



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ  
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

## **Ανάπτυξη Προγραμματιστικής Εφαρμογής για την Εκτίμηση του Συναισθήματος σε Καρέ Βίντεο**

### **ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

Βασιλική Η. Γιακουμή

**Επιβλέπων :** Στέφανος Δ. Κόλλιας  
Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2012





ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ  
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

## **Ανάπτυξη Προγραμματιστικής Εφαρμογής για την Εκτίμηση του Συναισθήματος σε Καρέ Βίντεο**

### **ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

Βασιλική Η. Γιακουμή

**Επιβλέπων :** Στέφανος Δ. Κόλλιας  
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 18<sup>η</sup> Ιουλίου 2012.

.....  
Στέφανος Κόλλιας  
Καθηγητής Ε.Μ.Π.

.....  
Ανδρέας Σταφυλοπάτης  
Καθηγητής Ε.Μ.Π.

.....  
Γιώργος Στάμου  
Λέκτορας Ε.Μ.Π.

Αθήνα, Ιούλιος 2012

.....  
Βασιλική Η. Γιακουμή

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Βασιλική Η. Γιακουμή, 2012.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

## Περίληψη

Ο σκοπός της παρούσας διπλωματικής εργασίας ήταν η ανάπτυξη μιας προγραμματιστικής εφαρμογής για την εκτίμηση του συναισθήματος σε εικόνες προσώπων προερχόμενες από ακολουθίες βίντεο.

Η μεθοδολογία που χρησιμοποιείται βασίζεται στον υπολογισμό των παραμέτρων κίνησης FAPs του προσώπου σύμφωνα με τον ορισμό τους από το διεθνές πρότυπο MPEG-4. Είσοδοι του συστήματος αναγνώρισης συναισθημάτων είναι οι συντεταγμένες συγκεκριμένων χαρακτηριστικών σημείων στο πρόσωπο. Η εκτίμηση για το συναίσθημα πραγματοποιείται με την αποτίμηση ενός συνόλου κανόνων που διατυπώνονται με τη βοήθεια των παραμέτρων FAPs. Η αποτίμηση των κανόνων αξιοποιεί βασικές αρχές της ασαφούς λογικής.

Η υλοποίηση έχει σχεδιαστεί με τέτοιο τρόπο ώστε να μπορεί να χρησιμοποιηθεί ως πρόσθετη συνάρτηση σε εφαρμογές πραγματικού χρόνου ή ανεξάρτητα ως εκτελέσιμο αρχείο. Το πρόγραμμα δε θέτει περιορισμούς για τη μέθοδο ανίχνευσης των σημείων εισόδου επιτρέποντας έτσι τη χρήση της εφαρμογής με δεδομένα εισόδου χειροκίνητης ή αυτόματης ανίχνευσης. Επιπλέον δεν τίθεται περιορισμός για το πλήθος ή το μέγεθος των κανόνων, παρά μόνο για τη μορφοποίηση της πληροφορίας, διατηρώντας έτσι την ευελιξία της εφαρμογής για χρήση με ποικίλους κανόνες και πιθανή μελλοντική επέκταση. Κατά την εκτέλεσή της, ο χρήστης της εφαρμογής έχει τη δυνατότητα να επιλέξει το πλήθος των πληροφοριών που εκτυπώνονται.

Το θεωρητικό υπόβαθρο της εργασίας περιλαμβάνει πληροφορίες για τις θεμελιώδεις θεωρίες συναισθημάτων, τις σημαντικότερες μεθόδους ποσοτικής αναπαράστασης συναισθημάτων και τις βασικές αρχές της ασαφούς λογικής. Επιπρόσθετα παρουσιάζονται πληροφορίες για το διεθνές πρότυπο MPEG-4. Η παρουσίαση και περιγραφή της υπολογιστικής υλοποίησης περιλαμβάνει παραδείγματα εικόνων με πραγματικά πρόσωπα και αποτελέσματα εκτύπωσης.

## Λέξεις κλειδιά

Θεωρίες συναισθημάτων, έκφραση προσώπου, ασαφής λογική, συνάρτηση συμμετοχής, πρότυπο MPEG-4, παράμετροι κίνησης προσώπου (FAP), χαρακτηριστικά σημεία προσώπου, εκτίμηση συναισθήματος, αποτίμηση κανόνων εάν – τότε.



# Abstract

The objective of this diploma thesis was to develop a program that would estimate a person's emotion demonstrated in picture frames taken from video sequences.

The implementation is based on the calculation of facial animation parameters (FAPs) as defined by the international MPEG-4 standard. The coordinates of specific feature points on the face are used as an input to the above program. The person's emotion is estimated through the evaluation of a set of rules which are formulated by means of FAP parameters. The evaluation process utilizes fundamental principles of fuzzy logic theory.

The program's design allows for use as an additional function to real-time applications or as a stand-alone executable file. The method used does not impose any limitations on the tracking method of the facial feature points. As a result, it is suitable for either manually or automatically detected input data. Furthermore, other than the formatting requirements, there is no restriction on the length or number of evaluation rules, which promotes flexibility and allows for future extensibility. As for the program's output, the user is given the option to choose the amount of information to be printed.

Finally, the theoretical background of this thesis includes information on the most notable theories of emotion, prominent methods for quantitative representation of emotions and basic principles of fuzzy logic. Moreover, information is provided on the MPEG-4 standard. The presentation and description of the computational implementation contains examples of natural faces and samples of printed results.

# Keywords

Emotion theories, facial expressions, fuzzy logic, membership function, MPEG-4 standard, facial animation parameters (FAP), facial feature points, emotion recognition, if-then rules.





## Ευχαριστίες

Θα ήθελα να ευχαριστήσω θερμά τον καθηγητή Ε.Μ.Π. Στέφανο Κόλλια για την ανάθεση και συνολική επίβλεψη της διπλωματικής εργασίας.

Ευχαριστώ τον διδάκτορα μηχανικό και κύριο ερευνητή ΕΠΙΣΕΥ Κώστα Καρπούζη για το ενδιαφέρον του, την καθοδήγηση που προσέφερε και την καθοριστική συμβολή του στην ολοκλήρωση της διπλωματικής εργασίας.

Τέλος θα ήθελα να ευχαριστήσω τη διδάκτορα μηχανικό Αμαρυλλίδα Ραουζαίου και τον διδάκτορα μηχανικό Στέλιο Αστεριάδη για την πολύτιμη βοήθειά τους κατά τη διάρκεια εκπόνησης της διπλωματικής εργασίας.



# ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

|   |    |
|---|----|
| Ευρετήριο Σχημάτων.....   | 15 |
| Ευρετήριο Πινάκων.....  | 16 |
| ΚΕΦΑΛΑΙΟ 1: Εισαγωγή.....   | 17 |
| 1.1 Αναγνώριση συναισθημάτων στην επικοινωνία ανθρώπου - μηχανής.....     | 17 |
| 1.2 Αντικείμενο διπλωματικής εργασίας .....                               | 18 |
| 1.2.1 Συνεισφορά εργασίας .....   | 19 |
| 1.3 Διάρθρωση Εργασίας .....  | 20 |
| ΚΕΦΑΛΑΙΟ 2: Θεωρίες Συναισθημάτων.....                                    | 21 |
| 2.1 Εισαγωγικά στοιχεία .....   | 21 |
| 2.1.1 Ορισμός συναισθήματος.....  | 21 |
| 2.1.2 Βασικότερες κατηγορίες θεωριών συναισθημάτων.....                   | 22 |
| 2.2 Θεωρία James-Lange.....   | 23 |
| 2.3 Θεωρία Cannon-Bard.....   | 24 |
| 2.4 Θεωρία Schachter-Singer .....   | 26 |
| 2.5 Θεωρία γνωσιακής αξιολόγησης .....                                    | 28 |
| 2.5.1 Δομικό μοντέλο γνωσιακής αξιολόγησης.....                           | 29 |
| 2.5.2 Διαδικαστικό μοντέλο γνωσιακής αξιολόγησης.....                     | 30 |
| 2.6 Υπόθεση ανατροφοδότησης της έκφρασης του προσώπου .....               | 32 |
| 2.6.1 Πείραμα του Strack .....  | 33 |
| 2.6.2 Πειράματα με βοτουλινική τοξίνη.....                                | 34 |
| 2.7 Θεωρία συναισθημάτων του Ekman.....                                   | 35 |
| 2.7.1 Καθολικότητα συναισθημάτων και αρχέτυπα συναισθήματα .....          | 35 |
| 2.7.2 Σύστημα κωδικοποίησης κινήσεων προσώπου FACS .....                  | 36 |
| 2.8 Θεωρία συναισθημάτων του Plutchik.....                                | 39 |
| 2.8.1 Συνδυασμοί βασικών συναισθημάτων .....                              | 41 |
| 2.8.2 Ο τροχός συναισθημάτων του Plutchik.....                            | 42 |
| 2.8.3 Συμπεριφορές που συνδέονται με τα βασικά συναισθήματα.....          | 44 |
| 2.9 Αναπαράσταση συναισθημάτων της Whissell.....                          | 45 |
| 2.10 Εργασίες στη συναισθηματική υπολογιστική .....                       | 47 |
| ΚΕΦΑΛΑΙΟ 3: Το πρότυπο MPEG-4.....  | 49 |
| 3.1 Εισαγωγικά στοιχεία .....   | 49 |
| 3.2 Τα χαρακτηριστικά και το πεδίο εφαρμογής του προτύπου MPEG-4.....     | 49 |
| 3.2.1 Κωδικοποιημένη αναπαράσταση πολυμεσικών αντικειμένων. ....          | 50 |
| 3.2.2 Σύνθεση πολυμεσικών αντικειμένων.....                               | 51 |
| 3.2.3 Περιγραφή, συγχρονισμός και μεταφορά πολυμεσικών αντικειμένων ..... | 52 |

|   |   |     |
|---|---|-----|
| 3.2.4   | Αλληλεπίδραση του χρήστη με τα πολυμεσικά αντικείμενα .....                                     | 53  |
| 3.2.5   | Αναγνώριση και διαχείριση της πνευματικής ιδιοκτησίας .....                                     | 53  |
| 3.3   | Εκδόσεις του MPEG-4.....  | 54  |
| 3.4   | Προφίλ και επίπεδα του MPEG-4 .....   | 55  |
| 3.5   | Τμήματα του MPEG-4.....   | 56  |
| 3.6   | Συνθετικά πρόσωπα στο MPEG-4.....   | 58  |
| 3.6.1   | Γενικά στοιχεία για τα συνθετικά πρόσωπα του MPEG-4 .....                                       | 58  |
| 3.6.2   | Μοντέλο προσώπου του MPEG-4 στην ουδέτερη κατάσταση.....  | 59  |
| 3.6.3   | Χαρακτηριστικά σημεία του μοντέλου προσώπου .....   | 61  |
| 3.6.4   | Μονάδες των FAPs .....  | 61  |
| 3.6.5   | Ομαδοποίηση των παραμέτρων FAP του MPEG-4.....  | 63  |
| 3.6.6   | Κωδικοποίηση FAPs με χρήση μάσκας.....  | 66  |
| 3.6.7   | Προσαρμοσμένα μοντέλα προσώπου στο MPEG-4 .....   | 67  |
| ΚΕΦΑΛΑΙΟ 4: Ασαφής Λογική και Ασαφή Σύνολα .....  |   | 69  |
| 4.1   | Εισαγωγικά Στοιχεία .....   | 69  |
| 4.2   | Όψεις της Ασαφούς Λογικής .....   | 69  |
| 4.2.1   | Όψη Ασαφών Συνόλων.....   | 71  |
| 4.2.2   | Όψη της Λογικής.....  | 71  |
| 4.2.3   | Επιστημική όψη.....   | 71  |
| 4.2.4   | Σχεσιακή όψη .....  | 72  |
| 4.3   | Ασαφή Σύνολα .....  | 73  |
| 4.3.1   | Συνήθεις συναρτήσεις συμμετοχής .....   | 74  |
| 4.3.2   | Πράξεις ασαφών συνόλων.....   | 80  |
| 4.3.3   | Τριγωνικές νόρμες <i>t-norms</i> , <i>t-conorms</i> .....                                       | 81  |
| 4.3.4   | Αρχή της επέκτασης.....   | 81  |
| ΚΕΦΑΛΑΙΟ 5: Υπολογιστική υλοποίηση συστήματος εκτίμησης<br>συναισθημάτων στη γλώσσα προγραμματισμού C ..... |   | 83  |
| 5.1   | Εισαγωγικά στοιχεία.....  | 83  |
| 5.2   | Είσοδοι του προγράμματος εκτίμησης συναισθημάτων .....  | 84  |
| 5.2.1   | Συντεταγμένες χαρακτηριστικών σημείων στο πρόσωπο.....  | 84  |
| 5.2.2   | Πληροφορίες μοντελοποιημένων προφίλ.....  | 88  |
| 5.3   | Μετασχηματισμός περιστροφής των χαρακτηριστικών σημείων.....                                    | 89  |
| 5.3.1   | Προσδιορισμός του προσανατολισμού του προσώπου.....   | 90  |
| 5.3.2   | Προσδιορισμός του κέντρου περιστροφής ( $x_0$ , $y_0$ ) και της γωνίας στροφής $\theta_0$ ..... | 96  |
| 5.3.3   | Οριακές θέσεις – Ειδικές περιπτώσεις ευθειών .....  | 98  |
| 5.3.4   | Μετασχηματισμός σημείων – Περιστροφή και μεταφορές.....   | 99  |
| 5.4   | Παράγωγα χαρακτηριστικά σημεία .....  | 102 |
| 5.5   | Υπολογισμός μονάδων των FAPs .....  | 105 |
| 5.6   | «Λεξιλόγιο» FAPs της εφαρμογής.....   | 106 |
| 5.7   | Υπολογισμός των FAPs.....   | 109 |
| 5.8   | Σύστημα ανάλυσης και εκτίμησης συναισθημάτων .....  | 113 |

|  |   |     |
|--|---|-----|
| 5.9  | Αποτίμηση Κανόνων.....                                    | 118 |
| 5.10   | Επιλογές εκτύπωσης και εμφάνιση αποτελεσμάτων.....        | 120 |
| 5.10.1   | Παράμετροι συναρτήσεων που ελέγχουν την εκτύπωση.....     | 120 |
| 5.10.2   | Παραδείγματα εκτύπωσης πληροφοριών.....                   | 121 |
| 5.11   | Χρήση προγράμματος.....                                   | 125 |
| 5.12   | Περιοχές τροποποίησης και επέκτασης του προγράμματος..... | 127 |
| ΚΕΦΑΛΑΙΟ 6: Συμπεράσματα και μελλοντικές επεκτάσεις..... |   | 129 |
| 6.1  | Αξιολόγηση της μεθόδου.....                               | 129 |
| 6.1.1  | Πλεονεκτήματα.....  | 129 |
| 6.1.2  | Μειονεκτήματα.....  | 130 |
| 6.2  | Μελλοντικές εφαρμογές και επεκτάσεις.....                 | 131 |
| 6.2.1  | Προτεινόμενες εναλλακτικές εφαρμογές.....                 | 131 |
| 6.2.2  | Προτεινόμενες επεκτάσεις.....                             | 132 |
| Παράρτημα .....  |   | 133 |
| Κανόνες για τα αρχέτυπα συναισθήματα .....               |   | 135 |
| Αναφορές.....  |   | 141 |



# Ευρετήριο Σχημάτων

|            |  |     |
|------------|--|-----|
| Σχήμα 2-1  | Θεωρία James-Lange .....   | 23  |
| Σχήμα 2-2  | Παράδειγμα θεωρίας James-Lange .....                                       | 24  |
| Σχήμα 2-3  | Θεωρία Cannon-Bard.....  | 25  |
| Σχήμα 2-4  | Παράδειγμα θεωρίας Cannon-Bard .....                                       | 25  |
| Σχήμα 2-5  | Θεωρία Schachter-Singer.....   | 27  |
| Σχήμα 2-6  | Παράδειγμα θεωρίας Schachter-Singer .....                                  | 28  |
| Σχήμα 2-7  | Θεωρία γνωσιακής αξιολόγησης.....  | 31  |
| Σχήμα 2-8  | Παράδειγμα θεωρίας γνωσιακής αξιολόγησης (α) .....                         | 31  |
| Σχήμα 2-9  | Παράδειγμα θεωρίας γνωσιακής αξιολόγησης (β) .....                         | 32  |
| Σχήμα 2-10 | Κώνος και τροχός συναισθημάτων του Plutchik .....                          | 43  |
| Σχήμα 2-11 | Προσδιορισμός γωνιακών θέσεων συναισθημάτων.....                           | 44  |
| Σχήμα 2-12 | Διδιάστατος χώρος αξιολόγησης – ενεργοποίησης .....                        | 46  |
| Σχήμα 3-1  | Εκδόσεις του MPEG-4 .....  | 54  |
| Σχήμα 3-2  | Χαρακτηριστικά σημεία στο μοντέλο προσώπου του MPEG-4 .....                | 60  |
| Σχήμα 3-3  | Αποστάσεις για τον ορισμό των FAPU.....                                    | 62  |
| Σχήμα 3-4  | Παράδειγμα εκφράσεων προσώπου για το FAP 2 .....                           | 65  |
| Σχήμα 4-1  | Κύριες Όψεις της Ασαφούς Λογικής.....                                      | 70  |
| Σχήμα 4-2  | Κοκκοποίηση μιας μεταβλητής .....  | 70  |
| Σχήμα 4-3  | Κοκκοποίηση μιας συνάρτησης.....   | 72  |
| Σχήμα 4-4  | Ασαφείς ποσότητες .....  | 75  |
| Σχήμα 4-5  | Τριγωνική συνάρτηση συμμετοχής .....                                       | 77  |
| Σχήμα 4-6  | Τραπεζοειδής συνάρτηση συμμετοχής .....                                    | 78  |
| Σχήμα 4-7  | Καμπανοειδής συνάρτηση συμμετοχής .....                                    | 78  |
| Σχήμα 4-8  | Συμμετρική συνάρτηση συμμετοχής Gauss .....                                | 79  |
| Σχήμα 4-9  | Σιγμοειδής συνάρτηση συμμετοχής.....                                       | 79  |
| Σχήμα 4-10 | Συνάρτηση συμμετοχής Z .....   | 79  |
| Σχήμα 5-1  | Μπλοκ διάγραμμα συστήματος εκτίμησης συναισθημάτων.....                    | 83  |
| Σχήμα 5-2  | Σημεία εισόδου στο σύστημα εκτίμησης συναισθημάτων .....                   | 86  |
| Σχήμα 5-3  | Χειροκίνητη ανίχνευση σημείων – έκφραση απέχθειας.....                     | 87  |
| Σχήμα 5-4  | Χειροκίνητη ανίχνευση σημείων – έκφραση φόβου.....                         | 87  |
| Σχήμα 5-5  | Αποκλίσεις στη μέθοδο ελαχίστων τετραγώνων .....                           | 91  |
| Σχήμα 5-6  | Υπολογισμός κάθετων στην ευθεία αποκλίσεων για τη συνάρτηση σφάλματος..... | 93  |
| Σχήμα 5-7  | Κέντρο και γωνία περιστροφής χαρακτηριστικών σημείων .....                 | 97  |
| Σχήμα 5-8  | Περιστροφή χαρακτηριστικών σημείων του προσώπου – έκφραση χαράς .....      | 101 |
| Σχήμα 5-9  | Περιστροφή χαρακτηριστικών σημείων του προσώπου – έκφραση θυμού .....      | 101 |
| Σχήμα 5-10 | Παράγωγα σημεία ενδιαφέροντος.....   | 103 |
| Σχήμα 5-11 | Αποστάσεις για τις μονάδες των FAPs .....                                  | 106 |
| Σχήμα 5-12 | Συνάρτηση συμμετοχής $\mu$ για την αναγνώριση συναισθημάτων .....          | 114 |
| Σχήμα 5-13 | Συνάρτηση συμμετοχής $\lambda$ για την αναγνώριση συναισθημάτων .....      | 115 |
| Σχήμα 5-14 | Συνάρτηση συμμετοχής $\nu$ για την αναγνώριση συναισθημάτων.....           | 115 |
| Σχήμα 5-15 | Διάγραμμα ροής της αποτίμησης κανόνων .....                                | 119 |

# Ευρετήριο Πινάκων

|              |  |     |
|--------------|--|-----|
| Πίνακας 2-1  | Παράδειγμα βασικών κωδικών AU του συστήματος FACS.....                               | 38  |
| Πίνακας 2-2  | Παράδειγμα κωδικών AU κίνησης κεφαλιού στο σύστημα FACS.....                         | 38  |
| Πίνακας 2-3  | Παράδειγμα κωδικών AU κίνησης ματιών στο σύστημα FACS.....                           | 39  |
| Πίνακας 2-4  | Παράδειγμα κωδικών AU ορατότητας και γενικής συμπεριφοράς.....                       | 39  |
| Πίνακας 2-5  | Συνδυασμοί βασικών συναισθημάτων κατά Plutchik .....                                 | 42  |
| Πίνακας 2-6  | Συμπεριφορές για τα βασικά συναισθήματα του Plutchik.....                            | 45  |
| Πίνακας 3-1  | Παράδειγμα οπτικών προφίλ του MPEG-4 .....   | 56  |
| Πίνακας 3-2  | Τμήματα του MPEG-4.....  | 57  |
| Πίνακας 3-3  | Ορισμοί Μονάδων FAP του MPEG-4 (FAPU).....   | 62  |
| Πίνακας 3-4  | Ομάδες των FAPs στο MPEG-4 .....   | 63  |
| Πίνακας 3-5  | FAP 1: Visemes και αντίστοιχα φωνήματα.....  | 64  |
| Πίνακας 3-6  | Βασικές εκφράσεις προσώπου όπως ορίζονται για το FAP 2 .....                         | 65  |
| Πίνακας 3-7  | Παράδειγμα αντιστοιχίας AU με FAPs .....   | 68  |
| Πίνακας 5-1  | Σημεία εισόδου στο σύστημα εκτίμησης συναισθημάτων.....                              | 85  |
| Πίνακας 5-2  | Βοηθητικό πρόγραμμα m-file για χειροκίνητη ανίχνευση σημείων στο MATLAB ....         | 88  |
| Πίνακας 5-3  | Παράδειγμα μοντελοποιημένου προφίλ όπως διαβάζεται από το αρχείο .txt .....          | 88  |
| Πίνακας 5-4  | Παράμετρος rotate για την επιλογή μεθόδου περιστροφής των σημείων .....              | 100 |
| Πίνακας 5-5  | Παράγωγα χαρακτηριστικά σημεία .....   | 102 |
| Πίνακας 5-6  | Αντιστοιχία χαρακτηριστικών σημείων με μεταβλητές αποθήκευσης του προγράμματος ..... | 104 |
| Πίνακας 5-7  | Ορισμοί Μονάδων FAPU: FAP Units .....  | 105 |
| Πίνακας 5-8  | FAPs που δεν μπορούν να υλοποιηθούν .....  | 107 |
| Πίνακας 5-9  | Λεξιλόγιο FAPs της εφαρμογής .....   | 108 |
| Πίνακας 5-10 | Αντιστοιχία FAPs $F_4 - F_{13}$ με $F_{51} - F_{60}$ .....                           | 109 |
| Πίνακας 5-11 | Υπολογισμός των FAPs .....   | 110 |
| Πίνακας 5-12 | Παράδειγμα υπολογισμού ενός FAP.....   | 111 |
| Πίνακας 5-13 | Υπολογισμός των FAPs $F_4 - F_{13}$ .....  | 112 |
| Πίνακας 5-14 | Επιλογές εκτύπωσης .....   | 121 |
| Πίνακας 5-15 | Λεπτομερής εκτύπωση συντεταγμένων σημείων ενδιαφέροντος. ....                        | 122 |
| Πίνακας 5-16 | Βασική εκτύπωση συντεταγμένων σημείων ενδιαφέροντος. ....                            | 123 |
| Πίνακας 5-17 | Λεπτομερής εκτύπωση υπολογισμένων FAPs. ....   | 124 |
| Πίνακας 5-18 | Βασική εκτύπωση υπολογισμένων FAPs.....  | 124 |
| Πίνακας 5-19 | Εκτύπωση πληροφοριών αποτίμησης κανόνων. ....  | 125 |



# ΚΕΦΑΛΑΙΟ 1: Εισαγωγή

## 1.1 Αναγνώριση συναισθημάτων στην επικοινωνία ανθρώπου - μηχανής

Τα τελευταία χρόνια υπάρχει αυξημένο ενδιαφέρον για τη βελτίωση όλων των πτυχών της αλληλεπίδρασης μεταξύ ανθρώπων και υπολογιστικών συστημάτων. Αυτός ο νέος τομέας έρευνας ελκύει το ενδιαφέρον από διάφορα επιστημονικά πεδία, όπως η ψυχολογία, η νευροεπιστήμη αλλά και η επιστήμη υπολογιστών. Η έρευνα εστιάζει στη βελτίωση της διεπαφής των υπολογιστών αλλά και στη βελτίωση της απόκρισης του υπολογιστή με βάση την ανατροφοδότηση από το χρήστη. Τα παραδοσιακά μέσα για την επικοινωνία με τον υπολογιστή ήταν τα περιφερειακά όπως το ποντίκι και το πληκτρολόγιο. Με την ταχεία εξέλιξη στην αύξηση της ταχύτητας και της απόδοσης των υπολογιστικών συστημάτων καθώς και τη μείωση του κόστους περιφερειακών όπως οι κάμερες και τα μικρόφωνα, παρέχεται η δυνατότητα εισόδου και επεξεργασίας οπτικής και ηχητικής πληροφορίας στον υπολογιστή.

Η ευφυής αλληλεπίδραση ανθρώπου-μηχανής είναι ένας ανερχόμενος τομέας που στοχεύει στην παροχή φυσικών τρόπων επικοινωνίας μεταξύ χρήστη και υπολογιστή. Για να μπορεί ο υπολογιστής να αλληλεπιδρά με τον άνθρωπο, χρειάζεται να έχει ή να προσομοιώνει τις επικοινωνιακές δεξιότητες ενός ανθρώπου. Μία από αυτές τις δεξιότητες είναι και η ικανότητα να αντιλαμβάνεται κανείς τη συναισθηματική κατάσταση ενός άλλου ατόμου με το οποίο συνομιλεί ή εν γένει αλληλεπιδρά. Ένας από τους πιο εκφραστικούς τρόπους που οι άνθρωποι εξωτερικεύουν τα συναισθήματά τους πέραν της ομιλίας, είναι μέσω των διαφόρων εκφράσεων του προσώπου και των κινήσεων του σώματος που συνοδεύουν τις εκφράσεις, ειδικά του κεφαλιού και των χεριών. Για το λόγο αυτό η αυτόματη αναγνώριση συναισθημάτων από υπολογιστές στα πλαίσια της συναισθηματικής υπολογιστικής συγκεντρώνει έντονο ερευνητικό ενδιαφέρον.

Ένα σύστημα αναγνώρισης συναισθημάτων μπορεί να δέχεται σαν είσοδο την πληροφορία της ομιλίας ενός χρήστη, την εικόνα του προσώπου του ή τη χρονική ακολουθία βίντεο των μεταβολών στις εκφράσεις και κινήσεις του. Τα ζητήματα που σχετίζονται με την επιτυχή και αποδοτική λειτουργία ενός τέτοιου συστήματος στη γενική περίπτωση είναι η ανίχνευση των ζητούμενων πληροφοριών ως δεδομένα εισόδου (εικόνα, φωνή, ακολουθία εικόνων βίντεο), η αναγνώριση των πληροφοριών (π.χ. αναγνώριση της έκφρασης του προσώπου) και εν τέλει η επιτυχής επεξεργασία των πληροφοριών ώστε να πραγματοποιείται η αναμενόμενη απόκριση του υπολογιστή σε σχέση με τα δεδομένα εισόδου.

Ένα αποτελεσματικό και πρακτικά αποδοτικό σύστημα αναγνώρισης ή εκτίμησης συναισθημάτων θα έβρισκε εφαρμογή σε διαδραστικές εφαρμογές μεταξύ ανθρώπου-μηχανής, ηλεκτρονικά παιχνίδια και ψυχολογικές μελέτες. Δεν λείπει και το ενδιαφέρον για τη χρήση αναγνώρισης συναισθημάτων σε άλλες εμπορικές εφαρμογές όπως μάρκετινγκ, διαδικτυακή μάθηση, τεχνική υποστήριξη, ψυχαγωγικές εφαρμογές κλπ.

Κάποιες από τις προκλήσεις που καλείται να αντιμετωπίσει η αναγνώριση συναισθημάτων για παράδειγμα από εικόνες προσώπου, είναι οι τεχνικές δυσκολίες που συνδέονται με την επιτυχημένη ανίχνευση εκφράσεων του προσώπου, τις ατομικές ιδιαιτερότητες στη φυσιογνωμία και τις αντιδράσεις των ανθρώπων, καθώς και η δυσχέρεια στη μοντελοποίηση και κατηγοριοποίηση ολόκληρου του φάσματος των ανθρώπινων εκφράσεων του προσώπου, συναισθηματικών και μη. Ιδιαίτερα αυτή η τελευταία δυσχέρεια είναι σημαντικό να συζητηθεί, καθώς η πλειοψηφία των ερευνών που έχουν γίνει με στόχο τη μοντελοποίηση και δομημένη περιγραφή των συναισθημάτων εστιάζει κατ'ανάγκη σε ένα υποσύνολο του φάσματος των ανθρώπινων εκφράσεων, και μάλιστα σε ένα συγκριτικά πολύ μικρό πλήθος, που είναι συνήθως τα αρχέτυπα ή αλλιώς βασικά συναισθήματα.

## 1.2 Αντικείμενο διπλωματικής εργασίας

Ο σκοπός της παρούσας διπλωματικής εργασίας ήταν η ανάπτυξη μιας υπολογιστικής εφαρμογής που θα χρησιμεύει για την εκτίμηση του συναισθήματος σε εικόνες προσώπων προερχόμενες από ακολουθίες βίντεο. Το πρόγραμμα γράφηκε στη γλώσσα C.

Ένα από τα ζητούμενα της εφαρμογής ήταν να μπορεί να χρησιμοποιηθεί σε πραγματικό χρόνο, σε συνδυασμό με ήδη υπάρχον πρόγραμμα ανίχνευσης χαρακτηριστικών σημείων στο πρόσωπο, το οποίο λειτουργεί επίσης σε πραγματικό χρόνο και υπολογίζει για μια ακολουθία βίντεο τη θέση των σημείων στο πρόσωπο και την κατεύθυνση του βλέμματος. Για να επιτευχθεί αυτός ο στόχος, το πρόγραμμα που αναπτύχθηκε στα πλαίσια της διπλωματικής εργασίας δέχεται ως είσοδο τις συντεταγμένες των σημείων που έχουν προκύψει από την ανίχνευση. Τα δεδομένα αυτά διαβάζονται από αρχείο κειμένου με τέτοιο τρόπο ώστε να είναι δυνατή η εκτέλεση σε πραγματικό χρόνο.

Ο χρησιμοποιούμενος αλγόριθμος στηρίζεται στον υπολογισμό των παραμέτρων FAP του διεθνούς προτύπου MPEG-4. Η επιλογή αυτή επιτρέπει τη χρήση παραμέτρων που συνδέονται στενά με τις μονάδες κίνησης AU (Action Units) του συστήματος FACS, οι οποίες αξιοποιούν τα αποτελέσματα νευροφυσιολογικών και ψυχολογικών ερευνών με μεγάλη επιρροή και αναγνωρισμένο κύρος στον τομέα της θεωρίας των ανθρώπινων συναισθημάτων.

Όπως αναφέρθηκε προηγουμένως, μία από τις προκλήσεις στην αναγνώριση συναισθημάτων με βάση την εικόνα του προσώπου είναι το πεπερασμένο πλήθος των συναισθημάτων που αναγνωρίζονται, που συνήθως είναι τα αρχέτυπα συναισθήματα ή κάποιες άλλες διαθέσιμες αρχετυπικές εκφράσεις. Η παρούσα εργασία επιχειρεί να ανεξαρτητοποιήσει τη διαδικασία της εκτίμησης από τη διαθεσιμότητα των προφίλ συναισθημάτων. Οι πληροφορίες σχετικά με τα μοντελοποιημένα προφίλ διαβάζονται κατά την εκτέλεση του προγράμματος και δεν ενσωματώνονται στην υλοποίηση, επιτρέποντας έτσι την επέκτασή τους ή την τροποποίηση κατά βούληση. Μία από τις βασικές επιδιώξεις ήταν η διατήρηση της ευελιξίας του προγράμματος ούτως ώστε να μεγιστοποιείται η χρησιμότητα για διάφορα σύνολα δεδομένων και να διευκολύνεται ο πειραματισμός με ποικίλα μοντελοποιημένα προφίλ συναισθημάτων.

### 1.2.1 Συνεισφορά εργασίας

Η συνεισφορά της διπλωματικής εργασίας συνοψίζεται στα εξής:

- Παρουσιάστηκε λεπτομερώς το θεωρητικό υπόβαθρο που είναι απαραίτητο για την κατανόηση του θέματος και τη θεωρητική θεμελίωση της εφαρμογής.
- Το πρόγραμμα υλοποιήθηκε με τέτοιο τρόπο ώστε να παρέχει μια ευέλικτη εφαρμογή με χαμηλό υπολογιστικό κόστος, κατάλληλη για εφαρμογές πραγματικού χρόνου. Η συνεισφορά της υπολογιστικής εφαρμογής περιλαμβάνει τα παρακάτω σημεία:
  - Εισήχθη μέθοδος προεπεξεργασίας των συντεταγμένων εισόδου ώστε να αντισταθμίζεται η στροφή του προσώπου συγκεκριμένης κατεύθυνσης. Υλοποιήθηκαν τρεις αλγόριθμοι για το μετασχηματισμό περιστροφής και συζητήθηκαν τα ιδιαίτερα χαρακτηριστικά κάθε αλγορίθμου.
  - Παρουσιάστηκε τρόπος υπολογισμού παράγωγων σημείων που απαιτούνται για τον υπολογισμό κάποιων FAPs.
  - Η εργασία προτείνει τρόπους υπολογισμού των FAPs και των μονάδων FAPU ώστε να εξαλείφεται κατά το δυνατόν η επίδραση ασυμμετριών στο πρόσωπο. Στόχος είναι η χρήση μεθόδων υπολογισμού που είναι κατάλληλες για πραγματικά πρόσωπα.
  - Εξετάστηκαν όλα τα FAPs ώστε να υλοποιείται το μέγιστο δυνατό λεξιλόγιο FAPs, έχοντας ως δεδομένο τη διαθεσιμότητα 19 χαρακτηριστικών σημείων του προσώπου. Το λεξιλόγιο FAPs παρουσιάζεται λεπτομερώς.
  - Προτάθηκε τρόπος για την ανεξαρτητοποίηση των μοντελοποιημένων προφίλ των συναισθημάτων από τα υπόλοιπα μέρη του συστήματος, ώστε να μην απαιτείται πλέον η ενσωμάτωση της πληροφορίας των προφίλ στον υπόλοιπο κώδικα και να διατηρείται η ευελιξία.
  - Συζητήθηκαν και προτάθηκαν κριτήρια για την επιλογή της τελικής εκτίμησης του προφίλ της έκφρασης του προσώπου.
- Αξιολογήθηκε η μέθοδος. Διερευνήθηκαν τα προτερήματα, τα πλεονεκτήματα και τα δυνατά σημεία της μεθόδου όπως επίσης οι πρακτικοί περιορισμοί, οι αδυναμίες της μεθόδου και της υλοποίησης.
- Προτάθηκαν περιοχές τροποποίησης και επέκτασης του προγράμματος. Επίσης συζητήθηκε η εναλλακτική χρήση του προγράμματος για διαφορετικές εφαρμογές που σχετίζονται με την ανίχνευση κινήσεων του προσώπου.

## 1.3 Διάρθρωση Εργασίας

- Το **1<sup>ο</sup> κεφάλαιο** της εργασίας περιέχει τα **εισαγωγικά στοιχεία** που είναι απαραίτητα για να κατατοπίσουν τον αναγνώστη σχετικά με το αντικείμενο και το πεδίο εφαρμογής της διπλωματικής.

- Το **2<sup>ο</sup> κεφάλαιο** εστιάζει στις **θεμελιώδεις θεωρίες συναισθημάτων** που αποτέλεσαν βάση για τις επιστημονικές μελέτες πάνω στο αντικείμενο της ανθρώπινης συναισθηματικής εμπειρίας. Αρχικά γίνεται παρουσίαση των πιο γνωστών θεωριών που είχαν μεγάλη επίδραση ιστορικά στη διαμόρφωση της γνώσης περί συναισθημάτων και έδωσαν σημαντική ώθηση στην έρευνα αυτού του τομέα. Συζητείται η σημασία που έχουν οι εκφράσεις του προσώπου στη συναισθηματική εμπειρία. Επίσης παρουσιάζονται εργασίες που εστιάζουν περισσότερο στη **μοντελοποίηση και ποσοτικοποίηση βασικών συναισθημάτων**. Αυτή η προσέγγιση τις καθιστά κατάλληλες για πρακτικές εφαρμογές. Το κεφάλαιο κλείνει με μια σύντομη παρουσίαση εργασιών στο πεδίο της συναισθηματικής υπολογιστικής.

- Το **3<sup>ο</sup> κεφάλαιο** περιέχει πληροφορίες σχετικά με το **διεθνές πρότυπο MPEG-4**, που είναι πρότυπο για την κωδικοποίηση και συμπίεση φυσικού και συνθετικού οπτικοακουστικού περιεχομένου με βάση τα αντικείμενα. Η βαρύτητα δίνεται στα **συνθετικά πρόσωπα** του MPEG-4 και τις **παραμέτρους κίνησης του προσώπου**, τα λεγόμενα **FAPs** (Facial Animation Parameters). Η χρησιμότητα των FAPs –και ο λόγος που χρησιμοποιήθηκαν στην παρούσα υλοποίηση– έγκειται στο ότι προσφέρουν τη δυνατότητα υπολογισμού μεταβολών σε φυσικά ή συνθετικά πρόσωπα, με εύχρηστο τρόπο και χαμηλό υπολογιστικό κόστος. Οι κανόνες που αποτιμούνται από την εφαρμογή έχουν διατυπωθεί ως διαστήματα των παραμέτρων FAP.

- Το **4<sup>ο</sup> κεφάλαιο** παρουσιάζει τις βασικές αρχές της **ασαφούς λογικής** και της **ασαφούς συνολοθεωρίας**. Οι έννοιες που αναλύονται είναι χρήσιμες για τη διαδικασία αποτίμησης των κανόνων, η οποία αξιοποιεί τα οφέλη της ασαφούς λογικής.

- Στο **5<sup>ο</sup> κεφάλαιο** παρουσιάζεται αναλυτικά η **υπολογιστική εφαρμογή** που αναπτύχθηκε κατά την εκπόνηση της διπλωματικής στη γλώσσα προγραμματισμού C. Αρχικά δίνονται τα βασικά **δομικά μέρη** του ευρύτερου συστήματος, και στη συνέχεια αναλύονται τα επιμέρους τμήματα. Αναλύονται **οι είσοδοι** του συστήματος, η **επεξεργασία των δεδομένων** εισόδου, οι **υπολογισμοί** των παραμέτρων **FAP** και το υποσύστημα που είναι υπεύθυνο για την **αποτίμηση των κανόνων**. Καταγράφονται οι **επιλογές εκτύπωσης** των δεδομένων εξόδου που είναι διαθέσιμες στο χρήστη, και δίνονται παραδείγματα εκτύπωσης. Παρουσιάζονται συνολικά όλες οι παράμετροι που μπορούν να δεχτούν τροποποίηση κατά τη χρήση του προγράμματος και όλες οι επιλογές που είναι στη διάθεση του χρήστη. Τέλος συζητούνται οι περιοχές του προγράμματος που μπορούν να υποστούν τροποποίηση με σκοπό φερ'επείν τον πειραματισμό με διαφορετικά δεδομένα, χωρίς να αλλάξει ριζικά η μεθοδολογία που υλοποιήθηκε.

- Το **τελευταίο κεφάλαιο (6<sup>ο</sup>)** αποτελεί τον **επίλογο** της εργασίας. Συνοψίζει τα βασικά συμπεράσματα της υλοποίησης, αξιολογεί εν συντομία τη μέθοδο που χρησιμοποιήθηκε και συζητά προτάσεις για μελλοντικές επεκτάσεις και εφαρμογές.

