$ . Εκτός από αυτές σε πολλές εφαρμογές χρειαζόμαστε τιμές στο διάστημα  $[-1, +1]$ , οπότε έχουμε τις συναρτήσεις της Εικόνας 23:



Εικόνα 23 Συναρτήσεις ενεργοποίησης με τιμές στο διάστημα  $[-1, +1]$ .

Οι συναρτήσεις αυτές περιγράφονται από τις ακόλουθες εξισώσεις.

- Συνάρτηση κατωφλίου {συνάρτηση προσήμου} (Εικόνα 23 α)

$$f(u) = \begin{cases} +1, & u > 0 \\ 0, & u = 0 \\ -1, & u < 0 \end{cases} \quad (\Gamma.8)$$

- Κατά τμήματα γραμμική συνάρτηση (Εικόνα 23 β)

$$f(u) = \begin{cases} 1, & u \geq +1 \\ u, & -1 < u < +1 \\ -1, & u \leq -1 \end{cases}$$

- Σιγμοειδής συνάρτηση (υπερβολική εφαπτομένη) (Εικόνα 23 γ)

$$f(u) = \tanh\left(\frac{u}{2}\right) = \frac{1 - \exp(-u)}{1 + \exp(-u)} \quad (\Gamma.9)$$

### 3.5 Λειτουργίες της Εκπαίδευσης (Modes of Learning)

Από μία ομοιόμορφη κατανομή τυχαίων αριθμών, χρησιμοποιούνται κάποια σετ προτύπων μέχρι να επιτευχθεί ο επιθυμητός στόχος. Το κάθε σετ που εξαντλείται αντιστοιχεί σε μία εποχή (epoch) της μάθησης. Η μάθηση θα τερματιστεί όταν το σφάλμα της τιμής της εξόδου μειωθεί στο καθορισμένο όριο. Η διαδικασία μάθησης διατηρείται σε μια κατά βάση εποχή μέχρι τα συναπτικά βάρη (synaptic weights) και τα επίπεδα πόλωσης (bias levels) του δικτύου να σταθεροποιηθούν. Ο μέσος όρος του μέσου τετραγωνικού σφάλματος πάνω σε ολόκληρο το σύνολο εκπαίδευσης πρέπει να συγκλίνει σε κάποια ελάχιστη τιμή. Είναι καλή πρακτική να τυχαιοποιηθεί η σειρά παρουσίασης των παραδειγμάτων εκπαίδευσης από μία εποχή στην επόμενη.

Για ένα δεδομένο σύνολο εκπαίδευσης, η μάθηση ανάστροφης διάδοσης μπορεί να προχωρήσει με ένα από τους δύο βασικούς τρόπους:

1. *Διαδοχική Λειτουργία (Sequential mode)*: Αναφέρεται, επίσης, ως on-line, πρότυπη ή στοχαστική. Σε αυτόν τον τρόπο λειτουργίας η ενημέρωση των βαρών πραγματοποιείται μετά την παρουσίαση του εκάστοτε παραδείγματος εκπαίδευσης.
2. *Σωρηδόν Τρόπος – Λειτουργία φουρνιάς προτύπων (Batch mode)*: Η ενημέρωση των βαρών πραγματοποιείται μετά την παρουσίαση όλων των παραδειγμάτων εκπαίδευσης

που αποτελούν μία εποχή. Ειδικότερα, η συνάρτηση κόστους ορίζεται ως το κατά μέσο όρο τετραγωνικό σφάλμα των εξισώσεων (Γ.2) και (Γ.3):

$$\bar{E} = 1/N \sum_{n=1}^N e_j(n) \quad (\Gamma.10)$$

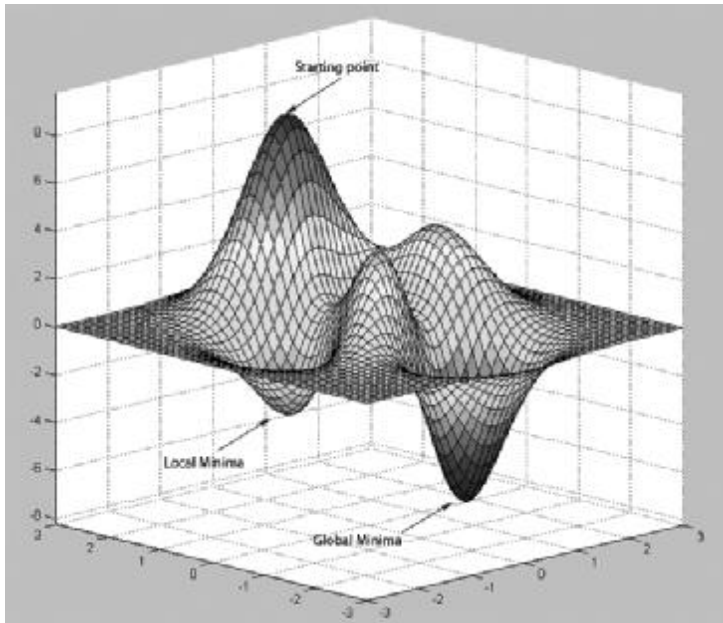
όπου το σήμα σφάλματος σχετίζεται με το νευρώνα εξόδου  $J$  για την εκπαίδευση παράδειγματος  $n$

Για μια on-line λειτουργική άποψη, η διαδοχική λειτουργία εκπαίδευσης προτιμάται έναντι της λειτουργίας φουρνιάς προτύπων, επειδή απαιτεί λιγότερη τοπική αποθήκευση για κάθε συναπτική σύνδεση. Επιπλέον, δεδομένου ότι τα πρότυπα παρουσιάζονται στο δίκτυο με τυχαίο τρόπο, η χρήση του προτύπου για την ενημέρωση των βαρών κάνει την αναζήτηση στο χώρο τους φυσικά στοχαστική. Αυτή με τη σειρά της καθιστά λιγότερο πιθανό τον αλγόριθμο οπισθοδιάδοσης να είναι παγιδευμένος σε ένα τοπικό ελάχιστο. (Baldi & Hornik, 2009)

Κατά τον ίδιο τρόπο, η στοχαστική φύση της διαδοχικής λειτουργίας καθιστά δύσκολη τη θέσπιση θεωρητικών προϋποθέσεων για τη σύγκλιση του αλγορίθμου. Σε αντίθεση, η χρήση της λειτουργίας φουρνιάς προτύπων παρέχει μια ακριβή εκτίμηση του διανύσματος κλίσης. Η σύγκλιση σε ένα τοπικό ελάχιστο είναι εγγυημένη κάτω από απλές συνθήκες.

### 3.6 Τοπικό Ελάχιστο της Επιφάνειας Σφάλματος (Local Minima of Error Surface)

Μια ιδιαιτερότητα της επιφάνειας σφάλματος (Εικόνα 24) που επηρεάζει την απόδοση του αλγόριθμου ανάστροφης διάδοσης είναι η παρουσία των τοπικών ελαχίστων επιπλέον των ολικών ελαχίστων. Διατρέχει τον κίνδυνο να παγιδευτεί σε ένα τοπικό ελάχιστο, όπου ακόμη και μικρή αλλαγή στα συναπτικά βάρη αυξάνει τη συνάρτηση κόστους. Αλλά κάπου αλλού στο χώρο των βαρών υπάρχει ένα άλλο σύνολό τους, για το οποίο η συνάρτηση κόστους είναι μικρότερη από το τοπικό ελάχιστο στο οποίο το δίκτυο έχει κολλήσει. Είναι σαφώς ανεπιθύμητο να τερματιστεί η διαδικασία μάθησης σε τοπικό ελάχιστο, ειδικά αν βρίσκεται πολύ πάνω από ένα ολικό ελάχιστο.



Εικόνα 24 Επιφάνεια Σφάλματος

### 3.7 Μηχανισμοί Διανυσμάτων Υποστήριξης (Support Vector Machines)

Οι μηχανισμοί διανυσμάτων υποστήριξης είναι καθολικά δίκτυα πρόσθιας τροφοδότησης, οι οποίοι μπορούν να χρησιμοποιηθούν για το πρότυπο ταξινόμησης και τη μη-γραμμική παλινδρόμηση. Βασικά ο μηχανισμός διανυσμάτων υποστήριξης είναι ένας γραμμικός μηχανισμός (linear machine) με κάποιες κατάλληλες ιδιότητες. Για να εξηγηθεί η λειτουργία του, είναι ίσως πιο εύκολο να πραγματοποιηθεί με την περίπτωση ξεχωριστών προτύπων που ενδέχεται να προκύψουν στο πλαίσιο της ταξινόμησης προτύπων. Στο πλαίσιο αυτό, η βασική ιδέα ενός μηχανισμού διανυσμάτων υποστήριξης είναι η κατασκευή μιας επιφάνειας κρίσης, που ονομάζεται υπερεπίπεδο (hyperplane), κατά τέτοιο τρόπο ώστε το περιθώριο του διαχωρισμού μεταξύ θετικών και αρνητικών παραδειγμάτων να μεγιστοποιείται.

Ο μηχανισμός επιτυγχάνει αυτή την επιθυμητή ιδιότητα, ακολουθώντας μια προσέγγιση βασισμένη σε αρχές με ρίζες στη στατιστική θεωρία μάθησης. Πιο συγκεκριμένα, ο μηχανισμός διανυσμάτων υποστήριξης είναι μία προσέγγιση της εφαρμογής της μεθόδου των διαρθρωτικών ελαχιστοποίησης του δομικού ρίσκου (method of structural risk minimization). Η αρχή της επαγωγής βασίζεται στο γεγονός ότι το ποσοστό λάθους ενός μηχανισμού μάθησης σε δεδομένα δοκιμών οριοθετείται από το

άθροισμα του ποσοστού εκπαίδευσης-λάθους και ένας όρος που εξαρτάται από τη διάσταση *Vapnik-Chervonenkis (VC)*.

Μια έννοια που έχει κεντρική σημασία για την κατασκευή του αλγορίθμου μάθησης του μηχανισμού διανυσμάτων υποστήριξης είναι ο πυρήνας του εσωτερικού γινομένου μεταξύ του «διανύσματος υποστήριξης»  $\mathbf{x}_i$  και του διανύσματος  $\mathbf{x}$  που προέρχεται από το χώρο εισόδου. Τα διανύσματα υποστήριξης αποτελούνται από ένα μικρό υποσύνολο των δεδομένων εκπαίδευσης που απορρέει από τον αλγόριθμο. Ανάλογα με το πώς αυτός ο πυρήνας εσωτερικού γινομένου παράγεται, η κατασκευή των διαφόρων μηχανισμών μάθησης χαρακτηρίζονται από τις αντίστοιχες μη γραμμικές επιφάνειες κρίσης τους. Ειδικότερα, ο αλγόριθμος εκπαίδευσης του διανύσματος υποστήριξης χρησιμοποιείται για την κατασκευή των ακόλουθων τριών τύπων μηχανισμών μάθησης:

- Πολυώνυμοι μηχανισμοί μάθησης
- Δίκτυα ακτινικών συναρτήσεων βάσης
- Προσομοιωτές διπλής στρώσης με ένα μονό κρυφό επίπεδο. (Payne, 2002)

Δηλαδή, για κάθε μία από αυτά τα δίκτυα εμπρόσθιας τροφοδότησης (feed forward neural network) χρησιμοποιείται ο αλγόριθμος μάθησης του διανύσματος υποστήριξης για να εφαρμόσει τη διαδικασία της εκπαίδευσης, χρησιμοποιώντας ένα συγκεκριμένο σύνολο δεδομένων εκπαίδευσης, καθορίζοντας αυτόματα τον απαιτούμενο αριθμό των κρυφών μονάδων. Ενώ στα δίκτυα οπισθοδιάδοσης, ο αλγόριθμος χρησιμοποιεί διανύσματα υποστήριξης γενικότερου χαρακτήρα λόγω ευρύτερης εφαρμογής.

Οι μηχανισμοί διανυσμάτων υποστήριξης είναι βασισμένοι στη στατιστική θεωρία μάθησης, δίνοντας μία οικογένεια ορίων που διέπουν την ικανότητα εκπαίδευσης του μηχανισμού:

$$R(\alpha) = \int 0,5|y - f(x, \alpha)|dP(x, y) \quad (Γ. 11)$$

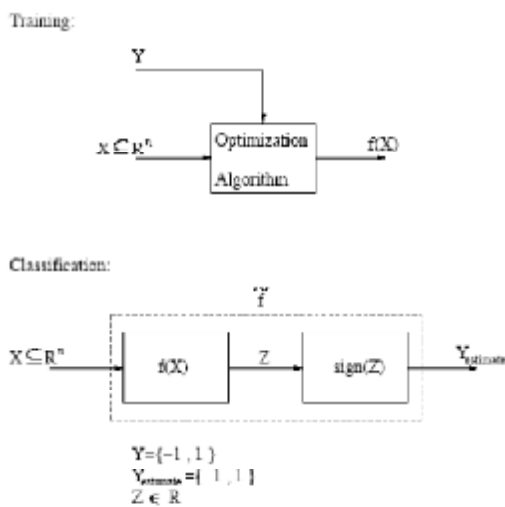
Ιδιότητες του ορίου:

1. Είναι ανεξάρτητη από το  $P(x,y)$ . Υποθέτει απλώς ότι τόσο τα δεδομένα εκπαίδευσης όσο και τα δεδομένα δοκιμών αντλούνται ανεξάρτητα σύμφωνα με κάποιο  $P(x,y)$ .

2. Συνήθως δεν είναι δυνατόν να υπολογιστεί η αριστερή πλευρά.

3. Εάν η  $h$  είναι γνωστή, η δεξιά πλευρά μπορεί να υπολογιστεί εύκολα.

Επομένως, δεδομένων πολλών διαφορετικών μηχανισμών εκ μάθησης, καθώς και της επιλογής μίας σταθεράς, επαρκώς μικρής  $\eta$ , επιλέγεται ο μηχανισμός που δίνει το χαμηλότερο ανώτατο όριο για τον πραγματικό κίνδυνο. Αυτό δίνει μια μέθοδο αρχών για την επιλογή ενός μηχανισμού εκπαίδευσης για μια δεδομένη εργασία, και είναι η βασική ιδέα των διαρθρωτικών ελαχιστοποίησης του κινδύνου. Η γενική δομή του μηχανισμού διανυσμάτων υποστήριξης SVM παρουσιάζεται στην *Εικόνα 25*.



Εικόνα 25 Γενική δομή του μηχανισμού διανυσμάτων υποστήριξης SVM

### 3.8 Συμπεράσματα

Αυτό το κεφάλαιο ασχολήθηκε με τεχνητά νευρωνικά δίκτυα. Αναλύεται διεξοδικά η έννοια των Τεχνητών Νευρωνικών Δικτύων, οι διάφοροι τύποι τους, τα πλεονεκτήματα και τα μειονεκτήματά τους και οι εφαρμογές τους, παρέχοντας μια λεπτομερή περιγραφή από τη μεθοδολογία που χρησιμοποιήθηκε σε αυτή τη διπλωματική εργασία. Ειδικότερα εισάγεται η αρχιτεκτονική των πολυεπίπεδων τεχνητών νευρωνικών δικτύων πρόσθιας τροφοδότησης. Η εποπτευόμενη εκπαίδευση μέσω του αλγόριθμου οπισθοδιάδοσης πίσω συζητείται μαζί με τους τρόπους της εκπαίδευσης, την επιλογή του ρυθμού μάθησης και τις πιθανές παγίδες από τον εγκλωβισμό σε τοπικά ελάχιστα. Επιπλέον, αναλύθηκαν οι μηχανισμοί διανύσματος υποστήριξης, καθώς και τα γραμμικά διαχωρίσιμα όσο και τα μη γραμμικά διαχωρίσιμα δεδομένα. Τέλος, πραγματοποιείται και μία εισαγωγή στη χρήση τεχνητών νευρωνικών δικτύων για την αναγνώριση προσώπου. Μία τεχνική που



χρησιμοποιείται συχνά για την αναγνώριση των προσώπων είναι η Ανάλυση των Κυρίων Συνιστωσών (Principal Components Analysis). Αυτό είναι το πρώτο θέμα του επόμενου κεφαλαίου.

## ΚΕΦΑΛΑΙΟ 4<sup>ο</sup>

### 4.1 Εισαγωγή

Το Τέταρτο κεφάλαιο αναφέρεται στα Τεχνητά Νευρωνικά Δίκτυα που χρησιμοποιούνται στην Αναγνώριση Προσώπων. Στο κεφάλαιο αυτό θα αναλυθούν οι μελέτες – προσεγγίσεις για την επίλυση του προβλήματος της Αναγνώρισης Προσώπων με τη χρήση Τεχνητών Νευρωνικών Δικτύων και θα σχολιαστούν τα αποτελέσματα της κάθε προσέγγισης.