# ΚΕΦΑΛΑΙΟ 2: Θεωρίες Συναισθημάτων

## 2.1 Εισαγωγικά στοιχεία

Οι περισσότεροι μεγάλοι φιλόσοφοι από την αρχαιότητα έως σήμερα έχουν διατυπώσει θεωρίες σχετικά με τα συναισθήματα. Ο Πλάτων, ο Αριστοτέλης, και μεταγενέστερα οι Descartes, Spinoza, Hume, Kant μεταξύ πολλών άλλων, διατύπωσαν τις απόψεις τους υπογραμμίζοντας με τον τρόπο αυτό τη σημασία των συναισθημάτων στην ανθρώπινη ύπαρξη [1].

Τα τελευταία χρόνια η μελέτη των συναισθημάτων έχει επανέλθει στο επίκεντρο του ενδιαφέροντος της ψυχολογίας, της φιλοσοφίας και των άλλων γνωστικών επιστημών. Η γνώση για το συναίσθημα έχει ενδιαφέρον και για περισσότερο πρακτικές επιστήμες όπως οι οικονομικές επιστήμες, η τεχνολογία υπολογιστών, η επικοινωνία ανθρώπου-μηχανής.

Στο παρόν κεφάλαιο εξετάζουμε κάποιες από τις βασικότερες σύγχρονες θεωρίες συναισθημάτων που έχουν διατυπωθεί τους τελευταίους δύο αιώνες, καθώς και τις σημαντικότερες εργασίες που έχουν πραγματοποιηθεί και χρησιμοποιούνται στα πλαίσια της παρούσας εργασίας.

### 2.1.1 Ορισμός συναισθήματος

Παρότι η έννοια των συναισθημάτων είναι αδιαμφισβήτητα οικεία σε όλους τους ανθρώπους, το εγχείρημα να δοθεί ένας πλήρης, κοινώς αποδεκτός ορισμός ενέχει ακόμη πολλές προκλήσεις. Στη βιβλιογραφία απαντώνται πολλοί ορισμοί με τον καθένα να δίνει έμφαση σε διαφορετικές πτυχές του συναισθήματος [2].

Από το γλωσσάρι της APA (American Psychological Association – Αμερικανική Ψυχολογική Εταιρεία) παρατίθεται ένας ορισμός του συναισθήματος σε ελληνική απόδοση:

***Συναίσθημα:** Ένα περίπλοκο σύνολο αλλαγών που περιλαμβάνει τη σωματική διέγερση, τη συγκίνηση, γνωσιακές διαδικασίες καθώς και συμπεριφορικές αντιδράσεις. Οι προαναφερθείσες αλλαγές πραγματοποιούνται ως απόκριση σε μία κατάσταση την οποία το άτομο αντιλαμβάνεται ως σημαντική για εκείνο. [3]*

Ο ορισμός της APA επιχειρεί να συμπεριλάβει τις βασικότερες πτυχές των συναισθημάτων που έχουν μελετηθεί ως τώρα, δηλαδή το ρόλο της σωματικής διέγερσης, της φυσιολογικής απόκρισης του ατόμου στο ερέθισμα, καθώς και τις γνωσιακές διαδικασίες που λαμβάνουν χώρα δεδομένου του ερεθίσματος. Το τρίπτυχο «σωματική διέγερση – συγκίνηση (δηλαδή η συναισθηματική εμπειρία) – γνωσιακές διαδικασίες» έχει κεντρικό ρόλο στη βιβλιογραφία είτε πρόκειται για τον ορισμό του συναισθήματος είτε τις διάφορες θεωρίες για το συναίσθημα, όπως θα φανεί και παρακάτω. Επίσης ορισμός αναφέρεται στις συμπεριφορικές αντιδράσεις, δίνοντας έτσι έμφαση στο λειτουργικό ρόλο του συναισθήματος. Η διαφοροποίηση των συναισθημάτων και ο υποκειμενικός χαρακτήρας τους αποδίδονται στον τρόπο που το άτομο αξιολογεί το ερέθισμα, μέσω συνειδητών ή υποσυνειδητών γνωστικών διαδικασιών. Αυτό διαφαίνεται και από την τελευταία πρόταση του παραπάνω ορισμού.

Τονίζεται ότι έχουν διατυπωθεί πολλοί ορισμοί για το συναίσθημα στη βιβλιογραφία. Ο παραπάνω ορισμός είναι ένα παράδειγμα πολύπτυχου ορισμού που περιλαμβάνει πολλές από τις αλληλένδετες συνιστώσες των συναισθημάτων. Κάποιοι άλλοι ορισμοί θα αναφερθούν παρακάτω, στις θεωρίες των συναισθημάτων.

### 2.1.2 Βασικότερες κατηγορίες θεωριών συναισθημάτων

Οι πρώτες θεωρίες σχετικές με τα συναισθήματα είχαν ήδη αρχίσει να διατυπώνονται από την αρχαιότητα, για παράδειγμα από τους στωικούς φιλόσοφους, τον Πλάτωνα, τον Αριστοτέλη, το Σωκράτη, το Δημόκριτο και πολλούς άλλους. Την εποχή εκείνη επρόκειτο κυρίως για στοχασμούς φιλοσοφικού περιεχομένου που αναφέρονταν σε θεωρίες για τη μεταφυσική προέλευση των συναισθημάτων, την ηθική κατάταξή τους, τη σχέση με την ψυχή του ανθρώπου και την ανθρώπινη φύση γενικότερα. Δεν έλειπε όμως και η περισσότερο συστηματοποιημένη προσέγγιση της έννοιας των συναισθημάτων καθώς και η στοχευμένη προσπάθεια σύνδεσης των συναισθημάτων με τη φυσιολογία του ανθρώπου.

Αναφέρουμε χαρακτηριστικά τη θεωρία των τεσσάρων ιδιοσυγκρασιών ή αλλιώς των τεσσάρων χυμών της σχολής του Ιπποκράτη, ο οποίος πίστευε ότι η διάθεση, η συμπεριφορά και η υγεία των ανθρώπων καθορίζεται από το μίγμα τεσσάρων χυμών στο ανθρώπινο σώμα. Οι τέσσερις χυμοί ήταν το αίμα, η κίτρινη χολή, η μέλαινα χολή και το φλέγμα [4].

Οι θεωρίες που έχουν διατυπωθεί σχετικά με τα συναισθήματα μπορούν να χωριστούν σε διάφορες κατηγορίες. Ως προς το μηχανισμό πρόκλησης συναισθημάτων στον άνθρωπο, διακρίνουμε τις παρακάτω βασικές κατηγορίες: Οι **σωματικές θεωρίες** δίνουν έμφαση στη φυσιολογία του ατόμου και υποστηρίζουν ότι κάποιες φυσιολογικές αποκρίσεις του σώματος είναι υπεύθυνες για τη δημιουργία συναισθημάτων. Οι **νευρολογικές θεωρίες** θεωρούν ότι η δραστηριότητα του εγκεφάλου οδηγεί σε συναισθηματικές αντιδράσεις. Οι **γνωσιακές θεωρίες** προτείνουν ότι οι σκέψεις και άλλες διανοητικές δραστηριότητες διαδραματίζουν κυρίαρχο ρόλο στην εμφάνιση των συναισθημάτων. Επίσης έχουν διατυπωθεί **υβριδικές θεωρίες** που δίνουν έμφαση σε συνδυασμούς των παραπάνω χαρακτηριστικών. Ως προς το λόγο ύπαρξης των συναισθημάτων, υπάρχουν **εξελικτικές θεωρίες** που εστιάζουν στα χαρακτηριστικά του περιβάλλοντος που οδήγησαν στην εμφάνιση συναισθημάτων με τη μέθοδο της φυσικής επιλογής, **κοινωνικοπολιτιστικές θεωρίες** που περιγράφουν τα συναισθήματα ως κοινωνικές συμβάσεις και κατασκευές.

Στις επόμενες παραγράφους περιγράφονται κάποιες από τις πιο γνωστές θεωρίες σχετικά με τα συναισθήματα που είχαν ιστορικά σημαντική επιρροή στην πορεία της έρευνας για το συναίσθημα.

Υπενθυμίζεται ότι οι βασικότερες κατηγορίες σχετικά με τους μηχανισμούς εμφάνισης των συναισθημάτων συνοψίζονται στις παρακάτω:

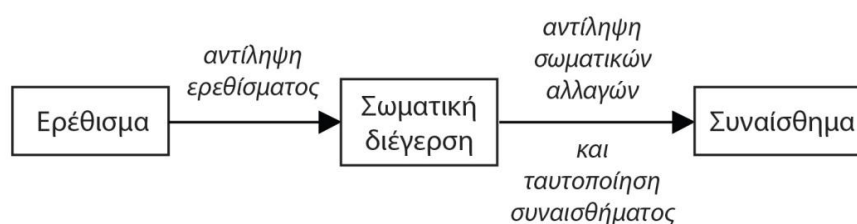
- Σωματικές θεωρίες συναισθημάτων
- Γνωσιακές θεωρίες συναισθημάτων
- Νευροβιολογικές θεωρίες συναισθημάτων

## 2.2 Θεωρία James-Lange

Το 1884 ο Αμερικανός ψυχολόγος William James και το 1885 ο Δανός παθολόγος Carl Lange διατύπωσαν ανεξάρτητα ο ένας από τον άλλον παραπλήσιες θεωρίες για το μηχανισμό δημιουργίας των συναισθημάτων. Για το λόγο αυτό η θεωρία που διατύπωσαν αναφέρεται στη βιβλιογραφία συνήθως με τα δύο τους ονόματα. Η θεωρία James-Lange αποτελεί χαρακτηριστικό παράδειγμα της κατηγορίας των σωματικών θεωριών συναισθημάτων [5].

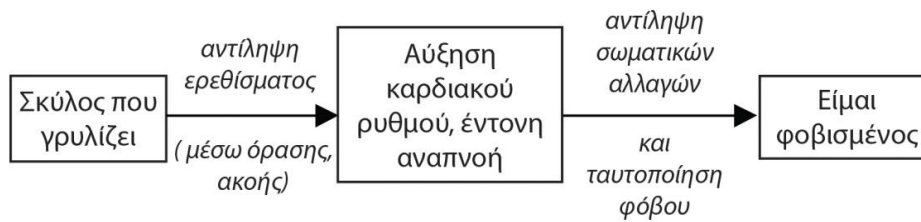
Σύμφωνα με τη θεωρία James-Lange, αρχικά το υποκείμενο άτομο αντιλαμβάνεται μία αλλαγή στο περιβάλλον. Η αλλαγή στο περιβάλλον προκαλεί αλλαγή στη φυσική κατάσταση του ατόμου. Για παράδειγμα αλλαγή στους χτύπους της καρδιάς, στην αρτηριακή πίεση, το ρυθμό αναπνοών, τα επίπεδα αδρεναλίνης, την εφίδρωση. Ο Lange μάλιστα διατύπωσε ότι τα συναισθήματα προκαλούνται συγκεκριμένα από αγγειοκινητικές αλλαγές στο υποκείμενο άτομο. Ο James ισχυρίστηκε ότι το συναίσθημα είναι ουσιαστικά η αντίληψη των σωματικών αλλαγών καθώς αυτές συμβαίνουν. Με άλλα λόγια, το μυαλό καταγράφει αυτές τις σωματικές δραστηριότητες και η νοητική κατάσταση που προκαλείται από τις σωματικές αλλαγές, είναι το συναίσθημα. Ο James έδωσε έμφαση στις αλλαγές των εσωτερικών οργάνων ως υπεύθυνες για την πρόκληση συναισθημάτων. Η θεωρία αντίκειται σε αυτό που πολλοί άνθρωποι πιστεύουν διαισθητικά, ότι τα συναισθήματα που νιώθουν προκαλούν συγκεκριμένες δράσεις. Η θεωρία James-Lange υποστηρίζει ότι πρώτα αντιδρούμε σωματικά σε μια αλλαγή στο περιβάλλον μας και στη συνέχεια ερμηνεύουμε τις ενέργειές μας με την ετικέτα μιας συναισθηματικής αντίδρασης. Με αυτόν τον τρόπο, τα συναισθήματα χρησιμεύουν ώστε το υποκείμενο άτομο να εξηγεί τις πράξεις του στον εαυτό του. Ο James περιέγραψε χαρακτηριστικά ότι «νιώθουμε θλίψη επειδή κλαίμε, οργή επειδή χτυπούμε, φόβο επειδή τρέμουμε, και δεν κλαίμε, χτυπούμε ή τρέμουμε επειδή είμαστε θλιμμένοι, οργισμένοι ή φοβισμένοι, ανάλογα με την περίπτωση» [6].

Σύμφωνα με τη θεωρία των James-Lange, ισχύει η εξής αλληλουχία:



**Σχήμα 2-1** Θεωρία James-Lange

Ένα παράδειγμα που χρησιμοποιείται συχνά στη βιβλιογραφία είναι αυτό για το συναίσθημα του φόβου:



**Σχήμα 2-2** Παράδειγμα θεωρίας James-Lange

Η θεωρία James-Lange είχε μεγάλη επιρροή και παραμένει στο ενδιαφέρον ακόμη και σήμερα, με πιο πρόσφατες έρευνες να καταλήγουν σε συμπεράσματα που δεν αποκλείουν την παραπάνω αλληλουχία. Η βασικότερη συνεισφορά της είναι η έμφαση που δίνεται στη σωματική συνιστώσα των συναισθημάτων. Ακόμα και σε θεωρίες που αμφισβητούν τη σχέση αιτίου και αιτιατού έτσι όπως περιγράφηκε παραπάνω, η επίδραση των σωματικών αποκρίσεων στην εμπειρία και την ένταση των συναισθημάτων είναι ευρέως αποδεκτή.

## 2.3 Θεωρία Cannon-Bard

Μερικές δεκαετίες αργότερα, τη δεκαετία του 1920, ο Αμερικανός παθολόγος Walter Bradford Cannon διατύπωσε τη δική του θεωρία για τα συναισθήματα, η οποία διαφωνούσε σε αρκετά σημεία με την επικρατούσα ως τότε θεωρία James-Lange. Η θεωρία του επεκτάθηκε τη δεκαετία του 1930 από τον Philip Bard, ο οποίος υπήρξε διδακτορικός φοιτητής του Cannon, οπότε η θεωρία έμεινε γνωστή ως Cannon-Bard [7].

Κάποια από τα βασικά επιχειρήματα που ασκούσαν κριτική στη θεωρία James-Lange ήταν τα εξής [7] [8]:

- Οι άνθρωποι μπορούν να βιώσουν φυσιολογική διέγερση χωρίς να αισθανθούν το αντίστοιχο συναίσθημα. Ο Cannon διεξήγαγε πειράματα όπου χορήγησε αδρεναλίνη στους συμμετέχοντες για να προσομοιώσει τη φυσιολογική σωματική απόκριση σε εξωτερικά ερεθίσματα. Αν ίσχυε η θεωρία James-Lange, η αδρεναλίνη θα προκαλούσε τεχνητά συναισθήματα, το οποίο όμως δεν παρατηρήθηκε.
- Οι σωματικές αλλαγές των εσωτερικών οργάνων –εκεί όπου ο James εντόπιζε την πηγή ανατροφοδότησης υπεύθυνη για την πρόκληση συναισθημάτων– συμβαίνουν πολύ αργά για να προκαλέσουν συναισθήματα, τα οποία δημιουργούνται πολύ πιο γρήγορα. Για παράδειγμα ένας ξαφνικός ήχος μπορεί να προκαλέσει το συναίσθημα του φόβου πριν εμφανιστεί η σωματική διέγερση του φόβου. Επίσης τα όργανα αυτά δεν χαρακτηρίζονται από μεγάλη ευαισθησία.
- Οι άνθρωποι μπορεί να βιώνουν πολύ διαφορετικά συναισθήματα με τα ίδια συμπτώματα σωματικής διέγερσης. Η αυξημένη καρδιακή συχνότητα και αυξημένος ρυθμός αναπνοής για παράδειγμα παρατηρούνται σε ένα άτομο όταν νιώθει οργή αλλά και φόβο. Επιπλέον συμπτώματα σωματικής διέγερσης

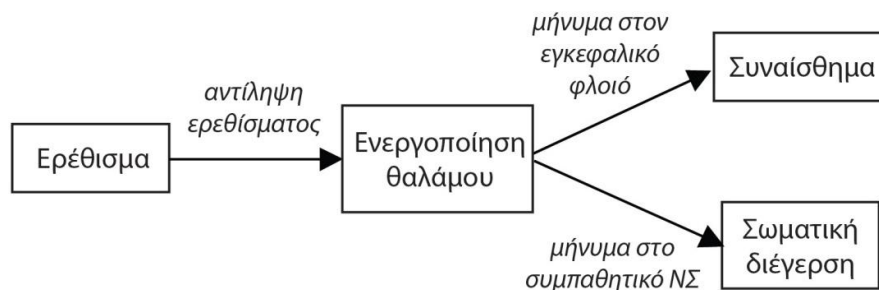


παρατηρούνται και σε περιπτώσεις όπου δεν υπάρχει συναισθηματική συνιστώσα. Για παράδειγμα έπειτα από σωματική άσκηση, σε εμπύρετη κατάσταση.

- Σε πείραμα που διεξήγαγε ο Cannon, ο πλήρης διαχωρισμός των εσωτερικών οργάνων από το κεντρικό νευρικό σύστημα μέσω κοπής των νεύρων δεν είχε επίδραση στη συναισθηματική αντίδραση των πειραματόζωνων.

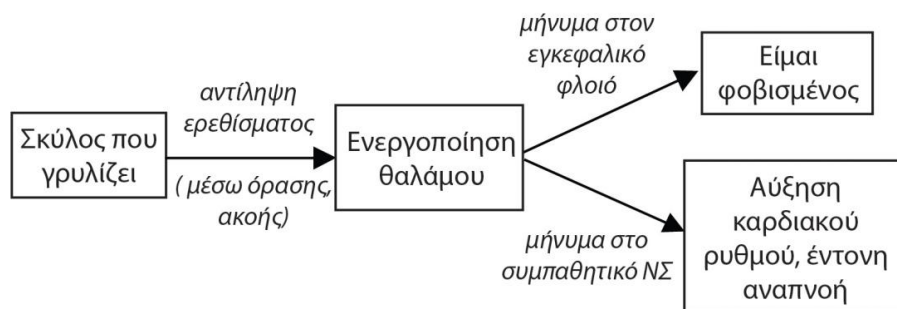
Σύμφωνα με τη θεωρία Cannon-Bard, η **σωματική διέγερση** και το **συναίσθημα** συμβαίνουν **ταυτόχρονα** και **ανεξάρτητα**, χωρίς η μία συνιστώσα να προκαλεί της άλλη. Ο **θάλαμος**, δομή του έσω εγκεφάλου, στέλνει μηνύματα στον εγκεφαλικό φλοιό για την ερμηνεία του συναισθήματος και ταυτόχρονα στέλνει μηνύματα στο συμπαθητικό νευρικό σύστημα για να προκληθούν οι κατάλληλες σωματικές αντιδράσεις. Η ταυτοποίηση και εμπειρία του συναισθήματος συμβαίνουν παράλληλα με την ενεργοποίηση της σωματικής διέγερσης και δε συνδέονται με σχέση αιτίου-αιτιατού όπως στη θεωρία James-Lange [9].

Κατά τους Cannon-Bard, ισχύει η εξής αλληλουχία:



**Σχήμα 2-3** Θεωρία Cannon-Bard

Το παράδειγμα του φόβου που χρησιμοποιήθηκε και στη θεωρία James-Lange διαμορφώνεται τώρα ως εξής:



**Σχήμα 2-4** Παράδειγμα θεωρίας Cannon-Bard

Ως προς τη συναισθηματική συμπεριφορά, η θεωρία Cannon-Bard υποστηρίζει ότι πρώτα το άτομο βιώνει το συναίσθημα, και στη συνέχεια δρα ανάλογα με αυτό.

Η θεωρία Cannon-Bard συμβαδίζει περισσότερο με την κοινή λογική και την πρακτική εμπειρία των περισσότερων ανθρώπων, δέχτηκε όμως αμφισβήτηση κυρίως για την έλλειψη ενσωμάτωσης κάποιας γνωσιακής συνιστώσας στον παραπάνω μηχανισμό.

## 2.4 Θεωρία Schachter-Singer

Το 1962, οι Αμερικανοί ψυχολόγοι Stanley Schachter και Jerome E. Singer διατύπωσαν μια διαφορετική θεωρία για το συναίσθημα η οποία λαμβάνει υπόψιν την επίδραση **γνωσιακών παραγόντων**. Στην ανάλυσή τους, σχολιάζοντας τις μέχρι τότε επικρατούσες θεωρίες κυρίως των James-Lange, σημειώνουν ότι το ευρύ φάσμα συναισθημάτων, διαθέσεων και συγκινησιακών καταστάσεων δε συνοδεύεται από αντίστοιχη ποικιλομορφία σωματικής διέγερσης εσωτερικών οργάνων. Η διαπίστωση αυτή οδηγεί στην υπόθεση ότι οι γνωσιακοί παράγοντες πιθανώς να έχουν σημαντικό ρόλο στον προσδιορισμό της συναισθηματικής κατάστασης [10].

Στη θεωρία που πρότειναν, διατύπωσαν την ιδέα ότι η εμπειρία ενός συναισθήματος είναι **συνάρτηση δύο παραγόντων**: της σωματικής διέγερσης και της γνωσιακής διεργασίας που σχετίζεται με αυτήν τη διέγερση, όπως είναι για παράδειγμα η συλλογιστική, η εκλογίκευση, η ανάκληση μνήμης. Για το λόγο αυτό η θεωρία Schachter-Singer είναι γνωστή και ως **θεωρία δύο παραγόντων** (Two-factor theory).

Σύμφωνα με τη θεωρία Schachter-Singer, όταν οι άνθρωποι αντιλαμβάνονται τα συμπτώματα της φυσιολογικής διέγερσης, **αναζητούν στο περιβάλλον** τους στοιχεία για την επεξήγηση της διέγερσης και στη συνέχεια την ταυτοποίηση του συναισθήματος. Η ετικέτα που δίνουν κατά την αναγνώριση του συναισθήματος εξαρτάται από τα στοιχεία που θα βρεθούν στο περιβάλλον τους. Το υποκείμενο άτομο συνήθως έχει επίγνωση της αιτίας της διεγερμένης συναισθηματικής του κατάστασης. Εάν η αιτία όμως δεν είναι προφανής, τότε αναζητά στοιχεία στο περιβάλλον του για την ερμηνεία του συναισθήματος.

Οι Schachter και Singer σχεδίασαν μια μελέτη όπου χορήγησαν επινεφρίνη στους συμμετέχοντες, για να μελετήσουν τις επιδράσεις της σωματικής διέγερσης και να επιβεβαιώσουν τις εξής **τρεις υποθέσεις**:

- Όταν το υποκείμενο άτομο δεν έχει άμεση εξήγηση για την κατάσταση σωματικής διέγερσης που βιώνει, θα ταυτοποιήσει την κατάστασή του και θα περιγράψει τα συναισθήματά του σε σχέση με τις γνωσιακές πληροφορίες που είναι διαθέσιμες στο περιβάλλον εκείνη τη στιγμή.
- Σε περίπτωση που το υποκείμενο άτομο έχει κατάλληλη εξήγηση για τη σωματική του διέγερση, είναι απίθανο να αναζητήσει εναλλακτική γνωσιακή ετικέτα για τη διεγερμένη κατάστασή του.
- Κάτω από τις ίδιες γνωσιακές περιβαλλοντικές συνθήκες, το υποκείμενο άτομο θα αντιδράσει συναισθηματικά μόνο στο βαθμό που βιώνει σωματική διέγερση.

Στα πλαίσια της παρούσας εργασίας δε θα περιγραφούν οι λεπτομέρειες του πειράματος. Σημειώνεται όμως ότι οι Schachter και Singer συμπέραναν ότι τα αποτελέσματα της μελέτης τους επιβεβαιώνουν και τις τρεις υποθέσεις της θεωρίας τους.

Η **κριτική** απέναντι στη θεωρία Schachter-Singer εστιάζει σε κάποιες από τις αδυναμίες της πειραματικής μεθόδου. Τα συναισθήματα προκλήθηκαν με τεχνητό τρόπο, που δεν αντιπροσωπεύει τη γέννηση συναισθημάτων στον πραγματικό κόσμο. Το δείγμα συμμετεχόντων που χρησιμοποιήθηκε δεν ήταν πλήρως αντιπροσωπευτικό (αποκλειστικά

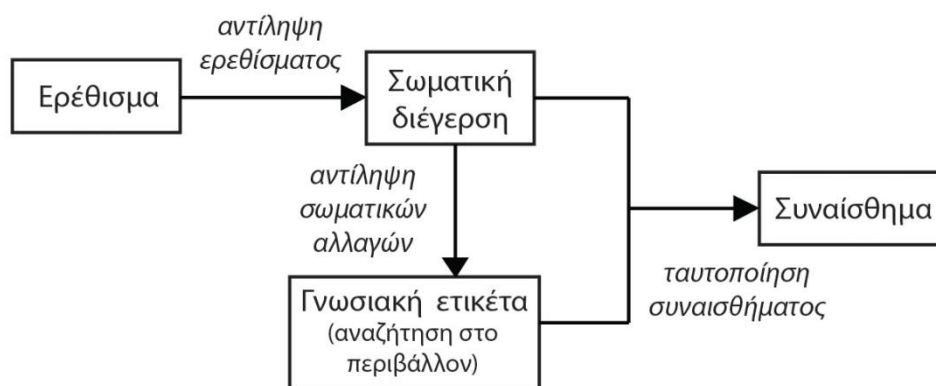
άντρες νεαρής ηλικίας) και δε λήφθηκε υπόψιν η πρότερη συναισθηματική κατάσταση των συμμετεχόντων [11].

Η παραπάνω θεωρία των Schachter και Singer συμφωνεί με τη θεωρία James-Lange στο ότι οι άνθρωποι συνάγουν το συμπέρασμα πως βιώνουν ένα συναίσθημα όταν αντιλαμβάνονται τη σωματική διέγερση. Επίσης οι Schachter και Singer συμφωνούν με τη θεωρία Cannon-Bard στο ότι το ίδιο μοτίβο σωματικής διέγερσης μπορεί να προκαλέσει διαφορετικά συναισθήματα, καθώς η εμπειρία του συναισθήματος εξαρτάται από τη γνωσιακή αποτίμηση του περιβάλλοντος [8].

Η θεωρία Schachter-Singer είχε σημαντική επιρροή καθώς στάθηκε αφορμή για την έρευνα του ρόλου της γνωσιακής αποτίμησης στη δημιουργία του συναισθήματος. Ωστόσο κάποιες άλλες έρευνες δεν επιβεβαίωσαν αλλά ούτε και διέψευσαν τη θεωρία δύο παραγόντων (για παράδειγμα οι μελέτες Lazarus et al., 1965; Maslach, 1979; Marshall & Zimbardo, 1979).

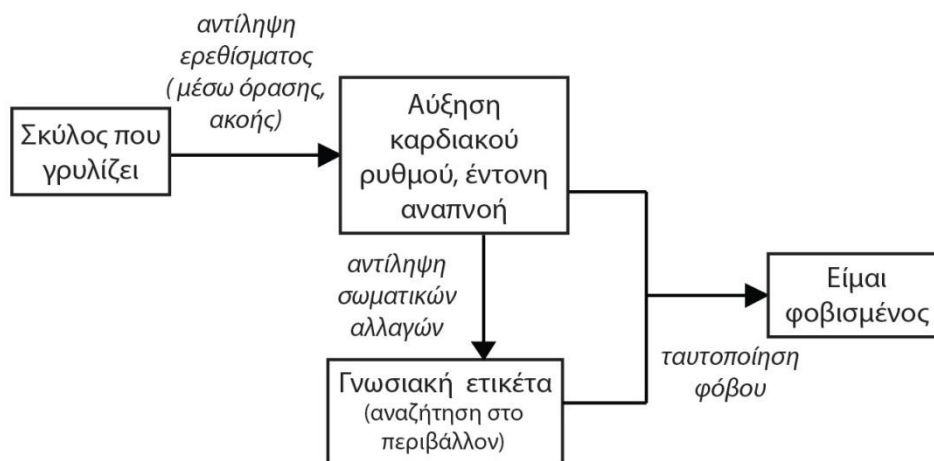
Μεταγενέστερες μελέτες έχουν δείξει ότι η σχέση για την ταυτοποίηση του συναισθήματος είναι πιο περίπλοκη από τους δύο παράγοντες της θεωρίας. Πολλοί ψυχολόγοι πιστεύουν ότι η διαδικασία για την επεξήγηση μιας σωματικής διέγερσης δεν περιορίζεται μόνο στην αναζήτηση στοιχείων στο άμεσο περιβάλλον, αλλά το υποκείμενο άτομο αναζητεί επεξήγηση και με άλλους μηχανισμούς, για παράδειγμα ανακαλεί παρόμοιες εμπειρίες που βίωσε στο παρελθόν [11].

Σύμφωνα με τα παραπάνω, ισχύει η εξής αλληλουχία:



**Σχήμα 2-5** Θεωρία Schachter-Singer

Το παράδειγμα του φόβου που χρησιμοποιήθηκε και παραπάνω διαμορφώνεται τώρα ως εξής:



Σχήμα 2-6 Παράδειγμα θεωρίας Schachter-Singer

## 2.5 Θεωρία γνωσιακής αξιολόγησης

Τα τελευταία 60 χρόνια αναπτύχθηκε ιδιαίτερα η θεωρία της γνωσιακής αξιολόγησης. Σημαντική επιρροή στην ανάπτυξή της είχε η δουλειά των Αμερικανών ψυχολόγων Magda B. Arnold και Richard S. Lazarus. Το ερώτημα που μελετάται από τις θεωρίες γνωσιακής αξιολόγησης είναι γιατί οι άνθρωποι αντιδρούν με διαφορετικούς τρόπους σε ένα γεγονός. Η θεωρία συναισθημάτων που αναπτύχθηκε γύρω από την ιδέα ότι η γνωσιακή αξιολόγηση είναι κεντρική στη δημιουργία συναισθημάτων, αναφέρεται ως θεωρία **γνωσιακής αξιολόγησης ή αποτίμησης** (cognitive appraisal theory) και άλλες φορές αναφέρεται ως **θεωρία γνωσιακής παρεμβολής ή μεσολάβησης** (cognitive mediational theory). Οι θεωρίες γνωσιακής αξιολόγησης κατηγοριοποιούνται σε δύο βασικά μοντέλα, το δομικό μοντέλο και το διαδικαστικό μοντέλο.

Η ιδέα πως η αξιολόγηση μιας κατάστασης οδηγεί στη δημιουργία των συναισθημάτων άρχισε να ερευνάται ήδη από τη δεκαετία του '40. Η Αμερικανή ψυχολόγος **Magda Arnold** ήταν από τους πρώτους θεωρητικούς επιστήμονες που απομάκρυναν το ενδιαφέρον από τις σωματικές και συμπεριφορικές θεωρίες συναισθημάτων, και **εστίασαν στη γνωσιακή προσέγγιση** των συναισθημάτων. Σύμφωνα με την Arnold, η αρχική γνωσιακή αξιολόγηση είναι αφετηρία για το ακολουθούμενο συναίσθημα, καθώς διεγείρει τη συναισθηματική εμπειρία αυτή καθ'εαυτή και προκαλεί τις κατάλληλες ενέργειες δράσης στο άτομο. Οι σωματικές αλλαγές, παρότι αναγνωρίζονται ως σημαντικές, συνοδεύουν τις ενέργειες και την εμπειρία του συναισθήματος αλλά δεν το προκαλούν. Μια αξιοσημείωτη εξέλιξη ήταν η ιδέα της Arnold σχετικά με τη διαισθητική αξιολόγηση (intuitive appraisal), που χρησιμοποιήθηκε για να περιγράψει συναισθήματα που είναι θετικά ή αρνητικά για το άτομο και το οδηγούν σε δράση [12].

Ο Richard Lazarus ξεκίνησε την έρευνα για τα συναισθήματα τη δεκαετία του '50 και συνέχισε να επεκτείνει τη θεωρία της γνωσιακής αξιολόγησης μέχρι και το θάνατό του το 2002. Αναγνώρισε δύο είδους διαδικασίες αξιολόγησης που βρίσκονται στον πυρήνα της γνωσιακής αποτίμησης: την **πρωτογενή αξιολόγηση** (primary appraisal) και τη **δευτερογενή αξιολόγηση** (secondary appraisal). Λεπτομέρειες δίνονται παρακάτω.

### 2.5.1 Δομικό μοντέλο γνωσιακής αξιολόγησης

Το δομικό μοντέλο αποτίμησης επιχειρεί να εξηγήσει τη σχέση ανάμεσα στη γνωσιακή αξιολόγηση και το συναίσθημα που προκαλείται. Το δομικό μοντέλο εξετάζει τη διαδικασία που ακολουθείται για την αξιολόγηση του ερεθίσματος και εξετάζει επίσης τον τρόπο που διαφορετικές αποτιμήσεις επηρεάζουν το συναίσθημα που βιώνεται. Σύμφωνα με τον Lazarus, η θεωρία για τα συναισθήματα περιλαμβάνει **τρεις συνιστώσες**. Η πρώτη είναι η σχέση του ατόμου με το περιβάλλον, καθώς θεωρείται ότι η συναισθηματική εμπειρία περιλαμβάνει πάντα μια αλληλεπίδραση ανάμεσα στον άνθρωπο και το περιβάλλον του. Η δεύτερη συνιστώσα αναφέρεται στο κίνητρο που παρέχουν τα συναισθήματα για συνεπακόλουθη δράση. Τέλος η τρίτη συνιστώσα, η γνωσιακή συνιστώσα, περιλαμβάνει την αξιολόγηση της κατάστασης. Γίνεται μια εκτίμηση της τρέχουσας κατάστασης των στόχων του, και το άτομο αξιολογεί μεταξύ άλλων πόσο σχετική είναι μια κατάσταση με τους προσωπικούς του στόχους, και σε ποιο βαθμό είναι σημαντική για τη ζωή του.

Ο Lazarus υποστηρίζει ότι η ποικιλία στα συναισθήματα οφείλεται στο διαφορετικό τρόπο αξιολόγησης των καταστάσεων, ανάλογα με τις προαναφερθείσες κατηγορίες. Το δομικό μοντέλο είναι απαραίτητο για να αξιολογηθεί κάθε συναίσθημα ξεχωριστά.

#### Πρωτογενής αξιολόγηση:

Στην πρωτογενή αξιολόγηση το άτομο αποτιμά δύο παραμέτρους μιας κατάστασης. Τη σχετικότητα των κινήτρων (motivational relevance) και τη συμβατότητα των κινήτρων (motivational congruence).

Κατά την αποτίμηση της **σχετικότητας των κινήτρων**, το άτομο απαντά στο ερώτημα «πόσο σχετική είναι η κατάσταση με τις ανάγκες μου;». Γίνεται αποτίμηση της σημασίας της κατάστασης σε σχέση με την ευημερία του ατόμου. Αυτή η διαδικασία έχει βρεθεί ότι επηρεάζει την ένταση του συναισθήματος που βιώνεται. Όταν μια κατάσταση αποτιμάται ως στενά συνδεδεμένη με την ευημερία του ατόμου, προκαλείται πιο έντονη συναισθηματική απόκριση.

Κατά την αποτίμηση της **συμβατότητας των κινήτρων**, το άτομο απαντά στο ερώτημα «αυτή η κατάσταση είναι σε συμφωνία ή όχι με τους στόχους μου;». Το υποκείμενο άτομο βιώνει διαφορετικά συναισθήματα ανάλογα με το αν αξιολογεί την κατάσταση στο περιβάλλον του ως συνεπή ή ασυνεπή με τους στόχους του.

#### Δευτερογενής αξιολόγηση:

Η δευτερογενής αξιολόγηση αποτιμά τις δυνατότητες του ατόμου για την αντιμετώπιση μιας κατάστασης και τις εναλλακτικές επιλογές αντιμετώπισης.

Μία πτυχή της δευτερογενούς αποτίμησης είναι η **εκτίμηση για την ευθύνη** μιας κατάστασης. Ένα άτομο θεωρεί είτε τον εαυτό του, είτε κάποιον άλλον ή ακόμη μια ολόκληρη ομάδα ανθρώπων υπεύθυνη για μια δεδομένη κατάσταση. Το άτομο μπορεί για παράδειγμα να αποδώσει ενοχή για μια ζημιογόνο κατάσταση ή να αποδώσει έπαινο για μια ωφέλιμη κατάσταση. Επιπρόσθετα, η ευθύνη για μια κατάσταση μπορεί να αποδοθεί απλά στην τύχη. Το αποτέλεσμα της απόδοσης της ευθύνης καθοδηγεί και κατευθύνει τις προσπάθειες να χειριστεί κανείς το συναίσθημα που βιώνει.

Η επόμενη πτυχή της δευτερογενούς αποτίμησης είναι το **δυναμικό διαχείρισης** (coping potential). Με τον όρο δυναμικό διαχείρισης εννοείται η ικανότητα του ατόμου να διαχειρίζεται μια συναισθηματική κατάσταση. Οι μηχανισμοί διαχείρισης μπορεί να

εστιάζουν στη διαχείριση του προβλήματος ή τη διαχείριση του συναισθήματος. Η **διαχείριση του προβλήματος** αναφέρεται στην ικανότητα του ατόμου να αναλαμβάνει δράση και να αλλάζει την κατάσταση ώστε να είναι πιο συμβατή με τους στόχους του. Η **διαχείριση του συναισθήματος** αναφέρεται στην ικανότητα να προσαρμόζεται το άτομο στην κατάσταση, εφόσον οι συνθήκες παραμείνουν ασύμβατες με τους στόχους του. Και στις δύο περιπτώσεις, το συναίσθημα που βιώνει το άτομο εξαρτάται από την εκτίμηση που έχει για τις ικανότητές του στη διαχείριση της κατάστασης.

Μια ακόμη πτυχή της δευτερογενούς αξιολόγησης είναι το η **προσδοκία για το μέλλον** (future expectancy). Ο όρος αναφέρεται στις προσδοκίες του ατόμου για πιθανές αλλαγές στην κατεύθυνση της συμβατότητας των κινήτρων. Το άτομο μπορεί να έχει την πεποίθηση, για οποιοδήποτε λόγο, ότι η κατάσταση θα αλλάξει προς όφελός του ή προς το χειρότερο. Η προσδοκία για το μέλλον επηρεάζει το προκληθέν συναίσθημα και τους μηχανισμούς διαχείρισης που χρησιμοποιούνται από το άτομο.

Το δομικό μοντέλο αξιολόγησης υπαινίσσεται ότι οι απαντήσεις στα επιμέρους ερωτήματα της πρωτογενούς και δευτερογενούς αξιολόγησης επιτρέπουν την πρόβλεψη του συναισθήματος που προκαλείται από ένα συγκεκριμένο σύνολο περιστάσεων. Ένα παράδειγμα είναι το συναίσθημα του θυμού. Αν ένα άτομο αξιολογεί μια κατάσταση ως σχετική με τα κίνητρά του, ασύμβατη με τους στόχους του, και εκτιμά κάποιον άλλον ως υπεύθυνο για την κατάσταση, τότε είναι πιο πιθανό να νιώσει θυμό ως απόκριση στο ερέθισμα. Ένα άλλο παράδειγμα είναι το άγχος. Όπως και στο θυμό, το άτομο αντιλαμβάνεται την κατάσταση ως σχετική με τα κίνητρά του και ασύμβατη με τους στόχους του. Η διαφορά έγκειται στην απόδοση της ευθύνης. Ενώ στο θυμό η ευθύνη αποδίδεται σε άλλο άτομο ή ομάδα ατόμων, στην περίπτωση του άγχους δεν είναι προφανές ποιο άτομο ή άτομα φέρουν την ευθύνη. Έτσι το συναίσθημα που βιώνει το άτομο διαφοροποιείται.

### 2.5.2 Διαδικαστικό μοντέλο γνωσιακής αξιολόγησης

Η δομική θεωρία γνωσιακής αποτίμησης δέχτηκε κριτική για την αδυναμία της να συλλάβει τη δυναμική φύση των συναισθημάτων. Μια προσέγγιση που προτάθηκε ήταν αυτή της κυκλικής διαδικασίας, που εναλλάσσεται μεταξύ αξιολόγησης και διαχείρισης και επαναξιολόγησης. Το μοντέλο αυτό απέτυχε να σταθεί στην επιστημονική κοινότητα καθώς δε μπορούσε να εξηγήσει τις ταχείες και αυτόματες συναισθηματικές αντιδράσεις. Παραδείγματα δύο θεωριών που προτάθηκαν ως απάντηση στο δομικό και κυκλικό μοντέλο, περιγράφονται παρακάτω [12].