### 4.2 Ανάλυση Κύριων Συνιστωσών (Principal Components Analysis- PCA)

Η Ανάλυση σε κύριες συνιστώσες μετατρέπει ένα σύνολο δεδομένων που λαμβάνει από πιθανες συσχετιζόμενες μεταβλητές σε ένα σύνολο των ασυσχέτιστων μεταβλητών που ονομάζονται κύριες συνιστώσες. Ο αριθμός των συνιστωσών μπορεί να είναι μικρότερος ή ίσος με τον αριθμό των αρχικών μεταβλητών. Η πρώτη κύρια συνιστώσα έχει την υψηλότερη δυνατή διακύμανση, και το καθένα από τα επόμενα στοιχεία έχει την υψηλότερη πιθανή διακύμανση με τον περιορισμό ότι πρέπει να είναι ορθογώνια προς το προηγούμενο στοιχείο. Θέλουμε να βρούμε τα κύρια συστατικά, σε αυτή την περίπτωση ιδιοδιανύσματα του πίνακα συνδιακύμανσης των εικόνων προσώπου.

Το πρώτο πράγμα που πρέπει να κάνουμε είναι να σχηματίσουμε ένα σύνολο δεδομένων εκπαίδευσης. 2D εικόνα  $I_i$  μπορεί να αναπαρασταθεί ως ένα 1D διάνυσμα με συνένωση γραμμών. (Perlibakas V. , 2004) Η εικόνα μετασχηματίζεται σε ένα διάνυσμα μήκους  $N=mn$ .

$$I = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1n} \\ X_{21} & X_{22} & \dots & X_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{m1} & X_{m2} & \dots & X_{mn} \end{bmatrix}_{m \times n} \xrightarrow{\text{CONCATENATION}} \begin{bmatrix} X_{11} \\ \vdots \\ X_{1n} \\ \vdots \\ X_{2n} \\ \vdots \\ X_{mn} \end{bmatrix}_{1 \times N} = X$$

Εστω  $M$  οι φορείς  $x_i$  ( $i = 1, 2, \dots, M$ ) μήκους  $N$  όπου σχηματίζουν ένα πίνακα της μάθησης των εικόνων,  $X$ . Για να εξασφαλιστεί ότι η πρώτη κύρια συνιστώσα περιγράφει την κατεύθυνση της μέγιστης διακύμανσης, είναι απαραίτητο να κεντράριστεί ο πίνακας.

Πρώτα καθορίζουμε το διάνυσμα των μέσων τιμών  $\Psi$ , και στη συνέχεια αφαιρούμε το διάνυσμα από κάθε διάνυσμα εικόνας.

$$\Psi = \frac{1}{M} \sum_{i=1}^M \mathbf{x}_i \quad (\Delta. 1)$$

$$\Phi_i = \mathbf{x}_i - \Psi \quad (\Delta. 2)$$

Ο μέσος όρος των διανυσματων σχηματίζουν ένα νέο πίνακα εκπαίδευσης (μεγέθους  $N \times M$ )

$$A = (\Phi_1, \Phi_2, \dots, \Phi_M)$$

Το επόμενο βήμα είναι να υπολογιστεί ο πίνακας συνδιασποράς  $C$ , και να βρει τα ιδιοδιανύσματα  $\mathbf{e}_i$  και ιδιοτιμές  $\lambda_i$

$$C = \frac{1}{M} \sum_{n=1}^M \varphi_n \varphi_n^T = AA^T \quad (\Delta. 3)$$

$$C\mathbf{e}_i = \lambda_i \mathbf{e}_i \quad (\Delta. 4)$$

Ο πίνακας συνδιασποράς  $C$  έχει διαστάσεις  $n \times n$ . Από τον οποίο έχουμε  $N$  ιδιοτιμές και ιδιοδιανύσματα. Για ένα μέγεθος εικόνας  $128 \times 128$ , θα πρέπει να υπολογίσει ο πίνακας διαστασεων  $16.384 \times 16.384$  και να βρει 16.384 ιδιοδιανύσματα. Δεν είναι πολύ αποτελεσματικό λόγο του ότι δεν χρειαζόμαστε περισσότερα από αυτά τα διανύσματα. Η τάξη του πίνακα συνδιασπορας περιορίζεται από τον αριθμό των εικόνων στο σύνολο εκμάθησης - αν έχουμε  $M$  εικόνες, θα έχουμε  $M-1$  ιδιοδιανύσματα που αντιστοιχούν σε μη-μηδενικές ιδιοτιμές.

Ένα από τα θεωρήματα στη γραμμική άλγεβρα αναφέρει ότι τα ιδιοδιανύσματα  $\mathbf{e}_i$  και ιδιοτιμές  $\lambda_i$  μπορούν να ληφθούν με την εύρεση ιδιοδιανυσμάτων και ιδιοτιμών του

πίνακα  $C_1 = A^T A$  ( $M \times M$  διάστασης) (Kim, 2008). Αν  $\mathbf{v}_i$  και  $\mu_i$  είναι ιδιοδιανύσματα και ιδιοτιμές του πίνακα  $A^T A$ , τότε:

$$A^T A \mathbf{v}_i = \mu_i \mathbf{v}_i \quad (\Delta. 5)$$

Πολλαπλασιάζοντας και τις δύο πλευρές της εξίσωσης ( $\Delta.5$ ) με ένα από τα αριστερά, θα έχουμε:

$$AA^T A \mathbf{v}_i = A \mu_i \mathbf{v}_i$$

$$AA^T(Av_i) = \mu_i(Av_i)$$

$$C(Av_i) = \mu_i(Av_i) \quad (\Delta.6)$$

Συγκρίνοντας τις εξισώσεις (Δ.4) και (Δ.6), μπορούμε να συμπεράνουμε ότι στο πρώτο  $M-1$

τα ιδιοδιανύσματα  $e_i$  και ιδιοτιμές  $\lambda_i$  του πίνακα  $C$  δίνονται από  $Av_i$  και  $\mu_i$ ,

αντίστοιχα. Τα ιδιοδιανύσματα που συνδέονται με την υψηλότερη ιδιοτιμή εκφράζει την υψηλότερη διακύμανση, και αντίστοιχα τα ιδιοδιανύσματα σχετίζεται με τη χαμηλότερη ιδιοτιμή, την μικρότερη διακύμανση. Στις ιδιοτιμές που μειώνονται εκθετικά, περίπου το 90% της συνολικής διακύμανσης περιέχεται στα πρώτα 5% έως 10% των ιδιοδιανύσματος. (Kim, 2008) Συνεπώς, τα διανύσματα θα πρέπει να ταξινομούνται κατά ιδιοτιμές έτσι ώστε το πρώτο διάνυσμα να αντιστοιχεί στην υψηλότερη ιδιοτιμή. Αυτά τα διανύσματα τότε είναι κανονικοποιημένα. Αυτά δημιουργούν ένα νέο πίνακα  $E$  έτσι ώστε κάθε διάνυσμα  $e_i$  είναι ένα διάνυσμα της στήλης. Οι διαστάσεις αυτού του πίνακα  $N \times D$ , όπου το  $D$  αντιπροσωπεύει τον επιθυμητό αριθμό των ιδιοδιανυσμάτων. Αυτό χρησιμοποιείται για την προβολή του πίνακα δεδομένων  $A$  και τον υπολογισμό των διανυσμάτων  $y_i$  του πίνακα  $Y = (y_1, \dots, \dots, y_M)$ :

$$Y = E^T A \quad (\Delta.7)$$

Κάθε αρχική εικόνα μπορεί να ανακατασκευαστεί από το άθροισμα των μέσων εικόνων  $\Psi$  για το σταθμισμένο άθροισμα όλων των διανυσμάτων  $e_i$ .

Το τελευταίο βήμα είναι η αναγνώριση προσώπων. Η εικόνα του ατόμου που θέλουμε να βρούμε σε σύνολο εκπαίδευσης μετατρέπεται σε ένα διάνυσμα  $P$ , μειωμένο κατά τη μέση τιμή  $\Psi$  και προκύπτει ένας πίνακας ιδιοδιανυσμάτων (eigenfaces):

$$\omega = E^T (P - \Psi) \quad (\Delta.8)$$

Η ταξινόμηση γίνεται με τον καθορισμό της απόστασης, ει, μεταξύ  $\omega$  και κάθε διάνυσμα  $y_i$  του πίνακα  $Y$  Η πιο κοινή είναι η Ευκλείδεια απόσταση, αλλά μπορούν να χρησιμοποιηθούν και άλλες. Η μελέτη αυτή παρουσιάζει τα αποτελέσματα για την

Ευκλείδεια απόσταση και την απόσταση Manhattan. Εάν A και B είναι δύο διανύσματα μήκους D, η απόσταση μεταξύ τους καθορίζεται ως εξής (Perlibakas V. , 2004):

Manhattan Απόσταση:

$$d(A, B) = \sum_{i=1}^D |a_i - b_i| \quad (\Delta. 9)$$

2) Ευκλείδεια απόσταση

$$d(A, B) = \sqrt{\sum_{i=1}^D (a_i - b_i)^2} = \|A - B\| \quad (\Delta. 10)$$

Αν η ελάχιστη απόσταση μεταξύ εικόνας δοκιμής και εικόνες εκπαίδευσης είναι υψηλότερη από ένα όριο  $\theta$ , η εικόνα δοκιμής θεωρείται ότι είναι άγνωστη, διαφορετικά είναι γνωστή και ανήκει στο πρόσωπο  $s = \operatorname{argmin}_i [\varepsilon_i]$  (Perlibakas V. , 2004)

Γιατί η δημιουργία ενός ορίου είναι σημαντική; Το πρόγραμμα απαιτεί μια ελάχιστη απόσταση μεταξύ της εικόνας της δοκιμής με τις εικόνες από τη βάση εκπαίδευσης. Ακόμη και αν το άτομο δεν είναι στη βάση δεδομένων, το πρόσωπο θα αναγνωρίζεται. Επομένως, είναι απαραίτητο να τεθεί ένα όριο το οποίο θα μας επιτρέπει να καθορίζουμε αν ένα πρόσωπο είναι στη βάση δεδομένων. Δεν υπάρχει κανένας τύπος για τον προσδιορισμό του ορίου. Ο πιο κοινός τρόπος είναι να υπολογιστεί πρώτα η ελάχιστη απόσταση της εικόνας από τις άλλες εικόνες στην βάση εκπαίδευσης, για να τοποθετηθεί αυτή η απόσταση σε ένα διάνυσμα RAST.

Το όριο υπολογίζεται ως 0,8 φορές της μέγιστης τιμής του διανύσματος rast (Gupta, Sahoo, Goel, & Gupta, 2010):

$$\theta = 0,8 * \max (\text{RAST}) \quad (\Delta. 11)$$

Εφαρμογές και αποτελέσματα

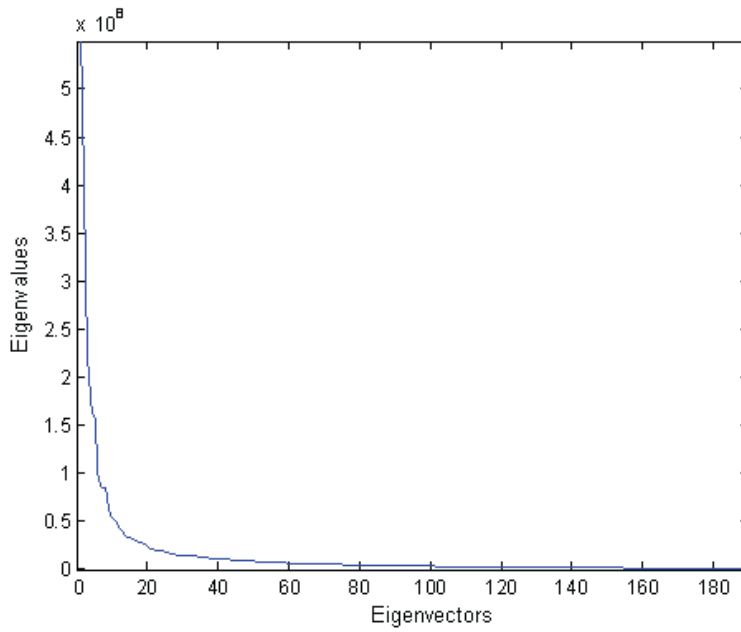
Το πείραμα διεξήχθη με τη χρήση της βάσης δεδομένων του ORL όψεων (<http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>, 2002). Η βάση δεδομένων εκπαίδευσης περιέχει 190 εικόνες 38 άτομα (5 εικόνες για κάθε άτομο), μια δοκιμαστική βάση δεδομένων έχει 40 εικόνες διαφορετικών ατόμων (38 γνωστά και 2 άγνωστα). Όλες οι φωτογραφίες έχουν διαστάσεις  $92 \times 112$  και ένα σκούρο φόντο και το

πρόσωπο φωτογραφίζεται σε όρθια, μετωπική στάση. Όλες οι εικόνες είναι στις αποχρώσεις του γκρι (στα επίπεδα έντασης του γκρι λαμβάνονται ως χαρακτηριστικά της εικόνας).

Παράδειγμα των εικόνων από τη βάση κατάρτισης δίνεται στην εικόνα 26. Η εικόνα 27 αναπαριστά όλες τις 190 ιδιοτιμές. Κάθε ιδιοτιμή αντιστοιχεί σε ένα μόνο ιδιοδιάνυσμα και μας δείχνει πόσο οι εικόνες από τη βάση εκπαίδευσης διαφέρουν από τη μέση εικόνα σε αυτό το πεδίο. Μπορούμε να δούμε ότι περίπου 10% των διανυσμάτων έχουν σημαντικές ιδιοτιμές, ενώ για τα υπόλοιπα διανύσματα οι τιμές είναι περίπου ίσες με το μηδέν. Δεν χρειάζεται να λαμβάνουμε υπόψη τα ιδιοδιανύσματα που αντιστοιχούν σε μικρές ιδιοτιμές επειδή δεν μεταφέρουν σημαντικές πληροφορίες σχετικά με την εικόνα.



Εικόνα 26 Εικόνες από την βάση εκπαίδευσης



Εικόνα 27 Ιδιοτιμές

Στις εικόνες 28 και 29 εμφανίζονται τα τρία πρώτα και τα τρία τελευταία ιδιοδιανύσματα , αντίστοιχα.Ενώ τα είδωλα στην εικόνα 28 μοιάζουν με τα πρόσωπα, στην εικόνα 29 δεν δίνουν σημαντικές πληροφορίες σχετικά με τις εικόνες από τη βάση εκπαίδευσης.



Εικόνα 28 Τρία πρώτα ιδιοδιανύσματα



Εικόνα 29 Τελευταία τρία ιδιοδιανύσματα

Ο Πίνακας 1 δείχνει ότι ο προτεινόμενος αλγόριθμος δίνει τα ίδια αποτελέσματα ανεξάρτητα από το αν θα χρησιμοποιηθούν μόνο για τα πρώτα 20 ή όλα τα 190 ιδιοδιανύσματα. Το ποσοστό αναγνώρισης είναι υψηλότερο για το Μανχάταν από ό, τι για την Ευκλείδεια απόσταση αν χρησιμοποιήσουμε 5 ή 10 ιδιοδιανύσματα, ενώ για 20 ιδιοδιανύσματα τα αποτελέσματα είναι ίδια και στις δύο περιπτώσεις.