#### Μοντέλο αξιολόγησης δύο διαδικασιών

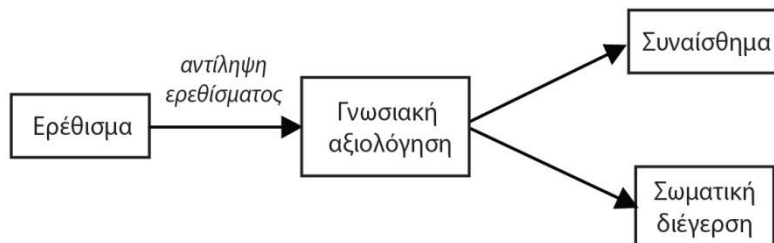
Οι Smith και Kirby διατύπωσαν ένα μοντέλο αξιολόγησης που επεκτείνει τη λειτουργία του δομικού μοντέλου, σε δύο διαδικασίες. Το δομικό μοντέλο εστιάζει στο ποια είναι τα ερωτήματα που αξιολογούνται ενώ το διαδικαστικό μοντέλο εστιάζει στον τρόπο που το άτομο αξιολογεί τα ερεθίσματα. Υπάρχουν τρεις βασικές συνιστώσες του διαδικαστικού μοντέλου αποτίμησης: αντιληπτά ερεθίσματα, συνειρμική επεξεργασία και συλλογιστική. Τα **αντιληπτά ερεθίσματα** είναι εκείνα που το άτομο αντιλαμβάνεται με τις αισθήσεις του από το περιβάλλον. Για παράδειγμα η αίσθηση του πόνου ή της ευχαρίστησης, η αντίληψη της έκφρασης του προσώπου. Η **συνειρμική επεξεργασία** είναι μια διαδικασία με βάση τη μνήμη που κάνει γρήγορες συσχετίσεις και παρέχει πληροφορίες αποτίμησης

βασισμένες σε ενεργές αναμνήσεις που συσχετίστηκαν με το δοσμένο ερέθισμα. Η **συλλογιστική** είναι μια σκόπιμη, αργή διαδικασία που περιλαμβάνει τη λογική, τη κριτική σκέψη σχετικά με το ερέθισμα και την κατάσταση (Marsella & Gratch 2009). Σε αυτό το μοντέλο, η συνειρμική επεξεργασία και η συλλογιστική εκτελούνται παράλληλα, ως αντίδραση στα αντιληπτά ερεθίσματα, και παρέχουν μια σύνθετη, γνωσιακή αξιολόγηση της κατάστασης.

### Πολυεπίπεδο σειριακό μοντέλο ελέγχου του Scherer

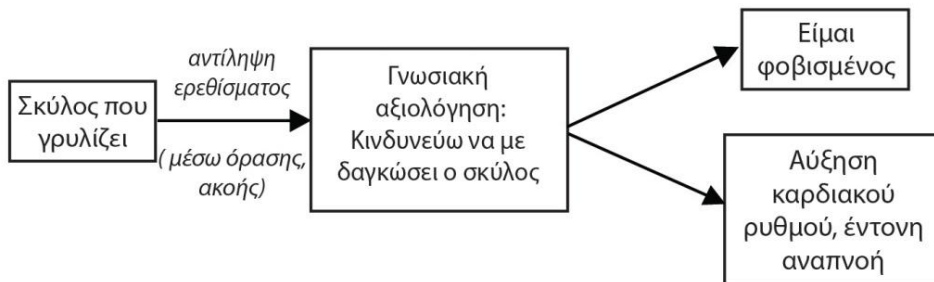
Το μοντέλο αυτό αποτελείται από τρία επίπεδα διαδικασιών αξιολόγησης, με σειριακούς περιορισμούς σε κάθε επίπεδο, ώστε να δημιουργείται μια συγκεκριμένη διάταξη επεξεργασίας. Τα τρία επίπεδα της επεξεργασίας είναι: **εγγενές επίπεδο** (αισθαντικό-κινησιακό), το **επίκτητο επίπεδο** (αποτέλεσμα μάθησης) και τέλος το **συνειδητό, εννοιολογικό επίπεδο**. Ο Scherer διατύπωσε μια αυστηρή διάταξη σύμφωνα με την οποία διενεργείται η διαδικασία αξιολόγησης. Υπάρχουν διάφοροι **έλεγχοι αξιολόγησης** κατά τη διάρκεια των διαδικασιών, ώστε να επιτρέπεται η παρατήρηση των ερεθισμάτων σε διάφορα σημεία της συνολικής διαδικασίας. Οι έλεγχοι αυτοί περιλαμβάνουν: έλεγχο σχετικότητας των στόχων, έλεγχο υπαινιγμών (αιτία, δεκτικότητα, το επείγον των στόχων), έλεγχο δυναμικού διαχείρισης, έλεγχος συμβατότητας με τις αξίες του ατόμου. Σε αντίθεση με το μοντέλο δύο διαδικασιών όπου έχουμε παράλληλες διαδικασίες αξιολόγησης, στο πολυεπίπεδο μοντέλο οι διαδικασίες πραγματοποιούνται σειριακά, σε μία συγκεκριμένη ακολουθία.

Παρατίθεται και πάλι ένα μπλοκ διάγραμμα που απεικονίζει τη θεωρία γνωσιακής αποτίμησης στη γενική περίπτωση:



**Σχήμα 2-7** Θεωρία γνωσιακής αξιολόγησης

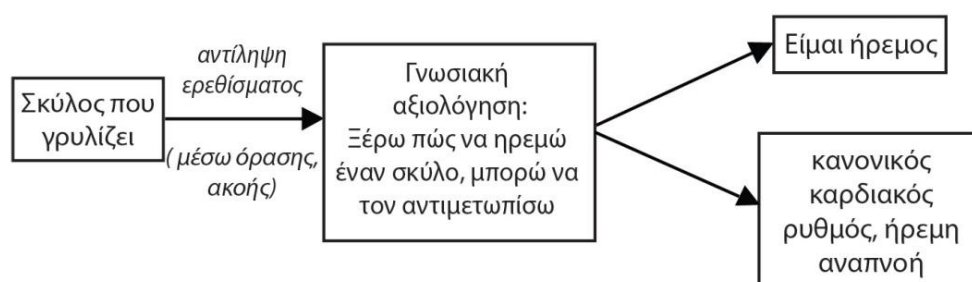
Το παράδειγμα του φόβου μπορεί να αποδοθεί με δύο τρόπους, ανάλογα με το αποτέλεσμα της γνωσιακής αξιολόγησης που πραγματοποιείται από το άτομο ως αντίδραση στο ερέθισμα (ο σκύλος που γρυλίζει) στο περιβάλλον του:



**Σχήμα 2-8** Παράδειγμα θεωρίας γνωσιακής αξιολόγησης (α)



Εάν η εκτίμηση του ατόμου είναι διαφορετική για το δυναμικό διαχείρισης, το αποτέλεσμα της αξιολόγησης θα προκαλούσε διαφορετική συναισθηματική εμπειρία:



Σχήμα 2-9 Παράδειγμα θεωρίας γνωσιακής αξιολόγησης (β)

## 2.6 Υπόθεση ανατροφοδότησης της έκφρασης του προσώπου

Σύμφωνα με την **υπόθεση της ανατροφοδότησης της έκφρασης του προσώπου** (facial feedback hypothesis) ή για **συντομία FFH**, οι μυοσκελετικές πληροφορίες του προσώπου ανατροφοδοτούνται και μπορούν να επηρεάσουν τη συναισθηματική κατάσταση του ατόμου. Δεν έχει διευκρινιστεί με βεβαιότητα ποια είναι η ακριβής σχέση ανατροφοδότησης ανάμεσα στην έκφραση του προσώπου και την εμπειρία του συναισθήματος, παρότι η αναγνώριση της ύπαρξης μιας τέτοιας επίδρασης είναι διαδεδομένη στη βιβλιογραφία.

Ο διάσημος Άγγλος θεμελιωτής της θεωρίας της εξέλιξης Charles Darwin ήταν ένας από τους πρώτους θεωρητικούς που υποστήριξαν ότι οι σωματικές αλλαγές έχουν άμεσο αντίκτυπο στη συναισθηματική εμπειρία, αντί να είναι απλά συνέπεια του συναισθήματος.

Παρομοίως, οι James-Lange μέσω της σωματικής θεωρίας τους διατυπώνουν την άποψη ότι οι σωματικές αλλαγές, που περιλαμβάνουν μεταξύ άλλων τις εκφράσεις του προσώπου, αποτελούν ουσιαστικά το συναίσθημα και συνεπώς επηρεάζουν τη συναισθηματική εμπειρία.

Διαφορετικές **παραλλαγές της FFH** αποδίδουν διαφορετική σημασία στην ανατροφοδότηση των εκφράσεων του προσώπου. Σύμφωνα με την **υπόθεση αναγκαιότητας**, δε μπορεί να υπάρξει συναισθηματική εμπειρία χωρίς ανατροφοδότηση έκφρασης. Η υπόθεση αυτή έχει αποκλειστεί επιτυχώς σε μια μελέτη από τον Keillor et al. το 2002 [13]. Σύμφωνα με την **υπόθεση επάρκειας**, για παράδειγμα από τον Ekman et al. το 1983, η κίνηση των μυών του προσώπου επαρκεί για την πρόκληση συναισθηματικής εμπειρίας. Η υπόθεση αυτή έχει βρει πειραματική υποστήριξη (Levenson & Ekman, 2002). Στη βιβλιογραφία αναφέρεται επίσης ως «ισχυρή εκδοχή» της FFH. Τέλος, η



**υπόθεση διαφοροποίησης** διατυπώνει την άποψη ότι οι εκφράσεις του προσώπου μπορούν να επηρεάσουν τη συναισθηματική εμπειρία που έχει προκληθεί από κάποιο άλλο, εξωτερικό ερέθισμα. Για παράδειγμα, η συνειδητή ή ασυνειδητή ενεργοποίηση μυών που συμμετέχουν σε συναισθηματικές εκφράσεις μπορεί να εντείνει τη συναισθηματική εμπειρία (Strack et al., 1988). Επιπρόσθετα η παρεμπόδιση ή αναστολή εκφράσεων του προσώπου μπορεί να μειώσει την ένταση της συναισθηματικής εμπειρίας. Και για αυτήν την εκδοχή υπάρχουν πειραματικά αποτελέσματα που τη στηρίζουν εν μέρει και δεν την αποκλείουν (Davis et al., 2009) [13]. Αυτή η υπόθεση αναφέρεται και ως «ασθενής εκδοχή» της FFH, την οποία υποστήριξε ο Darwin.

### 2.6.1 Πείραμα του Strack

Η **μελέτη των Strack, Martin & Stepper το 1988** είναι πλέον κλασσική στην έρευνα της FFH [14]. Το πείραμα στήθηκε με τέτοιο τρόπο ώστε οι συμμετέχοντες να υιοθετούν την επιθυμητή μυική ενεργοποίηση χωρίς να μουν στη διαδικασία να αναζητήσουν το συναίσθημα στο οποίο ανταποκρίνεται ή το πραγματικό κίνητρο των ερευνητών.

Η **πειραματική διαδικασία** σε αδρές γραμμές είχε ως εξής: Οι συμμετέχοντες ενημερώθηκαν ότι λαμβάνουν μέρος σε μια έρευνα σχετικά με ανθρώπους που δεν μπορούν να χρησιμοποιήσουν τα χέρια τους και τη δυσκολία που αντιμετωπίζουν σε καθημερινές εργασίες, όπως να συμπληρώσουν ένα ερωτηματολόγιο. Τους ζητήθηκε να κρατήσουν ένα στυλό με τρεις διαφορετικούς τρόπους. Ο πρώτος τρόπος, όπου οι συμμετέχοντες κρατούσαν το στυλό με τα χείλη, ενεργοποιεί τον σφιγκτήρα μυ του στόματος (orbicular oris), με αποτέλεσμα το άτομο να συνοφρυώνεται. Ο δεύτερος τρόπος, όπου οι συμμετέχοντες κράτησαν το στυλό με τα δόντια, ενεργοποιούσε τον μείζονα ζυγωματικό μυ (zygomaticus major) ή τον γελαστήριο μυ (risorius muscle), με αποτέλεσμα να προκαλεί έκφραση παρόμοια με το χαμόγελο. Στην ομάδα ελέγχου κράτησαν το στυλό με το μη-κυρίαρχο χέρι τους. Οι συμμετέχοντες συμπλήρωσαν ερωτηματολόγια σχετικά με τη δυσκολία των εργασιών που πραγματοποίησαν. Το πραγματικό **ζητούμενο** ήταν η τελευταία δοκιμή, δηλαδή η υποκειμενική βαθμολόγηση της κωμικότητας μιας ταινίας κινουμένων σχεδίων και του βαθμού της διασκέδασης που ένιωσαν.

Όπως είχε προβλεφθεί, οι συμμετέχοντες που κράτησαν το στυλό με τα δόντια και συνεπώς ενεργοποιούσαν μυς που ενεργοποιούνται και στο χαμόγελο, ανέφεραν σημαντικά **υψηλότερες βαθμολογίες διασκέδασης** από αυτούς που κράτησαν το στυλό με τα χείλη. Το πείραμα διέφερε από προηγούμενες μεθόδους στο ότι οι συμμετέχοντες δεν κλήθηκαν να προσομοιώσουν, να ενισχύσουν ή να καταπιέσουν κάποιο συναίσθημα. Το σενάριο κάλυψης (cover story) που χρησιμοποιήθηκε και η πειραματική διαδικασία θεωρήθηκαν πολύ επιτυχή στο να προκαλέσουν την επιθυμητή μυική ενεργοποίηση χωρίς να κινήσουν υποψίες για το σκοπό του πειράματος και χωρίς να επάγουν τη γνωσιακή ερμηνεία της έκφρασης του προσώπου.

Ως **κριτική** για την παραπάνω μέθοδο, έχει ειπωθεί ότι απαιτείται περισσότερη προσπάθεια για να κρατηθεί το στυλό με τα χείλη και συνεπώς αυτό θα έπρεπε να ληφθεί υπόψιν στην ερμηνεία των αποτελεσμάτων. Σε κάθε περίπτωση, το πείραμα έλυνε κάποια από τα προβλήματα της μεθοδολογίας για την έρευνα πάνω στη FFH.

## 2.6.2 Πειράματα με βοτουλινική τοξίνη

Τα πιο πρόσφατα πειράματα για την FFH στηρίχτηκαν στην επίδραση της **βοτουλινικής τοξίνης** (botulinum toxin), αλλιώς γνωστή και ως αλλαντική τοξίνη, και ευρέως γνωστή με το εμπορικό όνομα botox. Η τοξίνη χρησιμοποιείται για την προσωρινή παράλυση μυών του προσώπου, ώστε να μελετηθεί η επίδραση που έχει η επιλεκτική αναστολή της ανατροφοδότησης της μυικής δραστηριότητας στη συναισθηματική εμπειρία. Με χρήση της βοτουλινικής τοξίνης, οι κινητικές εντολές του απαγωγού νευρικού συστήματος παραμένουν ανεπηρέαστες, ενώ οι πληροφορίες των μυικών ινών που μεταφέρονται με το προσαγωγό νευρικό σύστημα μειώνονται.

Σε μια μελέτη που δημοσιεύτηκε το 2008 [15] καταγράφηκε η **ενεργοποίηση διαφόρων εγκεφαλικών δομών** με τη χρήση μαγνητικού τομογράφου fMRI, πριν και μετά τη χορήγηση βοτουλινικής τοξίνης στον σφιγκτήρα μυ των φρυδιών, που ενεργοποιείται κατά τη συνοφρύωση. Ζητήθηκε από τους συμμετέχοντες να προσομοιώσουν εκφράσεις θυμού ή λύπης στον τομογράφο πριν την ένεση, και δύο εβδομάδες μετά από την ένεση, όταν καταγράφεται η μέγιστη επίδραση της τοξίνης. Από τις μετρήσεις βρέθηκε ότι η χορήγηση βοτουλινικής τοξίνης **μείωσε την ενεργοποίηση της αμυγδαλής του εγκεφάλου** και του εγκεφαλικού στελέχους, δομές που έχουν συνδεθεί με τη συναισθηματική εμπειρία και επεξεργασία. Τα ευρήματα αυτά υποδεικνύουν ότι η ανατροφοδότηση της μυικής δραστηριότητας στο πρόσωπο επηρεάζει τη νευρωνική δραστηριότητα στον εγκέφαλο. Επιπλέον η χρήση βοτουλινικής τοξίνης αλλάζει την απόκριση του εγκεφάλου σε συναισθηματικές καταστάσεις.

Σε μια ακόμη έρευνα [16] όπου έγινε ένεση ως μέρος αισθητικής θεραπείας στον ίδιο μυ, δηλαδή το σφιγκτήρα των φρυδιών, οι συμμετέχοντες διάβασαν συναισθηματικά φορτισμένες προτάσεις (χαράς, λύπης, θυμού), πριν και δύο εβδομάδες μετά τη χορήγηση βοτουλινικής τοξίνης. Οι **χρόνοι ανάγνωσης** για τις προτάσεις του θυμού και της λύπης – συναισθήματα στα οποία εν γένει ενεργοποιείται ο σφιγκτήρας μυς των φρυδιών – ήταν μεγαλύτεροι μετά την ένεση της τοξίνης, ενώ οι χρόνοι ανάγνωσης για τις προτάσεις χαράς δεν επηρεάστηκαν. Τα ευρήματα αυτά υποδεικνύουν ότι η (προσωρινή) επιλεκτική **παράλυση συγκεκριμένων μυών** του προσώπου έχει επίδραση στην επεξεργασία συγκεκριμένου συναισθηματικού περιεχομένου. Επίσης είναι ένδειξη ότι η χρήση βοτουλινικής τοξίνης για αισθητικούς λόγους **επηρεάζει** κάποιες πτυχές της ανθρώπινης γνωσιακής λειτουργίας, εν προκειμένω **τη γλωσσική αντίληψη**.

Οι παραπάνω μελέτες δείχνουν ότι η χορήγηση βοτουλινικής τοξίνης είναι πιθανό να **μειώνει την ικανότητα** ενός ατόμου να **κατανοήσει τα συναισθήματα** κάποιου άλλου. Τα ευρήματα υποστηρίζουν την αρχική τοποθέτηση του Darwin, σύμφωνα με την οποία η καταπίεση της εξωτερικής έκφρασης των συναισθημάτων, στο βαθμό που αυτό είναι δυνατόν, αμβλύνει τη συναισθηματική εμπειρία μας [17].

Η πειραματική επαλήθευση της επίδρασης της θεωρίας FFH ανοίγει το δρόμο για νέες έρευνες πάνω στην αλληλεπίδραση των εκφράσεων του προσώπου με τη συναισθηματική εμπειρία.

## 2.7 Θεωρία συναισθημάτων του Ekman

Ο Αμερικανός ψυχολόγος Paul Ekman μελέτησε εκτενώς τη σχέση των συναισθημάτων με τις εκφράσεις του προσώπου και τις κινήσεις του σώματος. Διερεύνησε την ύπαρξη **καθολικών συναισθημάτων** που είναι κοινά σε όλους του ανθρώπους ανεξάρτητα από το κοινωνικοπολιτιστικό τους υπόβαθρο. Στα ενδιαφέροντά του περιλαμβάνονται επίσης η ανίχνευση ψεύδους και εξαπάτησης από τις κινήσεις του σώματος και τις λεγόμενες μικροεκφράσεις (microexpressions) του προσώπου.

### 2.7.1 Καθολικότητα συναισθημάτων και αρχέτυπα συναισθήματα

Κάποια από τα ερωτήματα που απασχόλησαν τον Ekman κατά τη διάρκεια της έρευνάς του, ήταν τα ακόλουθα: Μπορεί να υπάρξει συναίσθημα χωρίς έκφραση στο πρόσωπο; Και αντίστροφα, μπορεί να υπάρξει συναισθηματική έκφραση προσώπου χωρίς όμως τη συναισθηματική εμπειρία; Με ποιο τρόπο διαφέρουν οι προσωπικές συναισθηματικές εκφράσεις των ανθρώπων, και τί πληροφορία παρέχει η έκφραση του προσώπου; Τα συναισθήματα είναι καθολικά ή κοινωνικοπολιτισμικής προέλευσης [18];

Ο Ekman υποστήριξε ότι κάποιες πτυχές των συναισθηματικών εκφράσεων είναι όντως καθολικές και παγκόσμιες, ενώ άλλες είναι συγκεκριμένες για τον πολιτισμό στον οποίον απαντώνται. Αναγνώρισε μεν τις διαφορές στην έκφραση ανά πολιτισμό, αλλά διατύπωσε την άποψη ότι υπάρχουν κάποιες βασικές ομοιότητες οι οποίες μπορούν να ταξινομηθούν. Ως προς τα υπόλοιπα ερωτήματα, ο Ekman θεώρησε ότι μπορούν να υπάρξουν συναισθήματα χωρίς αντίστοιχη έκφραση στο πρόσωπο, και επίσης η έκφραση του προσώπου μπορεί να αναπαραχθεί χωρίς να βιώνεται το συναίσθημα. Βέβαια για τα βασικά συναισθήματα της χαράς, του φόβου, θυμού και λύπης, υπάρχουν κάποιες μυικές ενέργειες που σχετίζονται με την έκφραση του προσώπου και οι οποίες δεν μπορούν να εκτελεστούν εκούσια. Οι αντίστοιχοι μύες ενεργοποιούνται μόνο ακούσια, όταν το συναίσθημα είναι γνήσιο. Η παρατήρηση αυτή συνδέεται και με τη δουλειά του Ekman στον τομέα της ανίχνευσης ψεύδους.

Για τη μελέτη της **καθολικότητας των συναισθημάτων** ο Ekman επέλεξε μια απομονωμένη πολιτισμικά φυλή, τη φυλή Fore στην Παπούα Νέα Γουινέα, για να πραγματοποιήσει την έρευνά του. Η πειραματική διαδικασία περιλάμβανε μεταξύ άλλων την παρουσίαση μιας ιστορίας με συναισθηματικό περιεχόμενο κάποιο από τα έξι βασικά συναισθήματα που παρουσιάζονται παρακάτω, και την αναγνώριση του συναισθήματος από τους συμμετέχοντες σε φωτογραφίες ανθρώπων από άλλους πολιτισμούς, με τους οποίους οι Fore δεν είχαν έρθει ποτέ σε επαφή. Τα υψηλά ποσοστά επιτυχούς αναγνώρισης οδήγησαν τον Ekman να διατυπώσει τη θέση του σχετικά με την καθολική συνιστώσα των συναισθημάτων [19].

Το 1972 ο Ekman πρότεινε τα εξής **έξι αρχέτυπα συναισθήματα**:

- **Θυμός** (Anger)
- **Φόβος** (Fear)
- **Λύπη** (Sadness)
- **Αηδία - Απέχθεια** (Disgust)

- ο **Έκπληξη** (Surprise)
- ο **Χαρά** (Joy)

Το 1990 ο Ekman επέκτεινε το σύνολο των βασικών συναισθημάτων και περιέλαβε κάποια ακόμη συναισθήματα, χωρίς όλα όμως να κωδικοποιούνται απαραίτητα με αντίστοιχες μυικές κινήσεις. Τα πρόσθετα συναισθήματα περιλάμβαναν: περιφρόνηση, ικανοποίηση, ντροπή, αμηχανία, ενοχή, ενθουσιασμός, υπερηφάνεια της επίτευξης, ανακούφιση, ικανοποίηση, αισθητηριακή απόλαυση, διασκέδαση [20].

## 2.7.2 Σύστημα κωδικοποίησης κινήσεων προσώπου FACS

Το 1978 δημοσιεύτηκε για πρώτη φορά το **σύστημα κωδικοποίησης κινήσεων προσώπου** FACS (Facial Action Coding System) που είχε αναπτύξει ο Ekman μαζί με τους Wallace V. Friesen και Richard J. Davidson. Το **σύστημα FACS** καταγράφει όλες τις διακριτές ορατές κινήσεις του προσώπου. Πρόκειται για μια εξαντλητική λίστα όλων των **ανιχνεύσιμων μεταβολών** του προσώπου που είναι δυνατόν να ανιχνευθούν από ανθρώπινο παρατηρητή. Ως εκ τούτου δεν περιλαμβάνει μη-ορατές και μη-ανιχνεύσιμες αλλαγές όπως αλλαγές στους αδένες και στο μυϊκό τόνο [21].

Το **εγχειρίδιο FACS** [22] που βασίζεται στο ομώνυμο σύστημα ταξινόμησης, είναι ένας τεχνικός οδηγός ο οποίος χρησιμοποιείται από επαγγελματίες διαφόρων επιστημονικών πεδίων όπως η συμπεριφορική ψυχολογία, η ψυχοθεραπεία, η γραφική υπολογιστών, η επιστήμη υπολογιστών κλπ. Ο χρήστης εκπαιδεύεται ώστε να ανιχνεύει τις κινήσεις του προσώπου, να τις ταξινομεί και να μπορεί να τις αναπαράγει.

Το σύστημα FACS βασίστηκε στην **ανατομική ανάλυση** των κινήσεων του προσώπου. Η πλήρης περιγραφή του συστήματος επιτεύχθηκε από τη μελέτη της επίδρασης που έχει κάθε μυϊκή σύσπαση ή χαλάρωση στην εμφάνιση του προσώπου. Με δεδομένη αυτήν την ανατομική γνώση, είναι δυνατό να αναλυθεί κάθε κίνηση του προσώπου ως άθροισμα στοιχειωδών ανατομικών κινήσεων.

Η ανάπτυξη του FACS βασίστηκε επίσης στο **διαχωρισμό της περιγραφικής του λειτουργίας από την εννοιολογική επεξήγηση** των κινήσεων. Ο χρήστης της μεθόδου δε χρειάζεται να αξιολογεί την έννοια της κάθε κίνησης, γεγονός που καθιστά τη μέθοδο περισσότερο αντικειμενική και λιγότερο επισφαλής ή ευαίσθητη σε προσωπικές εκτιμήσεις. Ο χρήστης μαθαίνει τη μηχανική του προσώπου και εστιάζει στις μυϊκές κινήσεις και το αποτέλεσμά τους. Αυτό βοηθά επίσης ώστε να μη συγχέονται οι επιδράσεις των μυϊκών κινήσεων με τις φυσιολογικές ιδιαιτερότητες του κάθε προσώπου. Για την εννοιολογική ερμηνεία των κινήσεων που σχετίζονται με το πρόσωπο, έχει αναπτυχθεί το **σύστημα FACS-AID** (Facial Action Coding System Affect Interpretation Dictionary).

Κάθε κίνηση του προσώπου στο FACS αναλύεται σε άθροισμα κινήσεων που ονομάζονται **μονάδες κίνησης AU** (Action Units). Πρόκειται για στοιχειώδεις, ανεξάρτητα ανιχνεύσιμες κινήσεις του προσώπου, που οφείλονται κατά κανόνα στην ενεργοποίηση ενός μυός του προσώπου. Μικρότερο σύνολο AUs οφείλεται στην κίνηση

περισσότερων μυών ταυτόχρονα, ή το αντίστροφο, ένας μυς μπορεί να είναι υπεύθυνος για περισσότερα του ενός AUs. Επομένως τα AU έχουν οριστεί με βάση την ανιχνευσιμότητα των κινήσεων, και όχι καθαρά ανατομικά κριτήρια, παρότι υπάρχει σαφής αντιστοιχία ανάμεσα σε AUs και την ενεργοποίηση συγκεκριμένων μυών. Το εγχειρίδιο FACS περιγράφει κριτήρια για την αναγνώριση και κωδικοποίηση κάθε AU.

Μετά τον προσδιορισμό των επιμέρους AUs, εξετάστηκαν όλοι οι **πιθανοί συνδυασμοί** για το άνω πρόσωπο και σχεδόν όλοι οι συνδυασμοί για το κάτω πρόσωπο. Τα αποτελέσματα έδειξαν ότι οι περισσότερες αλλαγές στην εμφάνιση του προσώπου ήταν **προσθετικές**: κάθε AU ήταν σαφώς αναγνωρίσιμο και πρακτικά αναλλοίωτο. Σε μερικούς μόνο συνδυασμούς η εμφάνιση του προσώπου ήταν καινούρια. Όλοι αυτοί οι συνδυασμοί καταγράφονται στο σύστημα ταξινόμησης FACS. Το εγχειρίδιο FACS περιγράφει όλους τους συνδυασμούς στους οποίους εμφανίζεται ένα AU.

Πέρα από τα AUs υπάρχουν επίσης **περιγραφείς κινήσεων** με το όνομα Action Descriptors (AD). Τα ADs περιγράφουν ενιαίες κινήσεις που περιλαμβάνουν τις ενέργειες πολλών μυικών ομάδων, όπως για παράδειγμα η πρόσθια ώθηση του σαγονιού. Η μυική βάση αυτών των κινήσεων δεν έχει προσδιοριστεί, ούτε έχουν διευκρινιστεί λεπτομερείς συμπεριφορές όπως για τα υπόλοιπα AUs.

Η **ένταση** του κάθε AU καταγράφεται με την **προσθήκη** ενός **γράμματος A-E** της λατινικής αλφαβήτου στον κωδικό του AU. Η κλίμακα για την ένταση έχει ως εξής:

- A – Ίχνος (Trace)
- B – Ελαφρώς (Slight)
- C – Εμφανές ή Έντονο (Marked or Pronounced)
- D – Πολύ έντονο ή Ακραίο (Severe or Extreme)
- E – Μέγιστο (Maximum)

Για παράδειγμα ο κωδικός AU 12A αναφέρεται στο ασθενέστερο ίχνος του AU 12, ενώ ο κωδικός AU 12E είναι η μέγιστη δυνατή ένταση για το συγκεκριμένο άτομο.

Είναι σαφές από την παραπάνω περιγραφή ότι πρόκειται για ένα **ενδεδειγμένο, εξαντλητικό σύστημα** που περιλαμβάνει πληθώρα πληροφοριών. Δεν υπάρχει κίνηση που να περιγράφεται σε άλλο σύστημα η οποία να μη μπορεί να περιγραφεί και στο FACS. Το σύστημα προβλέπει και στην ταξινόμηση ασυμμετριών όσον αφορά τα AUs ή τις διαφορετικές εντάσεις. Επίσης υπάρχει τρόπος για την καταγραφή του χρονισμού των κινήσεων.

Παρακάτω δίνονται ενδεικτικά μερικά παραδείγματα μονάδων κινήσεων AU ανά κατηγορία.

| <b>Βασικοί κωδικοί (Main Codes)</b> |                        |  |
|-------------------------------------|------------------------|--|
| <b>Αριθμός AU</b>                   | <b>Όνομα FACS</b>      | <b>Μυική βάση (Muscular Basis)</b>                               |
| 0                                   | Neutral Face           |  |
| 1                                   | Inner Brow Raiser      | frontalis (pars medialis)  |
| 2                                   | Outer Brow Raiser      | frontalis (pars lateralis)                                       |
| 4                                   | Brow Lowerer           | depressor glabellae, depressor supercilii, corrugator supercilii |
| 5                                   | Upper Lid Raiser       | levator palpebrae superioris                                     |
| 6                                   | Cheek Raiser           | orbicularis oculi (pars orbitalis)                               |
| 7                                   | Lid Tightener          | orbicularis oculi (pars palpebralis)                             |
| 8                                   | Lips Toward Each Other | orbicularis oris   |
| 9                                   | Nose Wrinkler          | levator labii superioris alaeque nasi                            |
| 10                                  | Upper Lip Raiser       | levator labii superioris, caput infraorbitalis                   |

**Πίνακας 2-1** Παράδειγμα βασικών κωδικών AU του συστήματος FACS [23]

| <b>Κωδικοί κίνησης κεφαλιού (Head Movement Codes)</b> |                             |  |
|---|-----------------------------|--|
| <b>Αριθμός AU</b>                                     | <b>Όνομα FACS</b>           | <b>Κίνηση (Action)</b>   |
| M57   | Head Thrust Forward         | The onset of 17+24 is immediately preceded, accompanied, or followed by a head thrust forward.   |
| 58  | Head Back                   |  |
| M59   | Head Shake Up and Down      | The onset of 17+24 is immediately preceded, accompanied, or followed an up-down head shake (nod).  |
| M60   | Head Shake Side to Side     | The onset of 17+24 is immediately preceded, accompanied, or followed by a side to side head shake.   |
| M83   | Head Upward and to the Side | The onset of the symmetrical 14 is immediately preceded or accompanied by a movement of the head, upward and turned and/or tilted to either the left or right. |

**Πίνακας 2-2** Παράδειγμα κωδικών AU κίνησης κεφαλιού στο σύστημα FACS [23]

| Κωδικοί κίνησης ματιών (Eye Movement Codes) |                 |  |
|---|-----------------|--|
| Αριθμός AU                                  | Όνομα FACS      | Κίνηση (Action)  |
| 61  | Eyes Turn Left  |  |
| M61   | Eyes Left       | The onset of the symmetrical 14 is immediately preceded or accompanied by eye movement to the left.  |
| 62  | Eyes Turn Right |  |
| M62   | Eyes Right      | The onset of the symmetrical 14 is immediately preceded or accompanied by eye movement to the right. |
| 63  | Eyes Up         |  |

**Πίνακας 2-3** Παράδειγμα κωδικών AU κίνησης ματιών στο σύστημα FACS [23]

| Κωδικοί ορατότητας (Visibility Codes) |                                | Κωδικοί γενικής συμπεριφοράς (Gross Behavior Codes) |                      |
|---------------------------------------|--------------------------------|---|----------------------|
| Αριθμός AU                            | Όνομα FACS                     | Αριθμός AU  | Όνομα FACS           |
| 70                                    | Brows and forehead not visible | 40  | Sniff                |
| 71                                    | Eyes not visible               | 50  | Speech               |
| 72                                    | Lower face not visible         | 80  | Swallow              |
| 73                                    | Entire face not visible        | 81  | Chewing              |
| 74                                    | Unscorable                     | 85  | Head nod up and down |

**Πίνακας 2-4** Παράδειγμα κωδικών AU ορατότητας και γενικής συμπεριφοράς [23]

## 2.8 Θεωρία συναισθημάτων του Plutchik

Ο Αμερικανός ψυχολόγος Robert Plutchik είναι γνωστός για τη θεωρία που διατύπωσε σχετικά με τα συναισθήματα, η οποία έχει τη βάση της στην **εξελικτική ψυχολογία**.

Ο Plutchik υποστήριξε ότι υπάρχουν **οκτώ βασικά συναισθήματα**, το καθένα εκ των οποίων είναι μια προσαρμογή λόγω της εξέλιξης των ειδών. Υποστήριξε επίσης ότι και τα οκτώ συναισθήματα μπορούν να παρατηρηθούν **σε όλους τους οργανισμούς**. Σε αντίθεση δηλαδή με τις υπόλοιπες εξελικτικές θεωρίες συναισθημάτων που συνήθως διατυπώνουν ότι τα συναισθήματα είναι αποτέλεσμα της φυσικής επιλογής που πραγματοποιήθηκε στα πρώιμα ανθρωποειδή, ο Plutchik θεώρησε ότι η διαδικασία της φυσικής επιλογής άρχισε πολύ νωρίτερα, με αποτέλεσμα οι προσαρμογές-συναισθήματα να είναι κοινά σε περισσότερα είδη και όχι μόνο στον άνθρωπο. Τοποθετεί την εξέλιξη των συναισθημάτων στην Κάμβρια περίοδο, χονδρικά 550 εκατομμύρια χρόνια πριν [24]. Για να δώσει έμφαση

στον πρωτόγονο χαρακτήρα των συναισθημάτων, παραλλήλισε τα συναισθήματα με την ύπαρξη του DNA ή τους πνεύμονες, χαρακτηριστικά τόσο σημαντικά για τη ζωή που από τη στιγμή που εμφανίστηκαν, έχουν διατηρηθεί στα είδη που τα κατείχαν. Με τον ίδιο τρόπο θεωρεί ότι τα συναισθήματα είναι απαραίτητα στους οργανισμούς και έχουν εξελιχθεί ώστε να αυξάνουν την **ικανότητα ατομικής επιβίωσης** και την **αναπαραγωγική ικανότητα** των ζώων.

Η θεωρία του Plutchik βασίζεται στα εξής **δέκα αξιώματα** [25]:

1. Η έννοια του συναισθήματος μπορεί να εφαρμοστεί σε όλα τα επίπεδα της εξέλιξης και ισχύει για όλα τα ζώα και τον άνθρωπο.
2. Τα συναισθήματα έχουν εξελικτική προέλευση και έχουν εξελιχθεί σε ποικίλους τρόπους έκφρασης στα διαφορετικά είδη.
3. Τα συναισθήματα εξυπηρετούσαν την προσαρμογή του οργανισμού σε προβλήματα ζωτικής σημασίας που παρουσιάζονταν στο περιβάλλον του.
4. Παρά τους διαφορετικούς τρόπους έκφρασης των συναισθημάτων στα διάφορα είδη ζώων, υπάρχουν κάποια κοινά στοιχεία, κάποια πρωτογενή πρότυπα που μπορούν να προσδιοριστούν.
5. Υπάρχει ένας μικρός αριθμός βασικών, πρωταρχικών, πρότυπων συναισθημάτων.
6. Όλα τα υπόλοιπα συναισθήματα είναι μικτές ή παράγωγες καταστάσεις, δηλαδή παρουσιάζονται ως συνδυασμοί, μίξεις ή σύνθετες ενώσεις των βασικών συναισθημάτων.
7. Τα βασικά συναισθήματα είναι υποθετικά κατασκευάσματα ή εξιδανικευμένες καταστάσεις οι ιδιότητες και τα χαρακτηριστικά των οποίων μπορούν μόνο να συναχθούν από διάφορα είδη (πειραματικών) στοιχείων.
8. Τα βασικά συναισθήματα μπορούν να περιγραφούν εννοιολογικά ως ζεύγη διαμετρικά αντίθετων συναισθημάτων.
9. Όλα τα συναισθήματα έχουν διακυμάνσεις στους βαθμούς ομοιότητας μεταξύ τους.
10. Κάθε συναίσθημα μπορεί να παρατηρηθεί με κυμαινόμενη ένταση και επίπεδο διέγερσης.

Τα οκτώ συναισθήματα που ο Plutchik πρότεινε ως βασικά είναι:

- ο **Θυμός** (Anger)
- ο **Φόβος** (Fear)
- ο **Λύπη** (Sadness)
- ο **Αηδία** (Disgust)



- **Έκπληξη** (Surprise)
- **Προσμονή** (Anticipation , εναλλακτικά αναφέρεται και ως Expectation)
- **Εμπιστοσύνη, Αποδοχή** (Trust, εναλλακτικά αναφέρεται και ως Acceptance)
- **Χαρά** (Joy)

Όπως διατυπώθηκε στο 8<sup>ο</sup> αξίωμα της θεωρίας, τα συναισθήματα συνδυάζονται ανά δύο για να σχηματίσουν ζεύγη διπόλων, δηλαδή **ζεύγη εξ'ορισμού αντίθετων συναισθημάτων**.

Τα ζεύγη των αντίθετων συναισθημάτων είναι:

- Χαρά – Λύπη
- Εμπιστοσύνη/Αποδοχή – Αηδία
- Προσμονή – Έκπληξη
- Θυμός – Φόβος

Στη θεωρία του Plutchik γίνεται λόγος για προσαρμογές λόγω εξέλιξης, οπότε από αυτήν την οπτική, που εστιάζει στο λειτουργικό ρόλο των συναισθημάτων, τα συναισθήματα είναι προσαρμοσμένες συμπεριφορές των διαφόρων ειδών ζώων. Ο όρος «συναίσθημα» από τη στενή του έννοια χρησιμοποιείται για να περιγράψει αυτές τις συμπεριφορές στους ανθρώπους. Φυσικά ο Plutchik αναγνωρίζει ότι δεν εμφανίζουν όλα τα είδη τις ίδιες συμπεριφορές. Τα **συναισθήματα των ανθρώπων** είναι πολύ **πιο πολύπλοκα** από τις συμπεριφορές που εντοπίζονται σε άλλα είδη, που βρίσκονται χαμηλότερα στην εξελικτική κλίμακα. Οι βασικοί όμως μηχανισμοί λειτουργίας, υποστηρίζει, μένουν αναλλοίωτοι σε όλα τα ζώα, μέχρι και τους ανθρώπους.

### 2.8.1 Συνδυασμοί βασικών συναισθημάτων

Ο Plutchik ανέφερε στο 6<sup>ο</sup> αξίωμα την ύπαρξη πολλών ακόμη ενδιάμεσων συναισθημάτων, που υποστηρίζει ότι είναι συνδυασμοί των βασικών συναισθημάτων, ή παραλλαγές με αυξομειώσεις στην ένταση του συναισθήματος (6<sup>ο</sup> αξίωμα). Για παράδειγμα ο φόβος και η λύπη συνδυάζονται σε απόγνωση, η ενόχληση είναι πιο ήπια μορφή του θυμού.