Στην πράξη, λίγες μόνο εικόνες ανά άτομο είναι διαθέσιμες στην βάση εκπαίδευσης, έτσι είναι σημαντικό να σημειωθεί το αποτέλεσμα ενός αριθμού εικόνων ανά θέμα κατά την διαδικασία της αναγνώρισης. Τα αντίστοιχα ποσοστά αναγνώρισης για διαφορετικό αριθμό εικόνων εκπαίδευσης ανά θέμα δίνονται στον Πίνακα 2. Για να συγκριθούν, επιλέχθηκαν τις πρώτες 20 κύριες συνιστώσες για κάθε περίπτωση

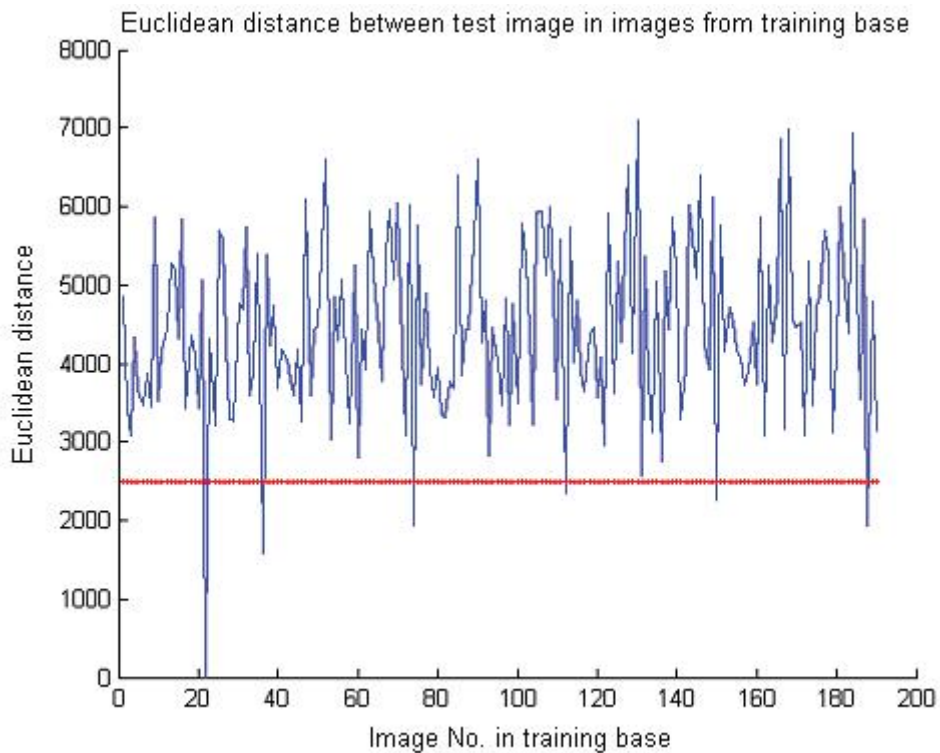
**Πίνακας 1 Το αποτέλεσμα της αναγνώρισης προσώπου, χρησιμοποιώντας ιδιοδιανύσματα**

Αριθμός κύριων Συνιστωσών	Ποσοστό Αναγνώρισης	
	Ευκλείδεια Απόσταση	Manhattan Απόσταση
5	77.5%	80%
10	92.5%	95%
20	97.5%	97.5%
190	97.5%	97.5%

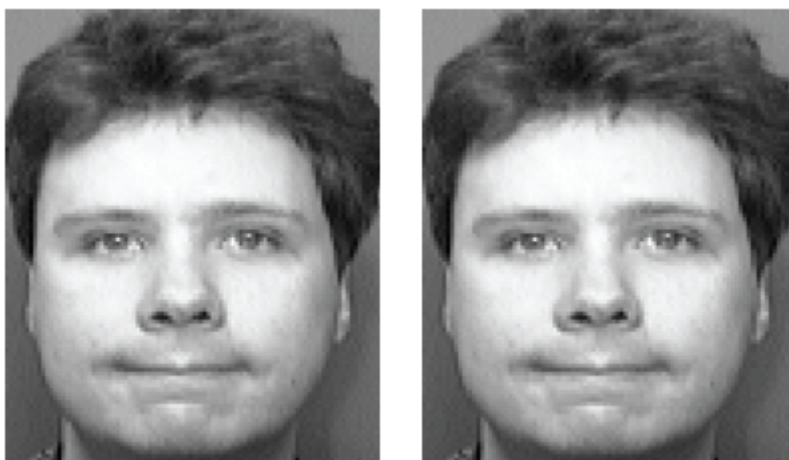
**Πίνακας 2 Ποσοστό αναγνώρισης για διαφορετικό αριθμό εικόνων ανά άτομο**

ΠΟΣΟΣΤΟ ΑΝΑΓΝΩΡΙΣΗΣ					
Αριθμός εικόνων ανά άτομο	1	2	3	4	5
Ευκλείδεια Απόσταση	70%	87,5%	87,5%	92,5%	97,5%
Manhattan απόσταση	67,5%	77,5%	82,5%	92,5%	97,5%





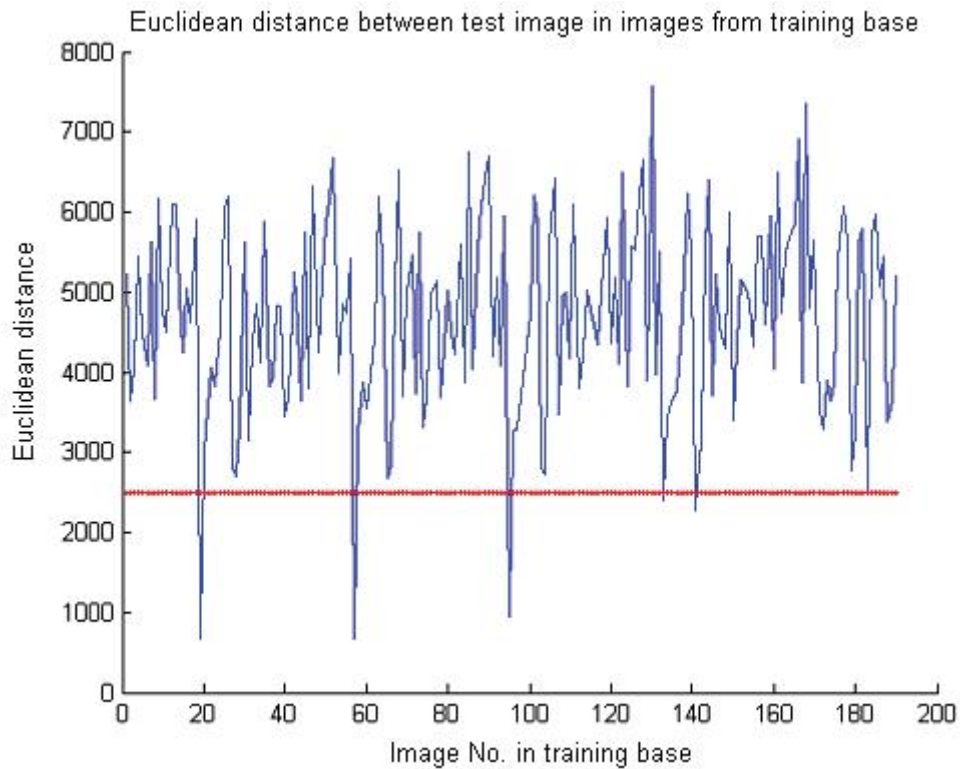
Εικόνα 30 Ευκλείδεια απόσταση ανάμεσα στην εικόνα δοκιμής και στις εικόνες από την βάση δεδομένων



Εικόνα 31 Εικόνα δοκιμής και αναγνωρισμένη εικόνα από την βάση εκπαίδευσης

Για την αναγνώριση του προσώπου υπολογίζουμε την απόσταση της εικόνας δοκιμής από την κάθε εικόνα στην βάση εκπαίδευσης. Η ελάχιστη απόσταση μας δείχνει ποια εικόνα από την βάση εκπαίδευσης ταιριάζει καλύτερα με την εικόνα δοκιμής. Η εικόνα 30 μας δείχνει την Ευκλείδεια απόσταση μεταξύ της εικόνας δοκιμής και όλων των 190 εικόνων από τη βάση εκπαίδευσης. Η απόσταση είναι η ελάχιστη για την εικόνα No 21 και είναι ίση με μηδέν. Αυτό σημαίνει ότι η εικόνα δοκιμής ταιριάζει απόλυτα με την εικόνα 21 από την βάση εκπαίδευσης (Εικόνα 31).

Η δεύτερη δοκιμή έγινε για ένα πρόσωπο που ήταν επίσης στην βάση δεδομένων, αλλά έχει μια διαφορετική έκφραση προσώπου. Η εικόνα δοκιμής έχει μια ελάχιστη Ευκλείδεια απόσταση η οποία είναι μικρότερη από το όριο (Εικόνα 32), οπότε αναγνωρίζει το πρόσωπο (Εικόνα 33).

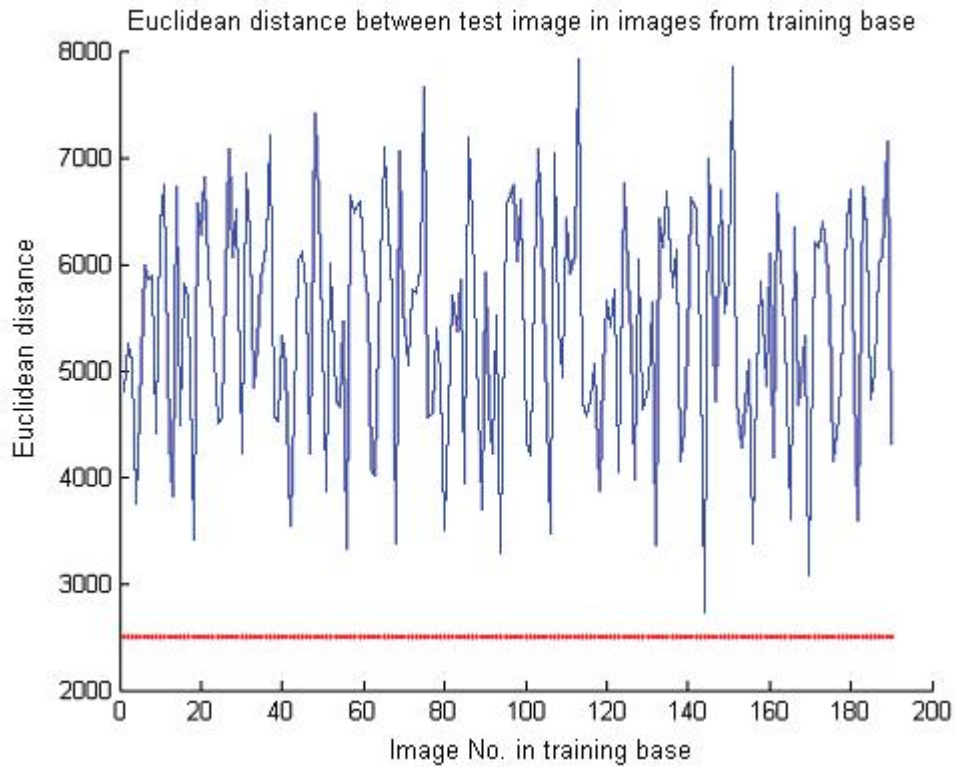


Εικόνα 32 απόσταση ανάμεσα στην εικόνα δοκιμής και στην βάση δεδομένων



Εικόνα 33 Εικόνα δοκιμής και αναγνωρισμένη εικόνα από την βάση εκπαίδευσης

Η ελάχιστη Ευκλείδεια Απόσταση της τρίτης δοκιμής (Εικόνα 34) είναι μεγαλύτερη από το όριο, οπότε η εικόνα δοκιμής παρουσιάζεται ως ένα άγνωστο πρόσωπο (Εικόνα 35).



Εικόνα 34 Ευκλείδεια απόσταση ανάμεσα σε εικόνα δοκιμής και στις εικόνες από την βάση δεδομένων



Εικόνα 35 Εικόνα από ένα άγνωστο πρόσωπο.

## ΣΥΜΠΕΡΑΣΜΑ

Η μέθοδος για την αναγνώριση προσώπων χρησιμοποιώντας τα ιδιοδυνασμάτα, προτείνεται. Χρησιμοποιήθηκε μια βάση δεδομένων από εικόνες προσώπων η οποία περιλάμβανε 190 εικόνες από 38 διαφορετικά πρόσωπα ( 5 εικόνες ανά πρόσωπο). Για τα αποτελέσματα, μπορούμε να συμπεράνουμε πως για την αναγνώριση είναι αρκετό να πάρουμε περίπου το 10% των ιδιοδυνασμάτων με τις υψηλότερες ιδιοτιμές. Είναι επίσης σαφές ότι το ποσοστό αναγνώρισης αυξάνεται από τον αριθμό των εικόνων εκπαίδευσης ανά πρόσωπο. Είναι προφανές πως όταν η ελάχιστη απόσταση μεταξύ της εικόνας δοκιμής με τις άλλες εικόνες από την βάση εκπαίδευσης είναι ίση με το 0(μηδέν), τότε, η εικόνα δοκιμής ταιριάζει απόλυτα με την εικόνα από την βάση εκπαίδευσης. Και τέλος σε περίπτωση που η απόσταση είναι μεγαλύτερη από μηδέν αλλά μικρότερη από ένα μέσο όριο, τότε έχουμε ένα γνωστό πρόσωπο με διαφορετική έκφραση προσώπου, διαφορετικά των παραπάνω το πρόσωπο είναι άγνωστο.

### 4.3 Νευρωνικό Δίκτυο Cottrell (Cottrell Neural Network).

Ένα νευρωνικό δίκτυο για την αναγνώριση προσώπου έχει αναπτυχθεί από τον Garrison Cottrell (1990) (Cottrell, 1990) στο Πανεπιστήμιο της Καλιφόρνια στο Σαν Ντιέγκο. Το δίκτυό του (Εικόνα 36) είναι εκπαιδευμένο με την ανάστροφη διάδοση και αποτελείται από τη συμπίεση της εικόνας του δικτύου (αυτόματη σύνδεση) και ένα ενιαίο στρώμα προσομοίωσης. Οι παράμετροι για τα δίκτυα στην (Εικόνα 36) είναι ως ακολούθως:

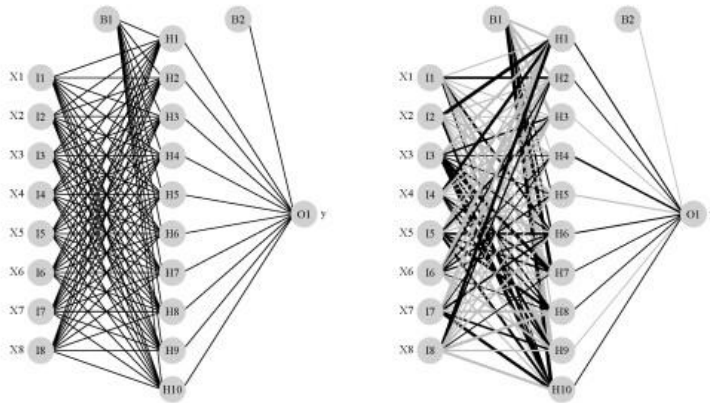
#### § Δίκτυο Αυτόματης Σύνδεσης

- Κόμβοι εισόδου: 4096
- Κρυφοί κόμβοι: 40, σειρά σιγμοειδούς ενεργοποίησης [-1,1]
- Κόμβοι εξόδου: 4096, σειρά σιγμοειδούς ενεργοποίησης [-1,1]
- Ταχύτητα κίνησης: 0
- Κρυφό στρώμα μάθησης: 0.0001
- Στρώμα εξόδου μάθησης: 0,1

#### § Δίκτυο Ταξινόμησης

- Κόμβοι εισόδου: 4096
- Κρυφοί κόμβοι: 40, σειρά σιγμοειδούς ενεργοποίησης [-1,1]

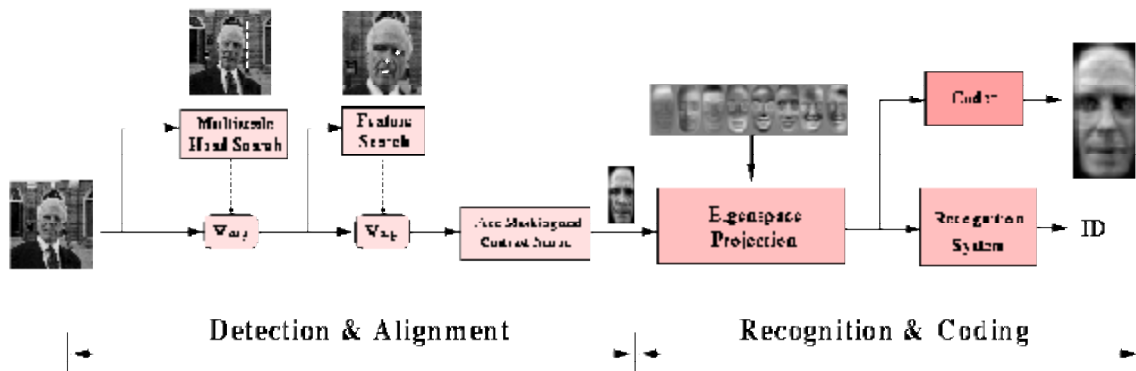
- Κόμβοι εξόδου: 20, σειρά σιγμοειδούς ενεργοποίησης [-1,1]
- Ταχύτητα Κίνησης: 0
- Κρυφό στρώμα μάθησης: Σταθερές αναμετρήσεις
- Στρώμα εξόδου μάθησης: 0,1



Εικόνα 36 Αριστερά: Δίκτυο αυτόματης σύνδεσης του στρώματος εισόδου για το δίκτυο στα δεξιά με τη χρήση της ανάστροφης διάδοσης.