Ο Plutchik συστηματοποίησε αυτή τη γνώση σε **πρωτογενείς, δευτερογενείς και τριτογενείς δυάδες**. Κατέγραψε τους συνδυασμούς των οκτώ βασικών συναισθημάτων και το παραγόμενο συναίσθημα για κάθε συνδυασμό. Αν τα συναισθήματα που συμμετέχουν στο συνδυασμό είναι γειτονικά πάνω στον τροχό συναισθημάτων (Σχήμα 2-10) τότε γίνεται λόγος για τις πρωτογενείς δυάδες. Αν παρεμβάλλεται ένα συναίσθημα μεταξύ τους τότε ονομάζονται δευτερογενείς δυάδες, ενώ όταν παρεμβάλλονται τρία συναισθήματα μεταξύ τους, ονομάζονται τριτογενείς δυάδες. Θεωρητικά οι πρωτογενείς δυάδες είναι συναισθήματα που βιώνουμε πιο συχνά από τις δευτερογενείς και κατ'επέκταση τριτογενείς δυάδες.

| Πίνακας 2-5 Συνδυασμοί βασικών συναισθημάτων κατά Plutchik [26] |   |  |
|---|---|--|
| Πρωτογενείς дуάδες<br>(συχνά)                                   | Δευτερογενείς дуάδες<br>(περιστασιακά)    | Τριτογενείς дуάδες<br>(σπάνια)             |
| Joy + Trust<br><b>Love</b>                                      | Joy + Fear<br><b>Guilt</b>                | Joy + Surprise<br><b>Delight</b>           |
| Trust + Fear<br><b>Submission</b>                               | Trust + Surprise<br><b>Curiosity</b>      | Trust + Sadness<br><b>Sentimentality</b>   |
| Fear + Surprise<br><b>Alarm</b>                                 | Fear + Sadness<br><b>Despair</b>          | Fear + Disgust<br><b>Shame</b>             |
| Surprise + Sadness<br><b>Disappointment</b>                     | Surprise + Disgust<br><b>?</b>            | Surprise + Anger<br><b>Outrage</b>         |
| Sadness + Disgust<br><b>Remorse</b>                             | Sadness + Anger<br><b>Envy</b>            | Sadness + Anticipation<br><b>Pessimism</b> |
| Disgust + Anger<br><b>Contempt</b>                              | Disgust + Anticipation<br><b>Cynicism</b> | Disgust + Joy<br><b>Morbidness</b>         |
| Anger + Anticipation<br><b>Aggression</b>                       | Anger + Joy<br><b>Pride</b>               | Anger + Trust<br><b>Dominance</b>          |
| Anticipation + Joy<br><b>Optimism</b>                           | Anticipation + Trust<br><b>Fatalism</b>   | Anticipation + Fear<br><b>Anxiety</b>      |

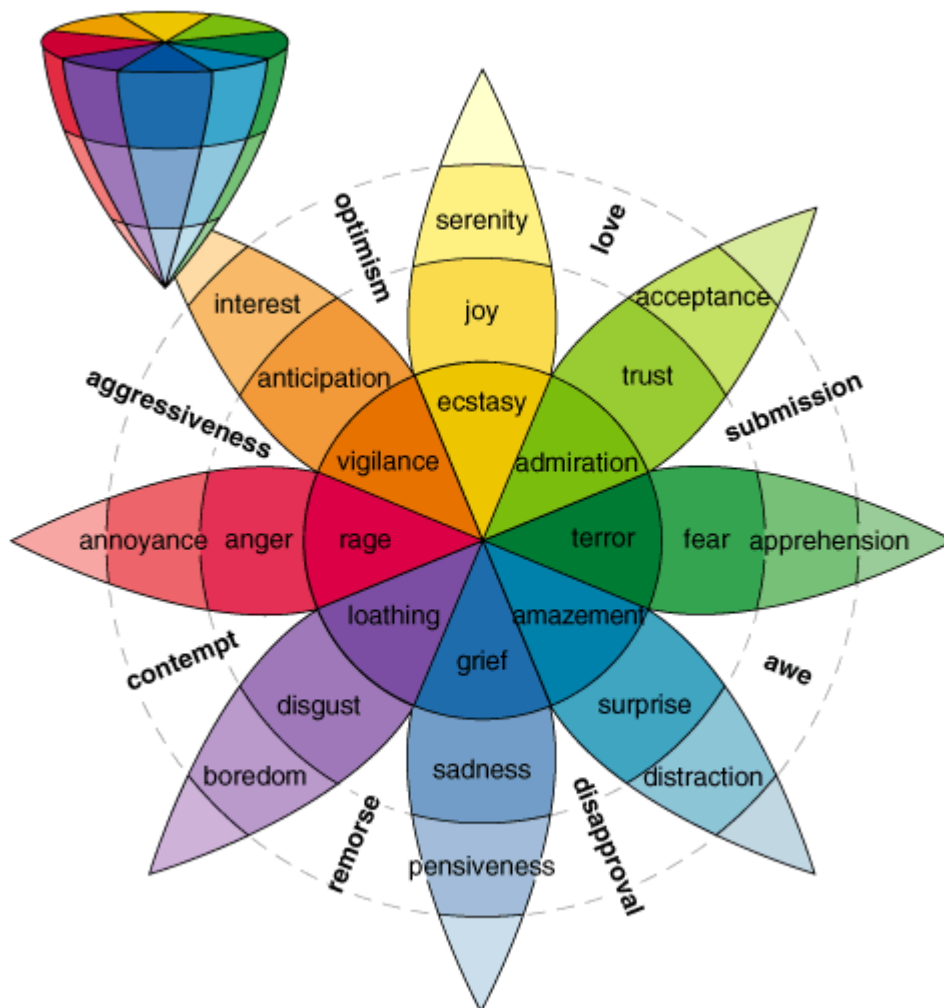
## 2.8.2 Ο τροχός συναισθημάτων του Plutchik

Ένα πολύ γνωστό στοιχείο της θεωρίας του Plutchik είναι η γραφική απεικόνιση της θεωρίας του. Πρόκειται για τον **‘τροχό συναισθημάτων του Plutchik’**, που οπτικοποιεί με κατανοητό τρόπο πολλές πτυχές της θεωρίας του. Πρότεινε ένα **κωνικό μοντέλο τριών διαστάσεων** και το **διδιάστατο ανάπτυγμα του κώνου** που είναι παρεμφερές με ένα **μοντέλο τροχού**, για να περιγράψει τις σχέσεις των συναισθημάτων.

Όπως ειπώθηκε ήδη ο Plutchik κατέγραψε πολλά ακόμη **ενδιάμεσα συναισθήματα**, που είναι συνδυασμοί δύο ή τριών βασικών συναισθημάτων και παραλλαγές των βασικών συναισθημάτων με αυξομειώσεις στην έντασή τους. Επίσης κεντρική έννοια στη θεωρία του είναι τα δίπολα συναισθημάτων. Τα στοιχεία αυτά αποδίδονται και οπτικά στον τροχό του.

Η πληροφορία για τους συνδυασμούς συναισθημάτων και κατ'επέκταση για ολόκληρο το φάσμα των συναισθημάτων αποτυπώνεται με πολύ αποδοτικό τρόπο καθώς ο τροχός συναισθημάτων συνδέεται με έναν **χρωματικό τροχό**. Όπως και στη θεωρία χρωμάτων, τα βασικά συναισθήματα μπορούν να εκφραστούν με **διαφορετικές εντάσεις**, ισοδύναμο των διαφορετικών φωτεινότητων, και διάφορα (πρωταρχικά) συναισθήματα μπορούν να αναμιχθούν για να δώσουν καινούριες, διαφορετικές χροιές συναισθημάτων. Οι οκτώ τομείς του κύκλου αποδίδουν προφανώς τα οκτώ βασικά συναισθήματα. Τα δίπολα

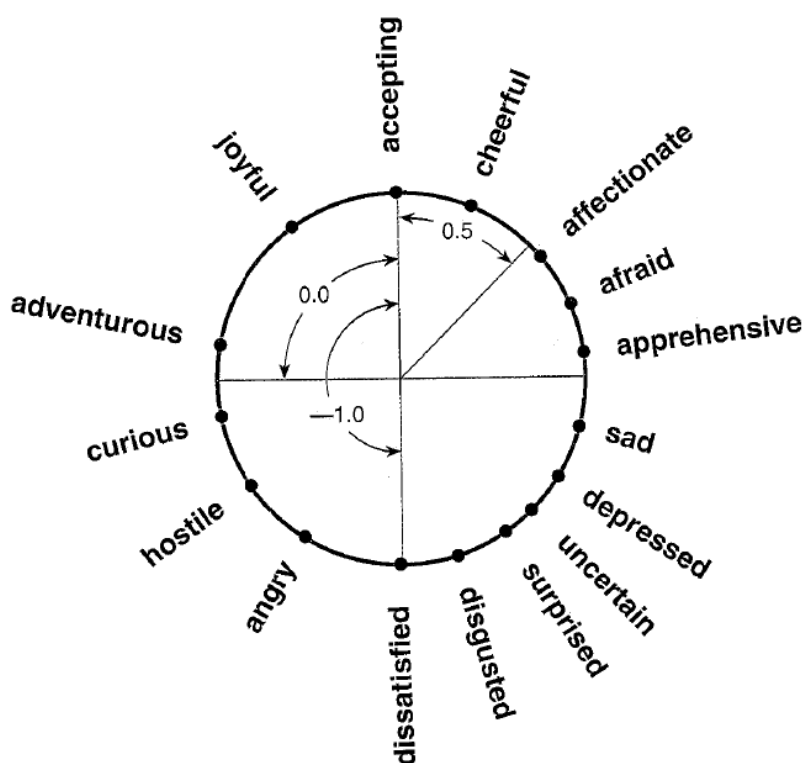
συναισθήματα αποδίδονται ως **αντιδιαμετρικές θέσεις** πάνω στον κύκλο. Στο διδιάστατο ανάπτυγμα, τα συναισθήματα που βρίσκονται στα λευκά διαστήματα αντιπροσωπεύουν τις πρωτογενείς δυνάδες, και βρίσκονται ανάμεσα στις δύο συνιστώσες τους [27].



**Σχήμα 2-10** Κώνος και τροχός συναισθημάτων του Plutchik [28]

Στο τρισδιάστατο μοντέλο, η κάθετη διάσταση αντιπροσωπεύει την ένταση του συναισθήματος και αποδίδεται με διαφορές στη φωτεινότητα του χρώματος. Όσο πιο ψηλά κινούμαστε στον κώνο, τόσο πιο έντονο γίνεται το συναίσθημα και το χρώμα. Η **κυκλική διάταξη** των συναισθημάτων αντιπροσωπεύει τη συνάφεια και το **βαθμό ομοιότητας** των συναισθημάτων. Αποδίδεται με τη μετάβαση από κάποιο χρώμα σε χρώμα που είναι κοντά του στο χρωματικό φάσμα. Συναισθήματα που γειτνιάζουν έχουν υψηλότερους βαθμούς ομοιότητας και μεγαλύτερη συνάφεια. Συναισθήματα που είναι αντιδιαμετρικά τοποθετημένα είναι τα διπολικά συναισθήματα που είναι αντιμαχόμενα και θεωρείται ότι δεν εμφανίζονται ταυτόχρονα. Τα αντίστοιχα χρώματα είναι συμπληρωματικά.

Για να ερευνηθούν οι σχέσεις ανάμεσα στα διάφορα συναισθήματα, έγιναν μελέτες στις οποίες ζητήθηκε από συμμετέχοντες να εκτιμήσουν ο βαθμό ομοιότητας μεταξύ συγκεκριμένων συναισθημάτων, ανά ζεύγη. Η χρήση μιας κλίμακας βαθμολόγησης για την ομοιότητα οδήγησε στον **εμπειρικό προσδιορισμό** των θέσεων των συναισθημάτων πάνω στο διδιάστατο επίπεδο. Οι **γωνιακές θέσεις** φαίνονται στο ακόλουθο σχήμα. Αντίθετα συναισθήματα βαθμολογούνται με βαθμό ομοιότητας -1.0, ενώ ταυτόσημες έννοιες βαθμολογούνται με +1.0 που είναι ο μέγιστος βαθμός. Εν ολίγοις, τα αποτελέσματα έδειξαν ότι τα συναισθήματα **δεν είναι ομοιοκατανεμημένα** στο χώρο, σε ίσες γωνίες όπως στο θεωρητικό μοντέλο αναπαράστασης. Αντίθετα υπάρχουν πιο πυκνές ή λιγότερο πυκνές περιοχές. Οι γωνιακές μετρήσεις βρίσκουν εφαρμογή σε πρακτικές υλοποιήσεις που χρειάζονται ποσοτική αναπαράσταση των συναισθημάτων.



**Σχήμα 2-11** Προσδιορισμός γωνιακών θέσεων συναισθημάτων [27]

### 2.8.3 Συμπεριφορές που συνδέονται με τα βασικά συναισθήματα

Για τον Plutchik, τα συναισθήματα είναι στενά συνδεδεμένα με συμπεριφορές που έχουν υψηλή σημασία για την επιβίωση του οργανισμού. Θεωρεί ότι το συναίσθημα είναι το έναυσμα που προκαλεί την απαραίτητη συμπεριφορά ή δράση για τον οργανισμό, δεδομένης μιας κατάστασης στο περιβάλλον του στην οποία πρέπει να ανταποκριθεί ή να προσαρμοστεί. Θεωρεί ότι ένα **ερέθισμα του περιβάλλοντος διαταράσσει την ισορροπία του οργανισμού η οποία πρέπει να αποκατασταθεί**. Το συναίσθημα είναι εκείνο που προκαλεί την κατάλληλη αλληλεπίδραση του οργανισμού με το περιβάλλον ούτως ώστε να αποκαθίσταται η ισορροπία του.

| Πίνακας 2-6 Συμπεριφορές για τα βασικά συναισθήματα του Plutchik [26] , [27] |  |   |  |                           |
|--|--|---|--|---------------------------|
| Ερέθισμα<br>(Stimulus Event)   | Γνωσιακή<br>αξιολόγηση<br>(Cognitive<br>Appraisal) | Συναίσθημα/<br>Υποκειμενική<br>αντίδραση<br>(Feeling state/<br>Subjective reaction) | Συμπεριφορά<br>αντίδρασης<br>(Behavioural<br>Reaction) | Σκοπιμότητα<br>(Function) |
| Gain of valued object  | “Possess”  | Joy   | Retain or repeat                                       | Gain resources            |
| Member of one’s group  | “Friend”   | Trust   | Groom  | Mutual support            |
| Threat   | “Danger”   | Fear  | Escape   | Safety                    |
| Unexpected event   | “What is it?”                                      | Surprise  | Stop   | Gain time to orient       |
| Loss of valued object  | “Abandonment”                                      | Sadness   | Cry  | Reattach to lost object   |
| Unpalatable object   | “Poison”   | Disgust   | Vomit  | Eject poison              |
| Obstacle   | “Enemy”  | Anger   | Attack   | Destroy obstacle          |
| New territory  | “Examine”  | Anticipation  | Map  | Knowledge of territory    |

## 2.9 Αναπαράσταση συναισθημάτων της Whissell

Η Cynthia Whissell υπέθεσε ότι τα συναισθήματα είναι σημεία σε ένα χώρο σχετικά λίγων διαστάσεων. Σε πρώτη προσέγγιση, αρκούν δύο διαστάσεις για την αναπαράσταση των συναισθημάτων. Εάν αναπαρασταθούν σε δύο διαστάσεις, οι άξονες του επιπέδου που προκύπτει είναι:

- ο άξονας της **ενεργοποίησης** (activation) και
- ο άξονας της **αξιολόγησης** (evaluation)

Πρόκειται για μια απλή αναπαράσταση, ικανή όμως να περιγράψει ένα σχετικά ευρύ φάσμα συναισθημάτων [29].

Η τιμή της ενεργοποίησης αναπαριστά το **βαθμό διέγερσης** που σχετίζεται με το συναίσθημα. Για παράδειγμα το συναίσθημα «έκπληκτος» (surprised) έχει υψηλή τιμή ενεργοποίησης ενώ χαρακτηρισμοί όπως το «ντροπαλός» (bashful), έχουν χαμηλή τιμή ενεργοποίησης.

Η παράμετρος της αξιολόγησης αναπαριστά **πόσο ευχάριστο** είναι το συναίσθημα για το υποκείμενο άτομο. Ο όρος «ενοχικός» (guilty) αποτελεί το αρνητικό ακρότατο με τιμή 1,1 ενώ ο όρος «πανευτυχής» (delighted) είναι το θετικό ακρότατο με τιμή 6,4.

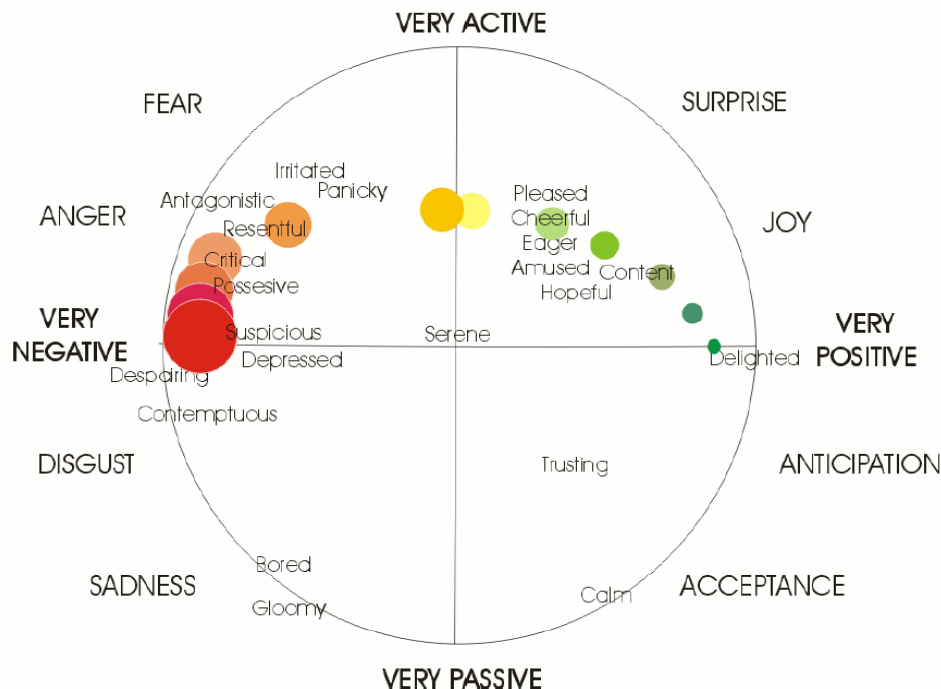
Από πρακτικής σκοπιάς, η παράμετρος αξιολόγησης εκφράζει τα **εσωτερικά αισθήματα** του ατόμου για ένα γεγονός και η εκτίμησή της μέσω εκφράσεων του

προσώπου είναι αδύνατη. Η παράμετρος της ενεργοποίησης όμως σχετίζεται με τις **κινήσεις των μυών** του προσώπου και μπορεί εύκολα να εκτιμηθεί από τις μεταβολές των χαρακτηριστικών του προσώπου.

Στο ακόλουθο σχήμα δίνεται ο διδιάστατος χώρος που περιγράφηκε παραπάνω. Ο οριζόντιος άξονας αναπαριστά την παράμετρο της αξιολόγησης ενώ ο κάθετος άξονας αναπαριστά την παράμετρο της ενεργοποίησης. Τα συναισθήματα κατανέμονται σε **ελλειπτικό σχήμα**, με ανόμοιες αποστάσεις μεταξύ τους. Επίσης σημειώνεται ότι υπάρχει μεγαλύτερος διαχωρισμός στα συναισθήματα με μεγαλύτερη ενεργοποίηση, παρά στα συναισθήματα χαμηλής ενεργοποίησης [30].

Σε αυτήν την αναπαράσταση, η **ένταση** του συναισθήματος μπορεί να ποσοτικοποιηθεί μέσω της **απόστασης από την αρχή των αξόνων**, που μπορεί να θεωρηθεί ότι είναι η ουδέτερη συναισθηματική κατάσταση του ατόμου, για τιμές παραμέτρων ίσες με μηδέν. Ένα συναίσθημα μπορεί να χαρακτηριστεί ως «έντονο» εάν η απόσταση από την αρχή των αξόνων περνά ένα κατώφλι που μπορεί να επιλεγεί ανάλογα με την εφαρμογή.

Επισημαίνεται ότι με τη σύμπτυξη μεγάλου πλήθους συναισθημάτων στις δύο διαστάσεις, είναι αναπόφευκτος ο κίνδυνος απώλειας πληροφορίας. Επίσης υπενθυμίζεται ότι τα αποτελέσματα διδιάστατων συστημάτων δεν ταυτίζονται απαραίτητα. Για παράδειγμα στο μοντέλο του Plutchik ο φόβος και ο θυμός είναι αντίθετα συναισθήματα και απεικονίζονται σε αντιδιαμετρικές θέσεις. Στο μοντέλο της Whissell, ο φόβος και ο θυμός βρίσκονται κοντά, στο ίδιο τεταρτημόριο, αφού και τα δύο είναι σχετικά δυσάρεστα για το άτομο και περίπου το ίδιο ενεργά.



**Σχήμα 2-12** Διδιάστατος χώρος αξιολόγησης – ενεργοποίησης [31]

## 2.10 Εργασίες στη συναισθηματική υπολογιστική

Η **συναισθηματική υπολογιστική** εστιάζει στη μελέτη και σχεδίαση συστημάτων που αναγνωρίζουν, ερμηνεύουν, επεξεργάζονται και προσομοιώνουν τις ανθρώπινες συναισθηματικές εμπειρίες. Πρόκειται για ένα σύγχρονο διεπιστημονικό πεδίο η απαρχή του οποίου θα μπορούσε να τοποθετηθεί το 1995, με τη δημοσίευση του άρθρου της R. Picard [32] σχετικά με τη σημασία της συναισθηματικής εμπειρίας και πώς αυτή θα πρέπει να ενσωματωθεί στη σύγχρονη επιστήμη υπολογιστών.

Οι εργασίες πάνω στη συναισθηματική υπολογιστική περιλαμβάνουν τα ακόλουθα **πεδία έρευνας**:

- Ανάλυση και αναγνώριση συναισθηματικού περιεχομένου ομιλίας.
- Αναγνώριση χειρονομιών και χειρομορφών, γενικά κινήσεων του σώματος.
- Καταγραφή και ανάλυση βιοσημάτων σχετικά με τη συναισθηματική κατάσταση του ατόμου.
- Αναγνώριση εκφράσεων προσώπου για την αναγνώριση του συναισθήματος.
- Σχεδίαση υπολογιστικών συστημάτων που αποκρίνονται συναισθηματικά, λαμβάνουν αποφάσεις με συναισθηματικά κριτήρια, προσομοιώνουν συναισθηματικές αντιδράσεις ή ακόμη «έχουν» συναισθήματα.

Ειδικότερα για την παρακολούθηση και επεξεργασία της έκφρασης του προσώπου έχουν χρησιμοποιηθεί **διάφορες μέθοδοι** όπως **κρυμμένα μοντέλα Markov, κυματίδια Gabor**, η μέθοδος της **οπτικής ροής**, επεξεργασία **νευρωνικών δικτύων** ή το **μοντέλο της ενεργούς εμφάνισης** (active appearance model).

Καταγράφονται εν συντομία εργασίες που σχετίζονται με τη συναισθηματική υπολογιστική της **οπτικής επεξεργασίας προσώπου** και παρουσιάζουν ενδιαφέρουσες προσεγγίσεις:

- Προηγμένες τεχνικές στην εξαγωγή και την ταξινόμηση χαρακτηριστικών, στην παρακολούθηση προσώπου και αντικειμένων επιτρέπουν τη σχεδίαση **εύρωστων εφαρμογών πραγματικού χρόνου**, όπως αυτό της Bartlett και συνεργατών [33], που υλοποιεί μέθοδο με **κυματίδια Gabor** για την ανίχνευση αυθόρμητων εκφράσεων.
- Οι Pantic και Patras [34] χρησιμοποιούν σύνολο χαρακτηριστικών σημείων γύρω από το στόμα, τα μάτια, τη μύτη και το πηγούνι. Στην εργασία τους διερευνούν την αυτόματη ανάλυση εκφράσεων του προσώπου σε **πλάγια όψη** αντί της συνήθους πρόσθιας, μετωπικής κατεύθυνσης.
- Σχετικά λίγες εργασίες συνδυάζουν την έκφραση του προσώπου με τις υπόλοιπες κινήσεις του κεφαλιού ή του σώματος. Σε μία από αυτές, οι Zhang και Ji [35] εξερευνούν το **συνδυασμό** εκφράσεων προσώπου με **κινήσεις του κεφαλιού**.

- Έχουν γίνει απόπειρες για την αυτόματη ανάλυση και τον αυτόματο **διαχωρισμό στημένων** (posed) και γνήσιων, **αυθόρμητων εκφράσεων**. Μια τέτοια προσπάθεια έγινε από τους Valstar και συνεργάτες [36], για το διαχωρισμό του αυθόρμητου χαμόγελου από το επιτηδευμένο. Χρησιμοποιούνται και πάλι χαρακτηριστικά σημεία του προσώπου, ενώ γίνεται χρήση της πληροφορίας για την κίνηση του κεφαλιού και των ώμων.
- Έχουν προταθεί μέθοδοι για την ανίχνευση υβριδικών συναισθηματικών/ συμπεριφορικών εκφράσεων σε αντιδιαστολή με τις αμιγώς συναισθηματικές εκφράσεις. Παράδειγμα είναι η κατάσταση της κούρασης, της απογοήτευσης, του πόνου, της συμφωνίας (agreement) ή της εκδήλωσης ενδιαφέροντος. Μια τέτοια πρόταση για την **αναγνώριση του ενδιαφέροντος** του χρήστη ανήκει στους Yeasin, Bulot και Sharma [37].
- Οι Cohn και συνεργάτες [38] αναλύουν αυθόρμητες συμπεριφορές με **χρήση τρισδιάστατων μοντέλων**, και το συνδυασμό πληροφοριών κίνησης του κεφαλιού. Επίσης χρησιμοποιούν στη μέθοδό τους κυματίδια Gabor.

Πρόσφατα έχει γίνει στροφή στα λεγόμενα **πολυτροπικά (multimodal) συστήματα** επικοινωνίας ανθρώπου-μηχανής, τα οποία συνδυάζουν αποτελέσματα περισσότερων του ενός πεδίων έρευνας, πχ. οπτικοακουστικές πληροφορίες ή το συνδυασμό με βιοσήματα.

Παρουσιάζουμε ενδεικτικά κάποιες εργασίες [39] σχετικές με τα multimodal συστήματα:

- Το σύστημα MAUI – Multimodal Affect User Interface των Lisetti και Nasoz [40] συνδυάζει την έκφραση του προσώπου με **βιοσήματα** για την αναγνώριση συναισθημάτων όπως ο φόβος ή ο θυμός. Στη συνέχεια προσαρμόζει έναν πράκτορα διεπαφής (animated interface agent) ώστε να **καθρεφτίζει το συναίσθημα του χρήστη**.
- Το δυναμικό σύστημα των Maat και Pantic [41] έχει τη δυνατότητα να μαθαίνει και να αναλύει μοτίβα συμπεριφοράς του χρήστη που εξαρτώνται από τα **συμφραζόμενα** χρησιμοποιώντας **πολυαισθητηριακά δεδομένα** (multisensory data). Στη συνέχεια προσαρμόζει την αλληλεπίδραση με το χρήστη.
- Ο **αυτοματοποιημένος οδηγός μάθησης** του Karoor και συνεργατών [42], συνδυάζει πληροφορίες από κάμερες, κάθισμα και ποντίκι με αισθητήρες, ασύρματο δερματικό αισθητήρα και την κατάσταση προόδου της εργασίας για να **ανιχνεύσει το αίσθημα της απογοήτευσης** του χρήστη και να προβλέψει πότε ο χρήστης χρειάζεται βοήθεια.
- Το υπολογιστικό **σύστημα υποβοηθούμενης μάθησης** του πανεπιστημίου του Illinois (UIUC) χρησιμοποιεί multimodal προσέγγιση [43]. Γίνεται χρήση ενός εικονικού προσώπου (avatar) που προτείνει κατάλληλη στρατηγική εκμάθησης, βασισμένη στην έκφραση προσώπου του χρήστη, την **κίνηση των ματιών** του, **λέξεις κλειδιά** και την κατάσταση προόδου της εργασίας.



## ΚΕΦΑΛΑΙΟ 3: Το πρότυπο MPEG-4

### 3.1 Εισαγωγικά στοιχεία

Το MPEG-4 είναι ένα πρότυπο ISO/IEC που αναπτύχθηκε από την ομάδα εμπειρογνομόνων MPEG (Moving Picture Experts Group), την ίδια επιτροπή που ανέπτυξε τα πρότυπα MPEG-1 και MPEG-2, που έκαναν πραγματικότητα τα διαδραστικά βίντεο σε CD-ROM, DVD και τη ψηφιακή τηλεόραση. Το πρότυπο MPEG-4, πρότυπο για την κωδικοποίηση και συμπίεση φυσικού και συνθετικού οπτικοακουστικού περιεχομένου με βάση τα αντικείμενα (object-based), είναι το αποτέλεσμα μιας διεθνούς προσπάθειας στην οποία συμμετείχαν εκατοντάδες ερευνητές και μηχανικοί από όλο τον κόσμο. Η πρώτη έκδοση του MPEG-4, με την επίσημη ονομασία 'ISO/IEC 14496', οριστικοποιήθηκε στα τέλη του 1998 και έγινε διεθνές πρότυπο στις αρχές του 1999. Στις αρχές του 2000 αναγνωρίστηκε ως διεθνές πρότυπο και η δεύτερη έκδοση του MPEG-4, που είναι επέκταση πλήρως συμβατή με την προηγούμενη έκδοση. Έκτοτε έχουν προστεθεί πολλές επεκτάσεις και πολλές εργασίες βρίσκονται ακόμη σε εξέλιξη.

Το πρότυπο MPEG-4 βασίζεται στην αποδεδειγμένη επιτυχία τριών πεδίων. Τα πεδία αυτά είναι

- Η ψηφιακή τηλεόραση.
- Οι διαδραστικές εφαρμογές γραφικών (συνθετικό περιεχόμενο).
- Διαδραστικά πολυμέσα (παγκόσμιος ιστός: διανομή και πρόσβαση σε περιεχόμενο).

Το MPEG-4 παρέχει όλα εκείνα τα τυποποιημένα τεχνολογικά στοιχεία που επιτρέπουν την ενσωμάτωση της παραγωγής, της διανομής και της πρόσβασης στο περιεχόμενο των τριών παραπάνω πεδίων.

Στο παρόν κεφάλαιο θα επιχειρηθεί μια επισκόπηση του προτύπου MPEG-4, με αναφορά στις διάφορες τεχνολογίες που περιλαμβάνει και τις εφαρμογές που υποστηρίζονται από αυτές [44]. Στη συνέχεια θα δοθεί έμφαση στις συνθετικές εικόνες και συγκεκριμένα στα συνθετικά πρόσωπα και τις παραμέτρους FAPs που σχετίζονται με τα μοντέλα συνθετικών προσώπων.

### 3.2 Τα χαρακτηριστικά και το πεδίο εφαρμογής του προτύπου MPEG-4

Το πρότυπο MPEG-4 παρέχει το σύνολο τεχνολογιών για να καλύψει τις ανάγκες των δημιουργών, παρόχων και τελικών χρηστών των εφαρμογών.

- Για τους **δημιουργούς εφαρμογών**, το MPEG-4 επιτρέπει την παραγωγή περιεχομένου που μπορεί να επαναχρησιμοποιηθεί πιο εύκολα και είναι πιο ευέλικτο από οποιαδήποτε άλλη επιμέρους τεχνολογία (ψηφιακή τηλεόραση,

παγκόσμιος ιστός www κλπ.). Επιπλέον, καθιστά πιο εύκολη τη διαχείριση και προστασία των δικαιωμάτων ιδιοκτησίας περιεχομένου.

- Για τους **παρόχους υπηρεσιών δικτύων**, το MPEG-4 προσφέρει διαφάνεια πληροφορίας, η οποία μπορεί να μεταφραστεί στα επιθυμητά μηνύματα σηματοδότησης κάθε δικτύου με τη βοήθεια των κατάλληλων προτύπων. Το MPEG-4 παρέχει έναν γενικά περιγραφέα ποιότητας υπηρεσίας QoS (Quality of Service). Η ακριβής μετάφραση των παραμέτρων του QoS στον αντίστοιχο περιγραφέα QoS του δικτύου είναι εκτός της εμβέλειας του MPEG-4 και εναπόκειται στους παρόχους υπηρεσιών δικτύου.
- Για τους **τελικούς χρήστες**, το MPEG-4 φέρει υψηλότερα επίπεδα αλληλεπίδρασης με το περιεχόμενο, εντός των ορίων που ορίζει ο δημιουργός του πολυμέσου. Επίσης επιτρέπει τη χρήση πολυμέσων σε νέα δίκτυα, συμπεριλαμβανομένων δικτύων με σχετικά χαμηλό ρυθμό μετάδοσης bitrate και δίκτυα κινητών επικοινωνιών.

Για όλες τις ομάδες χρηστών του, το MPEG-4 επιδιώκει να αποφευχθεί η πληθώρα ιδιοταγών, μη-συνεργαζόμενων φορμάτ και αναπαραγωγών πολυμέσων.

Για να επιτύχει τους παραπάνω στόχους, το MPEG-4 παρέχει τυποποιημένους τρόπους για τα εξής:

1. Αναπαράσταση μονάδων ακουστικού, οπτικού ή οπτικοακουστικού περιεχομένου, που καλούνται **«πολυμεσικά αντικείμενα»**. Τα πολυμεσικά αντικείμενα μπορεί να είναι φυσικής ή συνθετικής προέλευσης. Μπορεί να έχουν γραφτεί με μικρόφωνο ή κάμερα, ή μπορεί να έχουν συντεθεί με χρήση υπολογιστή.
2. Περιγραφή μιας σύνθεσης τέτοιων βασικών, πρωταρχικών αντικειμένων για τη δημιουργία **σύνθετων πολυμεσικών αντικειμένων** που σχηματίζουν **οπτικοακουστικές σκηνές**.
3. **Πολυπλεξία και συγχρονισμός δεδομένων** που σχετίζονται με τα πολυμεσικά αντικείμενα ώστε να μπορούν να μεταδοθούν από δίκτυα με περιγραφέα ποιότητας QoS κατάλληλο για τα συγκεκριμένα αντικείμενα.
4. **Αλληλεπίδραση του χρήστη** με την οπτικοακουστική σκηνή που δημιουργείται στην πλευρά του δέκτη.

Παρακάτω παρουσιάζονται κάποια από τα λειτουργικά χαρακτηριστικά του MPEG-4.

### 3.2.1 Κωδικοποιημένη αναπαράσταση πολυμεσικών αντικειμένων.

Οι οπτικοακουστικές σκηνές του MPEG-4 συντίθεται από διάφορα πολυμεσικά αντικείμενα που **οργανώνονται με ιεραρχικό τρόπο**. Στη βάση της ιεραρχίας βρίσκονται πρωταρχικά αντικείμενα (primitive objects) όπως:

- Στατικές εικόνες, για παράδειγμα ένα σταθερό φόντο
- Αντικείμενα βίντεο, π.χ. ένας ομιλητής χωρίς το φόντο
- Αντικείμενα ήχου, π.χ. η φωνή ενός ομιλητή, η μουσική υπόκρουση μιας σκηνής.

Το MPEG-4 τυποποιεί μια σειρά από τέτοια πρωταρχικά αντικείμενα που αναπαριστούν **φυσικό ή συνθετικό περιεχόμενο**, δύο ή τριών διαστάσεων. Επιπλέον, το MPEG-4 ορίζει την κωδικοποιημένη αναπαράσταση αντικειμένων όπως:

- Κείμενο και γραφικά
- Ομιλούντα συνθετικά πρόσωπα, κείμενο που χρησιμοποιείται για τη σύνθεση της ομιλίας και συνθετικά κινούμενα σώματα.
- Συνθετικός ήχος

Ένα πολυμεσικό αντικείμενο στην κωδικοποιημένη του μορφή αποτελείται από περιγραφικά στοιχεία που επιτρέπουν τη χρήση του αντικειμένου σε μια οπτικοακουστική σκηνή, καθώς και τη μετάδοσή του αν χρειαστεί. Σημειώνεται ότι ένα αντικείμενο όταν είναι κωδικοποιημένο μπορεί να αναπαρασταθεί ανεξάρτητα από το φόντο ή τον υπόλοιπο περίγυρό του. Η κωδικοποιημένη αναπαράσταση αντικειμένων είναι τόσο αποδοτική όσο επιτρέπουν οι πρόσθετες απαιτήσεις όπως είναι λόγου χάριν οι απαιτήσεις για ευρωστία σφάλματος, εύκολη εξόρυξη και επεξεργασία αντικειμένων, διαθεσιμότητα σε κλιμακούμενη μορφή.

### 3.2.2 Σύνθεση πολυμεσικών αντικειμένων

Μια οπτικοακουστική σκηνή του MPEG-4 συντίθεται από ένα πλήθος μεμονωμένων αντικειμένων. Δύο ή περισσότερα πρωταρχικά αντικείμενα μπορούν να ομαδοποιηθούν για να σχηματίσουν ένα **σύνθετο αντικείμενο**. Τα πρωταρχικά αντικείμενα αντιστοιχούν σε 'φύλλα' στο δέντρο περιγραφής ενώ τα σύνθετα αντικείμενα αντιστοιχούν σε ένα υπόδεντρο του δέντρου. Για παράδειγμα τα πρωταρχικά αντικείμενα της κινούμενης εικόνας ενός ομιλητή και της αντίστοιχης φωνής σχηματίζουν ένα νέο σύνθετο αντικείμενο που περιλαμβάνει την ακουστική και οπτική συνιστώσα του ομιλούντα.

Αυτή η **ομαδοποίηση των αντικειμένων** επιτρέπει στους συντάκτες να κατασκευάζουν περίπλοκες σκηνές και επιτρέπει στους τελικούς χρήστες να χειρίζονται αντικείμενα που συνδέονται νοηματικά σε μία οντότητα.

Γενικότερα, το MPEG-4 παρέχει έναν τυποποιημένο τρόπο για την περιγραφή μιας σκηνής με τις ακόλουθες δυνατότητες:

- Τοποθέτηση πολυμεσικών αντικειμένων οπουδήποτε σε ένα δοσμένο σύστημα αναφοράς.
- Εφαρμογή μετασχηματισμών για να αλλάξει η ακουστική ή γεωμετρική εμφάνιση ενός αντικειμένου.
- Ομαδοποίηση πρωταρχικών αντικειμένων σε νέα, σύνθετα αντικείμενα
- Εφαρμογή συνεχούς ροής δεδομένων σε πολυμεσικά αντικείμενα, ώστε να τροποποιούνται τα χαρακτηριστικά τους. Για παράδειγμα ένας ήχος, μια κινητή υφή που ανήκει σε ένα αντικείμενο, παράμετροι για την αναπαράσταση ενός συνθετικού προσώπου.
- Διαδραστική αλλαγή του σημείου θέασης ή ακρόασης του χρήστη σε οποιοδήποτε σημείο της σκηνής.

Η περιγραφή της σκηνής βασίζεται σε πολλές έννοιες της **γλώσσας VRML** (Virtual Reality Modelling Language) τόσο όσον αφορά τη δομή όσο και τη λειτουργικότητα των κόμβων/ σύνθετων αντικειμένων, και την επεκτείνει για να εκμεταλλευθεί πλήρως τα παραπάνω χαρακτηριστικά.

### 3.2.3 Περιγραφή, συγχρονισμός και μεταφορά πολυμεσικών αντικειμένων

Κάποια πολυμεσικά αντικείμενα μπορεί να χρειάζονται **δεδομένα συνεχούς ροής** (streaming data), τα οποία μεταφέρονται σε μία ή περισσότερες στοιχειώδεις ροές. Ένας περιγραφέας αντικειμένου αναγνωρίζει όλες τις ροές δεδομένων που συνδέονται με ένα αντικείμενο. Αυτό επιτρέπει τη διαχείριση ιεραρχικά κωδικοποιημένων δεδομένων, τη συσχέτιση μετα-δεδομένων σχετικά με το περιεχόμενο καθώς και τη συσχέτιση του αντικειμένου με τα δικαιώματα πνευματικής ιδιοκτησίας.

Κάθε ροή δεδομένων χαρακτηρίζεται από ένα **σύνολο περιγραφών** που φέρουν πληροφορίες διαμόρφωσης. Για παράδειγμα πληροφορίες για τον καθορισμό των χαρακτηριστικών του αποκωδικοποιητή, για την ακρίβεια πληροφοριών χρονισμού, πληροφορίες σχετικά με το απαιτούμενο QoS, όπως είναι μέγιστος ρυθμός μετάδοσης bit, ο ρυθμός σφάλματος, η προτεραιότητα, και άλλες παρεμφερείς απαιτήσεις.

Ο **συγχρονισμός στοιχειωδών ροών** επιτυγχάνεται μέσω της **χρονικής σήμανσης** (time stamping) μεμονωμένων μονάδων πρόσβασης που υπάρχουν στις στοιχειώδεις ροές. Το **επίπεδο συγχρονισμού** του MPEG-4 διαχειρίζεται τον εντοπισμό αυτών των μονάδων πρόσβασης και τη χρονική σήμανση. Ανεξάρτητα από τον τύπο του πολυμέσου, αυτό το επίπεδο επιτρέπει την αναγνώριση του τύπου της μονάδας πρόσβασης (καρέ βίντεο ή ήχου, εντολές περιγραφής σκηνής), την ανάκτηση της βάσης χρόνου και το συγχρονισμό μεταξύ τους. Η σύνταξη του επιπέδου συγχρονισμού μπορεί να διαμορφωθεί με πολλούς διαφορετικούς τρόπους ώστε να επιτρέπει τη χρήση ευρέως φάσματος συστημάτων.

Η συγχρονισμένη παράδοση δεδομένων συνεχούς ροής καθορίζεται από το **επίπεδο συγχρονισμού** και το **επίπεδο παράδοσης** που περιλαμβάνει έναν **πολυπλέκτη δύο επιπέδων**.

Το πρώτο από τα δύο επίπεδα πολύπλεξης ελέγχεται σύμφωνα με τις προδιαγραφές **DFIM** (Delivery Multimedia Integration Framework) του MPEG-4. Η πολύπλεξη μπορεί να πραγματοποιείται με το **εργαλείο FlexMux** του MPEG-4, η χρήση του ωστόσο είναι προαιρετική. Το δεύτερο επίπεδο πολύπλεξης καλείται **TransMux** (Transport Multiplexing) και έχει να κάνει με τις υπηρεσίες μεταφοράς. Το MPEG-4 καθορίζει μόνο τη διεπαφή του επιπέδου TransMux. Η επιλογή του πρωτοκόλλου μεταφοράς επαφίεται στον τελικό χρήστη ή τον πάροχο της υπηρεσίας μεταφοράς. Έτσι το MPEG-4 μπορεί να χρησιμοποιηθεί σε πολλά διαφορετικά περιβάλλοντα. Το επίπεδο πολύπλεξης FlexMux μπορεί να είναι κενό εάν το επίπεδο TransMux παρέχει όλα τα απαραίτητα λειτουργικά χαρακτηριστικά. Το επίπεδο συγχρονισμού απεναντίας είναι πάντα παρόν.

### 3.2.4 Αλληλεπίδραση του χρήστη με τα πολυμεσικά αντικείμενα

Στη γενική περίπτωση, ο χρήστης παρακολουθεί μια σκηνή που συντίθεται σύμφωνα με τη σχεδίαση του δημιουργού/συντάκτη της σκηνής. Ανάλογα με το βαθμό ελευθερίας που έχει επιτρέψει ο συντάκτης, ο χρήστης έχει τη **δυνατότητα αλληλεπίδρασης με το διαδραστικό περιεχόμενο της σκηνής**. Κάποιες από τις λειτουργίες που είναι δυνατόν να επιτρέπονται στο χρήστη περιλαμβάνουν τις εξής:

- Αλλαγή του σημείου θέασης ή ακρόασης της σκηνής, μέσω της πλοήγησης σε αυτήν
- Μετακίνηση αντικειμένων της σκηνής σε διαφορετικές θέσεις
- Έναρξη μιας λειτουργίας, για παράδειγμα κάνοντας κλικ σε ένα συγκεκριμένο αντικείμενο μπορεί να αρχίζει ή να σταματά η ροή βίντεο.
- Επιλογή της επιθυμητής γλώσσας όταν διατίθενται αρχεία ήχου σε διάφορες γλώσσες.

Το MPEG-4 υποστηρίζει και πιο πολύπλοκες λειτουργίες, όπως για παράδειγμα ένα εικονικό τηλέφωνο που χτυπά, ο χρήστης απαντά και εγκαθίσταται ένας διάυλος επικοινωνίας.

### 3.2.5 Αναγνώριση και διαχείριση της πνευματικής ιδιοκτησίας

Η ομάδα εργασίας του MPEG-4 έκρινε σημαντικό να υπάρχει η δυνατότητα για αναγνώριση της πνευματικής ιδιοκτησίας στα πολυμεσικά αντικείμενα του MPEG-4. Για το λόγο αυτό συνεργάστηκαν με εκπρόσωπους διαφόρων δημιουργικών κλάδων για τον ορισμό της σύνταξης και των εργαλείων που θα υποστήριζαν το στόχο τους.

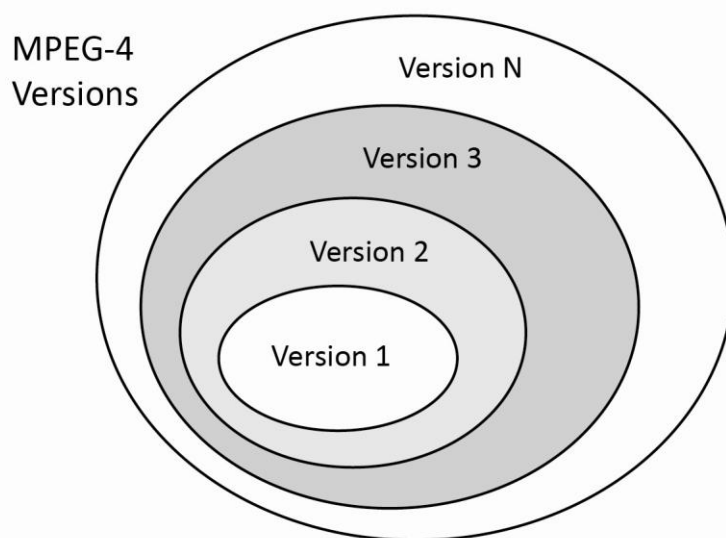
Το MPEG-4 ενσωματώνει την αναγνώριση της πνευματικής ιδιοκτησίας μέσω της **αποθήκευσης μοναδικών αναγνωριστικών** που έχουν εκδοθεί από διεθνή συστήματα αρίθμησης όπως για παράδειγμα το ISAN (International Standard Audiovisual Number), ISRC (International Standard Recording Code) κλπ. Αυτοί οι αριθμοί μπορούν να χρησιμοποιηθούν για να αναγνωριστεί ο τρέχων κάτοχος των δικαιωμάτων του πολυμεσικού αντικειμένου. Επειδή δεν αναγνωρίζεται όλο το περιεχόμενο από έναν τέτοιο αριθμό, η έκδοση 1 του MPEG-4 προσφέρει τη **δυνατότητα της αναγνώρισης μέσω ενός ζεύγους ‘ετικέτα-τιμή’** όπως για παράδειγμα »composer«/»John Smith«. Επιπρόσθετα, το MPEG-4 προσφέρει ένα τυποποιημένο περιβάλλον διεπαφής που είναι ενσωματωμένο στο επίπεδο των συστημάτων (systems layer) και απευθύνεται σε χρήστες που θέλουν να χρησιμοποιήσουν **συστήματα με ελεγχόμενη πρόσβαση** σε αντικείμενα πνευματικής ιδιοκτησίας. Με αυτή τη διεπαφή, τα ιδιοταγή συστήματα ελέγχου μπορούν εύκολα να συγχωνευθούν με το τυποποιημένο τμήμα του αποκωδικοποιητή.

### 3.3 Εκδόσεις του MPEG-4

Όπως αναφέρθηκε και στην εισαγωγή, η **πρώτη έκδοση** του MPEG-4 εγκρίθηκε στα τέλη του 1998 και η **δεύτερη έκδοση** του MPEG-4 αναγνωρίστηκε στις **αρχές του 2000**. Μετά από αυτές τις δύο εκδόσεις (versions) προστέθηκαν ακόμη πολλά εργαλεία σε μεταγενέστερες τροποποιήσεις που θα μπορούσαν να έχουν χαρακτηριστεί ως εκδόσεις, ακόμη και αν σε μερικές περιπτώσεις ήταν δυσδιάκριτο αν έπρεπε να γίνει λόγος για καινούρια έκδοση. Έκτοτε έγιναν πολλές προσθήκες στο MPEG-4, τα λεγόμενα **προφίλ**, και ο χαρακτηρισμός 'έκδοση' για ολόκληρο το πρότυπο άρχισε να εγκαταλείπεται.

Πλέον δε δίνεται έμφαση στην αναγνώριση καινούριων εκδόσεων, αλλά είναι πιο σημαντικό να γίνεται διάκριση ανάμεσα στα προφίλ του MPEG-4 (profiles). Υπαρκτά εργαλεία και προφίλ από οποιαδήποτε έκδοση δεν αντικαθίστανται ποτέ σε μεταγενέστερες εκδόσεις, αλλά η **προσθήκη νέων τεχνολογιών** στο MPEG-4 πραγματοποιείται με τη **προσθήκη νέων προφίλ**.

Παρακάτω στο Σχήμα 3-1 Εκδόσεις του MPEG-4 απεικονίζεται η σχέση ανάμεσα στις εκδόσεις (τμημάτων) του MPEG-4. Η έκδοση 2 έχει οπισθόδρομη συμβατότητα (backward compatibility) με την έκδοση 1, και αποτελεί επέκτασή της. Παρομοίως η έκδοση 3 είναι οπισθόδρομα συμβατή επέκταση της έκδοσης 2, και ούτω καθεξής. Οι εκδόσεις των σημαντικότερων τμημάτων (Systems, Audio, Video, DMIF) του MPEG-4 συγχρονίστηκαν στη 2<sup>η</sup> έκδοση και έκτοτε τα διάφορα μέρη του προτύπου ακολουθούν τη δική τους πορεία όσον αφορά στις προσθήκες, τις βελτιώσεις και τις επεκτάσεις σε καινούριες εκδόσεις.



Σχήμα 3-1 Εκδόσεις του MPEG-4 [44]

Το συστημικό επίπεδο (systems layer) του επιπέδου N είναι **οπισθόδρομα συμβατό** με όλες τις προγενέστερες εκδόσεις. Στους τομείς των συστημάτων, ήχου και εικόνας, **κάθε νέα έκδοση προσθέτει** στην ουσία νέα προφίλ στα ήδη υπάρχοντα, **χωρίς να τα τροποποιεί**. Είναι αξιοσημείωτο ότι τα υπάρχοντα συστήματα θα παραμείνουν πάντα συμβατά, επειδή τα προφίλ δεν αλλάζουν ποτέ αναδρομικά, ούτε πρόκειται ποτέ να αλλάξει η σύνταξη των συστημάτων με τρόπο που θα είναι οπισθόδρομα ασύμβατος.

### 3.4 Προφίλ και επίπεδα του MPEG-4

Το MPEG-4 προσφέρει ένα ευρύ και πλούσιο σύνολο εργαλείων για την κωδικοποίηση οπτικοακουστικών αντικειμένων. Προκειμένου να καταστούν δυνατές οι αποτελεσματικές υλοποιήσεις του προτύπου, έχουν αναγνωριστεί υποσύνολα των συνολικών εργαλείων των συστημάτων, ήχου και εικόνας για να χρησιμοποιηθούν σε συγκεκριμένες εφαρμογές. Αυτά τα υποσύνολα εργαλείων καλούνται **προφίλ του MPEG-4**, και περιορίζουν τα εργαλεία που χρειάζεται να υλοποιηθούν από τον αποκωδικοποιητή.

Για κάθε προφίλ, έχουν τεθεί ένα ή περισσότερα **επίπεδα** (levels), ώστε να **περιορίζεται η υπολογιστική πολυπλοκότητα**. Το επίπεδο ορίζει πληροφορίες και τεχνικές προδιαγραφές όπως ο μέγιστος ρυθμός μετάδοσης, μέγιστες διαστάσεις καρέ για συγκεκριμένη συχνότητα, μέγεθος buffer κλπ. Έτσι διατίθενται πολλοί συνδυασμοί προφίλ/επίπεδου, μια προσέγγιση παραπλήσια αυτή του MPEG-2, όπου ο πιο γνωστός συνδυασμός προφίλ/επίπεδου ήταν 'Main Profile @ Main Level'.

Οι συνδυασμοί 'Profile@Level':

- Επιτρέπουν τον έλεγχο συμμόρφωσης, δηλαδή ελέγχουν κατά πόσο συμμορφώνονται διάφορες συσκευές με το πρότυπο MPEG-4.
- Επιτρέπουν στον σχεδιαστή ενός κωδικοποιητή/αποκωδικοποιητή (codec) να υλοποιήσει μόνο το υποσύνολο του προτύπου που χρειάζεται, διατηρώντας ταυτόχρονα τη διασυνδεσιμότητα με συσκευές που σχεδιάστηκαν με τον ίδιο συνδυασμό.

Υπάρχουν προφίλ για διάφορα είδη πολυμεσικού περιεχομένου, δηλαδή ακουστικού, οπτικού περιεχομένου και γραφικών. Κάποια προφίλ ορίστηκαν στην έκδοση 1, κάποια προστέθηκαν με την έκδοση 2 και τα υπόλοιπα είναι μεταγενέστερα. Η ομάδα εργασίας του MPEG δεν υπαγορεύει ούτε προτείνει συγκεκριμένους συνδυασμούς των προφίλ, έχει ληφθεί όμως μέριμνα ώστε να υπάρχουν επιτυχημένοι συνδυασμοί ανάμεσα στα διάφορα είδη.

Τα προφίλ του MPEG-4 ανήκουν ως επί το πλείστον στις εξής κατηγορίες, χωρίς να είναι απαραίτητα εξαντλητική η παρακάτω λίστα:

- **Προφίλ εικόνας** (Visual profiles). Το οπτικό τμήμα του MPEG-4 περιέχει προφίλ για την κωδικοποίηση φυσικού, συνθετικού οπτικού υλικού ή υβριδικό περιεχόμενο, δηλαδή συνδυασμός φυσικού και συνθετικού περιεχομένου.
- **Προφίλ ήχου** (Audio profiles). Στα προφίλ αυτά περιλαμβάνονται προφίλ για ομιλία, μουσική, φυσικούς ή συνθετικούς ήχους, διεπαφή text-to-speech.
- **Προφίλ γραφικών** (Graphics profiles). Περιλαμβάνονται προφίλ για απλά ή σύνθετα 2D γραφικά, προφίλ για τον ήχο μια σκηνής γραφικών και άλλα σχετικά προφίλ.
- **Προφίλ περιγραφής σκηνών** (Scene Graph Profiles/Scene Description Profiles). Τα προφίλ αυτά χρησιμεύουν για οπτικοακουστικές σκηνές με αποκλειστικά ήχο (όπως το διαδικτυακό ραδιόφωνο ή σκηνές με 3D πηγές ήχων) σκηνές με περιεχόμενο 2D, 3D, ή ανάμεικτο 2D/3D (τηλεοπτική αναμετάδοση, gaming, 3D εικονικοί κόσμοι)

- **Προφίλ MPEG-J** (MPEG-Java profile). Για φορητές συσκευές, συσκευές κινητών επικοινωνιών, PDAs κλπ.
- **Προφίλ περιγραφέα αντικειμένων** (Object descriptor profile)

Ως παράδειγμα, δίνονται τα προφίλ για το αμιγώς συνθετικό ή υβριδικό (φυσικό/συνθετικό) περιεχόμενο της έκδοσης 1 στον παρακάτω πίνακα:

| Όνομα προφίλ                              | Περιγραφή   |
|---|---|
| Simple Facial Animation Visual Profile    | Απλό προφίλ για animation προσώπου, κατάλληλο για εφαρμογές όπως το βίντεο για άτομα με προβλήματα ακοής.   |
| Scalable Texture Visual Profile           | Παρέχει κωδικοποιημένες στατικές εικόνες με δυνατότητα κλιμάκωσης, χρήσιμο για εφαρμογές που χρειάζονται πολλαπλά επίπεδα κλιμάκωσης (π.χ. υφή σε αντικείμενα παιχνιδιών), υψηλής ανάλυσης ψηφιακές φωτογραφικές μηχανές  |
| Basic Animated 2-D Texture Visual Profile | Παρέχει χωρική κλιμάκωση, SNR κλιμάκωση, mesh-based animation για αντικείμενα στατικής εικόνας (υφής), απλό animation αντικειμένων προσώπων   |
| Hybrid Visual Profile                     | Συνδυάζει τη δυνατότητα αποκωδικοποίησης φυσικών αντικειμένων (αυθαίρετου σχήματος, κλιμακούμενα χρονικά) με τη δυνατότητα αποκωδικοποίησης συνθετικών και υβριδικών αντικειμένων, συμπεριλαμβανομένου αντικειμένων προσώπου και αντικειμένων στατικής εικόνας. Κατάλληλο για διάφορες πολυμεσικές εφαρμογές. |

**Πίνακας 3-1** Παράδειγμα οπτικών προφίλ του MPEG-4 [44]

## 3.5 Τμήματα του MPEG-4

Το πρότυπο MPEG-4, λόγω της έκτασής του, μπορεί να θεωρηθεί ότι αποτελείται από πληθώρα προτύπων, τα **τμήματα** (parts). Κάθε τμήμα καλύπτει μια συγκεκριμένη πτυχή ολόκληρου του προτύπου. Τα τμήματα παρουσιάζονται στον παρακάτω πίνακα.



| Αρίθμηση τμήματος | Ονομασία τμήματος   | Περιγραφή τμήματος  |
|-------------------|---|---|
| Part 1            | Systems   | Συγχρονισμός και πολυπλεξία εικόνας και ήχου  |
| Part 2            | Visual  | Πρότυπο για οπτικές πληροφορίες   |
| Part 3            | Audio   | Πρότυπο για ηχητικές πληροφορίες  |
| Part 4            | Conformance testing   | Έλεγχος συμμόρφωσης μεταξύ τμημάτων   |
| Part 5            | Reference software  | Λογισμικό αναφοράς για επίδειξη και αποσαφήνιση τμημάτων  |
| Part 6            | Delivery Multimedia Integration Framework (DMIF)  | Πρότυπο για τη διεπαφή μεταξύ εφαρμογής και διαύλου μεταφοράς   |
| Part 7            | Optimized reference software for coding of audio-visual objects                         | Παρέχει παραδείγματα για βελτιστοποιημένες εφαρμογές (αναφορά στο τμήμα 5)  |
| Part 8            | Carriage of ISO/IEC 14496 contents over IP networks                                     | Μεταφορά περιεχομένου MPEG-4 σε IP δίκτυα   |
| Part 9            | Reference hardware description  | Σχέδια υλικού για την υλοποίηση των υπόλοιπων τμημάτων του προτύπου   |
| Part 10           | Advanced Video Coding (AVC)   | Πρότυπο συμπίεσης βίντεο τεχνικά ισοδύναμο με το ITU-T H.264  |
| Part 11           | Scene description and application engine  | Διαδραστικό περιεχόμενο με πολλαπλά προφίλ, συμπεριλαμβανομένου 2D/3D εκδόσεων.                                   |
| Part 12           | ISO base media file format  | Φορμάτ για αποθήκευση time-based πολυμεσικού περιεχομένου.  |
| Part 13           | Intellectual Property Management and Protection (IPMP) extensions                       | Διαχείριση και προστασία δικαιωμάτων πνευματικής ιδιοκτησίας  |
| Part 14           | MP4 file format   | Το φορμάτ του αρχείου με MPEG-4 περιεχόμενο. Βασισμένο στο τμήμα 12   |
| Part 15           | Advanced Video Coding (AVC) file format   | Για αποθήκευση βίντεο του τμήματος 10   |
| Part 16           | Animation Framework eXtension (AFX)   | Αναπαράσταση περιεχομένου 3D γραφικών   |
| Part 17           | Streaming text format   | Φορμάτ κειμένων υπότιτλων   |
| Part 18           | Font compression and streaming  | Για τη γραμματοσειρά του τμήματος 22  |
| Part 19           | Synthesized texture stream  | Δημιουργία συνθετικών βίντεο κλιπ πολύ χαμηλού bit rate   |
| Part 20           | Lightweight Application Scene Representation (LSeR) and Simple Aggregation Format (SAF) | Μετάδοση και αναπαράσταση πολυμεσικών εφαρμογών σε συσκευές περιορισμένων δυνατοτήτων όπως τα κινητά τηλέφωνα     |
| Part 21           | MPEG-J Graphics Framework eXtensions (GFX)  | Ελαφρύ προγραμματιστικό περιβάλλον για προηγμένες διαδραστικές εφαρμογές πολυμέσων. Συνδυασμός MPEG-J με Java API |
| Part 22           | Open Font Format  | Γραμματοσειρά φορμάτ OpenType version 1.4   |
| Part 23           | Symbolic Music Representation (SMR)   | Μοντελοποιεί τα οπτικά και ακουστικά χαρακτηριστικά της μουσικής σημειογραφίας (π.χ. παρτιτούρας)                 |
| Part 24           | Audio and systems interaction   | Επιθυμητή συμπεριφορά συνδυασμού MPEG-4 File Format και MPEG-4 Audio.   |
| Part 25           | 3D Graphics Compression Model   | Συνδέει τα εργαλεία συμπίεσης 3D γραφικών του MPEG-4 με τα πρωταρχικά γραφικά άλλων προτύπων                      |
| Part 26           | Audio conformance   | Συμμόρφωση ηχητικού περιεχομένου  |
| Part 27           | 3D Graphics conformance   | Συμμόρφωση 3D γραφικών  |
| Part 28           | Composite font representation   | Αναπαράσταση σύνθετων γραμματοσειρών  |
| Part 29           | Web video coding  | Κωδικοποίηση βίντεο διαδικτύου  |

**Πίνακας 3-2** Τμήματα του MPEG-4 [45], [46]

## 3.6 Συνθετικά πρόσωπα στο MPEG-4

Τα συνθετικά αντικείμενα αποτελούν υποσύνολο του γενικότερου κλάδου της γραφικής με υπολογιστές. Το MPEG-4 υποστηρίζει τα ακόλουθα συνθετικά αντικείμενα [44]:

- Παραμετρική περιγραφή συνθετικών προσώπων και σώματος
- Παραμετρική περιγραφή στατικής και δυναμικής κωδικοποίησης πλέγματος (mesh coding) με απεικόνιση υφής
- Κωδικοποίηση υφής εξαρτώμενη από το σημείο θέασης (view dependent).

Το παρόν υποκεφάλαιο εστιάζει στα συνθετικά πρόσωπα, δηλαδή τα αντικείμενα ‘facial animation objects’ του MPEG-4.

### 3.6.1 Γενικά στοιχεία για τα συνθετικά πρόσωπα του MPEG-4

Η ομάδα του MPEG προέβλεψε για την **αυξανόμενη χρήση συνθετικών ομιλούντων προσώπων**, που αναμένεται να χρησιμοποιηθούν συχνότερα στο μέλλον σε διαδικτυακές εφαρμογές εξυπηρέτησης πελατών και στα παιχνίδια γραφικών. Για το σκοπό αυτό, το MPEG-4 επιτρέπει την ενσωμάτωση του animation προσώπων σε πολυμεσικές επικοινωνίες, παρουσιάσεις, και προβλέπει τη δυνατότητα για animation προσώπων ακόμα και σε διαύλους επικοινωνίας χαμηλού ρυθμού μεταφοράς bit [47].

Το MPEG-4 ορίζει ένα **γενικό μοντέλο προσώπου σε ουδέτερη κατάσταση** και ένα πλήθος σημείων ενδιαφέροντος πάνω στο ουδέτερο πρόσωπο που λειτουργούν ως σημεία αναφοράς. Επίσης ορίζεται ένα σύνολο από **παραμέτρους για την κίνηση του προσώπου**, που λέγονται **FAPs** (Facial Animation Parameters) και αντιστοιχούν η καθεμία σε μια συγκεκριμένη δράση του προσώπου, που παραμορφώνει το μοντέλο προσώπου της ουδέτερης κατάστασης. Το σχήμα, το χρώμα, η υφή και η έκφραση του κάθε προσώπου ελέγχεται από τις **παραμέτρους ορισμού προσώπου FDPs** (Facial Definition Parameters) και τις παραμέτρους FAPs. Εάν είναι επιθυμητό, μπορεί να οριστεί εξολοκλήρου από το μηδέν ένα μοντέλο προσώπου μέσω του συνόλου των παραμέτρων FDP.

Η αλληλουχία των παραμορφώσεων μέσω FAPs παράγει μια ακολουθία animation που αναπαριστά κινήσεις, εκφράσεις και ομιλία. Η τιμή μιας παραμέτρου FAP υποδεικνύει το μέγεθος της αντίστοιχης κίνησης, για παράδειγμα μικρό ή μεγάλο άνοιγμα του στόματος σε ένα χαμόγελο. Για να είναι δυνατή η ερμηνεία των τιμών των FAPs από ένα τερματικό τεχνολογίας MPEG-4, πρέπει να υπάρχουν προκαθορισμένοι κανόνες animation, συγκεκριμένοι για το κάθε μοντέλο, που παράγουν τη ζητούμενη κίνηση που αντιστοιχεί στο FAP. Το τερματικό μπορεί να χρησιμοποιήσει είτε τους δικούς του κανόνες ή να λάβει ένα μοντέλο προσώπου μαζί με τους **κανόνες animation οργανωμένους σε πίνακες FAT** (Face Animation Table) για προσαρμοσμένη συμπεριφορά. Καθώς τα FAPs χρησιμοποιούνται για πρόσωπα διαφορετικών διαστάσεων και αναλογιών, υπάρχει η ανάγκη για **ορισμό μονάδων FAPs**, τα λεγόμενα **FAPU**. Στις επόμενες υποπαραγράφους δίνονται περισσότερες λεπτομέρειες για τις άνωθι έννοιες.

Η αναπαράσταση συνθετικών προσώπων στην έκδοση 1 του MPEG-4 παρέχει έναν αποδοτικό τρόπο για την κωδικοποίηση των παραμέτρων σύνθεσης και υποστηρίζει ένα πολύ μεγάλο εύρος μοντέλων προσώπων. Τα **μοντέλα προσώπων** αυτά καθεαυτά όπως εξηγήθηκε **δεν είναι τυποποιημένα**, αλλά **τα εργαλεία για την περιγραφή τους** (παράμετροι FDP και FAP) **είναι τυποποιημένα**. Μπορεί να χρησιμοποιηθεί χρονικός διακριτός μετασχηματισμός συνημιτόνου για τη συμπίεση ενός μεγάλου πλήθους FAPs όταν χρειάζεται να αποδοθεί η ακριβής άρθρωση της ομιλίας.

Το **σύστημα BIFS** (Systems Binary Format for Scenes) διαθέτει κάποιες δυνατότητες για να υποστηρίξει εφαρμογές όπου χρειάζονται προσαρμοσμένα μοντέλα προσώπου και εξειδικευμένη ερμηνεία των FAPs. Τα χαρακτηριστικά αυτά περιλαμβάνουν [44]:

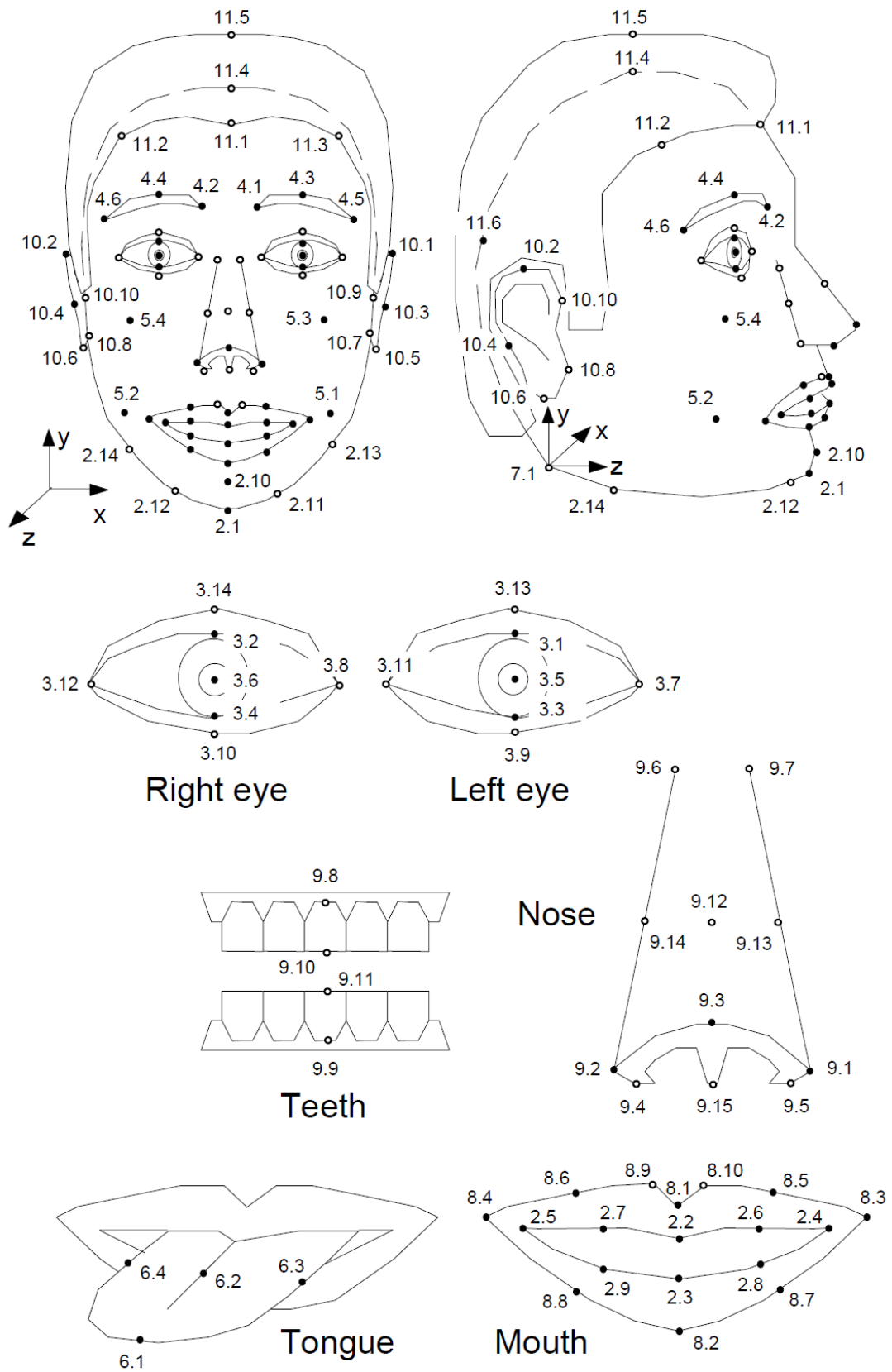
1. Τις **παραμέτρους FDPs** του συστήματος BIFS. Πρόκειται για σύνολο πληροφοριών που μπορεί να ανακτηθεί για τη διαμόρφωση του γενικού προσώπου στο τερματικό, ή για την εγκατάσταση ενός συγκεκριμένου μοντέλου προσώπου στην αρχή μιας εφαρμογής.
2. Τον **πίνακα FAT** (Face Animation Table) που χαρτογραφεί τα FAPs σε κινήσεις των σημείων του μοντέλου προσώπου. Το τερματικό μπορεί να χρησιμοποιήσει τους δικούς του κανόνες για το animation, ή να κατεβάσει έναν πίνακα FAT για προσαρμοσμένο animation.
3. **Τεχνική παρεμβολής FIT** (Face Interpolation Technique) χαρτογραφεί ένα υποσύνολο FAPs σε ολόκληρο το σύνολο, με τεχνικές παρεμβολής για τα FAPs που απουσιάζουν από το υποσύνολο εισόδου. Χρησιμοποιείται επίσης για να υπολογιστεί η συνδυασμένη επίδραση πολλών FAPs.

### 3.6.2 Μοντέλο προσώπου του MPEG-4 στην ουδέτερη κατάσταση

Ως πρώτο βήμα, το πρότυπο MPEG-4 ορίζει ένα **γενικό μοντέλο προσώπου σε ουδέτερη κατάσταση έκφρασης**. Το ουδέτερο πρόσωπο έχει τα εξής χαρακτηριστικά:

- Το βλέμμα έχει οριζόντια διεύθυνση, κατεύθυνση προς τα έξω, ακολουθεί δηλαδή την κατεύθυνση του z άξονα του σχήματος 3-3.
- Όλοι οι μύες του προσώπου είναι χαλαροί.
- Οι άκρες των βλεφάρων εφάπτονται στην ίριδα.
- Η κόρη του ματιού έχει διάμετρο ίση με το ένα τρίτο της διαμέτρου της ίριδας.
- Τα χείλη ακουμπούν, η γραμμή των χειλιών είναι οριζόντια και στο ίδιο ύψος με τις άκρες των χειλιών.
- Το στόμα είναι κλειστό και η επάνω οδοντοστοιχία ακουμπά την κάτω οδοντοστοιχία.
- Η γλώσσα είναι επίπεδη, οριζόντια, η άκρη της γλώσσας ακουμπά το όριο όπου εφάπτονται οι δύο οδοντοστοιχίες.

Όλα τα FAPs εκφράζονται ως μετατοπίσεις από τις θέσεις που ορίζονται στο ουδέτερο πρόσωπο. Το ουδέτερο πρόσωπο δίνεται και από το σχήμα 3-3: Αποστάσεις για τον ορισμό των FAPU.



**Σχήμα 3-2** Χαρακτηριστικά σημεία στο μοντέλο προσώπου του MPEG-4 [47]

### 3.6.3 Χαρακτηριστικά σημεία του μοντέλου προσώπου

Το MPEG-4 ορίζει **84 χαρακτηριστικά σημεία** στο ουδέτερο πρόσωπο, τα οποία φαίνονται στο σχήμα 3-2. Ο βασικός σκοπός των σημείων είναι η **παροχή μιας χωρικής αναφοράς** κατά τον ορισμό των FAPs. Κάποια σημεία όπως αυτά στη γραμμή των μαλλιών δεν επηρεάζονται από τις τιμές των FAPs. Απαιτούνται όμως για τον ορισμό του σχήματος ενός ιδιοταγούς μοντέλου προσώπου. Τα σημεία ενδιαφέροντος είναι **οργανωμένα σε ομάδες** ανάλογα με το χαρακτηριστικό προσώπου στο οποίο ανήκουν, για παράδειγμα τα μάγουλα, τα μάτια, το στόμα. Η τοποθεσία αυτών των σημείων πρέπει να είναι γνωστή για κάθε μοντέλο προσώπου που είναι συμβατό με το MPEG-4. Για την εύρεση της θέσης και των συντεταγμένων πολλών σημείων υπάρχουν κάποιες υποδείξεις με τη μορφή περιορισμών [47].

Για παράδειγμα, αν χρησιμοποιηθούν οι κωδικοί του σχήματος 3-2, το σημείο 3.1 που είναι το κέντρο του επάνω εσωτερικού βλεφάρου, έχει συντεταγμένη  $x$  με προτεινόμενο περιορισμό:

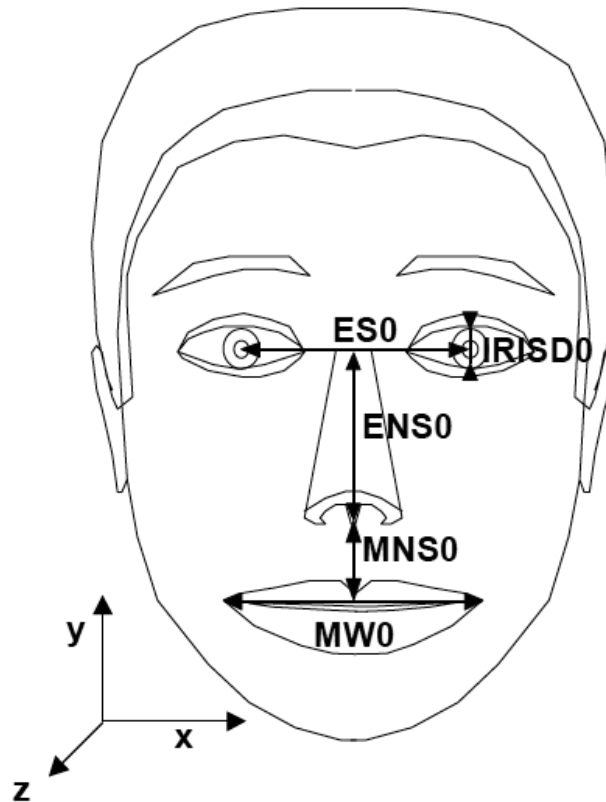
$$3.1.x = \frac{3.7.x + 3.11.x}{2}$$

Όπου τα σημεία 3.7 και 3.11 είναι τα άκρα του αριστερού ματιού (αριστερά κατά την αντίληψη του εικονιζόμενου προσώπου). Η παράθεση της κατάληξης  $.x$  μετά από τον κωδικό ενός σημείου συμβολίζει τη  $x$  συντεταγμένη. Παρομοίως συμβολίζεται και η  $y$  συντεταγμένη.

Οι περισσότεροι **προτεινόμενοι περιορισμοί συντεταγμένων** είναι για την  $x$  συντεταγμένη. Είτε την εξισώσουν με τη  $x$  συντεταγμένη ενός άλλου σημείου, με το μέσο όρο των συντ/νων δύο άλλων σημείων, δίνουν κάποιο διάστημα στο οποίο πρέπει να ανήκει με τη μορφή δύο ανισοτήτων, ή δίνουν μια απλή ανισότητα. Παρομοίως για την  $y$  συντεταγμένη, παρότι οι προτεινόμενοι περιορισμοί είναι πολύ λιγότεροι. Μόνο 4 σημεία έχουν περιορισμό στη  $z$  συντεταγμένη. Υπάρχουν σημεία στα οποία δεν τίθεται κανένας περιορισμός σε καμία συντεταγμένη (π.χ. 10.5 και 10.6, άκρα των λοβών των αυτιών) αλλά υπάρχουν και σημεία με περιορισμούς για δύο συντεταγμένες (π.χ. τα σημεία 5.3, 5.4, στα μάγουλα που έχουν περιορισμούς σε  $x$  και  $y$  συντεταγμένες). Ο ενδιαφερόμενος αναγνώστης μπορεί να ανατρέξει στο [47] για πίνακα που περιέχει όλους τους προτεινόμενους περιορισμούς.

### 3.6.4 Μονάδες των FAPs

Για να είναι εφικτός ο **ορισμός παραμέτρων FAPs για τυχαία μοντέλα προσώπων**, το MPEG-4 ορίζει μονάδες των FAPs που λέγονται FAPU (FAP Units) και χρησιμοποιούνται για την **κλιμάκωση των FAPs για οποιοδήποτε πρόσωπο**. Τα FAPUs ορίζονται ως κλάσματα αποστάσεων σημαντικών σημείων στο πρόσωπο. Οι εν λόγω αποστάσεις ορίζονται στην ουδέτερη κατάσταση ενός μοντέλου προσώπου. Η χρήση των μονάδων FAPU επιτρέπει την ερμηνεία ενός FAP σε οποιοδήποτε μοντέλο προσώπου με συνεπή τρόπο, για την παραγωγή εύλογων αποτελεσμάτων ως προς την έκφραση και την προφορά της ομιλίας.