Η βασική λειτουργία του δικτύου είναι ως ακολούθως: Πρώτον, το δίκτυο αυτόματης σύνδεσης έχει εκπαιδευτεί να ταιριάζει τις εικόνες εξόδου με τις εικόνες εισόδου. Ο Δρ. Cottrell υποθέτει ότι το κρυφό στρώμα συμπιέζει τα δεδομένα και εξάγει τα χαρακτηριστικά που αντιπροσωπεύουν μία κατανομημένη αναπαράσταση των κύριων συνιστωσών, οι οποίες είναι παρόμοιες με τις αντίστοιχες του συστήματος αναγνώρισης προσώπου του MIT. Υποστηρίζει αυτή τη θεωρία με τον σχηματισμό της μήτρας συνδιακύμανσης των ενεργοποιήσεων της κρυφής μονάδας πάνω από όλες τις εικόνες και εξάγοντας τα κύρια συστατικά που χρησιμοποιούν αυστηρή ανάλυση των κύριων συνιστωσών. Στη συνέχεια, τα κύρια συστατικά αποσυμπιέζονται με τη λειτουργία τους μέσα από το στρώμα εξόδου του δικτύου αυτόματης σύνδεσης. Η Εικόνα 37 δείχνει μία εικόνα που κατασκευάστηκε με αυτό τον τρόπο.

Μετά την εκπαίδευση του δικτύου αυτόματης σύνδεσης, οι αναμετρήσεις μεταξύ των εισροών και του κρυφού στρώματος είναι σταθερές και η έξοδος του κρυμμένου στρώματος συνδέεται με ένα μικρότερο (20 κόμβοι εξόδου), μονής στρώσης δίκτυο ταξινόμησης. Αυτό το μικρότερο δίκτυο καταρτίζεται για να ταξινομεί (προσδιορίζει) κάθε εικόνα. Το δίκτυο εκπαιδεύεται, επίσης, να εντοπίζει και να ταξινομεί ως «άγνωστες» εικόνες χωρίς πρόσωπο. Τα αποτελέσματα της δοκιμής αυτού του δικτύου είχαν 99 τοις εκατό ακρίβεια αναγνώρισης.



Εικόνα 37 Αναγνώριση Προσώπου

Ωστόσο, η βάση δεδομένων που χρησιμοποιήθηκε ήταν περιορισμένη σε 20 υποκείμενα και όλες οι εικόνες συλλήφθηκαν την ίδια ημέρα και ώρα. Οι δυνατότητες αυτού του δικτύου δεν έχουν ελεγχθεί για περισσότερες ημέρες ή για μεγαλύτερες βάσεις δεδομένων και πιστεύεται ότι αυτές οι δοκιμές θα προκαλέσουν προβλήματα στο δίκτυο.

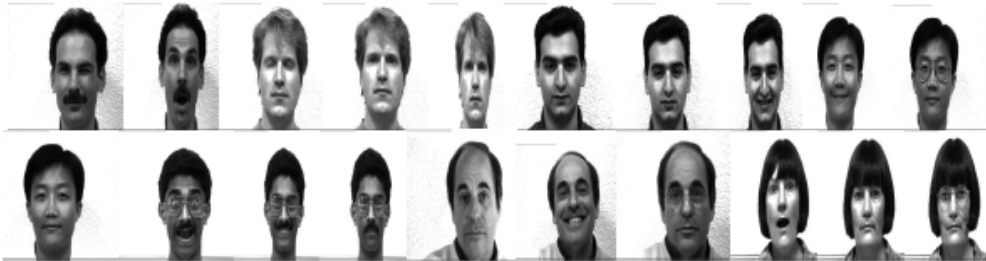
#### 4.4 Νευρωνικό Δίκτυο Επαλήθευσης Προσώπου με Ανάστροφη Λιάδοση (Backpropagation Neural Network for Face Verification).

Αναγνώριση προσώπου μέσω την οπιασθοδιάδοσης.

Ένα σύστημα αναγνώρισης προσώπου (Lawrence, Giles, Tsoi, & Back, 1993) (Kailash, Karande, & Talbar, 2009) είναι μια υπολογιστική οπτική και αναγνωρίζει αυτόματα ένα ανθρώπινο πρόσωπο από μια βάση δεδομένων με εικόνες. Το πρόβλημα στην αναγνώριση προσώπου είναι δύσκολο, καθώς χρειάζεται να υπολογίσει όλες τις πιθανές μεταβολές που προκαλούνται από το φωτισμό, τα χαρακτηριστικά του προσώπου κλπ. Εδώ θα δούμε 200 εικόνες προσώπων από την βάση δεδομένων του Yale και πως υπολογίστηκαν μερικές μετρήσεις απόδοσης, όπως το ποσοστό επίδοσης και ο χρόνος εκτέλεσης. Η αναγνώριση προσώπου με βάση τα νευρωνικά είναι αποτελεσματική γιατί η αναλογία επίδοσης και επιτυχίας είναι περισσότερη από 90%.

Σε αυτό το πείραμα, ελήφθησαν από την βάση δεδομένων του Yale 200 εικόνες και για δείγματα 20 εικόνες προσώπων όπως φαίνεται στην εικόνα 38. Μια από τις εικόνες, όπως βλέπουμε στην εικόνα 39(α) την πήραμε ως εικόνα εισόδου. Οι εικόνες 39(β) και 39(γ) δείχνουν την μέση εικόνα και την εικόνα εξόδου μέσω του PCA αντίστοιχα. Στα

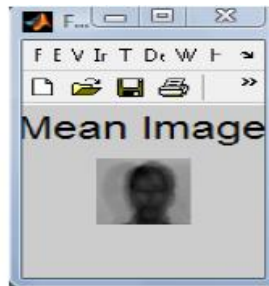
νευρωνικά δίκτυα οπισθοδιάδοσης η εικόνα 40(α) δείχνει ένα σύνολο εκπαίδευσης , η εικόνα 40(β) τα ιδιοδυναύσματα προσώπων και η εικόνα 40(γ) την αναγνωρισμένη εικόνα μέσω της μεθόδου οπισθοδιάδοσης.



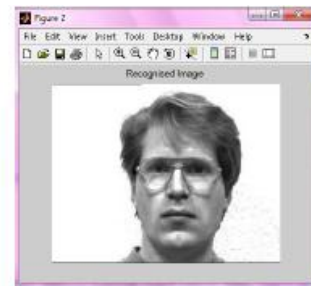
Εικόνα 38 Δείγματα Εικόνων από την βάση δεδομένων του Yale



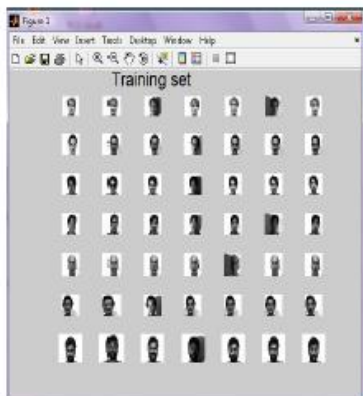
Εικόνα 39 (α) Εικόνα Εισόδου



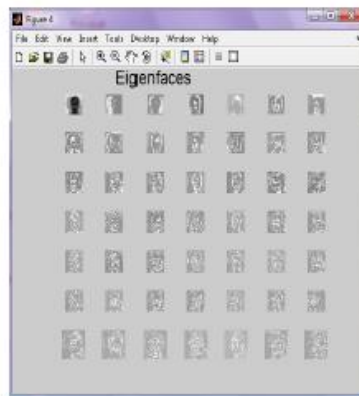
Εικόνα 39(β). Μέση Εικόνα



Εικόνα39(γ) Εικόνα εξόδου μέσω της PCA



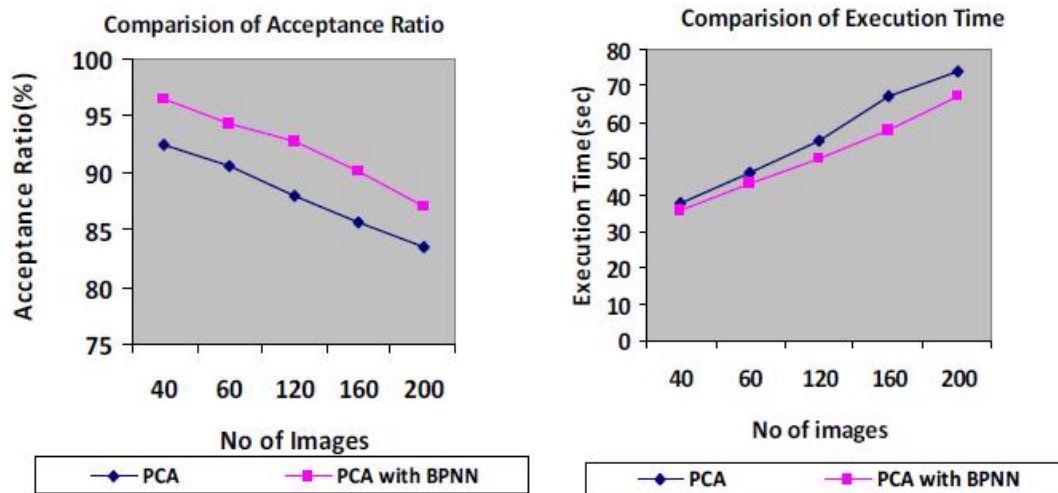
Εικόνα 40(α) Σύνολο Εκπαίδευσης 40(β) Ιδιοδυναύσματα προσώπων 40(γ) Αναγνωρισμένο πρόσωπο από την μέθοδο οπισθοδιάδοσης



Ο πίνακας 3 δείχνει την σύγκριση μεταξύ το ποσοστό επίδοσης και τον χρόνο εκτέλεσης για 40,80,120,160 και 200 εικόνες της βάσης δεδομένων του Yale και στην εικόνα 6 είναι η γραφική ανάλυση αυτών.

Πίνακας 3 Ποσοστό Επίδοσης & Χρόνος Εκτέλεσης

Αριθμός Εικόνων	(% ) Επίδοση		Χρόνος Εκτέλεσης (Λεπτά)	
	PCA	PCA με BPNN	PCA	PCA μεBPNN
40	92,4	96,5	38	36
60	90,6	94,3	46	43
120	87,9	92,8	55	50
160	85,7	90,2	67	58
200	83,5	87,1	74	67



Εικόνα 41 Γραμμική σύγκριση ποσοστού επίδοσης και χρόνου εκτέλεσης σε λεπτά σε σχέση με τον αριθμό των εικόνων



## Συμπέρασμα

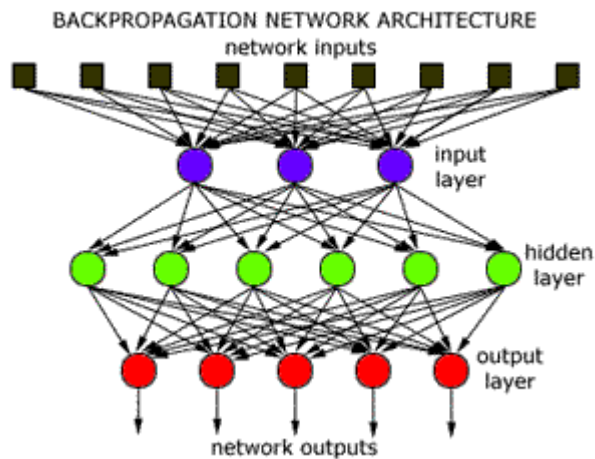
Εδώ είδαμε πως η αναγνώριση προσώπου χρησιμοποιώντας ιδιοδιανύσματα προσώπων λειτουργεί γρήγορα και με μεγάλη ακρίβεια. Όταν συνδυαστεί η BPNN (τεχνική οπισθοδιάδοσης) με την PCA, σε μη γραμμικές εικόνες προσώπων μπορεί να γίνει εύκολα η αναγνώριση. Ως εκ τούτου, συμπεραίνεται ότι η μέθοδος αυτή έχει ποσοστό επίδοσης μεγαλύτερο από 90% και ο χρόνος εκτέλεσης γίνεται σε πολύ λίγα λεπτά.

Επίσης οι ερευνητές από το Εθνικό Εργαστήριο του Los Alamos έχουν αναπτύξει ένα σύστημα επαλήθευσης προσώπου που χρησιμοποιεί ένα νευρωνικό δίκτυο τριπλού στρώματος με εκπαίδευση ανάστροφης διάδοσης. Οι εμπειρικές παράμετροι του δικτύου είναι οι εξής:

Κόμβοι εισόδου: 1400

- Κρυφοί κόμβοι: 20, σειρά σιγμοειδούς ενεργοποίησης [0,1]
- Κόμβους εξόδου: 1, σειρά σιγμοειδούς ενεργοποίησης [0,1]
- Ταχύτητα Κίνησης: 0,50
- Κρυφό στρώμα μάθησης: 0,15
- Στρώμα εξόδου μάθησης: 0.30

Το δίκτυο είναι πλήρως συνδεδεμένο και όλες οι αναμετρήσεις αρχικά κατανέμονται ομοιόμορφα σε τυχαίες τιμές μεταξύ -0.1 και 0.1. Τα διανύσματα εισόδου κλιμακώνονται στο εύρος [-1,1] και το δίκτυο καταρτίζεται για να εξάγει ένα 1.0 για τον στόχο και ένα 0.0 για τις εκροές εκτός από το στόχο. Η εκπαίδευση σταματά μετά από 15.000 επαναλήψεις. Σε όλες τις περιπτώσεις, ο αριθμός των εικόνων στόχου στη βάση δεδομένων αναγκάστηκε τεχνητά να είναι 10 τοις εκατό του συνολικού αναπαράγοντας τις εικόνες στόχου (*Εικόνα 42*).



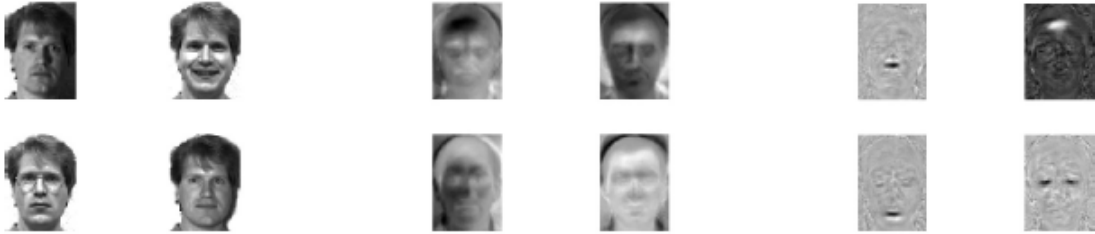
Εικόνα 42 Δίκτυο με Ανάστροφη Διάδοση

Το δίκτυο ελέγχεται με τη χρήση διαφόρων συνθέσεων των δεδομένων εκπαίδευσης. Τα σύνολα εκπαίδευσης αποτελούνται από 5, 10, 15, ή 20 τοις εκατό της βάσης δεδομένων, είτε τυχαία επιλεγμένες ή από συγκεκριμένες δημογραφικές ομάδες. Τα αποτελέσματα των δοκιμών επαλήθευσης κατά μέσο όρο σε όλα τα σενάρια εκπαίδευσης και τους αριθμούς δείχνουν μια 99,997 τοις εκατό σωστή απόρριψη των μη-στόχων και ένα 91,3 τοις εκατό σωστές αποδοχές των στόχων. Όπως στις δοκιμές του Cottrell, τα δεδομένα που χρησιμοποιήθηκαν για αυτή τη δοκιμή του δικτύου λήφθηκαν κατά τη διάρκεια μίας ημέρας. Για να είναι χρήσιμο ένα δίκτυο επαλήθευσης πρέπει να αποδίδει καλά σε πολλαπλές ημέρες των δεδομένων και το δίκτυο αυτό δεν αναμένεται να αποδίδει καλά σε πολλές ημέρες. Επιπλέον, το δίκτυο αυτό δεν έχει ελεγχθεί για τις κατηγορίες των δεδομένων μεγαλύτερες από ένα και, συνεπώς, τα ανώτατα όρια του δικτύου είναι άγνωστα.