**Σχήμα 3-3** Αποστάσεις για τον ορισμό των FAPU [47]

*Το πρόσωπο βρίσκεται σε ουδέτερη κατάσταση σύμφωνα με το MPEG-4.  
Τα FAPU υπολογίζονται ως κλάσματα των αποστάσεων του σχήματος.*

Οι πληροφορίες για τις μονάδες FAPU δίνονται στον παρακάτω πίνακα.

| <b>Πίνακας 3-3</b> Ορισμοί Μονάδων FAP του MPEG-4 (FAPU)  |  |  |  |
|---|--|--|--|
| Ονομασία και Περιγραφή  |  | Τρόπος υπολογισμού<br>(ουδέτερο καρέ) [48] | Κανονικοποιημένη μονάδα FAPU<br>Normalization factor n.f.=1024 |
| IRISD0  | Διάμετρος ίριδας<br>(Iris diameter)                | $3.1.y - 3.3.y = 3.2.y - 3.4.y$            | $IRISD = IRISD0 / n.f.$  |
| ES0   | Απόσταση ματιών<br>(Eye separation)                | $3.5.x - 3.6.x$                            | $ES = ES0 / n.f.$  |
| ENS0  | Απόσταση ματιών-μύτης<br>(Eye-nose separation)     | $3.5.y - 9.15.y$                           | $ENS = ENS0 / n.f.$  |
| MNS0  | Απόσταση μύτης-στόματος<br>(Mouth-nose separation) | $9.15.y - 2.2.y$                           | $MNS = MNS0 / n.f.$  |
| MW0   | Μήκος στόματος<br>(Mouth width)                    | $8.3.x - 8.4.x$                            | $MW = MW0 / n.f.$  |
| AU  | Μονάδα γωνίας<br>(Angle unit)                      | $10^{-5} \text{ rad}$                      | $10^{-5} \text{ rad}$  |
| <i>Εξ'ορισμού η απόσταση IRISD0, η διάμετρος της ίριδας, είναι ίση με την απόσταση ανάμεσα σε κάτω και πάνω βλέφαρο του ματιού στην ουδέτερη κατάσταση.</i> |  |  |  |

### 3.6.5 Ομαδοποίηση των παραμέτρων FAP του MPEG-4

Οι παράμετροι FAP, η λειτουργία των οποίων ήδη αναφέρθηκε παραπάνω, βασίζονται στις μελέτες των **ελάχιστων αισθητών δράσεων** και συνδέονται άμεσα με μυικές κινήσεις.

Οι **68 παράμετροι FAPs** κατηγοριοποιούνται σε **10 ομάδες** που σχετίζονται με μέρη ή χαρακτηριστικά του προσώπου [47].

| Group No. | Πλήθος FAPs στην ομάδα | Group Name                                    | Περιγραφή ομάδας   |
|-----------|------------------------|---|--|
| 1         | 2                      | Visemes and expressions                       | Οπτική αναπαράσταση φωνημάτων και αρχετυπικές εκφράσεις                |
| 2         | 16                     | Jaw, chin, inner lowerlip, cornerlips, midlip | Σαγόني, πηγούνι, εσωτερικό κάτω χείλος, γωνίες στόματος, μέσον χειλιών |
| 3         | 12                     | Eyeballs, pupils, eyelids                     | Μάτια (βολβοί, κόρες ματιών), βλέφαρα                                  |
| 4         | 8                      | Eyebrows                                      | Φρύδια   |
| 5         | 4                      | Cheeks  | Μάγουλα  |
| 6         | 5                      | Tongue  | Γλώσσα   |
| 7         | 3                      | Head rotation                                 | Περιστροφή κεφαλιού  |
| 8         | 10                     | Outer lip positions                           | Εξωτερικό περίγραμμα χειλιών   |
| 9         | 4                      | Nose  | Μύτη   |
| 10        | 4                      | Ears  | Αυτιά  |

**Πίνακας 3-4** Ομάδες των FAPs στο MPEG-4 [47]

Τα FAPs αντιπροσωπεύουν ένα **πλήρες σύνολο βασικών κινήσεων του προσώπου** που συμπεριλαμβάνει κινήσεις του κεφαλιού, της γλώσσας, των ματιών και έλεγχο του στόματος. Επιτρέπουν την αναπαράσταση φυσικών εκφράσεων του προσώπου. Κάθε FAP αριθμείται με έναν δείκτη που λαμβάνει τιμές από το 1 ως το 68. Για κάθε FAP, το πρότυπο του MPEG-4 ορίζει ένα περιγραφικό όνομα, την κατάλληλη μονάδα FAPU, την ομάδα στην οποία ανήκει, την κατεύθυνση της θετικής κίνησης και κατά πόσον η κίνηση του χαρακτηριστικού σημείου επιτρέπεται και προς τις δύο κατευθύνσεις ή μόνο προς μία κατεύθυνση. Τα FAPs μπορούν επίσης να ορίσουν μονάδες δράσεις του προσώπου (facial action units). Τιμές των FAPs με υπερβολικά μεγάλες απόλυτες τιμές ορίζουν κινήσεις που δεν είναι δυνατές για συνήθεις ανθρώπους, αλλά μπορεί να είναι επιθυμητές για χαρακτήρες κινουμένων σχεδίων.

Το σύνολο των FAPs περιλαμβάνει **δύο FAPs υψηλού επιπέδου**, τα FAPs **F<sub>1</sub>** και **F<sub>2</sub>**, που ονομάζονται **‘Viseme’** και **‘Expression’** αντίστοιχα. Μαζί αποτελούν την 1<sup>η</sup> ομάδα των FAPs.

Ο όρος viseme προέκυψε από το συνδυασμό των λέξεων **visual phoneme**. Πρόκειται ουσιαστικά για το **οπτικό ισοδύναμο ενός φωνήματος** μιας γλώσσας (phoneme). Χρησιμοποιείται για να περιγράψει τη στοιχειώδη μορφολογία του προσώπου κατά την

εκφώνηση ενός συγκεκριμένου φωνήματος. Η σωστή χρήση των visemes επιτρέπει στα άτομα με προβλήματα ακοής να διαβάζουν τους ήχους αποτελεσματικά από ένα συνθετικό πρόσωπο, παρακολουθώντας τις μορφολογικές αλλαγές σε ολόκληρο το πρόσωπο και διαβάζοντας τα χείλη. Στο τυποποιημένο σύνολο του MPEG-4 περιλαμβάνονται μέσω του **FAP F<sub>1</sub>** μόνο **14 στατικά visemes** που διακρίνονται σαφώς μεταξύ τους. Για να καταστεί δυνατή η σωστή άρθρωση του λόγου σε συνδυασμό με την κίνηση του στόματος, το σχήμα του στόματος επηρεάζεται όχι μόνο από το τρέχον φώνημα, αλλά και από το προηγούμενο και το επόμενο φώνημα. Στο MPEG-4 οι μεταβάσεις μεταξύ των visemes ορίζονται από το συνδυασμό δύο φωνημάτων κάθε φορά, με κατάλληλο συντελεστή στάθμισης. Δεν είναι βέβαιο αν αυτή η τεχνική μπορεί να χρησιμοποιηθεί για animation ομιλίας, επιτυγχάνοντας υψηλής ποιότητας οπτική αναπαράσταση.

| Viseme no. | Αντίστοιχο φώνημα (phoneme) | Παραδείγματα σε λέξεις                  |
|------------|-----------------------------|---|
| 0          | -                           | -                                       |
| 1          | p, b, m                     | <u>put</u> , <u>bed</u> , <u>mill</u>   |
| 2          | f, v                        | <u>far</u> , <u>voice</u>               |
| 3          | T, D                        | <u>think</u> , <u>that</u>              |
| 4          | t, d                        | <u>tip</u> , <u>doll</u>                |
| 5          | k, g                        | <u>call</u> , <u>gas</u>                |
| 6          | tS, dZ, S                   | <u>chair</u> , <u>join</u> , <u>she</u> |
| 7          | s, z                        | <u>sir</u> , <u>zeal</u>                |
| 8          | n, l                        | <u>lot</u> , <u>not</u>                 |
| 9          | r                           | <u>red</u>                              |
| 10         | A:                          | <u>car</u>                              |
| 11         | e                           | <u>bed</u>                              |
| 12         | l                           | <u>tip</u>                              |
| 13         | Q                           | <u>top</u>                              |
| 14         | U                           | <u>book</u>                             |

**Πίνακας 3-5** FAP 1: Visemes και αντίστοιχα φωνήματα [47]

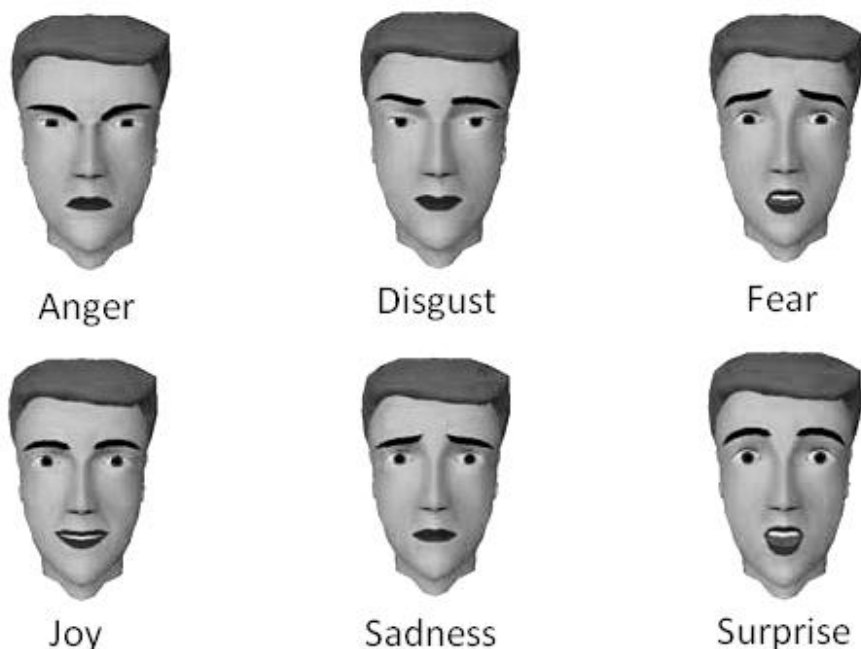
Η παράμετρος **FAP F<sub>2</sub>**, ορίζει τις **έξι αρχετυπικές εκφράσεις προσώπου** που αντιστοιχούν στα έξι αρχετυπικά συναισθήματα. Σε αντιδιαστολή με τα visemes του F<sub>1</sub>, οι εκφράσεις του προσώπου αποδίδονται με μια τιμή που ορίζει το μέγεθος της έντασης της έκφρασης. **Δύο εκφράσεις μπορούν να αποδοθούν ταυτόχρονα** με ένταση που κυμαίνεται στο διάστημα [0-63] για κάθε έκφραση. Οι **προδιαγραφές** για τις παραμέτρους των εκφράσεων του προσώπου αποδίδονται **με λεκτικές περιγραφές**. Η παράμετρος F<sub>2</sub> για την έκφραση επιτρέπει την αποδοτική απεικόνιση προσώπων. Τόσο το F<sub>1</sub> όσο και το F<sub>2</sub>, είναι παράμετροι animation υψηλού επιπέδου. Ο σχεδιαστής ενός μοντέλου προσώπου πρέπει να ορίζει τις παραμέτρους και για κάθε διαφορετικό μοντέλο προσώπου. Δεδομένου ότι σχεδιάζονται για ολόκληρη έκφραση, επιτρέπουν την απόδοση άγνωστων μοντέλων προσώπων με άκρως υποκειμενικά κριτήρια ποιότητας.



| α/α | Όνομα έκφρασης               | Περιγραφή   |
|-----|------------------------------|---|
| 1   | Joy<br>(Χαρά)                | Τα φρύδια είναι χαλαρά, Το στόμα είναι ανοιχτό και οι γωνίες του στόματος είναι τραβηγμένες προς τα πίσω, με κατεύθυνση προς τα αυτιά.                          |
| 2   | Sadness (Λύπη)               | Το εσωτερικό των φρυδιών έχει κλίση προς τα πάνω. Τα μάτια είναι ελαφρώς κλειστά, Το στόμα είναι χαλαρό.  |
| 3   | Anger<br>(Θυμός)             | Το εσωτερικό των φρυδιών τραβιέται προς τα κάτω και σμίγει. Τα μάτια είναι ορθάνοιχτα. Τα χείλη πιέζονται μεταξύ τους ή είναι ανοιχτά για να δείξουν τα δόντια. |
| 4   | Fear<br>(Φόβος)              | Τα φρύδια είναι σηκωμένα και σμίγουν. Το εσωτερικό των φρυδιών έχει κλίση προς τα πάνω. Τα μάτια έχουν ένταση και είναι σε εγρήγορση.                           |
| 5   | Disgust<br>(Αηδία -Απέχθεια) | Τα φρύδια και τα βλέφαρα είναι χαλαρά. Το επάνω χείλος είναι σηκωμένο και κυρτό, συχνά με ασύμμετρο σχήμα.  |
| 6   | Surprise (Εκπληξη)           | Τα φρύδια είναι σηκωμένα. Τα επάνω βλέφαρα είναι ορθάνοιχτα, τα κάτω βλέφαρα είναι χαλαρά. Το σαγόνι είναι ανοιχτό.   |

**Πίνακας 3-6** Βασικές εκφράσεις προσώπου όπως ορίζονται για το FAP 2 [47]

Παράδειγμα των παραπάνω εκφράσεων δίνεται στο ακόλουθο σχήμα.



**Σχήμα 3-4** Παράδειγμα εκφράσεων προσώπου για το FAP 2 [47]

Τα FAPs  $F_3$ - $F_{68}$  είναι χαμηλού επιπέδου, χωρίζονται στις υπόλοιπες 9 ομάδες και το καθένα αντιπροσωπεύει κίνηση συγκεκριμένου σημείου ή χαρακτηριστικού του μοντέλου προσώπου. Οι κινήσεις αυτές είναι οριζόντιες ή κάθετες μετατοπίσεις, αλλαγές στο βάθος, περιστροφές του κεφαλιού, διαστολή των κορών των ματιών, κινήσεις τις γλώσσας.

Η χρήση των FAPs  $F_1$  και  $F_2$  σε συνδυασμό με FAPs χαμηλού επιπέδου ( $F_3$ - $F_{68}$ ) που επιδρούν στα ίδια χαρακτηριστικά του προσώπου, μπορεί να έχει απρόσμενα αποτελέσματα στην οπτική αναπαράσταση του προσώπου. Σε γενικές γραμμές τα FAPs χαμηλού επιπέδου έχουν προτεραιότητα έναντι των FAPs  $F_1$  και  $F_2$ . Όταν προσδιορίζεται η έκφραση με τη βοήθεια του FAP  $F_2$ , ο κωδικοποιητής μπορεί να στείλει ένα bit με όνομα `init_face`, που είναι υπεύθυνο για την παραμόρφωση του ουδέτερου προσώπου, πριν εφαρμοστούν οι παραμορφώσεις των FAPs χαμηλού επιπέδου. Η παραμόρφωση του FAP  $F_2$  εφαρμόζεται με τους περιορισμούς του ουδέτερου προσώπου. Επειδή όμως ο κωδικοποιητής δε γνωρίζει κατ'ανάγκη τον τρόπο με τον οποίο έχουν υλοποιηθεί τα FAPs  $F_1$  και  $F_2$ , συνίσταται η χρήση μόνο των FAPs που δεν παρεμβαίνουν με τις αναπαραστάσεις των  $F_1$  και  $F_2$ .

### 3.6.6 Κωδικοποίηση FAPs με χρήση μάσκας

Όταν τα FAPs κωδικοποιούνται σε αρχείο FBA ή σε ροή δεδομένων χρησιμοποιείται **μάσκα για την επιλογή των FAPs που θα κωδικοποιηθούν**, και για τις οδηγίες για τον αποκωδικοποιητή. Ο σκοπός του μηχανισμού είναι η αποδοτική κωδικοποίηση. Η επιλογή των FAPs γίνεται με **ιεραρχία μάσκας δύο επιπέδων**. Το πρώτο επίπεδο είναι ο τύπος της μάσκας (mask type). Το δεύτερο επίπεδο είναι η μάσκα ομάδας (group mask).

Υπάρχει μία **παράμετρος mask type** για κάθε μία από τις 10 ομάδες των FAPs. Η τιμή της παραμέτρου είναι 0, 1, 2 ή 3. Η ερμηνεία της είναι η ακόλουθη:

- Τιμή παραμέτρου mask type = 0: Δεν κωδικοποιείται κανένα FAP της ομάδας.
- Τιμή παραμέτρου mask type = 1: Τίθεται η μάσκα ομάδας για να δείξει ποια FAPs της ομάδας κωδικοποιούνται. Τα FAPs που δεν επιλέγονται μέσω της μάσκας ομάδας **κρατούν τις προηγούμενες τιμές τους**.
- Τιμή παραμέτρου mask type = 2: Τίθεται και πάλι η μάσκα ομάδας για να δείξει ποια FAPs της ομάδας κωδικοποιούνται. Τα FAPs που δεν επιλέγονται μέσω της μάσκας ομάδας πρέπει να υπολογιστούν **με παρεμβολή** από τον αποκωδικοποιητή.
- Τιμή παραμέτρου mask type = 3: Κωδικοποιούνται όλα τα FAPs της ομάδας.

Εάν η τιμή της mask type είναι 1 ή 2, πρέπει να οριστεί και η **μάσκα ομάδας**. Η μάσκα ομάδας κρατά μία τιμή για κάθε FAP της ομάδας, οπότε έχει **μεταβλητό μήκος**, που ισούται με το πλήθος των FAPs που βρίσκονται σε κάθε ομάδα. Για παράδειγμα το μήκος για την ομάδα 2 είναι 16, όσα και τα FAPs που βρίσκονται σε αυτήν. Η τιμή που κρατείται για το FAP μπορεί να είναι 0 ή 1.

- Τιμή 0 στη μάσκα ομάδας: το FAP δεν κωδικοποιείται. Ανάλογα με την τιμή του mask type η τιμή του παραμένει ανεπηρέαστη ή προκύπτει από διαδικασία παρεμβολής.
- Τιμή 1 στη μάσκα ομάδας: το FAP κωδικοποιείται και χρησιμοποιείται.

Εάν τα FAPs κωδικοποιούνται από αρχείο ASCII, οι μάσκες των FAPs τίθενται στο αρχείο EPF (Encoder Parameter File).

### 3.6.7 Προσαρμοσμένα μοντέλα προσώπου στο MPEG-4

Κάθε τερματικό που είναι σε θέση να αποκωδικοποιήσει ροή πληροφοριών που περιέχει FAPs, πρέπει να παρέχει ένα μοντέλο προσώπου που είναι συμβατό με το MPEG-4. Συνήθως αυτό είναι ένα **ιδιοταγές μοντέλο προσώπου** στον αποκωδικοποιητή (**proprietary model**). Ο κωδικοποιητής δεν έχει γνώση για την εμφάνιση του προσώπου. Το MPEG-4 επιτρέπει στον κωδικοποιητή να ορίσει εξ'ολοκλήρου το μοντέλο προσώπου μέσω του συνόλου FDP. Η διαδικασία περιλαμβάνει τον ορισμό της στατικής γεωμετρίας του προσώπου στην ουδέτερη κατάσταση με τη χρήση γράφου σκηνής (scene graph), τον ορισμό ιδιοτήτων επιφάνειας και τον ορισμό των κανόνων animation με χρήση πινάκων FAT που αντιστοιχίζουν τις παραμέτρους FAP σε κινήσεις του μοντέλου. Ο κόμβος FDP μπορεί να χρησιμοποιηθεί και για την τροποποίηση του ιδιοταγούς μοντέλου του αποκωδικοποιητή. Ωστόσο το MPEG-4 δε διευκρινίζει με ποιους κανόνες μπορεί να τροποποιηθεί το ιδιοταγές μοντέλο.

Για να είναι ένα μοντέλο προσώπου συμβατό με το MPEG-4, πρέπει να μπορεί να εκτελεί όλα τα FAPs σύμφωνα με τις προδιαγραφές του προτύπου. Το μοντέλο πρέπει να έχει τουλάχιστον ισάριθμες κορυφές (vertices) με το πλήθος των χαρακτηριστικών σημείων που επιδέχονται κινήσεις μέσω FAP. Ένα συμβατό μοντέλο απαιτεί λοιπόν ως **ελάχιστο πλήθος κορυφών** τις 50 κορυφές. Ένα τέτοιο μοντέλο δε θα ήταν ικανοποιητικό. Για να είναι αισθητικά αποδεκτό και λειτουργικό το αποτέλεσμα, απαιτούνται τουλάχιστον 500 κόμβοι.

Τα βήματα για την κατασκευή ενός ιδιοταγούς μοντέλου προσώπου είναι τα εξής:

- Ορίζεται το **σχήμα του μοντέλου** και ορίζονται τα **σημεία ενδιαφέροντος** του προσώπου σύμφωνα με το σχήμα 3-2.
- Για κάθε FAP, ορίζεται η **κίνηση του αντίστοιχου σημείου**. Για τα περισσότερα σημεία το MPEG-4 ορίζει κίνηση κατά μία μόνο διάσταση. Για παράδειγμα το FAP  $F_{54}$ , ονόματι stretch\_r\_cornerlip\_o, μετατοπίζει την εξωτερική γωνία του στόματος οριζοντίως. Σε ένα ανθρώπινο πρόσωπο όμως αυτή η κίνηση συνοδεύεται και από κίνηση προς τα πίσω, που ακολουθεί την κυρτότητα του προσώπου. Επαφίεται στον προγραμματιστή/σχεδιαστή του μοντέλου προσώπου να ορίσει μια υποκειμενικά επιθυμητή παραμόρφωση για κάθε FAP.
- Ακολούθως πρέπει να οριστεί πώς η κίνηση ενός σημείου **επηρεάζει τις γειτονικές κορυφές**. Η χαρτογράφηση της κίνησης των χαρακτηριστικών σημείων σε κίνηση των κορυφών του μοντέλου μπορεί να γίνει με χρήση πινάκων FAT, πληροφορίες για τις μυικές παραμορφώσεις ή μετασχηματισμούς απόστασης.
- **Σχεδίαση των FAPs  $F_1$  και  $F_2$** . Για τις εκφράσεις του FAP  $F_2$ , το MPEG-4 παρέχει μόνο ποιοτικές υποδείξεις (Πίνακας 3-6) για τη σχεδιάσή τους. Παρομοίως τα visemes

του F<sub>1</sub> ορίζονται μόνο από τα αντίστοιχα φωνήματα που πρέπει να αναπαριστούν, αφήνοντας την επιλογή της σχεδιάσής τους στον προγραμματιστή.

Σύμφωνα με τα παραπάνω τα FAPs F<sub>1</sub> και F<sub>2</sub> πρέπει να σχεδιάζονται με προσοχή αφού κατά κανόνα θα χρησιμοποιηθούν σε εφαρμογές όπου χρειάζεται να ικανοποιούν **αισθητικά και λειτουργικά κριτήρια**. Όταν απεικονίζεται ένα πρόσωπο στο τερματικό, μπορεί να σταλεί και **πληροφορία για το φύλο** του προσώπου. Το MPEG-4 δεν απαιτεί τη χρήση διαφορετικού μοντέλου για αντρικό ή γυναικείο πρόσωπο. Συνίσταται όμως η χρήση πληροφορίας για το φύλο ώστε να τροποποιείται το πρόσωπο σε αρσενικό ή θηλυκό, για να αποφεύγεται η παρουσίαση ενός γυναικείου προσώπου με αντρική φωνή, ή το αντίστροφο.

Για την παρούσα εφαρμογή, έχει ενδιαφέρον ότι μπορεί να παραχθεί πίνακας αντιστοιχίας ανάμεσα στις μονάδες κίνησης AU του συστήματος FACS και τις παραμέτρους FAPs του MPEG-4. Απόσπασμα ενός τέτοιου πίνακα δίνεται παρακάτω:

| Αντιστοιχία AUs με FAPs |                        |  |
|-------------------------|------------------------|--|
| Αριθμός AU              | Όνομα FACS             | Αντίστοιχα FAPs  |
| 0                       | Neutral Face           | -  |
| 1                       | Inner Brow Raiser      | raise_l_i_eyebrow + raise_r_i_eyebrow  |
| 2                       | Outer Brow Raiser      | raise_l_o_eyebrow + raise_r_o_eyebrow  |
| 4                       | Brow Lowerer           | raise_l_o_eyebrow + raise_r_o_eyebrow +<br>raise_l_m_eyebrow + raise_r_m_eyebrow +<br>raise_l_i_eyebrow + raise_r_i_eyebrow +<br>squeeze_l_eyebrow + squeeze_r_eyebrow |
| 5                       | Upper Lid Raiser       | close_t_l_eyelid + close_t_r_eyelid  |
| 6                       | Cheek Raiser           | lift_l_cheek + lift_r_cheek  |
| 7                       | Lid Tightener          | close_b_l_eyelid + close_b_r_eyelid  |
| 8                       | Lips Toward Each Other | -  |
| 9                       | Nose Wrinkler          | lower_t_midlip + raise_nose +<br>stretch_l_nose + stretch_r_nose   |
| 10                      | Upper Lip Raiser       | raise_nose ( + stretch_l_nose + stretch_r_nose )<br>+ lower_t_midlip   |

**Πίνακας 3-7** Παράδειγμα αντιστοιχίας AU με FAPs [30]

# ΚΕΦΑΛΑΙΟ 4: Ασαφής Λογική και Ασαφή Σύνολα

## 4.1 Εισαγωγικά Στοιχεία

Η ασαφής λογική μπορεί να θεωρηθεί επέκταση της κλασσικής αριστοτέλειας λογικής. Στην ασαφή λογική μια πρόταση είναι αληθής "με κάποιο βαθμό αληθείας", και όχι απλά αληθής ή ψευδής. [49]

Ένας από τους κύριους στόχους της ασαφούς λογικής είναι η τυποποίηση και αυτοματοποίηση της ικανότητας των ανθρώπων να εξάγουν λογικά συμπεράσματα και να λαμβάνουν αποφάσεις σε περιβάλλοντα αβεβαιότητας, ανακρίβειας, ασάφειας, ελλειπών ή ανεπαρκούς πληροφορίας, μεροληπτικής πληροφορίας ή γνώσης και μεροληπτικής κατηγοριοποίησης σε κλάσεις ή συστάδες. [50]

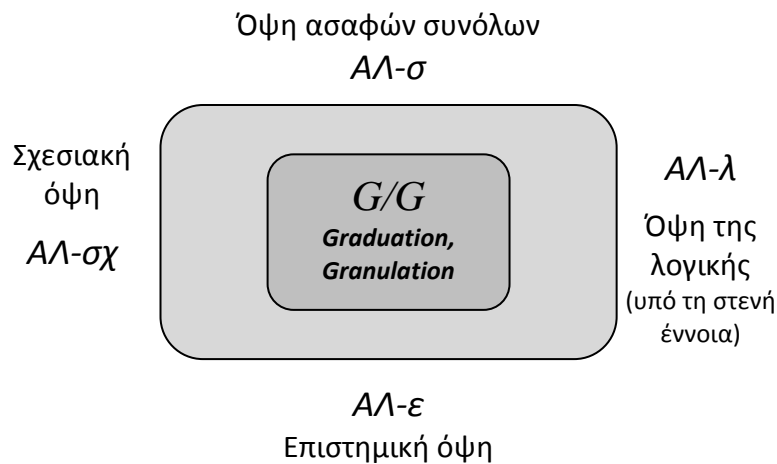
Η γέννηση της ασαφούς λογικής αποδίδεται στον Lotfi A. Zadeh, καθηγητή στο πανεπιστήμιο Berkeley της California, και ξεκίνησε από τη δουλειά του Zadeh το 1965 σχετικά με τη θεωρία ασαφών συνόλων. Ο Zadeh παρουσίασε μια εναλλακτική προσέγγιση για την παρουσίαση και επεξεργασία δεδομένων η οποία επιτρέπει μερική συμμετοχή σε ένα σύνολο, αντί για τη δικατάστατη λογική στην οποία το αντικείμενο είτε ανήκει, είτε δεν ανήκει στο σύνολο. [51]

Υπάρχουν πολλές παρερμηνείες του όρου ασαφής λογική. Η ασαφής λογική δεν είναι ασαφής με την έννοια ότι δεν δίνει σαφή απάντηση για ένα σύνολο δεδομένων. Σε μεγάλο βαθμό η ασαφής λογική είναι ακριβής [50]. Έχει καθιερωθεί ο όρος «ασαφής» (αντίστοιχος αγγλικός όρος fuzzy) επειδή ο πραγματικός κόσμος είναι ιδιαίτερα πολύπλοκος και υπάρχει εγγενής αβεβαιότητα ή ασάφεια στα δεδομένα. Σύμφωνα με τον Zadeh [52], η πηγή της αβεβαιότητας είναι η έλλειψη σαφώς ορισμένων κριτηρίων που καθορίζουν αυστηρά τη συμμετοχή σε κλάσεις, παρά η ύπαρξη τυχαίων μεταβλητών.

## 4.2 Όψεις της Ασαφούς Λογικής

Όπως αναφέρθηκε προηγουμένως ο όρος «ασαφής» χρησιμοποιείται για να επισημανθεί η ασάφεια και αβεβαιότητα που υπάρχει σε πολλά χαρακτηριστικά του πραγματικού κόσμου. Ένα ακόμη σημείο που προκαλεί συχνά σύγχυση είναι η διπλή χρήση του όρου «ασαφής λογική». Χρησιμοποιείται υπό τη στενή έννοια για να περιγράψει ένα λογικό σύστημα, αλλά με την ευρεία έννοιά του ο όρος χρησιμοποιείται συχνά σήμερα για να περιγράψει συστήματα γενικότερα από ένα σύστημα λογικής. Η ασαφής λογική, που εφεξής μπορεί να αναφέρεται στο κείμενο και ως ΑΛ, έχει σύμφωνα με τον Zadeh [50] τέσσερις κύριες όψεις που την περιγράφουν.

1. Η όψη της θεωρίας των ασαφών συνόλων ΑΛ-σ (Fuzzy-set-theoretic facet)
2. Η όψη της λογικής ΑΛ-λ (Logical facet)
3. Η επιστημική όψη ΑΛ-ε (Epistemic facet)
4. Η σχεσιακή όψη ΑΛ-σχ (Relational facet)



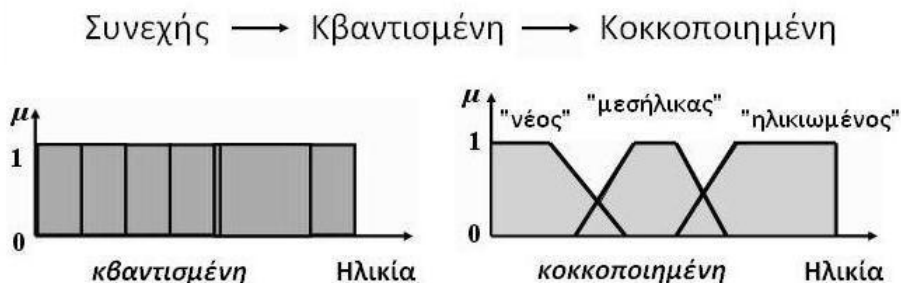
**Σχήμα 4-1** Κύριες Όψεις της Ασαφούς Λογικής (ΑΛ) [50]

Ο πυρήνας της ΑΛ είναι η διαβάθμιση/κοκκοποίηση  
(Graduation/Granulation – G/G)

Οι βασικές έννοιες της **διαβάθμισης** (graduation) και **κοκκοποίησης** (granulation) βρίσκονται στον πυρήνα της ΑΛ και είναι τα κύρια χαρακτηριστικά γνωρίσματα της ΑΛ. Πιο συγκεκριμένα στην ασαφή λογική όλα τα μεγέθη είναι ή επιτρέπεται να είναι αποτέλεσμα διαβάθμισης δηλαδή χαρακτηρίζονται από βαθμό αλήθειας που τα καθιστά «ασαφή» καθώς δεν χρειάζεται να είναι είτε αληθή ή ψευδή, αλλά λαμβάνουν μια ενδιάμεση τιμή αλήθειας. Επιπρόσθετα, στην ΑΛ τα μεγέθη είναι ή επιτρέπεται να είναι κοκκοποιημένα. Ο **κόκκος** (granule) είναι ένα συνονθύλευμα χαρακτηριστικών ή τιμών που συνδέονται λόγω ομοιότητας, εγγύτητας, λειτουργικότητας ή επειδή είναι αδύνατον να πραγματοποιηθεί διάκριση μεταξύ τους.

Για παράδειγμα η μεταβλητή «Ηλικία» είναι κοκκοποιημένο μέγεθος όταν οι τιμές της περιγράφονται με γλωσσικές μεταβλητές όπως: νέος, μεσήλικας, ηλικιωμένος (Σχήμα 4-2).

Μια γλωσσική μεταβλητή μπορεί να θεωρηθεί ως μια κοκκοποιημένη μεταβλητή οι τιμές της οποίας είναι γλωσσικές ετικέτες των κόκκων της μεταβλητής. Σε ποιοτικό επίπεδο, η διαβάθμιση και η κοκκοποίηση παίζουν καίριο ρόλο στην ανθρώπινη νόηση.



"νέος", "μεσήλικας" και "ηλικιωμένος" είναι Ασαφή Σύνολα

**Σχήμα 4-2** Κοκκοποίηση μιας μεταβλητής (π.χ. «Ηλικία») [50]

Τα ιδιαίτερα χαρακτηριστικά γνωρίσματα των τεσσάρων κυρίως όψεων της ΑΛ περιγράφονται παρακάτω.

#### 4.2.1 Όψη Ασαφών Συνόλων

Η όψη της ασαφούς λογικής συνόλων (Όψη Ασαφών Συνόλων - ΑΛσ) εστιάζει στα λεγόμενα ασαφή σύνολα, δηλαδή σε σύνολα με ασαφή όρια. Για παράδειγμα το σύνολο των όμορφων γυναικών, των έντιμων αντρών, ή των ψηλών βουνών. Η έννοια των ασαφών συνόλων προτάθηκε από τον Zadeh το 1965. Η θεωρία των ασαφών συνόλων είναι κεντρική για την ασαφή λογική. [53]

Πιο συγκεκριμένα, τα ασαφή σύνολα είναι σύνολα διαβάθμισης, όπου δηλαδή η συμμετοχή στο σύνολο υπόκειται σε διαβάθμιση. Δεν απαντούμε στο ερώτημα της κλασσικής λογικής ανήκει ή δεν ανήκει το στοιχείο στο σύνολο, αλλά τίθεται το ερώτημα σε τί βαθμό το στοιχείο ανήκει στο σύνολο. Κεντρική έννοια εδώ είναι ο **βαθμός συμμετοχής** κάθε στοιχείου σε ένα σύνολο, παρά η δικατάστατη συμμετοχή ή όχι σε ένα σύνολο το οποίο έχει συγκεκριμένα μέλη. Ένα ασαφές σύνολο  $A$ , υποσύνολο του συνόλου αναφοράς  $U$  (universe of discourse), ορίζεται από τη **συνάρτηση συμμετοχής**  $\mu_A$  (membership function) που συσχετίζει κάθε αντικείμενο  $u$  του συνόλου αναφοράς  $U$  με το βαθμό συμμετοχής του  $u$  στο σύνολο  $A$ . Ένα ασαφές σύνολο είναι θεμελιώδες αν η συνάρτηση συμμετοχής λαμβάνει τιμές στο μοναδιαίο διάστημα  $[0,1]$ . Γενικότερα η συνάρτηση συμμετοχής μπορεί να λάβει τιμές και σε ένα μερικώς διατεταγμένο σύνολο. Υπάρχουν πολλοί τύποι ασαφών συνόλων, για παράδειγμα ασαφή σύνολα τύπου 2 [54], ασαφή σύνολα-L [55], διπολικά ασαφή σύνολα [56] [57] και διαισθητικά ασαφή σύνολα (intuitionistic fuzzy sets) [58]. Στη γενική περίπτωση η τομή και η ένωση ασαφών συνόλων ορίζονται με τη βοήθεια κατάλληλης νόρμας. Συνήθεις νόρμες είναι οι τριγωνικές:  $t$ -norm και  $t$ -conorm [59].

#### 4.2.2 Όψη της Λογικής

Η λογική όψη (ΑΛλ) περιλαμβάνει την ασαφή λογική υπό τη στενή έννοια της λογικής. Η ΑΛλ μπορεί να θεωρηθεί ως γενίκευση της λογικής πολλαπλών τιμών ή αλλιώς της πλειότιμης λογικής. Οι επιδιώξεις της ΑΛλ είναι συναφείς με τους στόχους της κλασσικής λογικής [60]. Οι τιμές αληθείας στην ΑΛλ επιτρέπεται να είναι ασαφή σύνολα.

#### 4.2.3 Επιστημική όψη

Η επιστημική όψη της ΑΛε ασχολείται με την αναπαράσταση γνώσης, τη σημασιολογία των φυσικών γλωσσών και την ανάλυση πληροφοριών. Η ΑΛε αντιμετωπίζει τις φυσικές γλώσσες ως ένα σύστημα για την περιγραφή αντιλήψεων. Ένας σημαντικός κλάδος της επιστημικής ασαφούς λογικής είναι η θεωρία πιθανοτήτων όπως διατυπώθηκε από τους Zadeh [61], Dubois και Prade [62]. Ένας άλλος σημαντικός κλάδος της ΑΛε είναι η υπολογιστική θεωρία των αντιλήψεων όπως διατυπώθηκε από τον Zadeh το 1999.

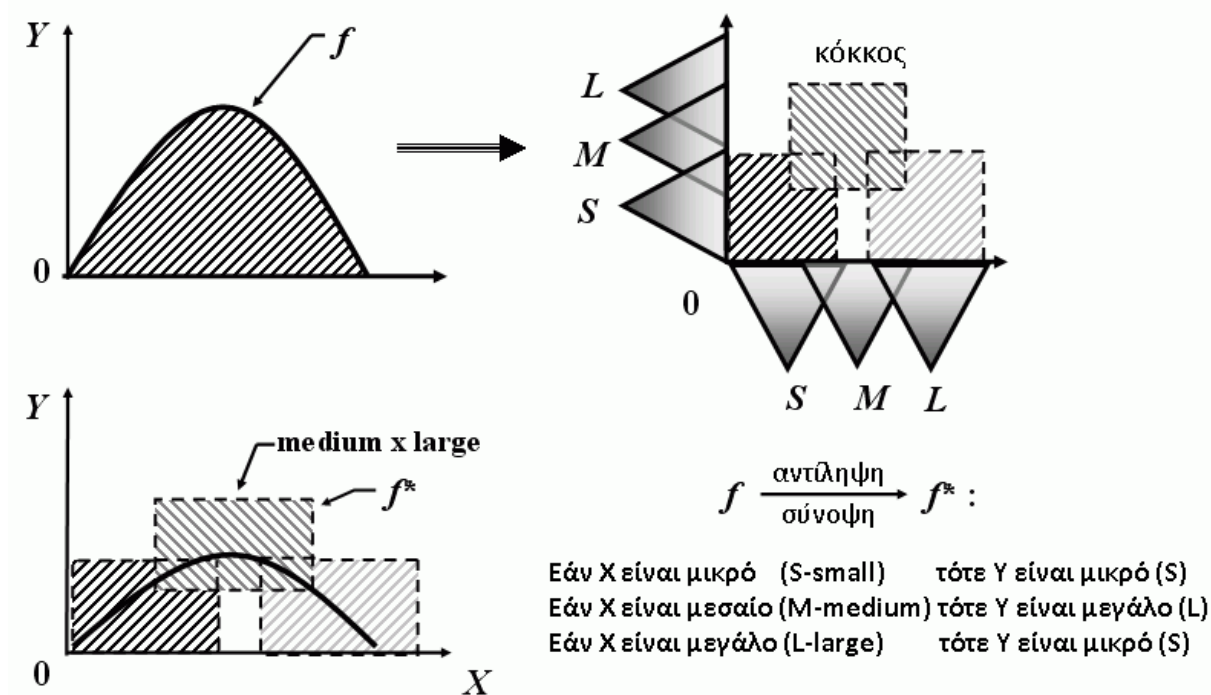
#### 4.2.4 Σχεσιακή όψη

Η σχεσιακή πτυχή της ασαφούς λογικής (ΑΛσχ) επικεντρώνεται στις ασαφείς σχέσεις (fuzzy relations) και στις ασαφείς εξαρτήσεις (fuzzy dependencies). Στη σχεσιακή ασαφή λογική, μια κοκκοποιημένη συνάρτηση  $f^*$  περιγράφεται ως μια συλλογή κανόνων ασαφούς λογικής τύπου Εάν-Τότε (fuzzy if-then rules).

Για παράδειγμα ένας **κανόνας τύπου Εάν-Τότε** μπορεί να έχει τη μορφή:

*Εάν  $X$  είναι  $A$ , τότε  $Y$  είναι  $B$*

Όπου τα  $A$  και  $B$  είναι ασαφή σύνολα με γλωσσικές ετικέτες όπως για παράδειγμα μικρό, μεσαίο, μεγάλο (Σχήμα 4-3). Τα  $X$  και  $Y$  είναι σε αυτήν την περίπτωση γλωσσικές μεταβλητές. Η έννοια της γλωσσικής μεταβλητής και του λογισμού που σχετίζεται με την αποτίμηση των κανόνων τύπου Εάν-Τότε, παίζουν καθοριστικό ρόλο σε όλες σχεδόν τις εφαρμογές της ασαφούς λογικής.