#### 4.5 Αναγνώριση Προσώπου βάση την Ανάλυση Ανεξαρτήτων Συνιστωσών (Face Recognition Based on Independent Component Analysis- ICA):

Η τεχνική ICA έχει ως στόχο να βρει ένα γραμμικό μετασχηματισμό για τα δεδομένα εισόδου χρησιμοποιώντας ως βάση, όσο το δυνατόν, στατιστικά ανεξάρτητες μεταβλητές. Έτσι η ICA μπορεί να θεωρηθεί ως γενίκευση της PCA. Η PCA προσπαθεί να επιτύχει μια αναπαράσταση των εισροών με βάση ασυσχέτιστες μεταβλητές, ενώ η

ICA προσφέρει μια αναπαράσταση που βασίζεται σε στατιστικά ανεξάρτητες μεταβλητές. Το Σχήμα 1 δείχνει τη διαφορά μεταξύ της PCA και ICA εικόνες .



Εικόνα 43 Αριστερά, Αρχικές εικόνες, Κεντρικά PCA και δεξιά ICA εικόνες από την Yale βάση δεδομένων

Στο πλαίσιο της αναγνώρισης προσώπου, η ICA έχει δείξει να παράγει καλύτερα αποτελέσματα από εκείνα που λαμβάνονται με την PCA (Chengjun & Wechsler, 1999). Σε αντίθεση με την PCA, η ICA δεν παρέχει μια ενδογενή αναπαράσταση για τους συντελεστές του προσώπου.

Προκειμένου να καθοριστεί η απόδοση της ICA / SVM, σε σύγκριση με άλλα συστήματα, πραγματοποιήθηκαν πειράματα σε δύο ανεξάρτητες βάσεις δεδομένων προσώπου, η βάση δεδομένων Yale (Belhumeur, Hespanha, & Kriegman, 1997), και ένα τυχαία επιλεγμένο υποσύνολο της βάσης δεδομένων AR (Martinez & Benavente, 1998). Η βάση δεδομένων του Yale περιέχει 165 εικόνες (11 ανά άτομο), με αλλαγές στην έκφραση του προσώπου και τις συνθήκες φωτισμού. Από τη βάση δεδομένων AR που χρησιμοποιήθηκαν 300 εικόνες προσώπου (12 ανά άτομο), με αλλαγές στην έκφραση του προσώπου και των συνθηκών φωτισμού, και οι εικόνες που λήφθηκαν σε δύο σκέλη σε διάστημα δύο εβδομάδων. Όλα τα αποτελέσματα που λήφθηκαν χρησιμοποιώντας 2-φορές (AR) και 5 φορές (Yale) διασταυρώθηκαν και μεταβλήθηκε ο αριθμός των συντελεστών που χρησιμοποιούνται στην περιοχή 1-N, όπου N είναι ο αριθμός των εικόνων εκπαίδευσης ICA συντελεστών που διατάχθηκαν. Όλες οι εικόνες είχαν προηγουμένως μετατραπεί σε 256 αποχρώσεις του γκρι χρώματος και εφαρμόστηκε εξισορρόπηση ιστογράμματος. Το φόντο στις εικόνες Yale αφαιρέθηκε χειροκίνητα. Για τις εικόνες της βάσης δεδομένων AR επιτεύχθηκε μια πιο εύλογη ομαλοποίηση. Εκτός από την εξισορρόπηση ιστογράμματος, πραγματοποιήθηκε επίσης και γεωμετρική ομαλοποίηση. Στις εικόνες έγινε περικοπή, αφαιρέθηκαν τα μαλλιά και οι ώμοι. Τα μάτια και το στόμα τοποθετούνται χειροκίνητα και στη συνέχεια στις εικόνες μετατοπίζονται σε  $x$  και  $y$  στρεβλωτά, ώστε να έχουν τα μάτια και το στόμα την ίδια θέση για όλες τις εικόνες.

Ο αλγόριθμος ICA που χρησιμοποιήθηκε στα πειράματα ήταν FastICA ([www.cis.hut.fi/projects/ica/fastica/](http://www.cis.hut.fi/projects/ica/fastica/)). Η FastICA παρέχει ταχεία σύγκλιση και οι εκτιμήσεις των ανεξάρτητων συνιστωσών μεγιστοποιώντας ένα μετρο θεμελίωσης των αρχικών εκτιμώμενων συνιστωσών.

Η ταξινόμηση SVM είναι μια ταξινόμηση δύο κατηγοριών. Οι δύο επιλογές είναι: χρησιμοποιώντας N SVMs (N είναι ο αριθμός των κλάσεων) διαχωρίζοντας μία κλάση από τις υπόλοιπες, ή χρησιμοποιώντας  $N(N-1) / 2$  SVMs, έναν για κάθε ζεύγος των κλάσεων. Δεδομένου ότι η ακρίβεια και στις δύο επιλογές είναι σχεδόν ίδια (Chapelle, Haffner, & Vapnik, 1999), επιλέγουμε αυτή που είναι λιγότερο περίπλοκη. Ο αλγόριθμος SVM που χρησιμοποιείται στα πειράματα παρουσιάζει προβλήματα σύγκλισης όταν οι συντελεστές εισόδου έχουν σχετικά υψηλό μέγεθος. Αυτό μπορεί να μετριαστεί διαιρώντας τους συντελεστές εισόδου με μία σταθερή τιμή.

Τα αποτελέσματα εμφανίζονται στον Πίνακα 4. Οι SVM χρησιμοποιήθηκαν μόνο με πολυώνυμα (μέχρι 3 %) και πυρήνες Gaussian (με την παράμετρο  $\sigma$  του πυρήνα να ποικίλει).

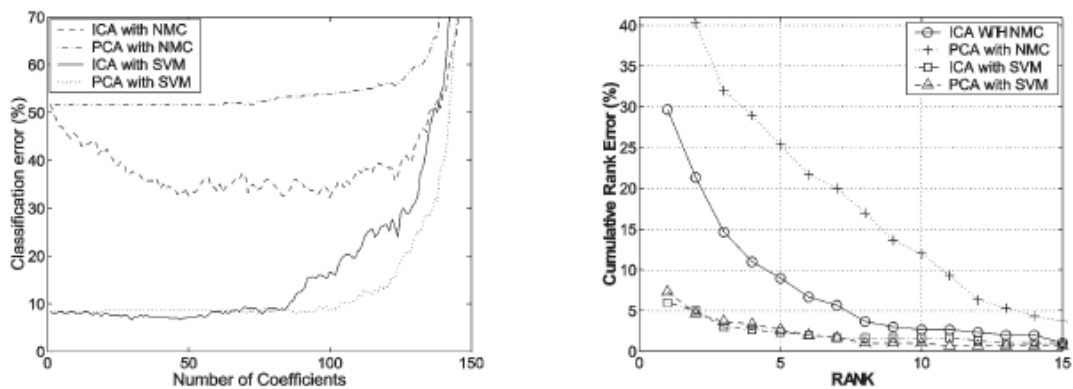
Πίνακας 4 Το ποσοστό αναγνώρισης για τις Yale και AR εικόνες χρησιμοποιούν το πλησιέστερο μέσο ταξινομητή (nearest mean classifier -NMC) και SVM

	NMC using		SVM			
	Euclidean distance		p=1	p=2	p=3	Gaussian
Yale	PCA	92,73%	98,79%	98,79%	98,79%	99,39%
	ICA	95,76%	99,39%	99,39%	99,39%	99,39%
AR	PCA	48,33%	92%	91,67%	91%	92,67%
	ICA	70,33%	93,33%	93,33%	92,67%	94%

Για τους SVM, ένα πλήθος από 1000 εικόνες χρησιμοποιήθηκε ως εσφαλμένη ταξινόμηση. Η τελευταία στήλη μας δείχνει τα καλύτερα αποτελέσματα που λήφθηκαν από τα ποικίλα  $\sigma$ . (ο Συνδυασμός PCA-NMC απευθύνεται στην μεθοδο των ιδιοδυναυσμάτων)

Για την Βάση Δεδομένων του Yale, δεν υπάρχει σαφής διαφορά μεταξύ ICA / SVM και PCA / SVM. Θεωρούμε, πως το γεγονός του ότι η εσφαλμένη ταξινόμηση είναι κοντά στο μηδέν δεν επιτρέπει στις διαφορές να φανούν καθαρά. Όσο για τις AR εικόνες, αν

και τα πιο απόλυτα αποτελέσματα είναι με την ICA και τα SVM, η αναπαράσταση δεν είναι μακριά από αυτά που λαμβάνονται με την PCA και τα SVM. (Jonsson, 2000)



Εικόνα 44 Παρουσιάζει το συσσωρευμένο βαθμό σφάλματος και την ταξινόμηση του σφάλματος ως συνάρτηση του αριθμού των συντελεστών για το σύνολο προσώπων από την AR.

## Συμπέρασμα

Τα αποτελέσματα των πειραμάτων έδειξαν πως έχει μεγάλο ποσοστό αναγνώρισης και αυτό επιτυγχάνεται χρησιμοποιώντας την ICA/SVM αν και η PCA/SVM έδωσε επίσης καλά αποτελέσματα. Έτσι, τα αποδεικτικά στοιχεία για το γεγονός αυτό, δεδομένου ότι SVMs είναι σχετικά ανεξάρτητοι στο χώρο εμφάνισης, (Jonsson, 2000) δίνουν μεγαλύτερη σημασία ανάμεσα στην πολυπλοκότητα και την απόδοση.

## 4.6 Οι Μηχανισμοί Διανύσματος Υποστήριξης στην Αναγνώριση Προσώπου (Support Vector Machines for Face Recognition)

Υπάρχουν 2 παγκόσμιες μέθοδοι για την αναγνώριση προσώπου όπου έχουν την ισχύ να διαχωρίζουν μια εικόνα ανάλογα την θέση του προσώπου.

Η πρώτη παγκόσμια μέθοδος αποτελείται από έναν ανιχνευτή προσώπου, που εξάγει το τμήμα ενός προσώπου από μια εικόνα και το διαδίδει σε σύνολο από SVM ταξινομητές όπου εκτελούν την αναγνώριση του προσώπου. Χρησιμοποιώντας έναν ανιχνευτή προσώπου, μπορούμε να πετύχουμε την μετάφραση και την μη αλλοίωση της κλίμακας.

Στη δεύτερη παγκόσμια μέθοδο, χωρίζουμε τις εικόνες του κάθε προσώπου σε διαφορετικές ειδικές ομάδες. Στην συνέχεια εκπαιδεύουμε τους SVM ταξινομητές σε κάθε ένα ενιαίο σύνολο.

Σε αντίθεση με τις παγκόσμιες μεθόδους, το σύστημα των συνιστωσών χρησιμοποιεί ένα ανιχνευτή προσώπου όπου ανιχνεύει και εξάγει τα τμήματα του προσώπου. Ο ανιχνευτής αποτελείται από ένα σύνολο SVM ταξινομητών που εντοπίζει τα τμήματα του προσώπου και ταξινομεί με βάση την γεωμετρία του προσώπου και ελέγχει αν οι παράμετροι των τμημάτων ταιριάζουν με το εκπαιδευμένο γεωμετρικό πρότυπο. Τα ανιχνευόμενα τμήματα εξάγονται από την εικόνα, μορφοποιούνται ως προς το μέγεθος και μπαίνουν σε ένα σύνολο SVM ταξινομητών.

Και οι δύο παγκόσμιες μέθοδοι που αναφέραμε αποτελούνται από ένα στάδιο ανίχνευσης προσώπου, όπου το πρόσωπο ανιχνεύεται και εξάγεται από μια εικόνα εισόδου και ένα στάδιο αναγνώρισης όπου καθορίζεται η ταυτότητα του προσώπου.

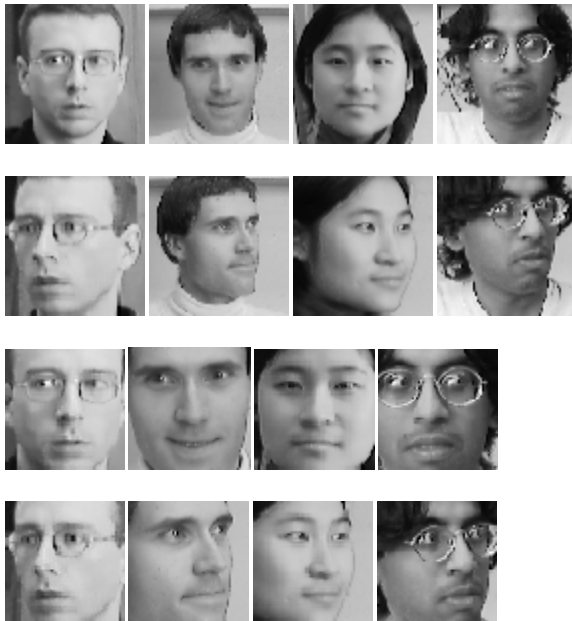
### **Ανίχνευση προσώπου**

Αναπτύχθηκε ένας ανιχνευτής προσώπου, παρόμοιος με εκείνον που περιγράφεται στο (Heisele, Poggio, & Pontil, 2000).

Με την προϋπόθεση να ανιχνεύει πρόσωπα σε διαφορετικές κλίμακες, αρχικά υπολογίζει μια πυραμίδα ανάλυσης των εικόνων εισόδου και τότε δημιουργεί ένα παράθυρο 58x58 πάνω από κάθε εικόνα στην πυραμίδα.

Εφαρμόζονται σε δύο στάδια προεπεξεργασίας των γκριζών εικόνων για να αντισταθμιστούν ορισμένες διακυμάνσεις μιας εικόνας (Sung, 1996). Για την αντιστάθμιση των σκιών πρέπει να αφαιρεθούν τα επίπεδα έντασης των αποχρώσεων του γκρι. Στην συνέχεια εφαρμόζεται η ανάλυση για την ρύθμιση της φωτεινότητας και της αντίθεσης της εικόνας. Οι προκύπτουσες τιμές του γκρι ομαλοποιήθηκαν για να είναι σε μία περιοχή μεταξύ 0 και 1 και χρησιμοποιήθηκαν ως στοιχεία εισόδου σε ένα γραμμικό ταξινομητή SVM. Ορισμένα αποτελέσματα ανίχνευσης φαίνονται στην εικόνα 45.

Τα δεδομένα εκπαίδευσης για τον ανιχνευτή προσώπου δημιουργήθηκαν από την τοποθέτηση 7 3D μοντέλων κεφαλιού (Vetter, 1998). Τα κεφάλια αυτά περιστρέφονται μεταξύ -30 και 30 σε πλάτος και φωτίζονται από ένα κίτρινο φως που έχει πάντα μια συγκεκριμένη κατεύθυνση και συγκεκριμένα στο κέντρο του προσώπου. Δημιουργήθηκαν 3590 εικόνες προσώπου μεγέθους 58x58 pixels. Το αρνητικό σύνολο εκπαίδευσης ατελείται από 10209 πρότυπα μεγέθους 58x58 χωρίς πρόσωπο (Sung, 1996) τα οποία εξήχθησαν τυχαία από 502 εικόνες χωρίς πρόσωπο.



Εικόνα 45 Αναγνώριση

Το πρώτο σύστημα έχει ένα γραμμικό SVM για κάθε πρόσωπο στη βάση δεδομένων. Κάθε SVM είχε εκπαιδευτεί να διακρίνει μεταξύ όλων των εικόνων από ένα μόνο πρόσωπο (με την ένδειξη +1) και όλες τις άλλες εικόνες στο σύνολο εκπαίδευσης (με ένδειξη -1). Και για τα δύο, εκπαίδευση και δοκιμή, «τρέχει» τον ανιχνευτή προσώπου στην εικόνα εισόδου για να εξάγει το πρόσωπο. Επαναδημιουργείτε η εικόνα του προσώπου σε 40x40 pixels και μετατρέπει τιμές του γκρι σε ένα διάνυσμα χαρακτηριστικών. Έχοντας ένα σύνολο από  $q$  άτομα και ένα σύνολο από  $q$  SVMs, όπου ο καθένας συνδέεται με ένα άτομο, το  $y$  από το πρότυπο ενός προσώπου  $x$  υπολογίζεται ως:

$$y = \begin{cases} n & \text{if } d_n(x) + t > 0 \\ 0 & \text{if } d_n(x) + t \leq 0 \end{cases} \quad (\Delta.12)$$

$$\text{με} \quad d_n(x) = \max\{d_i(x)\}_{i=1}^q$$

όπου  $d_i(x)$  υπολογίζεται σύμφωνα με την εξίσωση (Δ.12) για τους SVM που εκπαιδεύτηκαν για την αναγνώριση του προσώπου  $i$ . Το όριο ταξινόμησης συμβολίζεται με  $t$ .

Το 0 σημαίνει απόρριψη.

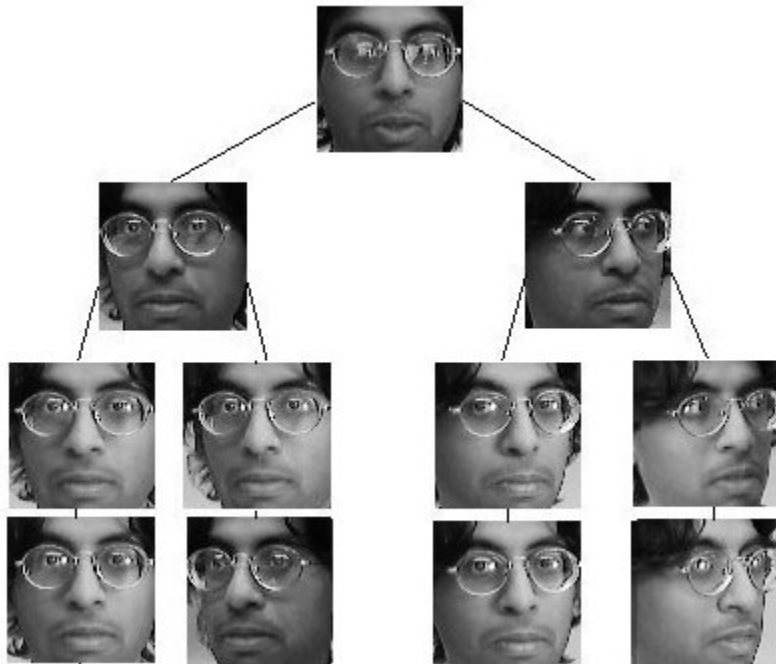
Οι αλλαγές στην θέση του κεφαλιού οδηγούν σε σημαντικές διαφοροποιήσεις στην εικόνα του προσώπου ενός ατόμου. Αυτές οι αλλαγές περιπλέκουν την διαδικασία

αναγνώρισης. Για το λόγο αυτό αναπτύχθηκε μια δεύτερη μέθοδος όπου διαχωρίζει τις εικόνες εκπαίδευσης του κάθε ατόμου σε ομάδες από διαιρετική τεχνική ομάδων (Linde, Buzo, & Gray, 1980). Ο αλγόριθμος ξεκίνησε ένα αρχικό σύνολο που περιέχει όλες τις εικόνες προσώπων από τα άτομα μετά από την προεπεξεργασία. Το σύνολο με την υψηλότερη διακύμανση χωρίζεται στα δύο από ένα υπερεπίπεδο. Η διακύμανση του συνόλου υπολογίζεται ως:

$$\sigma^2 = \min \left\{ \frac{1}{N} \sum_{m=1}^N \|x_n - x_m\|^2 \right\}_{n=1}^N \quad (\Delta.13)$$

Όπου N είναι ο αριθμός των προσώπων στο σύνολο. Μετά από τον διαχωρισμό που πραγματοποιήθηκε, το πρόσωπο με την μικρότερη απόσταση με όλα τα άλλα πρόσωπα στο ίδιο σύνολο, επιλέγεται ως το μέσο πρόσωπο του συνόλου. Η επανάληψη της ομοδοποίησης σταματάει όταν ολοκληρωθεί ο μέγιστος αριθμός των συνόλων. Τα μέσα πρόσωπα μπορούν να τοποθετηθούν σε ένα δυαδικό δέντρο. Η εικόνα 46 δείχνει τα αποτελέσματα της ομαδοποίησης που εφαρμόζεται στις εικόνες εκπαίδευσης ενός ατόμου στην βάση δεδομένων. Οι κόμβοι μας δείχνουν το μέσο πρόσωπο, τα φύλλα του δέντρου είναι μερικά παραδείγματα προσώπων των τελικών συνόλων. Όπως ήταν αναμενόμενο η διαιρετική ομαδοποίηση εκτελεί σε μια μόνο συγκεκριμένη ομάδα προσώπων. Εκπαιδεύτηκε ένα γραμμικός SVM για να διακρίνει ανάμεσα σε όλες τις εικόνες σε ένα σύνολο (με ένδειξη +1) και όλες τις εικόνες άλλων ανθρώπων στο σύνολο εκπαίδευσης (με ένδειξη -1). Η ομαδοποίηση πραγματοποιήθηκε χρησιμοποιώντας την συνάρτηση (Δ.13) με q τον αρχικό αριθμό των συνόλων από όλα τα άτομα στο σύνολο εκπαίδευσης.





Εικόνα 46 Δυαδικό δέντρο προσώπων που δημιουργείται από την διαιρετική ομαδοποίηση

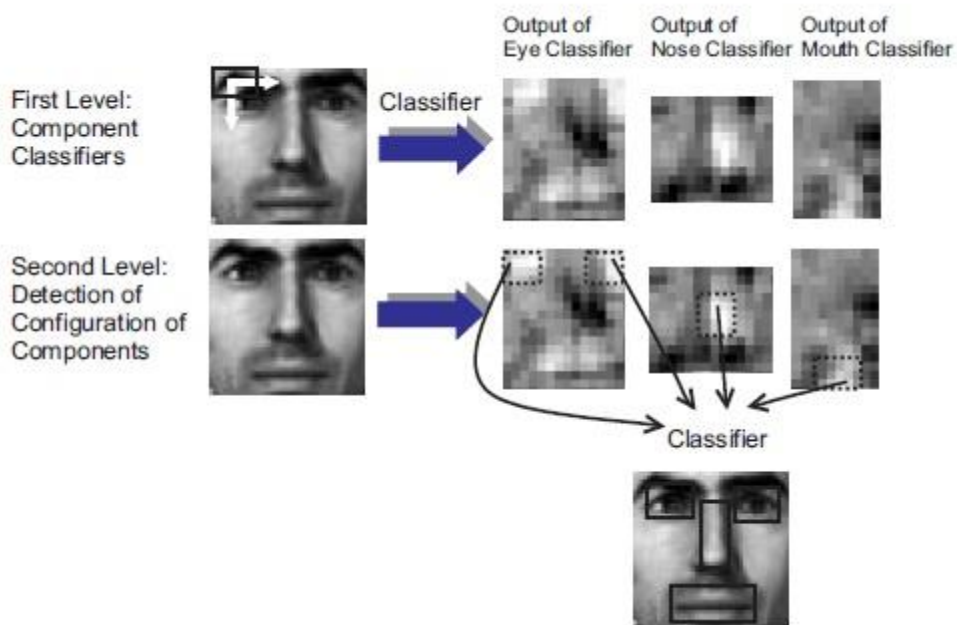
#### 4.6.1 Component-based προσέγγιση

Η παγκόσμια προσέγγιση είναι ιδιαίτερα ευαίσθητη στις διακυμάνσεις εικόνων που προκαλείται από τις αλλαγές στη στάση του προσώπου. Η προσέγγιση η οποία είναι βασισμένη στις συνιστώσες αποφεύγει αυτό το πρόβλημα, ανεξάρτητα από την ανίχνευση των χαρακτηριστικών του προσώπου. Για μικρές περιστροφές, οι αλλαγές στα χαρακτηριστικά είναι σχετικά μικρές σε σύγκριση με τις αλλαγές σε όλο το πρότυπο προσώπων. Αλλαγές στην 2D διάσταση των χαρακτηριστικών οφείλονται σε αλλαγές που δημιουργούνται από ένα εκπαιδευμένο, ευέλικτο μοντέλο προσώπου.

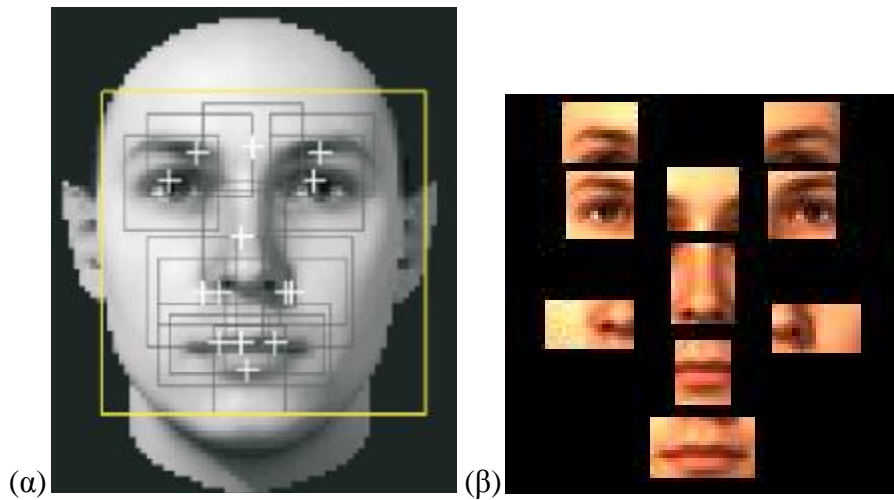
#### 4.6.2. Ανίχνευση

Σε μια 2D διάσταση υλοποιήθηκε μια συνιστώσα που βασίζεται στην ανίχνευση προσώπου, η οποία περιγράφεται λεπτομερώς στο (Heisele, Roggio, & Pontil, 2000). Οι αρχές του συστήματος που απεικονίζονται στην εικόνα 47. Στο πρώτο επίπεδο, οι ταξινομητές συνιστωσών ανιχνεύουν ξεχωριστά συνιστώσες προσώπων. Στο δεύτερο επίπεδο, μια γεωμετρική διαμόρφωση του ταξινομητή πραγματοποιεί την τελική ανίχνευση προσώπου συνδυάζοντας τα αποτελέσματα των ταξινομημένων συνιστωσών. Δίνεται σε ένα 58x58 παράθυρο, οι μέγιστες συνεχείς έξοδοι των ταξινομημένων συνιστωσών μέσα σε ορθογώνιες περιοχές αναζήτησης γύρω από τις αναμενόμενες

θέσεις των συνιστωσών που χρησιμοποιήθηκαν ως είσοδο σε ένα γεωμετρικά διαμορφωμένο ταξινομητή. Οι περιοχές αναζήτησης έχουν υπολογιστεί από τη μέση τιμή και τυπική απόκλιση των συνιστωσών, στην τοποθεσία των εικόνων εκπαίδευσης. Επίσης υπάρχει ένα γεωμετρικός ταξινομητής με συγκεκριμένη θέση των ανιχνευόμενων συνιστωσών σε σχέση με την υψηλότερη αριστερή γωνία του 58x58 παραθύρου. Οι 14 συνιστώσες προσώπου που χρησιμοποιήθηκαν στο σύστημα ανίχνευσης παρουσιάζονται στην εικόνα 48(α). Τα σχήματα και οι θέσεις των συνιστωσών έχουν καθοριστεί αυτόματα από τα δεδομένα εκπαίδευσης έτσι ώστε να βρεθεί η μέγιστη διαφορά μεταξύ εικόνας με πρόσωπο και εικόνας χωρίς πρόσωπο. Το σύνολο εκπαίδευσης είναι το ίδιο όπως και για τον ανιχνευτή προσώπου, που περιγράψαμε παραπάνω.



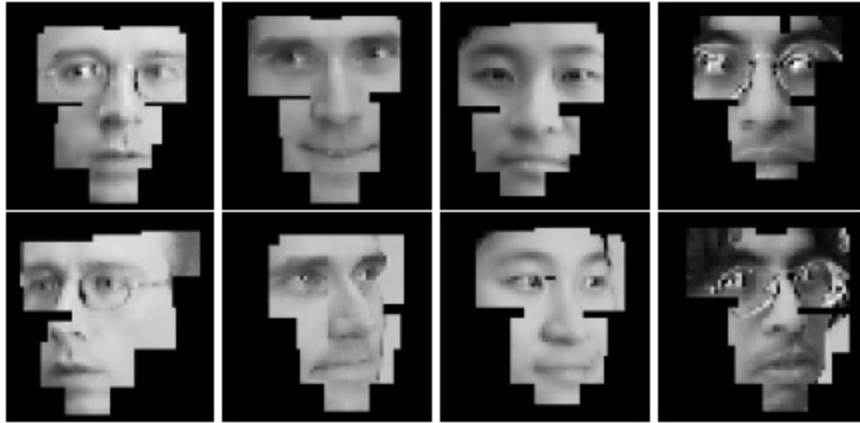
Εικόνα 47 Στο σύστημα φαίνονται οι τέσσερις συνιστώσες όπου χρησιμοποιεί ο ανιχνευτής προσώπου. Στο πρώτο επίπεδο, τα παράθυρα του μεγέθους των συνιστωσών αλλάζουν την εικόνα του προσώπου και την ταξινομούν σύμφωνα με τους ταξινομητές των συνιστωσών. Στο δεύτερο επίπεδο, οι μέγιστοι εξοδοι των ταξινομητών των συνιστωσών, στην προκαθορισμένη περιοχή αναζήτησης και οι θέσεις των ανιχνευσιμων συνιστωσών τροφοδοτούνται από την γεωμετρική διάταξη του ταξινομητή.



Εικόνα 48 Η εικόνα (α) δείχνει τις 14 συνιστώσες του ανιχνευτή προσώπων. Τα κέντρα των συνιστωσών είναι μαρκαρισμένα με ένα λευκό σταυρό. Οι 10 συνιστώσες που χρησιμοποιήθηκαν για την αναγνώριση του προσώπου, φαίνονται στο (β)

### 4.6.3. Αναγνώριση

Για να εκπαιδευτεί ένα πρόγραμμα αναγνώρισης προσώπου θα πρέπει αρχικά να τρέξει ο βάση συνιστωσών ανιχνευτής πάνω από κάθε εικόνα στο σύνολο εκπαίδευσης και να εξάγει τις συνιστώσες. Από τις 14 αρχικές συνιστώσες κρατήθηκαν 10 για την αναγνώριση προσώπου, αφαιρώντας αυτές τις οποίες περιείχαν λίγα γκρι ίχνη (πχ μάγουλα) ή πιο έντονα ίχνη που καλύπτουν αυτό. Οι 10 επιλεγμένες συνιστώσες φαίνονται στην εικόνα 48(β). Παραδείγματα του βάσης συνιστωσών ανιχνευτή που εφαρμόζονται σε εικόνες του συνόλου εκπαίδευσης φαίνονται στην εικόνα 49. Για να δημιουργηθεί η είσοδος του ταξινομητή της αναγνώρισης προσώπου μας κανονικοποιούνται κάθε μία από τις συνιστώσες στο μέγεθος και συνδιάζονται οι γκρι τιμές σε ένα ενιαίο χαρακτηριστικό διανύσματος. Όπως στο πρώτο παγκόσμιο σύστημα χρησιμοποιήθηκε μια προσέγγιση ένα προς όλα (one-vs-all) με ένα γραμμικό SVM για κάθε άτομο στην βάση, το αποτέλεσμα της ταξινόμησης προκύπτει από την εξίσωση (Δ.12)



Εικόνα 49 Παραδείγματα με βάση συνιστωσών ανίχνευση προσώπου. Παρουσιάζονται τα τμήματα του προσώπου καλυμμένα με τις 10 συνιστώσες που χρησιμοποιήθηκαν για την αναγνώριση προσώπου.