**Σχήμα 4-3** Κοκκοποίηση μιας συνάρτησης [50]

Μια κοκκοποιημένη συνάρτηση  $f^*$  μπορεί να θεωρηθεί ως σύνοψη (summary) της συνάρτησης  $f$ , όπου  $f^*$  είναι η τιμή-κόκκος της  $f$ .

Αξιοσημείωτη ειδική περίπτωση κοκκοποιημένης συνάρτησης είναι η κοκκοποιημένη συνάρτηση πυκνότητας πιθανότητας που μπορεί να περιγραφεί ως μια κοκκοποιημένη συνάρτηση κατανομής πιθανοτήτων.

Από τις τέσσερις όψεις της ασαφούς λογικής που περιγράφηκαν παραπάνω, η όψη των ασαφών συνόλων και οι κανόνες της σχεσιακής όψης σχετίζονται περισσότερο με την παρούσα εργασία. Λεπτομέρειες αναπτύσσονται στα παρακάτω κεφάλαια.



## 4.3 Ασαφή Σύνολα

Ένα κλασσικό σύνολο  $A$ , υποσύνολο του συνόλου αναφοράς  $U$ , μπορεί να παρασταθεί με τη βοήθεια της **χαρακτηριστικής ή αλλιώς δείκτριας συνάρτησης**  $\varphi_A$  η οποία αντιστοιχίζει κάθε στοιχείο  $x$  του συνόλου  $U$  σε ένα στοιχείο του συνόλου  $R=\{0,1\}$  ανάλογα με το αν ανήκει ή όχι στο σύνολο  $A$  [63]. Πιο συγκεκριμένα:

$$\text{Για κάθε } x \in U, \quad \varphi_A(x) = \begin{cases} 0, & x \notin A \\ 1, & x \in A \end{cases} \quad \text{Εξ. 4-1}$$

Κατ'αναλογία ένα ασαφές σύνολο  $A$  μπορεί να παρασταθεί με τη λεγόμενη **συνάρτηση συμμετοχής** (membership function)  $\mu_A(x)$  η οποία λαμβάνει τιμές στο κλειστό διάστημα

$[0, 1]$  αντιστοιχίζοντας δηλαδή κάθε στοιχείο  $x$  του συνόλου  $U$  σε μία τιμή που ανήκει στο κλειστό μοναδιαίο διάστημα. Έχουμε:

$$\text{Για κάθε } x \in U, \quad \mu_A(x) = \begin{cases} 0, & x \notin A \\ 1, & x \in A \\ \in (0,1), & x \text{ πιθανώς ανήκει στο } A \end{cases}$$

Εδώ η αβεβαιότητα ως προς τη συμμετοχή ή όχι ενός στοιχείου στο σύνολο  $A$  αντιμετωπίζεται με το να επιτρέψουμε στη συνάρτηση  $\mu_A(x)$  να λάβει τιμές εντός του συνεχούς διαστήματος  $(0, 1)$  αντί του δίτιμου πεδίου τιμών  $\{0, 1\}$ . Όσο η τιμή  $\mu_A(x)$  πλησιάζει τη μονάδα  $\mu_A(x) \rightarrow 1$ , τόσο μεγαλύτερη είναι η πιθανότητα το  $x$  να ανήκει στο σύνολο  $A$ , δηλαδή  $x \in A$ .

Ανάλογα με την εφαρμογή, η δυνατότητα να χρησιμοποιήσουμε τη συνάρτηση  $\mu_A(x)$  με πεδίο τιμών το  $R=[0,1]$  δεν χρειάζεται να ερμηνευθεί απαραίτητα ως αποτέλεσμα της έλλειψης βεβαιότητας δηλαδή της ύπαρξης ασάφειας. Ο ορισμός ενός συνόλου με συνάρτηση συμμετοχής  $\mu_A: x \in U \rightarrow [0,1]$  μπορεί να θεωρηθεί ότι ποσοτικοποιεί το βαθμό συμμετοχής στο σύνολο  $A$ . Έτσι επεκτείνεται η έννοια της συμμετοχής σε ένα σύνολο για να απαντήσει στο ερώτημα σε τί βαθμό ανήκει το κάθε στοιχείο στο σύνολο και όχι απλά αν ανήκει ή δεν ανήκει. Στα παρακάτω, τα κλασσικά σύνολα θα αναφέρονται στο κείμενο και ως «σαφή» σύνολα για να πραγματοποιηθεί η διάκριση με τα ασαφή σύνολα.

Ένα **ασαφές σύνολο**  $A$  συμβολίζεται και ως  $A = \{ (x, \mu_A(x)) \mid x \in U \}$ .

Το σύνολο χαρακτηρίζεται πλήρως από τα διατεταγμένα ζεύγη που αντιπροσωπεύουν την τιμή κάθε στοιχείου του συνόλου αναφοράς μαζί με το αντίστοιχο βαθμό συμμετοχής του στοιχείου, που υπολογίζεται από τη συνάρτηση συμμετοχής.

Όταν το σύνολο αναφοράς περιέχει **στοιχεία πεπερασμένου πλήθους**, δηλαδή

$U = \{x_1, x_2, \dots, x_n\}$ , τότε το ασαφές υποσύνολο  $A$  μπορεί να γραφεί ως:

$$A = \mu_A(x_1)/x_1 + \mu_A(x_2)/x_2 + \dots + \mu_A(x_n)/x_n$$

ή πιο συμπυκνωμένα:

$$A = \sum_{i=1}^n \mu_A(x_i)/x_i$$

Κάθε όρος  $\mu_A(x_i)/x_i$ ,  $i = 1, 2, \dots, n$  αναπαριστά το στοιχείο  $x_i$  του  $A$  με αντίστοιχο βαθμό συμμετοχής  $\mu_A(x_i)$ . Το σύμβολο  $+$  χρησιμοποιείται για να συμβολίσει την ένωση των επιμέρους στοιχείων σε ένα σύνολο [64].

Όταν το  $U$  είναι **μη-πεπερασμένο σύνολο**, το ασαφές υποσύνολο  $A$  γράφεται ως:

$$A = \int_U \mu_A(x)/x, \quad x \in U$$

Στον παραπάνω συμβολισμό το ολοκλήρωμα δεν υποδηλώνει την πράξη της ολοκλήρωσης, αλλά χρησιμοποιείται για να συμβολίσει την ένωση των στοιχείων του μη-πεπερασμένου συνόλου σε αντιδιαστολή με το σύμβολο άθροισης  $\Sigma$  που χρησιμοποιήθηκε για τα πεπερασμένα σύνολα.

Το **κενό ασαφές σύνολο**  $\emptyset$  ορίζεται από τη συνάρτηση συμμετοχής

$$\mu_{\emptyset}(x) = 0, \quad \forall x \in U$$

ενώ το **σύνολο αναφοράς**  $U$  ορίζεται ως

$$\mu_U(x) = 1, \quad \forall x \in U.$$

Το σύνολο αναφοράς αναφέρεται συχνά στη βιβλιογραφία και ως **υπερσύνολο αναφοράς** ή **χώρος ορισμού** ή **καθολικό σύνολο** αφού θεωρείται ότι τα υπό εξέταση ασαφή σύνολα της εκάστοτε εφαρμογής είναι πάντοτε υποσύνολα ενός συνόλου αναφοράς.

Το  $A$  είναι **ασαφές υποσύνολο** του  $B$  ή ισοδύναμα το  $A$  περιέχεται στο  $B$  ή το  $A$  είναι μικρότερο ή ίσο του  $B$  εάν  $\mu_A(x) \leq \mu_B(x)$ . Δηλαδή:

$$A \subset B \Leftrightarrow \mu_A(x) \leq \mu_B(x)$$

Δυο ασαφή σύνολα θεωρούνται ίσα και γράφονται ως  $A = B$  όταν ικανοποιούν τη σχέση

$$\mu_A(x) = \mu_B(x) \text{ για κάθε } x \in U. \text{ Για τα } \textbf{ίσα ασαφή σύνολα} \text{ ισχύει:}$$

$$A = B \Leftrightarrow \mu_A(x) = \mu_B(x)$$

#### 4.3.1 Συνήθειες συναρτήσεις συμμετοχής

Σε αυτό το υποκεφάλαιο εξετάζονται κάποιες συνήθειες συναρτήσεις συμμετοχής που συναντώνται συχνά στη βιβλιογραφία ή σε πρακτικές εφαρμογές.

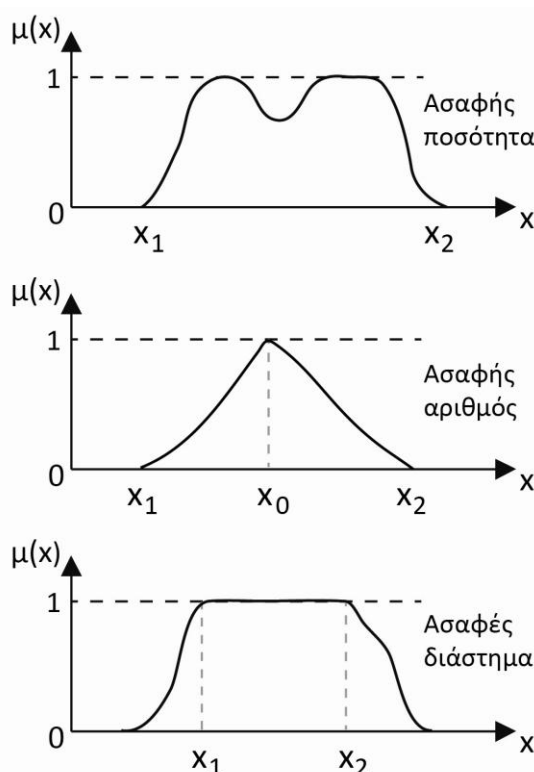
Στις υπολογιστικές μεθόδους με ασαφή στοιχεία εισόδου κάθε ασαφής ποσότητα αναπαρίσταται ως ένα ασαφές υποσύνολο των πραγματικών αριθμών  $\mathbb{R}$ .

Ένα ασαφές υποσύνολο  $A$  του  $\mathbb{R}$  καλείται **ασαφής ποσότητα** (fuzzy quantity) εάν υπάρχουν  $x_1 < x_0 < x_2 \in \mathbb{R}$  τέτοια ώστε για τη συνάρτηση συμμετοχής να ισχύει:

$$\mu_A(x_0) = 1 \text{ και } \mu_A(x) = 0 \text{ για κάθε } x \notin [x_1, x_2].$$

Εάν η συνάρτηση συμμετοχής  $\mu_A(x)$  έχει τριγωνοειδή μορφή όπως φαίνεται στο σχήμα 4-4, τότε το ασαφές υποσύνολο  $A$  καλείται **ασαφής αριθμός** (fuzzy number).

Εάν έχει τραπεζοειδή μορφή, τότε καλείται **ασαφές διάστημα** (fuzzy interval) [63] .



**Σχήμα 4-4** Ασαφείς ποσότητες [63]

Μερικές χρήσιμες έννοιες σχετικά με τα ασαφή σύνολα και τις συναρτήσεις συμμετοχής τους:

Το **σύνολο στήριξης**  $A_S$  (support) ή αλλιώς  $\text{supp}A$  ενός ασαφούς συνόλου  $A$  είναι το σαφές (crisp) σύνολο όλων των στοιχείων του συνόλου αναφοράς για τα οποία  $\mu_A(x) \neq 0$ . Η δείκτρια συνάρτηση του  $A_S$  μπορεί να γραφεί ως

$$\varphi_{A_S}(x) = \begin{cases} 0, & \mu_A(x) = 0 \\ 1, & \mu_A(x) \neq 0 \end{cases}$$

Ο **πυρήνας**  $A_K$  (kernel) ενός ασαφούς συνόλου είναι το σαφές σύνολο:

$$A_K = \{ x \mid x \in U, \mu_A(x) = 1 \}$$

Ένα ασαφές σύνολο λέγεται **κανονικό ασαφές σύνολο** εάν ο πυρήνας του δεν είναι κενό σύνολο. Με άλλα λόγια, το σύνολο  $A$  είναι κανονικό ασαφές σύνολο όταν έχει τουλάχιστον ένα στοιχείο, έστω  $x_0$ , το οποίο έχει βαθμό συμμετοχής  $\mu_A(x_0) = 1$ .

Το **σύνολο στήριξης**  $A_S$  αναφέρεται στη βιβλιογραφία και ως στήριξη του  $A$  και είναι το σαφές σύνολο που δίνεται από τη σχέση:

$$\text{supp}A = A_S = \{ x \mid x \in U, \mu_A(x) > 0 \}$$

**Ασαφές σημείο** (fuzzy singleton) καλείται το ασαφές σύνολο  $x_s$  το οποίο έχει σύνολο στήριξης  $supp x_s$  που αποτελείται από ένα μόνο σημείο του συνόλου αναφοράς  $U$  με τιμή συμμετοχής ίση με 1. Η συνάρτηση συμμετοχής για το ασαφές σημείο διαμορφώνεται συνεπώς ως εξής:

$$\mu_{x_s}(x) = \begin{cases} 0, & x \neq x_s \\ 1, & x = x_s \end{cases}$$

**Σημείο περάσματος** (crossover point) ενός ασαφούς συνόλου  $A$  είναι κάθε στοιχείο  $x_c$  του συνόλου αναφοράς  $U$  για το οποίο ισχύει  $\mu_A(x_c) = 0.5$

Η  **$\alpha$ -τομή** ενός **ασαφούς συνόλου**  $A$  είναι το σαφές σύνολο στοιχείων που περιέχει όλα τα στοιχεία του  $A$  για τα οποία ο βαθμός συμμετοχής είναι μεγαλύτερος ή ίσος από το όριο  $\alpha$ :

$$A_\alpha = \{ x \mid x \in U, \mu_A(x) \geq \alpha \} \quad \text{Εξ. 4-2}$$

Η συνάρτηση συμμετοχής –εφεξής μπορεί να αναφέρεται στο κείμενο και με το ακρωνύμιο ΣΣ– καθορίζει σε ουσιαστικό βαθμό το ασαφές σύνολο, συνεπώς η επιλογή της συνάρτησης συμμετοχής είναι σημαντική και εξαρτάται από την εκάστοτε εφαρμογή. Η επιλογή της συνάρτησης συμμετοχής είναι κατά κάποιο τρόπο η υποκειμενική πτυχή της ασαφούς λογικής καθώς επηρεάζει την ανάθεση και πιθανώς ερμηνεία των βαθμών συμμετοχής κάθε στοιχείου του συνόλου. [65]

Στη βιβλιογραφία υπάρχει πληθώρα διαφορετικών συναρτήσεων συμμετοχής. Τα κριτήρια που επελέγησαν στη σχεδίαση καθεμιάς από αυτές διαφέρουν και βοηθούν στο να κατηγοριοποιήσουμε τις συναρτήσεις συμμετοχής στις εξής κατηγορίες [66]:

1. **Ευριστικές** συναρτήσεις συμμετοχής.
2. Συναρτήσεις συμμετοχής που επελέγησαν με βάση την **αξιοπιστία** σε σχέση με την εφαρμογή για την οποία σχεδιάστηκαν.
3. Συναρτήσεις συμμετοχής που σχεδιάστηκαν ξεκινώντας με κάποιες **θεωρητικές απαιτήσεις**. Για παράδειγμα στη βιβλιογραφία απαντώνται συναρτήσεις συμμετοχής που σχεδιάστηκαν με βάση συναρτήσεις spline, τμηματικά πολυωνυμικές συναρτήσεις, συναρτήσεις πυκνότητας πιθανότητας. Παραδείγματα επιθυμητών ιδιοτήτων είναι η συνέχεια, αύξουσα ή φθίνουσα μονοτονία, δοσμένες οριακές συνθήκες και η γραμμικότητα.
4. Συναρτήσεις συμμετοχής σχεδιασμένες για ασαφή **συστήματα ελέγχου**. Δοσμένου ενός συστήματος προσδιορίζονται οι παράμετροι της ΣΣ ή το αντίστροφο, ταυτοποιούνται οι παράμετροι του συστήματος βάσει του ορισμού της ΣΣ.
5. Συναρτήσεις συμμετοχής που **μοντελοποιούν την ανθρώπινη αντίληψη** ή προκύπτουν από εμπειρική έρευνα πάνω σε έννοιες της ανθρώπινης επικοινωνίας.

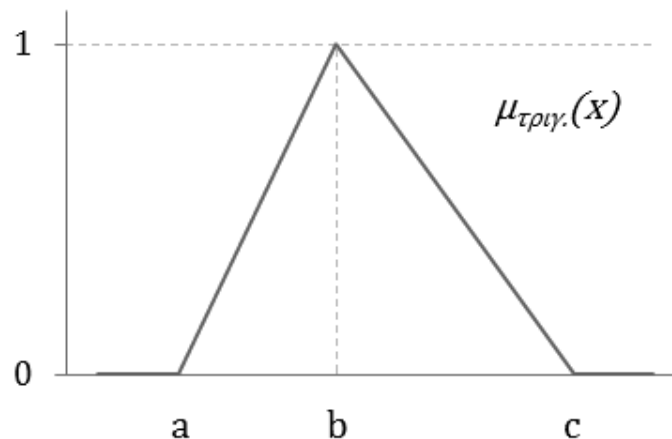
Η λίστα αυτή δεν είναι εξαντλητική και μπορεί να παρουσιαστεί με περισσότερο ή λιγότερο εκτενή τρόπο. Η παραπάνω ανάλυση καλύπτει τις ανάγκες της παρούσας εργασίας. Συνοπτικά, κάποιες **βασικές απαιτήσεις ή ιδιότητες των συναρτήσεων συμμετοχής** είναι οι εξής [66]:

- Θεωρητικές απαιτήσεις που εκφράζονται με τη μορφή μαθηματικών περιορισμών.
- Μικρή υπολογιστική πολυπλοκότητα.
- Περιγράφονται με μικρό πλήθος παραμέτρων.
- Οι παράμετροι έχουν νόημα.
- Οι συναρτήσεις συμμετοχής συνοδεύονται από τελεστές για τις βασικές πράξεις.
- Μπορούν να εκφραστούν και με γραμμική μορφή ώστε να καθίστανται κατάλληλες για πρακτικές εφαρμογές.
- Αξιοπιστία κατά την εφαρμογή τους.
- Μπορούν να γενικευθούν σε άλλες περιπτώσεις.

Παρακάτω δίνονται κάποιες βασικές συναρτήσεις συμμετοχής μαζί με τη σχηματική τους αναπαράσταση [64].

Η **τριγωνική συνάρτηση συμμετοχής** έχει ως παραμέτρους τρεις πραγματικούς αριθμούς  $a < b < c$ . Ορίζεται ως:

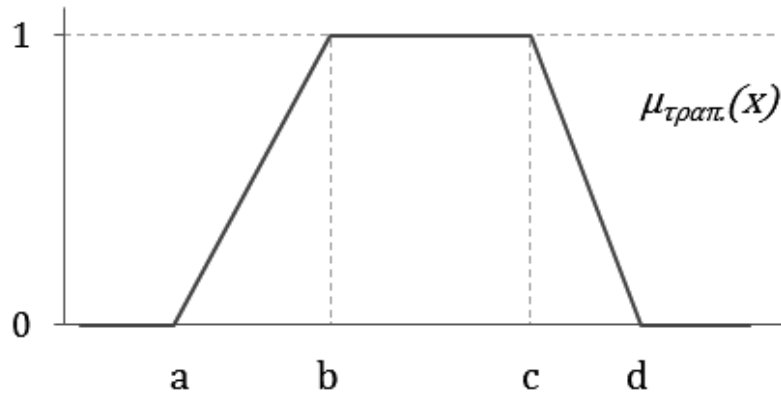
$$\mu_{\text{τριγ.}}(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x < b \\ \frac{c-x}{c-b}, & b \leq x < c \\ 0, & x \geq c \end{cases}$$



**Σχήμα 4-5** Τριγωνική συνάρτηση συμμετοχής

Η **τραπεζοειδής συνάρτηση συμμετοχής** έχει ως παραμέτρους τέσσερεις πραγματικούς αριθμούς  $a < b < c < d$ . Ορίζεται ως:

$$\mu_{\text{τραπ.}}(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x < b \\ 1, & b \leq x < c \\ \frac{d-x}{d-c}, & c \leq x < d \\ 0, & x \geq d \end{cases}$$

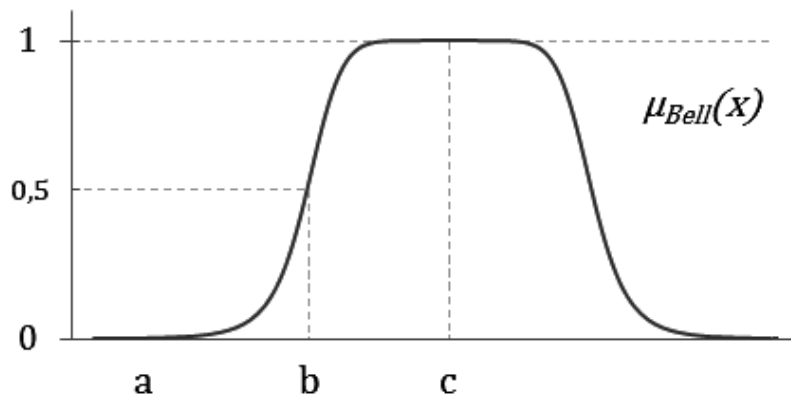


Σχήμα 4-6 Τραπεζοειδής συνάρτηση συμμετοχής

Η **καμπανοειδής συνάρτηση συμμετοχής** έχει ως παραμέτρους τρεις πραγματικούς αριθμούς  $a, b, c$ . Ορίζεται ως:

$$\mu_{Bell}(x) = \frac{1}{1 + \left| \frac{x-c}{a} \right|^{2b}}$$

Η παράμετρος  $c$  δείχνει το κέντρο ή αλλιώς τον άξονα συμμετρίας της καμπύλης, η παράμετρος  $b$  το σημείο περάσματος, δηλαδή το σημείο  $x_c$  για το οποίο  $\mu_k(x_c) = 0.5$ . Τέλος η παράμετρος  $a$  είναι το σημείο από όπου η  $\mu_{Bell}(x)$  από 0 γίνεται αυστηρά θετική.

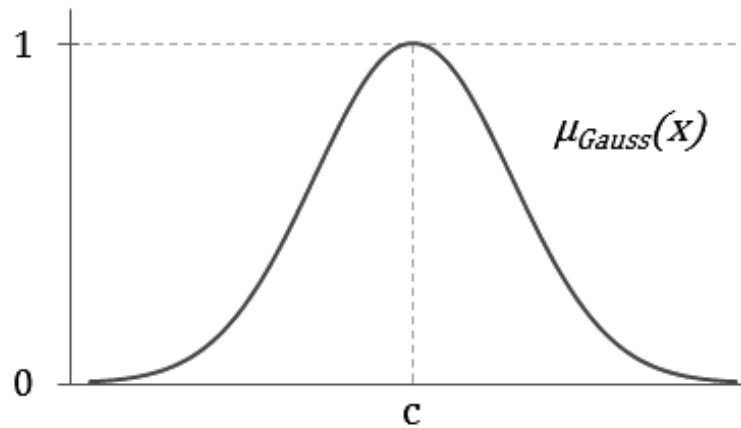


Σχήμα 4-7 Καμπανοειδής συνάρτηση συμμετοχής

Η **συμμετρική συνάρτηση συμμετοχής του Gauss** έχει ως παραμέτρους δύο πραγματικούς αριθμούς  $a, c$ .

$$\mu_{Gauss}(x) = e^{-\frac{(x-c)^2}{2a^2}}$$

Η παράμετρος  $c$  δείχνει το κέντρο της καμπύλης.



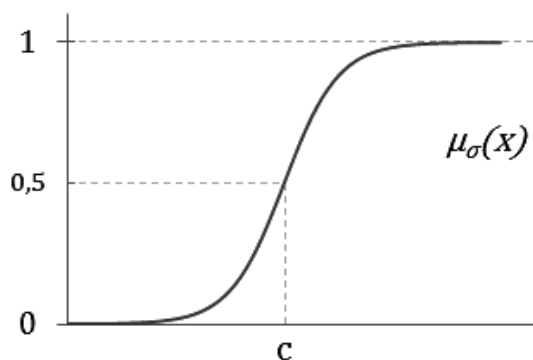
**Σχήμα 4-8** Συμμετρική συνάρτηση συμμετοχής Gauss

Η **σιγμοειδής συνάρτηση συμμετοχής** έχει ως παραμέτρους δύο πραγματικούς  $a, c$ . Η παράμετρος  $c$  δείχνει το σημείο περάσματος, δηλαδή το σημείο  $x_c$  για το οποίο  $\mu_\sigma(x_c) = 0.5$ .

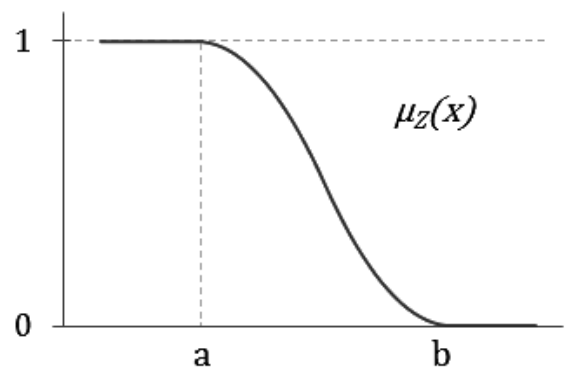
$$\mu_\sigma(x) = \frac{1}{1 + e^{-a(x-c)}}$$

Η **συνάρτηση συμμετοχής Z** έχει ως παραμέτρους δύο πραγματικούς αριθμούς  $a, b$  οι οποίες δείχνουν τα άκρα της καμπύλης.

$$\mu_Z(x) = \begin{cases} 1 & , \quad x < a \\ 1 - 2 \left( \frac{x-a}{b-a} \right)^2 & , \quad a \leq x < \frac{a+b}{2} \\ 2 \left( \frac{b-x}{b-a} \right)^2 & , \quad \frac{a+b}{2} \leq x < b \\ 0 & , \quad x \geq b \end{cases}$$



**Σχήμα 4-9** Σιγμοειδής συνάρτηση συμμετοχής



**Σχήμα 4-10** Συνάρτηση συμμετοχής Z

### 4.3.2 Πράξεις ασαφών συνόλων

Όπως και στην κλασσική συνολοθεωρία, έτσι και στη θεωρία ασαφών συνόλων μπορούμε να ορίσουμε τις έννοιες του συμπληρώματος, της ένωσης και της τομής δύο (ή και περισσότερων) ασαφών συνόλων με τη βοήθεια των αντίστοιχων συναρτήσεων συμμετοχής. [63]

Εάν  $A$  και  $B$  είναι δύο ασαφή σύνολα με συναρτήσεις συμμετοχής  $\mu_A(x)$  και  $\mu_B(x)$  αντίστοιχα, τότε το συμπλήρωμα  $A'$ , η ένωση  $A \cup B$  και η τομή  $A \cap B$  είναι επίσης ασαφή σύνολα με συναρτήσεις συμμετοχής:

**Συμπλήρωμα  $A'$ :**  $\mu_{A'}(x) = 1 - \mu_A(x)$ , όπου  $x \in U$ .

**Ένωση  $A \cup B$ :**  $\mu_{A \cup B}(x) = \max(\mu_A(x), \mu_B(x))$ ,  $x \in U$ .

**Τομή  $A \cap B$ :**  $\mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x))$ ,  $x \in U$ .

Παρότι οι παραπάνω πράξεις μπορούν να θεωρηθούν ως επέκταση της κλασσικής θεωρίας που γενικεύουν την εφαρμογή των αντίστοιχων τελεστών, οι ιδιότητες των ασαφών συνόλων δεν είναι κατ'ανάγκην ταυτόσημες με αυτές των κλασσικών συνόλων.

Για παράδειγμα μπορεί να υπάρξει ασαφές σύνολο  $A$  το οποίο να ικανοποιεί τη σχέση  $A \cap A' \neq \emptyset$ . Συνεπώς η αρχή της αντίφασης δεν ισχύει στα ασαφή σύνολα. Επίσης δεν ισχύει η αρχή του αποκλειόμενου τρίτου, καθώς  $A \cup A'$  δεν είναι απαραίτητα ίσο με το σύνολο αναφοράς  $U$ . Στα ασαφή σύνολα μπορεί να ισχύει ακόμη  $A \subset A'$  κάτι που δεν ισχύει στα κλασσικά, σαφή σύνολα [63].

Κάποιες βασικές ιδιότητες που ισχύουν τόσο στην κλασσική θεωρία συνόλων όσο και στη θεωρία ασαφών συνόλων είναι οι εξής:

**Αντιμεταθετική ιδιότητα:**

$$A \cap B = B \cap A \text{ και } A \cup B = B \cup A$$

**Προσεταιριστική ιδιότητα:**

$$(A \cap B) \cap C = A \cap (B \cap C) \text{ και } (A \cup B) \cup C = A \cup (B \cup C)$$

**Επιμεριστική ιδιότητα:**

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \text{ και } A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

**Κανόνας De Morgan:**

$$(A \cap B)' = A' \cup B' \text{ και } (A \cup B)' = A' \cap B'$$

Η έννοια των συνόλων βρίσκει εφαρμογή σε πολλά πεδία μαθηματικών ή άλλων μοντέλων, κατά συνέπεια η θεωρία των ασαφών συνόλων μπορεί να εφαρμοστεί στα αντίστοιχα πεδία. Για παράδειγμα στην προτασιακή λογική κάθε πρόταση μπορεί να αναπαρασταθεί ως ένα ασαφές σύνολο στοιχείων του εκάστοτε πεδίου εφαρμογής.



### 4.3.3 Τριγωνικές νόρμες t-norms, t-conorms.

Οι τριγωνικές νόρμες t-norms και t-conorms είναι πράξεις οι οποίες γενικεύουν την έννοια της κλασσικής λογικής σύζευξης και διάζευξης στην ασαφή λογική. Αποτελούν φυσική ερμηνεία της σύζευξης και διάζευξης σύμφωνα με τη σημασιολογία των μαθηματικών της ασαφούς λογικής. Οι τριγωνικές νόρμες χρησιμοποιούνται για το συνδυασμό πολλών κριτηρίων στη λήψη αποφάσεων πολλαπλών κριτηρίων. [67]

Μια **τριγωνική νόρμα t-norm** (triangular norm, εξού και η συντομογραφία t-norm) είναι μια δυαδική λειτουργία  $T$  στο διάστημα  $[0, 1]$  η οποία ικανοποιεί τις ακόλουθες προϋποθέσεις:

$$\begin{aligned} \text{Αντιμεταθετικότητα:} & T(x, y) = T(y, x) \\ \text{Προσεταιριστικότητα:} & T(x, T(y, z)) = T(T(x, y), z) \\ \text{Μονοτονία:} & y \leq z \Rightarrow T(x, y) = T(x, z) \\ \text{Ουδέτερο στοιχείο η μονάδα:} & T(x, 1) = x \end{aligned}$$

Συνήθεις **τριγωνικές νόρμες t-norm** είναι οι εξής:

$$\begin{aligned} \text{Νόρμα ελαχίστου, ή Gödel t-norm} & T_{\min}(x, y) = \min(x, y) \\ \text{Νόρμα γινομένου (product norm)} & T_{\text{prod}}(x, y) = x \cdot y \\ \text{Νόρμα t-norm Lukasiewicz} & T_L(x, y) = \max(x + y - 1, 0) \end{aligned}$$

Καμία t-norm δε μπορεί να λάβει τιμές μεγαλύτερες από την  $T_{\min}(x, y)$ .

Η δυαδική έννοια της τριγωνικής νόρμας t-norm είναι η t-conorm ή αλλιώς s-norm,  $S$ . Το ουδέτερο στοιχείο εδώ είναι το μηδέν:  $S(x, 0) = x$ . Οι λοιπές ιδιότητες ισχύουν κατ'αναλογία με τις t-norm.

Συνήθεις **τριγωνικές νόρμες s-norm** είναι οι εξής:

$$\begin{aligned} \text{Νόρμα μεγίστου, ή Gödel t-conorm} & S_{\max}(x, y) = \max(x, y) \\ \text{Νόρμα γινομένου (product t-conorm)} & S_{\text{prod}}(x, y) = x + y - x \cdot y \\ \text{Νόρμα t-conorm Lukasiewicz} & S_L(x, y) = \min(x + y, 1) \end{aligned}$$

Καμία t-conorm δε μπορεί να λάβει τιμές μικρότερες από την  $S_{\max}(x, y)$ .

Μπορούμε να σχηματίσουμε δυαδικά ζευγάρια  $(T, S)$  με τη βοήθεια της σχέσης:

$$S(x, y) = 1 - T(1 - x, 1 - y)$$

Αν  $T$  είναι μια τριγωνική νόρμα t-norm, τότε η  $S$  όπως προκύπτει παραπάνω είναι μια νόρμα t-conorm.

### 4.3.4 Αρχή της επέκτασης

Μια δυαδική αλγεβρική πράξη, έστω η  $x \star y$ , μπορεί να επεκταθεί σε ασαφείς ποσότητες με βάση την αρχή της επέκτασης. Αν  $a$  και  $b$  είναι ασαφή σύνολα με

συναρτήσεις συμμετοχής  $\mu_a$  και  $\mu_b$  αντίστοιχα, το αποτέλεσμα της  $a \star b$  είναι ασαφές σύνολο με συνάρτηση συμμετοχής

$$\mu_{a \star b}(x) = \sup[\min(\mu_a(y), \mu_b(z)) \mid y, z \in \mathbb{R}, x = y \star z] , x \in \mathbb{R}$$

Οι αλγεβρικές ιδιότητες των επεκτεταμένων πράξεων είναι πιο αδύναμες από τις αντίστοιχες ιδιότητες των πράξεων στο σύνολο των πραγματικών αριθμών. Η διαφορά αυτή οφείλεται κυρίως στην ασάφεια στον ορισμό του ασαφούς μηδενός ή της ασαφούς μονάδας καθώς και στην ασάφεια των σχέσεων ισότητας [68] .

# ΚΕΦΑΛΑΙΟ 5: Υπολογιστική υλοποίηση συστήματος εκτίμησης συναισθημάτων στη γλώσσα προγραμματισμού C

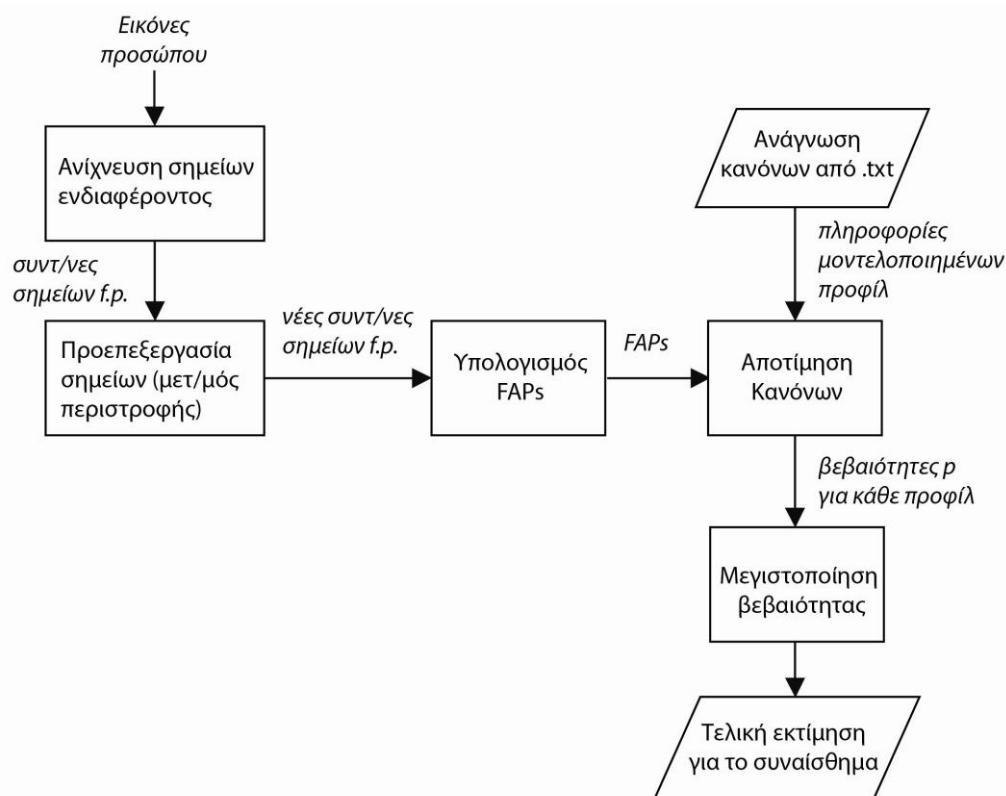
## 5.1 Εισαγωγικά στοιχεία

Στο παρόν κεφάλαιο παρουσιάζονται αναλυτικά τα κυριότερα στοιχεία σχετικά με την υπολογιστική υλοποίηση στη γλώσσα προγραμματισμού C, του συστήματος εκτίμησης συναισθημάτων βασισμένο στις παραμέτρους FAP του MPEG-4.

Περιγράφονται τα δεδομένα εισόδου, οι κύριες αρχές στις οποίες βασίζεται το σύστημα εκτίμησης του συναισθήματος, τα βήματα και οι σχετικοί αλγόριθμοι για την επεξεργασία των δεδομένων καθώς και οι έξοδοι του συστήματος.

Επίσης συζητείται με ποιον τρόπο μπορεί ο χρήστης να προσαρμόσει τις εισόδους του προγράμματος για νέα δεδομένα και πώς μπορεί να τροποποιήσει το πρόγραμμα αν χρειαστούν αλλαγές στην υλοποίηση ή για να πειραματιστεί με άλλες επιλογές στις διάφορες παραμέτρους.

Το παρακάτω μπλοκ διάγραμμα περιλαμβάνει τα κύρια μέρη του συστήματος. Λεπτομέρειες για τα επιμέρους τμήματα συζητούνται στα επόμενα υποκεφάλαια.



Σχήμα 5-1 Μπλοκ διάγραμμα συστήματος εκτίμησης συναισθημάτων.

## 5.2 Είσοδοι του προγράμματος εκτίμησης συναισθημάτων

Με τον όρο «είσοδοι» του προγράμματος αναφερόμαστε στα δεδομένα που διαβάζονται από αρχεία κειμένου .txt και που επηρεάζουν τα αποτελέσματα, δηλαδή την έξοδο του προγράμματος εκτίμησης συναισθήματος. Στις εισόδους πρέπει προφανώς να περιλαμβάνονται και οι πληροφορίες σχετικές με το εικονιζόμενο πρόσωπο. Οι πληροφορίες αυτές δίνονται με τη μορφή συντεταγμένων σημείων του προσώπου. Λεπτομέρειες αναπτύσσονται στις παρακάτω υποπαραγράφους.

### 5.2.1 Συντεταγμένες χαρακτηριστικών σημείων στο πρόσωπο

Οι είσοδοι του συστήματος εκτίμησης συναισθήματος που αφορούν στο εικονιζόμενο πρόσωπο είναι οι  $x, y$  συντεταγμένες 19 χαρακτηριστικών σημείων (feature points) του προσώπου για δύο καρέ του βίντεο εισόδου. Το ένα καρέ θεωρείται το ουδέτερο καρέ στο οποίο το εικονιζόμενο πρόσωπο βρίσκεται σε ουδέτερη συναισθηματική κατάσταση (Η παράμετρος activation του τροχού συναισθημάτων της Whissel θα είναι κοντά στο μηδέν). Το δεύτερο καρέ, που εφ'εξής αναφέρεται στο κείμενο ως τρέχον ή υπό εξέταση καρέ, θεωρείται ότι απεικονίζει το πρόσωπο σε μια ενεργή κατάσταση, πιθανώς υπό την επίδραση κάποιου συναισθήματος. Φυσικά το πρόσωπο μπορεί να διατηρήσει την ουδέτερη έκφραση ανάμεσα στα δύο καρέ.

Τα 19 σημεία προκύπτουν από διαδικασία παρακολούθησης (tracking), αλλά μπορούν να υπολογιστούν και με οποιονδήποτε άλλο τρόπο. Η επεξεργασία των δεδομένων εισόδου είναι ανεξάρτητη από τη μέθοδο που χρησιμοποιείται για την ανίχνευση, συνεπώς μπορεί να χρησιμοποιηθεί αυτόματη ή χειροκίνητη ανίχνευση. Η απόδοση του συστήματος εκτίμησης συναισθημάτων εξαρτάται βέβαια άμεσα από την ακρίβεια και επιτυχία της διαδικασίας ανίχνευσης.