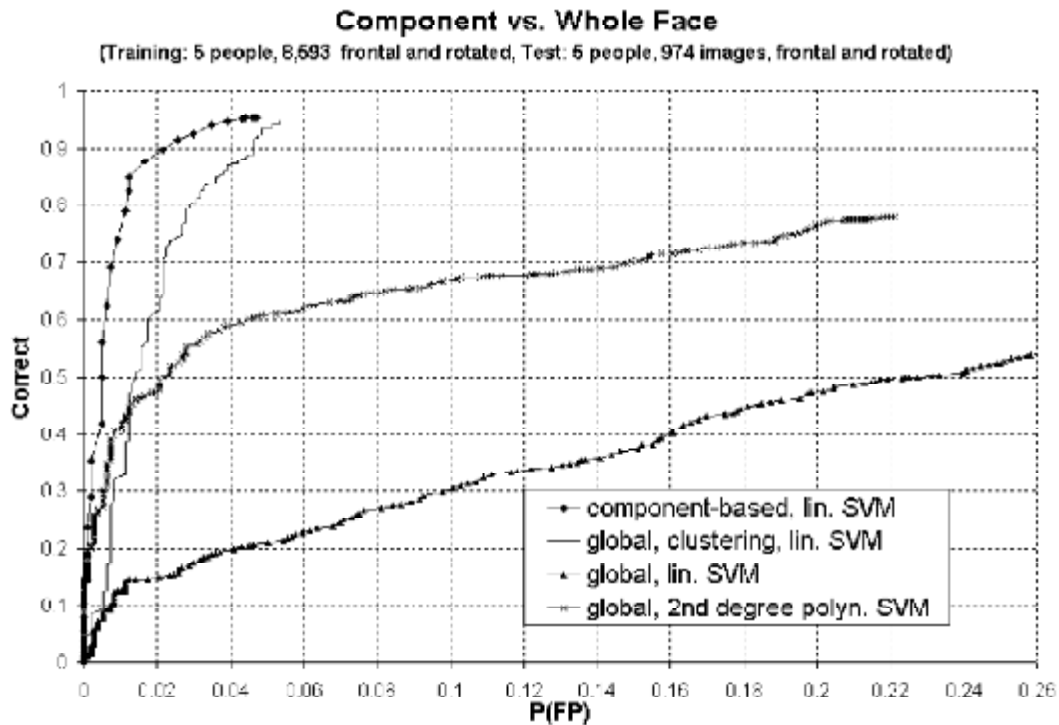
#### 4.6.4. Πειράματα

Τα δεδομένα εκπαίδευσης για το σύστημα αναγνώρισης προσώπου καταγράφηκαν με μια ψηφιακή φωτογραφική βίντεοκάμερα με ρυθμό καρέ περίπου τα 5 Hz. Το σύνολο εκπαίδευσης αποτελείται από 8.593 γκρι εικόνες προσώπου πέντε ατόμων από τις οποίες 1.383 ήταν ανφας. Η ανάλυση, των εικόνων των προσώπων, κυμάνθηκε μεταξύ 80x80 και 130x130 pixels με περιστροφές σε αζιμούθιο μέχρι  $\pm 40$  περίπου. Το σύνολο της δοκιμής καταγράφηκε με την ίδια κάμερα, αλλά σε διαφορετική ημέρα κάτω από διαφορετικές συνθήκες φωτισμού και με διαφορετικό φόντο. Το σύνολο περιλάμβανε 974 φωτογραφίες από τα πέντε άτομα στη βάση δεδομένων μας. Οι εναλλαγές στο βάθος ήταν και πάλι μέχρι  $\pm 40$  περίπου.

Δύο πειράματα διεξήχθησαν. Στο πρώτο πείραμα έχουμε εκπαιδευτεί πάνω σε όλες τις 8.593 περιστραμμένες και μετωπικές εικόνες προσώπων στο σύνολο εκπαίδευσης όπου δοκιμάστηκαν στο σύνολο της δοκιμής. Αυτό το πείραμα περιείχε τέσσερις διαφορετικές δοκιμές:

Σφαιρική προσέγγιση χρησιμοποιώντας ένα γραμμικό SVM ταξινομητή για κάθε άτομο, χρησιμοποιώντας ένα γραμμικό ταξινομητή SVM για κάθε ομάδα, χρησιμοποιώντας ένα δευτέρου βαθμού πολυωνυμικό SVM ταξινομητή για κάθε άτομο

Και προσέγγιση βάση συνιστώσας, χρησιμοποιώντας ένα γραμμικό SVM ταξινομητή για κάθε άτομο.

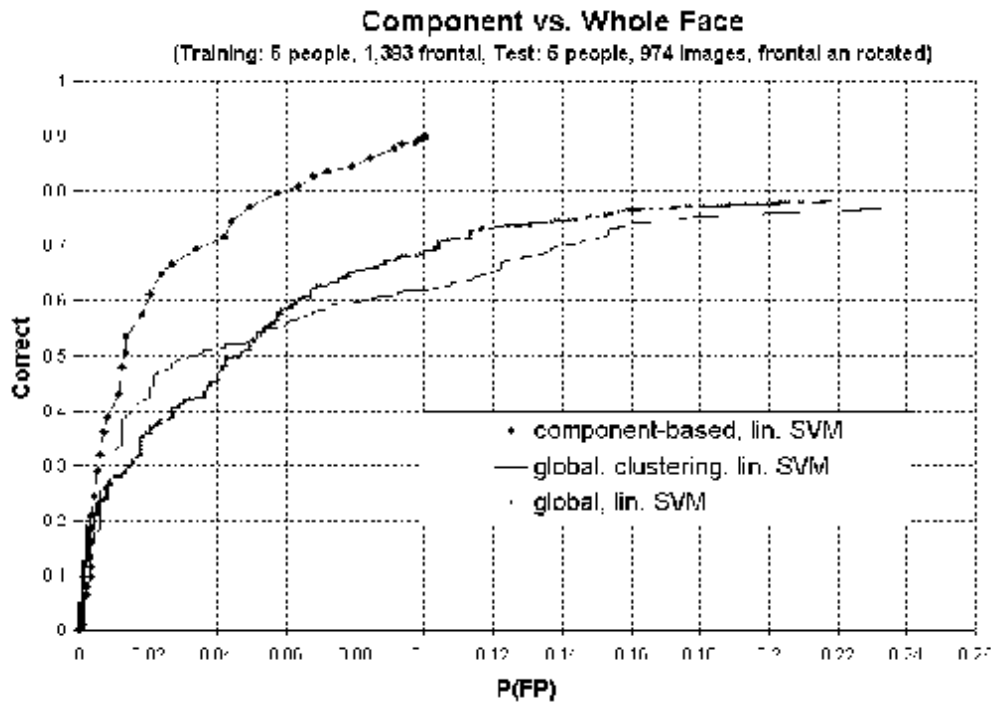


Εικόνα 50 Οι καμπύλες ROC όταν εκπαιδευτούν και δοκιμαστούν σε μετωπικά και περιστραμμένα πρόσωπα.

Στο δεύτερο πείραμα εκπαιδεύτηκαν μόνο το 1383 μετωπικών εικόνων προσώπου στο σύνολο εκπαίδευσης που όμως δοκιμάστηκε με το σύνολο δοκιμής. Αυτό το πείραμα περιελάμβανε τρεις διαφορετικές δοκιμές:

Η παγκόσμια προσέγγιση, χρησιμοποιώντας ένα γραμμικό SVM ταξινομητή για κάθε πρόσωπο, χρησιμοποιώντας ένα γραμμικό SVM ταξινομητή για κάθε ομάδα, και προσέγγιση βάση συνιστώσας χρησιμοποιώντας ένα γραμμικό SVM ταξινομητή για κάθε άτομο.

Οι καμπύλες ROC αυτών των δύο πειραμάτων φαίνονται στην εικόνα 50 και στην εικόνα 51, αντίστοιχα. Κάθε σημείο στην καμπύλη ROC αντιστοιχεί σε μια διαφορετική τιμή του ορίου κατάταξης  $t$  από την Εξίσωση (Δ.12). Στα τελικά σημεία των καμπυλών ROC το ποσοστό απόρριψης είναι 0.



Εικόνα 51 Οι καμπύλες ROC όταν εκπαιδεύονται σε μετωπικά πρόσωπα και δοκιμαστούν σε μετωπικά και περιστραμμένα πρόσωπα.

Ορισμένα αποτελέσματα του βάσης συνιστώσων συστήματος αναγνώρισης φαίνονται στην εικόνα 52.

Υπάρχουν τρεις ενδιαφέρουσες παρατηρήσεις:

- Και στα δύο πειράματα το σύστημα με τις συνιστώσες έχει σαφώς καλύτερες επιδόσεις στα παγκόσμια συστήματα. Γι αυτό, το πρόσωπο ταξινομητής (με 5 γραμμικούς SVMs) ήταν λιγότερο ισχυρό από ό, τι οι ταξινομητές που χρησιμοποιούνται σε παγκόσμιες μεθόδους (με 5 μη-γραμμικούς SVMs στην παγκόσμια μέθοδο χωρίς ομαδοποίηση, και 20 γραμμικούς SVMs στη μέθοδο με ομαδοποίηση).
- Η συμμετοχή της ομαδοποίησης έχει οδηγήσει σε σημαντική βελτίωση της παγκόσμιας μεθόδου όταν στο σύνολο της εκπαίδευσης περιλαμβάνονται περιστρεφόμενα πρόσωπα. Αυτό συμβαίνει επειδή ο αλγόριθμος δημιουργεί ειδικές ομάδες που έχουν μικρότερες στην κατηγορία διακυμάνσεις από ό, τι το

σύνολο των εικόνων ενός προσώπου. Η παγκόσμια μέθοδος με την ομαδοποίηση και γραμμικούς SVMs ήταν ανώτερη από το παγκόσμιο σύστημα, χωρίς ομαδοποίηση και με μη-γραμμικό SVM (βλέπε Εικ. 50). Αυτό δείχνει ότι ο συνδυασμός των αδύναμων ταξινομητών εκπαιδεύτηκαν σωστά σε επιλεγμένα υποσύνολα των δεδομένων που μπορούν να ξεπεράσουν ένα ενιαίο, και πιο ισχυρό ταξινομητή εκπαιδευμένο στο σύνολο δεδομένων.

- Προσθέτοντας περιστραμμένα πρόσωπα για το σύνολο εκπαίδευσης βελτιώνονται τα αποτελέσματα της παγκόσμιας μεθόδου με την ομαδοποίηση και τα στοιχεία της μεθόδου. Παραδόξως, τα αποτελέσματα για την παγκόσμια μέθοδο χωρίς ομαδοποίηση χειροτέρευαν. Αυτό δείχνει ότι το πρόβλημα της ταξινόμησης των όψεων ενός ατόμου σε ένα μεγάλο εύρος εικόνων είναι υπερβολικά πολύπλοκο για έναν γραμμικό ταξινομητή. Πράγματι, η επίδοση βελτιώθηκε σημαντικά με τη χρήση μη-γραμμικών SVMs με διστρωματικό πολυωνυμικό πυρήνα.



Εικόνα 52 Παραδείγματα των με βάση συνιστωσών αναγνώριση προσώπων. Οι τρεις πρώτες στήλες και η πρώτη εικόνα από την τελευταία στήλη, δείχνουν μια σωστή αναγνώριση. Οι δύο τελευταίες εικόνες στην τελευταία στήλη δείχνουν μια λάθος ταξινόμηση λόγω την μεγάλης περιστροφής και έκφρασης του προσώπου.

#### 4.6.6. Συμπέρασμα

Παρουσιάστηκε μια βάση συνιστωσών τεχνική και δύο παγκόσμιες τεχνικές για την αναγνώριση προσώπου και αξιολογήθηκαν οι αποδόσεις τους ανάλογα με τις αλλαγές στις εκφράσεις και τις θέσεις των προσώπων. Το βάση συνιστωσών σύστημα ανιχνεύει και εξάγει ένα σύνολο από δέκα συνιστώσες του προσώπου και ταξινομεί σε ένα ενιαίο διάνυσμα το οποίο ταξινομείται με γραμμικούς SVMs. Και στα δύο παγκόσμια συστήματα ανιχνεύθηκε ολόκληρο το πρόσωπο, εξήχθει από την εικόνα και χρησιμοποιήθηκε ως είσοδο στους ταξινομητές. Το πρώτο παγκόσμιο σύστημα αποτελείται από ένα μονο SVM για κάθε άτομο στην βάση δεδομένων. Στο δεύτερο σύστημα ομοδοποιήθηκε η βάση δεδομένων από κάθε άτομο και εκπαιδεύτηκε σε ένα σύνολο από εστιασμένους SVM ταξινομητές.

Δοκιμάστηκαν τα συστήματα σε μία βάση δεδομένων όπου περιλάμβανε περιστραμμένα πρόσωπα σε πλάτος πάνω από 40°. Σε όλα τα πειράματα τα με βάση συνιστωσών συστήματα απέδωσαν καλύτερα από ότι τα παγκόσμια συστήματα παρόλο που χρησιμοποιούνται πιο ισχυροί ταξινομητές (π.χ. μη γραμμικοί αντί των γραμμικών SVMs) για το παγκόσμιο σύστημα. Αυτό δείχνει πως χρησιμοποιώντας συνιστώσες προσώπου αντί για ολόκληρο το πρότυπο του προσώπου απλοποιεί σημαντικά την διαδικασία της αναγνώρισης του προσώπου.

#### 4.7 Συμπεράσματα

Το κεφάλαιο αυτό αναφέρθηκε στα Τεχνητά Νευρωνικά Δίκτυα που χρησιμοποιούνται στην Αναγνώριση Προσώπων. Αναλυθούν οι μελέτες – προσεγγίσεις που έχουν ήδη αναφερθεί και προταθεί στο Κεφάλαιο 1<sup>ο</sup> για την επίλυση του προβλήματος της Αναγνώρισης Προσώπων με τη χρήση Τεχνητών Νευρωνικών Δικτύων και σχολιάστηκαν τα αποτελέσματα της κάθε προσέγγισης. Πιο συγκεκριμένα, εξετάστηκε η προσέγγιση των χαρακτηριστικών του προσώπου για την αναγνώρισή του. Αυτή είναι η πρώτη τεχνική αναγνώρισης προσώπου με βάση τα στατιστικά στοιχεία που έχει προταθεί από τους ερευνητές. Ένα κύριο πλεονέκτημα αυτής είναι η επιδεκτικότητά της σε πραγματικό χρόνο αναγνώρισης προσώπου. Επίσης, παρουσιάστηκε η προσέγγιση της ανάλυσης των κύριων συνιστωσών και των πλησιέστερων τεχνικών ταξινόμησης.



Ειδικότερα, σε σχέση με την αναγνώριση προσώπου, συζητήθηκαν ένα δίκτυο ταυτότητας, ένα δίκτυο ταξινόμησης Cottrell και ένα δίκτυο ανάστροφης διάδοσης.

Επιπλέον, αναλύθηκαν οι μηχανισμοί διανύσματος υποστήριξης, καθώς και τα γραμμικά διαχωρίσιμα όσο και τα μη γραμμικά διαχωρίσιμα δεδομένα. Η αναγνώριση προσώπου είναι γενικά ένα πρόβλημα πολυταξινόμησης. Εφόσον οι μηχανισμοί διανύσματος υποστήριξης είναι τυπικά μία δυαδική τεχνική ταξινόμησης, περιγράφεται επιπλέον η λύση για την πολυ-ταξινόμηση χρησιμοποιώντας τους μηχανισμούς διανύσματος υποστήριξης.

## ΚΕΦΑΛΑΙΟ 5<sup>ο</sup>

### 5.1 Γενικά

Το Πέμπτο κεφάλαιο αναφέρεται στα συμπεράσματα με βάση τη βιβλιογραφική ανασκόπηση και προβαίνει σε συστάσεις για τη μελλοντική μελέτη.

Η παρούσα διπλωματική εργασία επικεντρώθηκε στη σημαντικότητα ενός αυτοματοποιημένου συστήματος αναγνώρισης προσώπου και στη μελλοντική προσφορά στην ανθρωπότητα, καθώς και στις τεχνολογικές προκλήσεις που θέτει στους ερευνητές η εφαρμογή του συστήματος αυτού. Ο σκοπός αυτής της διπλωματικής εργασίας ήταν αφενός η κατανόηση και η όσο το δυνατό ενδελεχής περιγραφή του προβλήματος της Αναγνώρισης Προσώπου, καθώς και η παρουσίαση των έως σήμερα προτεινόμενων μεθόδων και αφετέρου οι τρόποι εφαρμογής των Τεχνητών Νευρωνικών Δικτύων, τονίζοντας τη συμβολή τους στην επίλυση του εν λόγω προβλήματος. Οι στόχοι της παρούσας διπλωματικής εργασίας ήταν η κατανόηση των εννοιών των Τεχνητών Νευρωνικών Δικτύων, καθώς και του τρόπου εφαρμογής τους σε προβλήματα που άπτονται της Αναγνώρισης Προτύπων και συγκεκριμένα της Αναγνώρισης Προσώπου. Ως Μεθοδολογία της Διπλωματικής Εργασίας χρησιμοποιήθηκε η διεθνής και ελληνική βιβλιογραφία.

## 5.2 Συμπεράσματα

Η παρούσα διπλωματική εργασία προσπάθησε να παρουσιάσει με επάρκεια την Εφαρμογή των Τεχνητών Νευρωνικών Δικτύων για την Αναγνώριση Προσώπων. Πραγματοποιήθηκε λεπτομερής ανάλυση της αναγνώρισης των προσώπων μέσω των Τεχνητών Νευρωνικών Δικτύων, της Ανάλυσης Κύριων Συνιστωσών και των Μηχανισμών Διανύσματος Υποστήριξης. Οι εικόνες προσώπου είναι οι είσοδοι του συστήματος αναγνώρισης προσώπου. Οι εντάσεις των εικονοκυττάρων χρησιμοποιούνται ως εισροές για τις αντίστοιχες μεθόδους. Η προ-επεξεργασία των επιφανειών πραγματοποιείται και λαμβάνονται τα αποτελέσματα. Για κάθε μέθοδο οι αλλαγές της προ-επεξεργασίας γίνονται με τη χρήση εξίσωσης ιστογράμματος και το διάμεσο φιλτράρισμα των εικόνων προσώπου.

Το Κεφάλαιο 2 εισήγαγε τα Τεχνητά Νευρωνικά Δίκτυα και τον αλγόριθμο ανάστροφης διάδοσης και χρησιμοποίησε αυτή την εισαγωγή για την αναγνώριση προσώπων. Η αναγνώριση επιφανειών με την χρήση τεχνητών νευρωνικών δικτύων μπορεί να γίνει με δύο τρόπους. Η πρώτη μέθοδος είναι μια άμεση προσέγγιση βασισμένη στο τεχνητό νευρωνικό δίκτυο και το άλλο γίνεται με την απόκτηση των βαρών των χαρακτηριστικών του προσώπου, και, στη συνέχεια, με την κατάταξη των βαρών στη χρήση των τεχνητών νευρωνικών δικτύων. Η παράμετρος των Τεχνητών Νευρωνικών Δικτύων, η οποία μεταβάλλεται σε κάθε προσομοίωση, είναι ο αριθμός των κρυμμένων νευρώνων. Η αλλαγή του ρυθμού αναγνώρισης δεν είναι αισθητή όταν ο αριθμός των κρυφών νευρώνων είναι περισσότερο από 40. Επίσης, το τεχνητό νευρωνικό δίκτυο δεν αποδίδει κατάλληλα όταν η ανάλυση είναι 50x40, αλλά αποδίδει πολύ καλά για χαμηλότερες αναλύσεις. Αυτό υποδηλώνει ότι το δίκτυο μάθαινε ανεπιθύμητες παραμέτρους που οδηγούν σε φτωχές δυνατότητες γενίκευσης.

Το Κεφάλαιο 3 εισήγαγε μία θεωρία τεχνικής της πληροφορίας, δηλαδή την προσέγγιση χαρακτηριστικών προσώπου που χρησιμοποιεί τις κύριες συνιστώσες. Ακόμη και σε αυτή τη μέθοδο πραγματοποιείται διαφορετικός συνδυασμός προ-επεξεργασίας και η ανάλυση των προσώπων μεταβάλλεται. Η απόδοση είναι πολύ αξιόπαινη, όταν ο αριθμός των χαρακτηριστικών προσώπου είναι πάνω από 25. Υψηλότερη ακρίβεια 93% επιτυγχάνεται όταν ο αριθμός των χαρακτηριστικών προσώπου είναι 26 και η ανάλυση είναι 10x10.

Στο Κεφάλαιο 4 συζητήθηκαν οι μηχανισμοί διανύσματοςυποστήριξης και μέσω αυτών αναφέρθηκε η μέθοδος της αναγνώρισης προσώπου. Συνάγεται το συμπέρασμα ότι η απόδοση των μηχανισμών διανύσματος υποστήριξης είναι σχεδόν αμετάβλητες σε αλλαγές στην ανάλυση. Περαιτέρω, η προεπεξεργασία δεν φαίνεται να έχει οποιαδήποτε επίδραση στην ακρίβεια αναγνώρισης. Αυτό υποδηλώνει ότι οι μηχανισμοί διανύσματος υποστήριξης είναι πιο ισχυρή σε σχέση με την προεπεξεργασία. Τα απαιτούμενα χαρακτηριστικά εξάγονται παρά την προεπεξεργασία των εικόνων, η οποία είναι διαφορετική από εκείνη που παρατηρείται όταν χρησιμοποιούνται οι προαναφερθείσες τεχνικές. Επιπλέον, οι μηχανισμοί διανύσματος υποστήριξης ξεπερνούν τις προαναφερθείσες τεχνικές για να δώσουν τις καλύτερες ακρίβειες αναγνώρισης. Η ακρίβεια είναι περίπου 96%.

### 5.3 Μελλοντική Εργασία

Πολλές βελτιώσεις μπορούν να γίνουν στις μεθόδους που συζητούνται στην παρούσα διπλωματική εργασία. Οι αλγόριθμοι που εισάγονται στη διπλωματική εργασία μπορούν να τελειοποιηθούν με διάφορους τρόπους, σε βάρος περισσότερων υπολογισμών. Μία τέτοια προοπτική είναι η χρήση του πυρήνα της ανάλυσης κύριων συνιστωσών αντί τηςανάλυσης κύριων συνιστωσών. Ο πυρήνας της ανάλυσης κύριων συνιστωσώνχαρτογραφεί τις μη γραμμικότητες καλύτερα από την ανάλυση κύριων συνιστωσών. Μια άλλη καλύτερη εναλλακτική λύση είναι η χρήση αυτο-οργανωμένων χαρτών, η οποία χρησιμοποιεί αλγορίθμους εκπαίδευσης χωρίς επίβλεψη. Αυτοί εξαλείφουν εντελώς την ανάγκη για συνολική εκπαίδευση. Η ακρίβεια της αναγνώρισης πρέπει να βελτιωθεί. Συστήματα που βασίζονται στο σύνολογια την αναγνώριση προσώπου είναι ακόμα σε φάση εν τη γενέσει. Προσφέρουν μια καλύτερη ακρίβεια.

Η περαιτέρω επέκταση των τρεχόντων μπορεί να είναι η εφαρμογή σε ένα μεγαλύτερο πρόβλημα της ανίχνευσης και αναγνώρισης προσώπου. Η ανίχνευση προσώπου είναι το έργο της εξεύρεσης προσώπου δοθείσας μιας εικόνας. Μετά από αυτόν τον τρόπο, το όλο σύστημα μπορεί να επεκταθεί και σε βίντεο. Αυτό θα είναι ένα παρόμοιο έργο, δεδομένου ότι το βίντεο είναι μια ακολουθία εικόνων. Αν και πολλές επεκτάσεις μπορούν να γίνουν με το ισχύον σύστημα, για το σκοπό της αναγνώρισης προσώπου οι αλγόριθμοι που συζητήθηκαν στοχεύουν να είναι επαρκείς.

# Βιβλιογραφία

- Baldi, P., & Hornik, K. (2009). Neural Networks and Principal Component Analysis: Learning from Examples Without Local Minima. *Neural Networks* , 53-58.
- Belhumeur, P. N., Hespanha, J. P., & Kriegman, D. J. (1997). Eigenfaces vs Fisherfaces: Recognition using class specific linear projection. *Transactions on Pattern Analysis and Machine Intelligence* , 19 (7), 711-720.
- Bouattour, M. (2002). Neural Nets for Human Face Recognition. 700-704.
- Chapelle, O., Haffner, P., & Vapnik, V. (1999). SVMs for Histogram-based Image Classification. *Trans on Neural Networks* , 10 (5), 1055-1065.
- Chen, E. (2000). Computational Intelligence for Decision Support. CRC Press.
- Chengjun, L., & Wechsler, H. (1999). Comparative Assessment of Independent Component Analysis (ICA) for Face Recognition. *Second International Conference on Audio and Video-based Biometric Person Authentication* .
- Cottrell, G. (1990). Categorization of Faces Using Unsupervised Feature Extraction. *International Joint Conference on Neural Networks* , 2:65-70.
- Damasio, A. (2005). Prosopagnosia. *Trends in Neuroscience* , 8:132-135.
- Eberhart, R., & Yuhui, S. (2007). *Computational Intelligence: Concepts To Implementations*. Morgan Kaufmann Publishers.
- Engelbrecht, A. (2007). Computational Intelligence. *An Introduction* , Wiley.
- Farahati, N. (2002). Real-time recognition using novel infrared illumination. *Optical Engineering* , 1658-1662.
- Foldiak, P. (2009). Adaptive Network for Optimal Linear Feature Extraction. *International Joint Conference on Neural Networks* , 401-405.
- Geschwind, N. (2009). Specializations of the Human Brain. *Scientific American* , 107-120.
- Gupta, S., Sahoo, O., Goel, A., & Gupta, R. (2010). A New Optimized Approach to Face Recognition using EigenFaces. *Global Journal of Computer Science and Technology* , 10 (1), 15 – 17.
- Harandi, M. (2004). Feature selection using genetic algorithm and its application to face recognition. *Cybernetics and Intelligence Systems* , 2, 1368-1373.
- Harnon, R. (2011). Machine Identification of Faces. *Pattern Recognition* , 13:97-110.
- Haykin, S. (2008). Neural Networks and Learning Machines. *Prentice Hall 3rd edition* .
- Heisele, B., Poggio, T., & Pontil, M. (2000). *Face detection in still gray images*. MA: Center for Biological and Computational Learning.

<http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>. (2002).

Jonsson, K. T. (2000). *Robust Correlation and Support Vector Machines for Face Identification*. UK: Information Technology and Mathematics.

Kailash, J., Karande, S., & Talbar, K. (2009). Independent Component Analysis of Edge Information for Face Recognition. *International Journal of Image Processing* , 3 (3), 120 -131.

Kim, K. (2008). Face Recognition using Principal Component Analysis. *National Institute of Technology* .

Kresimir, D. (2008). Recent Advances In Face Recognition. *In Teh* .

Kung, S., & Diamantaras, K. (1990). A Neural Network Learning Algorithm for Adaptive Principal Component Extraction. *APEX* , 861-864.

Lawrence, S., Giles, C., Tsoi, A., & Back, A. (1993). Transactions of Neural Networks. 8 (1), 98-113.

Linde, Y., Buzo, A., & Gray, R. (1980). An algorithm for vector quantizer design. 28 (1), 84–95.

Martinez, A., & Benavente, R. (1998). *The AR face database*. CVC: Technical Report 24.

Meadows, J. C. (2004). Varieties of Prosopagnosia. *Journal of Neurology, Neurosurgery, and Psychiatry* , 498-501.

Nan, L., & W, H. (2009). Modeling Images with Multiple Trace Transforms for Pattern Analysis. *Signal Processing Letters* , 6 (5), 394-397.

Oja, P. (2011). Learning in Nonlinear Constrained Hebbian Networks. *Artificial Neural Networks* , 385-389.

Payne, T. (2002). Backpropagation Neural Networks for Facial Verification Update. *LosAlamos National Laboratory* .

Perlibakas, V. (2005). Distance Measures for PCA-based Face Recognition. *Pattern Recognition Letters* , 25 (6), 711 – 724.

Perlibakas, V. (2004). Face Recognition using Principal Component Analysis and Wavelet Packet Decomposition. *Informatica* , 15 (2), 243 – 250.

Rizon, M. (2006). Face Recognition using Eigenfaces and Neural Networks. *American Journal of Applied Sciences* 2 , 1872-1875.

Rolls, A. (2009). The effect of learning on the face selective responses of neurons in the cortex in the superior temporal sulcus of the monkey. *Experimental Brain Research* , 76:153-164.

Samal, A., & Prasana, A. (2012). Automatic Recognition and Analysis of Human Faces and Facial Expressions. *Pattern Recognition* , 25:65-77.

Sanger, T. (2009). Optimal Unsupervised Learning in a Single-Layer Linear Feedforward NeuralNetwork. *Neural Networks* , 459-473.

- Suarez, P. (2001). Face Recognition with the Karhunen-Loeve Transform. (W.-P. A. OH, Ed.) *School of Engineering* , AFrT/GE/ENG/91D-54.
- Sung, K. (1996). *Learning and Example Selection for Object and Pattern Recognition*. Artificial Intelligence Laboratory and Center for Biological and Computational Learning. MA: Cambridge.
- Tarr, G. (1991). Multi-Layered Feedforward Neural Networks for Image Segmentation. In A. F. (AU), *School of Engineering*. Wright-Patterson AFB OH.
- Tou, J., & Rafel, C. (2004). *Pattern Recognition Principles Reading*. Addison -Wesley Publisind.
- Turk, M., & Pentland, A. (2010). Recognition in Face Space. *SPIE Intelligent Robots and Computer Vision IX: Algorithms and Techniques* , 1381:43-54.
- Valentine, T., & Ferrara, A. (2012). Typicality in Categorization, recognition and identification: Evidence from face recognition. *British Journal of Psychology* , 82:87-102.
- Vetter, T. (1998). Synthesis of novel views from a single face. *International Journal of Computer Vision* , 28 (2), 103–116.
- Wright, J. (2009). Robust Face Recognition via Sparce Representation. *Pattern Analysis and Machine Intelligence* , 31 (2), 210-217.
- [www.cis.hut.fi/projects/ica/fastica/](http://www.cis.hut.fi/projects/ica/fastica/). (n.d.).

## ΠΗΓΕΣ ΕΙΚΟΝΩΝ

Εικόνα 1 : <http://faceattendance.blogspot.gr/2009/12/examples-of-biometrics-wwwinttelixcom.html>

Εικόνα 2 :  
[http://docs.opencv.org/trunk/doc/py\\_tutorials/py\\_feature2d/py\\_features\\_harris/py\\_features\\_harris.html](http://docs.opencv.org/trunk/doc/py_tutorials/py_feature2d/py_features_harris/py_features_harris.html)

Εικόνα 3: <http://www.extremetech.com/extreme/178777-facebooks-facial-recognition-software-is-now-as-accurate-as-the-human-brain-but-what-now>

Εικόνα 4: <http://www.couriermail.com.au/news/super-spy-camera-surveillance-network-to-fight-crime/story-e6frep2f-1226081971440>

Εικόνα 5& 6: <https://statsguys.wordpress.com/2014/03/15/data-analytics-for-beginners-march-machine-learning-mania-part-ii/>

Εικόνα 7: <http://blogs.msdn.com/b/socaldevgal/archive/2008/03/17/why-data-mining-matters.aspx>

Εικόνα 8: <https://inquiry.ucsd.edu/Reviewer/flashComponents.php>

Εικόνα 9: [http://openi.nlm.nih.gov/detailedresult.php?img=3082789\\_JRMS-15-84-g001&req=4](http://openi.nlm.nih.gov/detailedresult.php?img=3082789_JRMS-15-84-g001&req=4)

Εικόνα 10: <http://automotive.electronicsspecifier.com/driver-assistance-systems/image-sensors-for-advanced-driver-assistance-systems>

Εικόνα 11: <http://kishordgupta.wordpress.com/2010/12/26/analysis-and-proposal-of-development-on-line-edge-method-algorithm-of-face-recognition/>

Εικόνα 12: <http://www1.i2r.a-star.edu.sg/~hhzhang/PhDThesis/KAA.htm>

Εικόνα 13: [http://www.eecs.berkeley.edu/~yang/software/face\\_recognition/](http://www.eecs.berkeley.edu/~yang/software/face_recognition/)

Εικόνα 14:  
[http://docs.opencv.org/master/dc/d0d/tutorial\\_py\\_features\\_harris.html#gsc.tab=0](http://docs.opencv.org/master/dc/d0d/tutorial_py_features_harris.html#gsc.tab=0)

Εικόνα 15: <http://nextbigfuture.com/2013/07/quantum-computing-boosted-artificial.html>

Εικόνα 16: <http://spie.org/x25547.xml>

Εικόνα 17: <http://www.bloggang.com/viewblog.php?id=dimiasyou&date=03-09-2008&group=4&gblog=1>

Εικόνα 18:  
<http://www.sparknotes.com/testprep/books/sat2/biology/chapter9section1.rhtml>

Εικόνα 19: <http://www.bordalierinstitute.com/target11.html>



Εικόνα 20: <http://www.learnartificialneuralnetworks.com/introduction-to-neural-networks.html>

Εικόνα 21: <http://www.hindawi.com/journals/tswj/2012/185085>

Εικόνα 22 &23: [http://artemis.cslab.ntua.gr/el\\_thesis/artemis.ntua.ece/DT2004-0127/DT2004-0127.doc](http://artemis.cslab.ntua.gr/el_thesis/artemis.ntua.ece/DT2004-0127/DT2004-0127.doc)

Εικόνα 24: <http://www.codeproject.com/Articles/175777/Financial-predictor-via-neural-network>

Εικόνα 25: <http://proj.ncku.edu.tw/research/articles/e/20081205/3.html>

Εικόνα 26-35: <http://www.doiserbia.nb.rs/img/doi/1451-4869/2012/1451-48691201121S.pdf>

Εικόνα 36: <https://beckmw.wordpress.com/2013/03/04/visualizing-neural-networks-from-the-nnet-package/>

Εικόνα 37: <http://vismod.media.mit.edu/vismod/demos/facerec/>

Εικόνα 38-44:

<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.170.3965&rep=rep1&type=pdf>

Εικόνα 45-52: <http://cbcl.mit.edu/publications/ps/iccv2001.pdf>