Τα σημεία του προσώπου που ανιχνεύονται περιγράφονται και ονομάζονται με κωδικούς σε συμφωνία με το πρότυπο MPEG-4 (δες επίσης Σχήμα 3-2, στη σελίδα 60 και [47]).

Ο χαρακτηρισμός «δεξιά» ή «αριστερά» έχει ως σημείο αναφοράς το εικονιζόμενο πρόσωπο. Δηλαδή «αριστερό μάτι» θα είναι εκείνο που το εικονιζόμενο άτομο αντιλαμβάνεται ότι βρίσκεται στην αριστερή του πλευρά. Για τον θεατή της εικόνας, το «αριστερό» μάτι βρίσκεται στα δεξιά της εικόνας. Οι χαρακτηρισμοί «εσωτερικό» και «εξωτερικό» αναφέρονται στην απόσταση από τον κάθετο άξονα συμμετρίας του προσώπου. «Εσωτερικό» θεωρείται ένα σημείο που πλησιάζει τον άξονα συμμετρίας, ο οποίος διέρχεται χονδρικά από το κέντρο του μετώπου, τη μύτη, το μέσον του στόματος. «Εξωτερικό» είναι ένα σημείο που απομακρύνεται από αυτόν τον άξονα συμμετρίας.

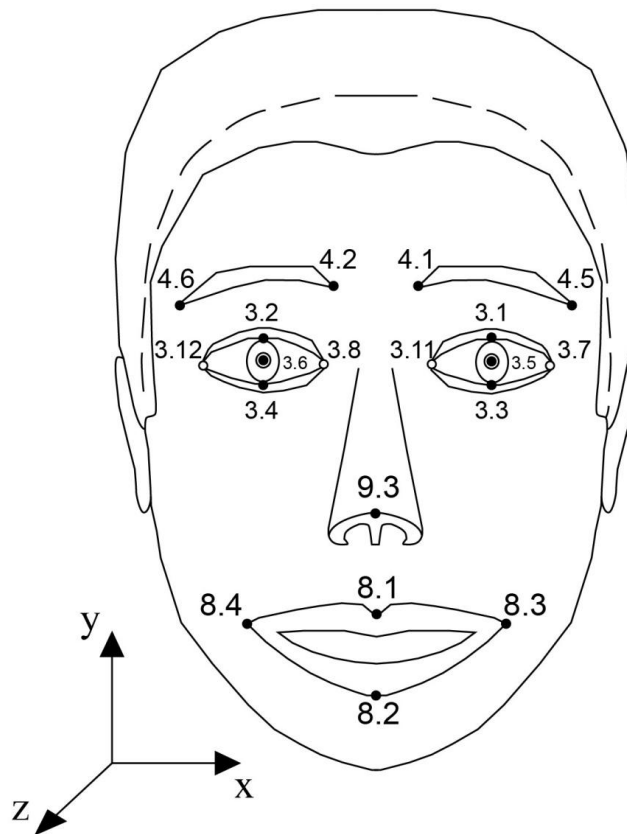
Εάν θέλουμε να αναφερθούμε σε μία από τις συντεταγμένες  $x, y$  των σημείων, προσθέτουμε τον συμβολισμό  $.x$  ή  $.y$  στο τέλος του κωδικού που δίνεται στον πίνακα 5-1. Για παράδειγμα, οι συντεταγμένες του σημείου 9.3 (άκρη μύτης) συμβολίζονται ως 9.3.x και 9.3.y, σε συμφωνία με το [47].

| Περιγραφή Χαρακτηριστικού Σημείου                 | Feature Point Code<br>( f.p. ) |
|---|--------------------------------|
| <b>Μάτια:</b>                                     |                                |
| Αριστερή γωνία (εσωτερική γωνία) δεξιού ματιού    | 3.8                            |
| Δεξιά γωνία (εξωτερική γωνία) δεξιού ματιού       | 3.12                           |
| Αριστερή γωνία (εξωτερική γωνία) αριστερού ματιού | 3.7                            |
| Δεξιά γωνία (εσωτερική γωνία) αριστερού ματιού    | 3.11                           |
| Κέντρο ίριδας δεξιού ματιού                       | 3.6                            |
| Κέντρο ίριδας αριστερού ματιού                    | 3.5                            |
| Κέντρο κάτω βλεφάρου του δεξιού ματιού            | 3.4                            |
| Κέντρο πάνω βλεφάρου του δεξιού ματιού            | 3.2                            |
| Κέντρο κάτω βλεφάρου του αριστερού ματιού         | 3.3                            |
| Κέντρο πάνω βλεφάρου του αριστερού ματιού         | 3.1                            |
| <b>Στόμα:</b>                                     |                                |
| Δεξιά γωνία στόματος (εξωτερικό περίγραμμα)       | 8.4                            |
| Αριστερή γωνία στόματος (εξωτερικό περίγραμμα)    | 8.3                            |
| Κάτω χείλος, κέντρο εξωτερικού περιγράμματος      | 8.2                            |
| Πάνω χείλος, κέντρο εξωτερικού περιγράμματος      | 8.1                            |
| <b>Φρύδια:</b>                                    |                                |
| Αριστερή (εσωτερική) γωνία δεξιού φρυδιού         | 4.2                            |
| Δεξιά (εξωτερική) γωνία δεξιού φρυδιού            | 4.6                            |
| Αριστερή (εξωτερική) γωνία αριστερού φρυδιού      | 4.5                            |
| Δεξιά (εσωτερική) γωνία αριστερού φρυδιού         | 4.1                            |
| <b>Μύτη:</b>                                      |                                |
| Άκρη μύτης  | 9.3                            |

**Πίνακας 5-1** Σημεία εισόδου στο σύστημα εκτίμησης συναισθημάτων

Η σειρά με την οποία καταγράφονται τα σημεία στον παραπάνω πίνακα είναι επίσης η σειρά με την οποία δίνονται στο αντίστοιχο αρχείο κειμένου από όπου διαβάζονται. Η σειρά του πίνακα 5-1 διατηρείται και κατά την αποθήκευση των συντεταγμένων των παραπάνω σημείων στα αντίστοιχα διανύσματα `fp_x_neutr[FPMAX]`, `fp_y_neutr[FPMAX]`, `fp_x[FPMAX]`, `fp_y[FPMAX]` όπου η σταθερά `FPMAX` ισούται με το πλήθος των διαθέσιμων σημείων (δες επίσης τον πίνακα 5-6, αντιστοιχία με μεταβλητές αποθήκευσης).

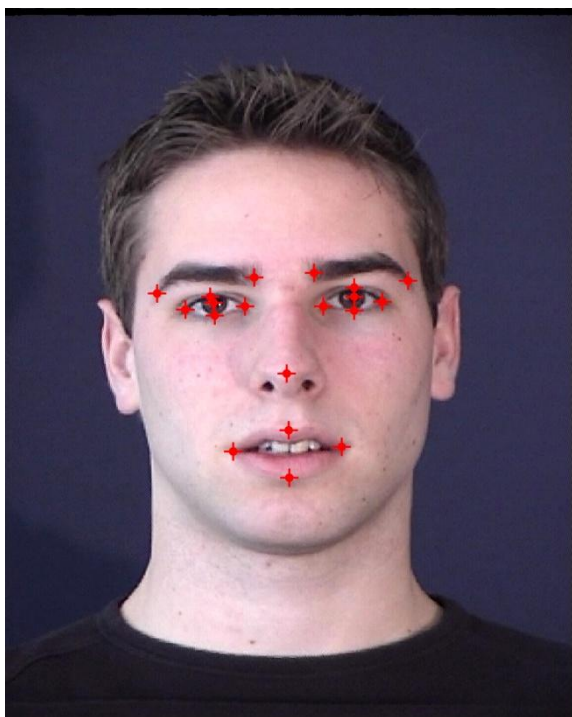
Τα παραπάνω σημεία φαίνονται και στο σχήμα:



**Σχήμα 5-2** Σημεία εισόδου στο σύστημα εκτίμησης συναισθημάτων  
Βάσει του τρισδιάστατου μοντέλου προσώπου [47]

Η ανάγνωση των συντεταγμένων του ουδέτερου καρέ και του τρέχοντος καρέ γίνεται με τη βοήθεια της συνάρτησης `read_frames_realtime` και πραγματοποιείται κάθε φορά που καλείται η συνάρτηση αποτίμησης συναισθήματος `analyze_emotions`. Οι συντεταγμένες διαβάζονται από ένα αρχείο κειμένου (.txt) και αποθηκεύονται σε μονοδιάστατους πίνακες/διανύσματα. Η συνάρτηση `read_frames_realtime` διαβάζει αρχικά την πρώτη γραμμή του αρχείου που είναι το καρέ αναφοράς και θεωρείται ότι είναι το ουδέτερο καρέ, και αποθηκεύει τις συντεταγμένες σε δύο διανύσματα `fp_x_neutr` και `fp_y_neutr`. Στη συνέχεια απομονώνει και διαβάζει την τελευταία γραμμή αποθηκεύοντας τις συντεταγμένες στα διανύσματα `fp_x` και `fp_y`. Για εφαρμογές πραγματικού χρόνου το αρχείο κειμένου θα μπορούσε να ενημερώνεται με την πάροδο του χρόνου, με αποτέλεσμα η συνάρτηση αποτίμησης συναισθήματος `analyze_emotions` να εξετάζει το πιο πρόσφατο καρέ κάθε φορά.

Παρακάτω δίνονται κάποια παραδείγματα φυσικών προσώπων στα οποία έχει γίνει χειροκίνητη ανίχνευση των 19 σημείων ενδιαφέροντος. Παρουσιάζονται το ουδέτερο καρέ και το υπό εξέταση καρέ. Οι εικόνες έχουν παρθεί από βίντεο της βάσης δεδομένων εκφράσεων προσώπου MMI Facial Expression Database [69].



**Σχήμα 5-3** Χειροκίνητη ανίχνευση σημείων – έκφραση απέχθειας [69]



**Σχήμα 5-4** Χειροκίνητη ανίχνευση σημείων – έκφραση φόβου [69]

Για τις ανάγκες της εργασίας, η χειροκίνητη ανίχνευση σημείων έγινε με τη βοήθεια του MATLAB, χρησιμοποιώντας τον παρακάτω κώδικα:

```
filename = 'Picture_name';
A = imread(filename,'jpg'); % use the correct file extension
figure, image(A)
%axis off           % Remove axis ticks and numbers
axis image         % Set aspect ratio to obtain square pixels

% Insert a breakpoint here in order to resize/zoom.
% Click continue button for execution to resume.

[x1,x2] = ginput(19);

% [x,y] = ginput(n) enables you to select n points
% from the current axes and returns the x- and y-coordinates
% in the column vectors x and y respectively.

fid = fopen('frames_coordinates.txt', 'a');

fprintf(fid, '\n');

for i=1:18
    fprintf(fid, '%.6f\t%.6f\t', x2(i), x1(i)) ;
end

fprintf(fid, '%.6f\t%.6f', x2(19),x1(19)); % print without \t

fprintf(fid, '\n%s\n', filename);

fclose(fid);
```

**Πίνακας 5-2** Βοηθητικό πρόγραμμα m-file για χειροκίνητη ανίχνευση σημείων στο MATLAB

### 5.2.2 Πληροφορίες μοντελοποιημένων προφίλ

Οι πληροφορίες για τα μοντελοποιημένα προφίλ είναι επίσης είσοδοι για το πρόγραμμα και πρόκειται να αποτιμηθούν από τη συνάρτηση `analyze_emotions`. Οι πληροφορίες των προφίλ βρίσκονται σε ένα αρχείο τύπου `.txt` με όνομα `emotion profiles.txt`. Οι πληροφορίες είναι οργανωμένες με τέτοιο τρόπο, ώστε κάθε γραμμή (διαχωρίζονται με χαρακτήρες `<enter>`, αλλαγής γραμμής) να περιέχει όλες τις απαραίτητες πληροφορίες για ένα μοντελοποιημένο προφίλ. Ακολουθεί παράδειγμα ενός μοντελοποιημένου προφίλ για να καταστεί περισσότερο κατανοητή η περιγραφή του:

**Πίνακας 5-3** Παράδειγμα μοντελοποιημένου προφίλ όπως διαβάζεται από το αρχείο `.txt`

```
Anger2 fap[19]-[-330 -200]-[1 1] fap[20]-[-335 -205]-[1 1] fap[21]-
[200 330]-[1 1] fap[22]-[205 335]-[1 1] fap[31]-[-200 -80]-[1 1]
fap[32]-[-194 -74]-[1 1] fap[33]-[70 190]-[1 1] fap[34]-[70 190]-[1 1]
fap[0]-[0 0]-[0 0]
```



- Σε κάθε προφίλ προηγείται μια αλφαριθμητική συμβολοσειρά (string) που είναι ετικέτα του εξεταζόμενου προφίλ (π.χ. Anger2).
  - Ακολουθούν τα δεδομένα για κάθε FAP που συμμετέχει σε αυτό το προφίλ. Η σειρά των FAPs δεν έχει σημασία. Διαβάζεται κάθε φορά μια ακολουθία χαρακτήρων της μορφής " fap[%d]-[%f %f]-[%f %f]", για παράδειγμα fap[19]-[-330 -200]-[1 1]. Ο ακέραιος, έστω  $j$ , στην 1<sup>η</sup> παρένθεση έχει το ρόλο δείκτη (index) και δηλώνει για ποιο FAP  $F_j$  πρόκειται. Οι δύο επόμενοι αριθμοί (κινητής υποδιαστολής) εντός της παρένθεσης είναι τα άκρα του διαστήματος  $X_{i,j}^{(k)}$  που δείχνει το εύρος τιμών που λαμβάνει το  $F_j$  στο προφίλ  $P_i^{(k)}$  ( $k$ -προφίλ του συναισθήματος  $i$ ).
- Εδώ για παράδειγμα  $i = Anger$ ,  $k = 2$  και  $j = 19$ , οπότε στο προφίλ  $P_{Anger}^{(2)}$ ,  $X_{Anger,19}^{(2)} = [-330, -200]$  για το  $F_{19}$ .
- Οι επόμενοι δύο αριθμοί έχουν σχέση με τον ορισμό ασαφούς κλάσης για το FAP  $F_j$  με βάση τραπεζοειδείς συναρτήσεις συμμετοχής. Οι δύο αριθμοί (που κυμαίνονται συνήθως από 0,3-1,0) καθορίζουν το λόγο της προβολής της πλάγιας πλευράς του τραπεζίου προς το μήκος του διαστήματος  $X_{i,j}^{(k)}$ . Περισσότερες λεπτομέρειες δίνονται στην παράγραφο 5.8 στη σελίδα 113 αυτού του κεφαλαίου.
  - Η τελευταία ακολουθία, fap[0]-[0 0]-[0 0], χρησιμοποιείται για να δηλώσει το τέλος του συγκεκριμένου προφίλ.
  - Το τέλος του αρχείου με τα προφίλ δηλώνεται με την ετικέτα End η οποία δεν ακολουθείται από άλλες πληροφορίες.
  - Οι ενδιάμεσες κενές γραμμές (χαρακτήρες αλλαγής γραμμής enter) δεν επηρεάζουν την ανάγνωση των προφίλ από το αρχείο.
  - Στις ακολουθίες χαρακτήρων " fap[%d]-[%f %f]-[%f %f]" στο αρχείο, είναι σημαντικό να μην υπάρχει χαρακτήρας κενού (space) σε σημεία που δεν προβλέπεται.

## 5.3 Μετασχηματισμός περιστροφής των χαρακτηριστικών σημείων

Πριν εφαρμοστούν οποιαδήποτε βήματα για τον υπολογισμό των FAPs, πραγματοποιείται ένας μετασχηματισμός περιστροφής των σημείων ενδιαφέροντος του προσώπου του ουδέτερου και του τρέχοντος καρέ, γύρω από κατάλληλο σημείο.

Οι υπολογισμοί κάποιων FAPs (λεπτομέρειες αναπτύσσονται στην παράγραφο 5.7) χρησιμοποιούν τον υπολογισμό κάθετων ή οριζοντίων αποστάσεων ανάμεσα σε σημεία, αντί για την ευκλείδεια απόστασή τους. Αυτό συμβαίνει στα FAPs  $F_{53} - F_{60}$  ( $F_6 - F_{13}$ ).

Η **ευκλείδεια απόσταση** δύο σημείων **δεν επηρεάζεται** από τυχόν περιστροφές του προσώπου γύρω από τον  $z$  άξονα του τρισδιάστατου μοντέλου που πιθανώς πραγματοποιείται ανάμεσα σε δύο καρέ (οι άξονες δίνονται στο Σχήμα 5-2, στη σελίδα 86). Αντίθετα η **κάθετη ή οριζόντια απόσταση** ανάμεσα σε δύο σημεία, δηλαδή η απόσταση κατά  $x$  ή  $y$  άξονα, **επηρεάζεται από την περιστροφή** των σημείων γύρω από τον  **$z$  άξονα**. Αυτό μας οδηγεί στην απαίτηση να βρεθεί κατάλληλος τρόπος επεξεργασίας των συντεταγμένων των σημείων που προέκυψαν από την ανίχνευση ώστε να ελαχιστοποιηθούν τα σφάλματα υπολογισμού των FAPs που θα οφείλονται σε τυχαίες περιστροφές του προσώπου γύρω από τον  $z$  άξονα.

Ως γνωστόν η περιστροφή είναι **μετασχηματισμός ομοιότητας**, δηλαδή δεν αλλοιώνει τις αποστάσεις μετά την εφαρμογή του. Το ίδιο ισχύει και για το μετασχηματισμό μεταφοράς και οποιαδήποτε ακολουθία μεταφορών ή/και στροφών. Η γενική ιδέα είναι να εφαρμοστεί κάποιος μετασχηματισμός περιστροφής κατάλληλης γωνίας και γύρω από κατάλληλο σημείο στις συντεταγμένες του υπό εξέταση καρέ, ούτως ώστε ο προσανατολισμός του προσώπου να προσεγγίσει όσο γίνεται τον προσανατολισμό του προσώπου στο ουδέτερο καρέ, μειώνοντας έτσι τα σφάλματα υπολογισμού των FAPs που θα οφείλονταν στις τυχαίες περιστροφές του προσώπου γύρω από τον  $z$  άξονα.

Ισοδύναμα, μπορούμε να περιστρέψουμε τα σημεία του ουδέτερου καρέ ώστε να αποκτήσουν συγκεκριμένο προσανατολισμό, για παράδειγμα οριζόντιο προσανατολισμό, και στη συνέχεια να περιστρέψουμε τα σημεία του υπό εξέταση καρέ ώστε να αποκτήσουν τον ίδιο προσανατολισμό με το ουδέτερο καρέ, στην προκειμένη περίπτωση οριζόντιο προσανατολισμό.

Για την υλοποίηση των παραπάνω στον κώδικα, πρέπει να προσδιοριστεί η **γωνία περιστροφής** και το **κέντρο περιστροφής** στο διδιάστατο επίπεδο της εικόνας. Η μεθοδολογία που υλοποιήθηκε στον κώδικα περιγράφεται λεπτομερώς στις παρακάτω παραγράφους και συνοψίζεται στα εξής βήματα:

- Προσδιορισμός του προσανατολισμού του προσώπου μέσω εύρεσης της ευθείας ελαχίστων τετραγώνων που προσαρμόζεται καλύτερα στα άκρα των ματιών. Υλοποιήθηκαν τρεις μέθοδοι εύρεσης της ευθείας.
- Προσδιορισμός του σημείου  $(x_0, y_0)$  και της γωνίας στροφής  $\theta_0$ .
- Μεταφορά ώστε το σημείο  $(x_0, y_0)$  να συμπίσει με την αρχή των αξόνων.
- Περιστροφή των σημείων κατά κατάλληλη γωνία  $\theta$  ώστε η ευθεία να συμπίσει με τον οριζόντιο άξονα.
- Μεταφορά των σημείων κατά  $(-x_0, -y_0)$ .

### 5.3.1 Προσδιορισμός του προσανατολισμού του προσώπου

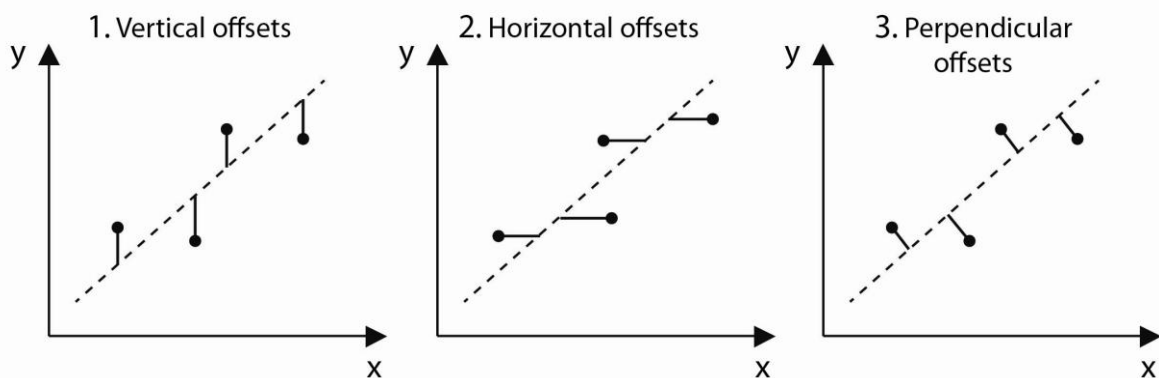
Για να προσδιοριστεί ο προσανατολισμός στο διδιάστατο επίπεδο της εικόνας, υπολογίζουμε την ευθεία που προσαρμόζεται καλύτερα στα άκρα των ματιών. Επιλέξαμε τα τέσσερα αυτά σημεία με την παραδοχή ότι πρόκειται για σταθερά σημεία πάνω στο πρόσωπο, που δεν επηρεάζονται (τουλάχιστον όχι αισθητά) από τις εκφράσεις

συναισθήματος του προσώπου. Επιπλέον θεωρούμε ότι όταν το πρόσωπο βρίσκεται σε ουδέτερη κατάσταση, η ευθεία αυτή τείνει να είναι οριζόντια, δηλαδή παράλληλη με τον οριζόντιο άξονα. Δεν χρησιμοποιείται η ευθεία που ορίζουν δύο από τα σημεία αυτά, αλλά προτιμάται η ευθεία που προσαρμόζεται καλύτερα στα τέσσερα άκρα των ματιών, ώστε να εξαλείφεται όσο γίνεται τυχόν σφάλματα της διαδικασίας ανίχνευσης. Ένας ακόμη λόγος που ωθεί στη χρήση ευθείας γραμμικής παλινδρόμησης είναι ότι ενώ γίνεται η παραδοχή ότι η ευθεία τείνει να είναι οριζόντια, δε θεωρείται δεδομένο ότι τα άκρα των ματιών σε ουδέτερη κατάσταση βρίσκονται κατ'ανάγκη πάνω σε μια οριζόντια ευθεία.

Η μέθοδος που προσφέρεται για αυτό το σκοπό είναι η **Μέθοδος Ελαχίστων Τετραγώνων** (εφεξής θα χρησιμοποιείται και το **ακρωνύμιο MET**).

#### Κλασσική μέθοδος ελαχίστων τετραγώνων με κάθετες αποκλίσεις

Θεωρώντας ως  $x$  την ανεξάρτητη μεταβλητή, η κλασσική μέθοδος ελαχίστων τετραγώνων ελαχιστοποιεί τις αποκλίσεις κατά  $y$  άξονα, που είναι δηλαδή παράλληλες με τον κάθετο άξονα  $y$  του καρτεσιανού επιπέδου. (Σχήμα 5-5, 1.)



**Σχήμα 5-5** Αποκλίσεις στη μέθοδο ελαχίστων τετραγώνων [69]

1. Αποκλίσεις παράλληλες στον  $y$  άξονα.
2. Αποκλίσεις παράλληλες στον  $x$  άξονα.
3. Αποκλίσεις κάθετες στην ευθεία παλινδρόμησης.

Η ευθεία γραμμικής παλινδρόμησης  $y = a + bx$  προσεγγίζει βέλτιστα  $n$  σημεία με συντεταγμένες  $(x_i, y_i)$ , με κριτήριο την ελαχιστοποίηση του σφάλματος  $R^2(a, b)$  που δίνεται από τη σχέση:

$$R^2(a, b) = \sum_i [y_i - (a + bx_i)]^2$$

Ως γνωστόν [70] οι σχέσεις που δίνουν τις παραμέτρους  $a, b$  της ευθείας, προκύπτουν από την ελαχιστοποίηση της συνάρτησης του σφάλματος ως προς τις παραμέτρους  $a, b$  και είναι οι εξής:

$$b = \frac{nS_{xy} - S_x S_y}{nS_{xx} - (S_x)^2} \quad \text{Εξ. 5-1}$$

$$a = \bar{y} - b\bar{x} = \frac{S_{xx}S_y - S_xS_{xy}}{nS_{xx} - (S_x)^2} \quad \text{Εξ. 5-2}$$

όπου

$$S_x = \sum_{i=1}^n x_i, \quad S_y = \sum_{i=1}^n y_i, \quad S_{xx} = \sum_{i=1}^n x_i^2, \quad S_{xy} = \sum_{i=1}^n x_i y_i, \quad \bar{x} = \frac{S_x}{n}, \quad \bar{y} = \frac{S_y}{n}$$

Εξ. 5-3

Οι αποκλίσεις  $y_i - \hat{y}_i = y_i - (a + bx_i)$  φαίνονται στο σχήμα 5-5. Σημειώνεται επίσης ότι το σημείο με συντεταγμένες  $(\bar{x}, \bar{y})$  ανήκει στην ευθεία, όπως φαίνεται από την Εξ. 5-2.

#### Μέθοδος ελαχίστων τετραγώνων με οριζόντιες αποκλίσεις

Εάν θεωρηθεί ως ανεξάρτητη μεταβλητή η  $y$  αντί της  $x$ , ουσιαστικά ελαχιστοποιούνται οι αποκλίσεις που είναι παράλληλες με τον  $x$  άξονα, όπως φαίνεται στο σχήμα 5-5. Η ευθεία που προσαρμόζεται βέλτιστα στα σημεία είναι της μορφής  $x = c + dy$ . Οι εξισώσεις για τις παραμέτρους της ευθείας προκύπτουν με απλή αντιμετάθεση των  $x \leftrightarrow y$  στις παραπάνω εξισώσεις.

Προκύπτουν συνεπώς οι σχέσεις:

$$d = \frac{nS_{xy} - S_xS_y}{nS_{yy} - (S_y)^2} \quad \text{Εξ. 5-4}$$

$$c = \bar{x} - d\bar{y} = \frac{S_{yy}S_x - S_yS_{xy}}{nS_{yy} - (S_y)^2} \quad \text{Εξ. 5-5}$$

Όπου  $S_x, S_y, S_{xy}, \bar{x}, \bar{y}$  έχουν την ίδια σημασία όπως στα προηγούμενα και δίνονται από τις ίδιες σχέσεις (εξισώσεις Εξ. 5-3).

Κατ'αναλογία με τα προηγούμενα, ο συμβολισμός  $S_{yy}$  αντιστοιχεί σε  $S_{yy} = \sum_{i=1}^n y_i^2$ . Όπως και προηγουμένως, το σημείο με συντεταγμένες  $(\bar{x}, \bar{y})$  ανήκει στην ευθεία (Εξ. 5-5).

Για την υπολογιστική υλοποίηση των παραπάνω προτιμάται μια εναλλακτική έκφραση των σχέσεων Εξ. 5-1 έως Εξ. 5-5, η οποία είναι ισοδύναμη αλγεβρικά με τα προηγούμενα, είναι όμως λιγότερο επισφαλής σε πρακτικές εφαρμογές όπου η ακρίβεια αναπαράστασης αριθμητικών δεδομένων είναι πεπερασμένη [71].

Ο όρος που βρίσκεται στον παρονομαστή των εξισώσεων, για τις Εξ. 5-1 και Εξ. 5-2 είναι  $\Delta_x = nS_{xx} - (S_x)^2$  και για τις Εξ. 5-4, Εξ. 5-5 είναι  $\Delta_y = nS_{yy} - (S_y)^2$ , χάνει σε ακρίβεια όταν οι τιμές των  $x_i$  (ή  $y_i$  αντίστοιχα) έχουν μεγάλη απόλυτη τιμή σε σχέση με τη διακύμανσή τους. Σε μια τέτοια περίπτωση ο υπολογισμός των  $\Delta_x, \Delta_y$  ανάγεται στην αφαίρεση δύο αριθμών που είναι σχεδόν ίσοι, και θα συμφωνούν σε μεγάλο πλήθος σημαντικών ψηφίων. Συνεπώς η διαφορά τους μπορεί να παρασταθεί μόνο με πολύ μικρό πλήθος σημαντικών ψηφίων και έτσι προκαλεί την απώλεια ακρίβειας.

Για τη διατύπωση των αλγεβρικά ισοδύναμων σχέσεων ορίζουμε ως:

$$S_{tt} = \sum_{i=1}^n (x_i - \bar{x})^2 \quad \text{και} \quad S_{ts} = \sum_{i=1}^n (x_i - \bar{x}) y_i \quad \text{Εξ. 5-6}$$

Οι σχέσεις Εξ. 5-1 και Εξ. 5-2 γράφονται ως:

$$b = \frac{S_{ts}}{S_{tt}} = \frac{\sum_{i=1}^n (x_i - \bar{x}) y_i}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad a = \bar{y} - b\bar{x} = \frac{S_y}{n} - \frac{S_x S_{ts}}{n S_{tt}} \quad \text{Εξ. 5-7}$$

Παρομοίως για τις Εξ. 5-4, Εξ. 5-5 ορίζουμε πρώτα:

$$S_{ss} = \sum_{i=1}^n (y_i - \bar{y})^2 \quad \text{και} \quad S_{st} = \sum_{i=1}^n (y_i - \bar{y}) x_i \quad \text{Εξ. 5-8}$$

Οπότε οι Εξ. 5-4 και Εξ. 5-5 μετατρέπονται σε:

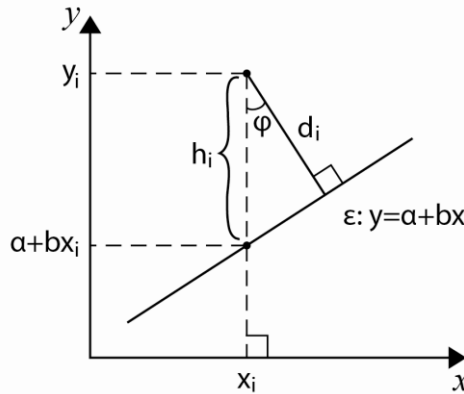
$$d = \frac{S_{st}}{S_{ss}} \quad c = \bar{x} - d\bar{y} = \frac{S_x}{n} - \frac{S_y S_{st}}{n S_{ss}} \quad \text{Εξ. 5-9}$$

Οι σχέσεις Εξ. 5-6 έως Εξ. 5-9 υλοποιήθηκαν στις συναρτήσεις `rotate_feature_points_dy` και `rotate_feature_points_dx`. Οι συναρτήσεις αυτές εκτός από τον υπολογισμό των ευθειών, υπολογίζουν τις παραμέτρους για την περιστροφή των σημείων και υλοποιούν το μετασχηματισμό ομοιότητας. Λεπτομέρειες συζητούνται παρακάτω.

#### Μέθοδος ελαχίστων τετραγώνων με αποκλίσεις κάθετες στην ευθεία παλινδρόμησης

Η τελευταία μέθοδος που παρουσιάζεται [69], βασίζεται στην ελαχιστοποίηση των τετραγώνων των αποκλίσεων των σημείων από την ευθεία, όταν οι αποκλίσεις είναι κάθετες στην ευθεία παλινδρόμησης (perpendicular offsets) και όχι παράλληλες σε έναν από τους δύο άξονες. Η παραλλαγή αυτή φαίνεται στο σχήμα 5-5 (3), στη σελίδα 91.

Η κάθετη απόσταση του σημείου  $(x_i, y_i)$  από την ευθεία παλινδρόμησης  $y = a + bx$  φαίνεται στο σχήμα 5-6:



**Σχήμα 5-6** Υπολογισμός κάθετων στην ευθεία αποκλίσεων για τη συνάρτηση σφάλματος.

Η απόσταση  $d_i$  δίνεται από τη σχέση:

$$d_i = \cos \varphi \cdot h_i = \frac{1}{\sqrt{\tan^2 \varphi + 1}} \cdot h_i = \frac{h_i}{\sqrt{b^2 + 1}} = \frac{|y_i - (a + bx_i)|}{\sqrt{b^2 + 1}}$$

Η συνάρτηση σφάλματος που ελαχιστοποιείται είναι το άθροισμα των τετραγώνων των  $d_i$ .

$$R^2(a, b) = \sum_i (d_i)^2 = \sum_i \frac{[y_i - (a + bx_i)]^2}{b^2 + 1}$$

Οι παράμετροι  $a, b$  ελαχιστοποιούν την  $R^2(a, b)$  όταν ικανοποιούν την απαίτηση:

$$\frac{\partial R^2}{\partial a} = 0 \quad \text{και} \quad \frac{\partial R^2}{\partial b} = 0 \quad \text{Εξ. 5-10}$$

Από τις Εξ. 5-10 προκύπτει μετά από αλγεβρική επίλυση το σύστημα εξισώσεων:

$$\begin{cases} a = \bar{y} - b\bar{x} & \text{Εξ. 5-11} \\ b^2 + 2Bb - 1 = 0 & \text{Εξ. 5-12} \end{cases}$$

Όπου

$$\bar{x} = \frac{S_x}{n} \quad \bar{y} = \frac{S_y}{n}$$

$$\text{και} \quad B = \frac{1}{2} \frac{(S_{xx} - n\bar{x}^2) - (S_{yy} - n\bar{y}^2)}{S_{xy} - n\bar{x}\bar{y}} \quad \text{Εξ. 5-13}$$

Όπως και προηγουμένως,

$$S_x = \sum_{i=1}^n x_i \quad S_y = \sum_{i=1}^n y_i \quad S_{xx} = \sum_{i=1}^n x_i^2 \quad S_{yy} = \sum_{i=1}^n y_i^2 \quad S_{xy} = \sum_{i=1}^n x_i y_i$$

Επίλυση ως προς  $b$  της δευτεροβάθμιας εξίσωσης  $b^2 + 2Bb - 1 = 0$  (Εξ. 5-12) του συστήματος εξισώσεων δίνει

$$b_{1,2} = -B \pm \sqrt{B^2 + 1} \quad \text{Εξ. 5-14}$$

Αντικαθιστώντας τις δύο τιμές στην Εξ. 5-11 προκύπτουν δύο τιμές για την παράμετρο  $a$ :

$$a_{1,2} = \bar{y} - b\bar{x} = \bar{y} - (-B \pm \sqrt{B^2 + 1})\bar{x} \quad \text{Εξ. 5-15}$$

Από τις εξισώσεις Εξ. 5-13, Εξ. 5-14 και Εξ. 5-15 είναι σαφές ότι οι υπολογισμοί αυτής της μεθόδου είναι πιο πολύπλοκοι σε σύγκριση με την κλασσική μέθοδο ελαχίστων τετραγώνων. Αυτή η υλοποίηση προσθέτει υπολογιστικό κόστος. Δεδομένου όμως του μικρού όγκου δεδομένων (4 σημεία) η διαφορά είναι αμελητέα.

Κατ'αναλογία με τις προηγούμενες δύο μεθόδους, ο τύπος για το  $B$  μπορεί να γραφεί σε αλγεβρικά ισοδύναμη μορφή που είναι καταλληλότερη για την υπολογιστική υλοποίηση σε πρακτικές εφαρμογές. Η Εξ. 5-13 γράφεται ως:

$$B = \frac{1}{2} \frac{S_{tt} - S_{ss}}{S_{ts}} = \frac{1}{2} \frac{\sum_{i=1}^n (x_i - \bar{x})^2 - \sum_{i=1}^n (y_i - \bar{y})^2}{\sum_{i=1}^n (x_i - \bar{x}) y_i} \quad \text{Εξ. 5-16}$$

Όπου τα  $S_{tt}$ ,  $S_{ss}$ ,  $S_{ts}$  ορίζονται με τον ίδιο τρόπο όπως προηγουμένως (Εξ. 5-6 και Εξ. 5-8).

Οι παράμετροι  $a_{1,2}$  και  $b_{1,2}$  υπολογίζονται και πάλι από τις ίδιες σχέσεις, δηλαδή τις Εξ. 5-14 και Εξ. 5-15.

Από τη διαδικασία ελαχιστοποίησης του κάθετου στην ευθεία σφάλματος προκύπτει λοιπόν το ζεύγος ευθειών:

$$\begin{aligned} \varepsilon_1: y &= a_1 + b_1 x \\ \varepsilon_2: y &= a_2 + b_2 x \end{aligned} \quad \text{Εξ. 5-17}$$

Σημειώνεται επίσης ότι το σημείο με συντεταγμένες  $(\bar{x}, \bar{y})$  ανήκει και σε αυτές τις ευθείες (δες Εξ. 5-11).

Εύκολα διαπιστώνουμε πως οι δύο ευθείες είναι κάθετες μεταξύ τους. Πολλαπλασιάζοντας τις παραμέτρους  $b_1, b_2$  (η παράμετρος  $b_i$  αντιπροσωπεύει την κλίση της ευθείας) έχουμε:

$$\begin{aligned} b_1 \cdot b_2 &= (-B + \sqrt{B^2 + 1}) \cdot (-B - \sqrt{B^2 + 1}) \\ &= (-B)^2 - (\sqrt{B^2 + 1})^2 = B^2 - B^2 - 1 = -1 \end{aligned}$$

Από αυτές τις ευθείες πρέπει να επιλεγεί μία, καθώς η άλλη είναι μεν μαθηματική λύση στο πρόβλημα αλλά δεν αποτελεί τη βέλτιστη ευθεία παλινδρόμησης των σημείων  $(x_i, y_i)$ .

Δεν είναι δυνατόν να απορριφθεί μία από τις δύο ευθείες με αναλυτικό τρόπο [72]. Προτείνεται η επιλογή της ευθείας που βρίσκεται πιο κοντά στην ευθεία που προκύπτει με την κλασσική μέθοδο ελαχίστων τετραγώνων, με οριζόντιες (ή κάθετες) αποκλίσεις.

Ένας τρόπος για να επιλεγεί η ευθεία από το ζεύγος  $\varepsilon_1, \varepsilon_2$  είναι να χρησιμοποιηθεί ως κριτήριο η οξεία γωνία  $\varphi_1, \varphi_2$  που σχηματίζουν καθεμιά εκ των  $\varepsilon_1, \varepsilon_2$  με την κλασσική ευθεία ελαχίστων τετραγώνων (κάθετες ή οριζόντιες αποκλίσεις)  $\varepsilon$ . Επιλέγεται εκείνη που σχηματίζει τη μικρότερη γωνία.

Η γωνία  $\varphi$  που σχηματίζουν δύο ευθείες  $\varepsilon_a, \varepsilon_b$  μπορεί να εκφραστεί ως η διαφορά των γωνιών  $\varphi_a, \varphi_b$  των ευθειών με έναν από τους δύο άξονες:  $\varphi = |\varphi_a - \varphi_b|$ . Αν το αποτέλεσμα είναι  $\varphi > \frac{\pi}{2}$  τότε υπολογίστηκε η αμβλεία γωνία που σχηματίζουν οι ευθείες, ειδάλλως (αν  $\varphi < \frac{\pi}{2}$ ) το αποτέλεσμα είναι η οξεία γωνία. Δεδομένης της αμβλείας γωνίας  $\varphi$  μεταξύ δύο ευθειών, η οξεία γωνία  $\varphi'$  υπολογίζεται εύκολα ως  $\varphi' = \pi - \varphi$ .

Μετά τον υπολογισμό των ευθειών  $\varepsilon_1, \varepsilon_2$  (Εξ. 5-17) υπολογίζουμε τις γωνίες  $\theta_1, \theta_2$  που σχηματίζουν οι ευθείες με τον  $y$  άξονα της εικόνας. Υπολογίζεται επίσης με τη MET με αποκλίσεις κατά  $x$  άξονα η ευθεία  $\varepsilon: x = c + dy$  που θα χρησιμοποιηθεί για την επιλογή της τελικής ευθείας.

Βρίσκουμε τη γωνία της ευθείας  $\varepsilon$  με τον άξονα  $y$ :  $\theta_\varepsilon = \arctan(d)$ .

Στη συνέχεια υπολογίζεται η γωνία  $\varphi = |\theta_\varepsilon - \theta_1|$ .

Υπενθυμίζεται ότι οι ευθείες  $\varepsilon_1, \varepsilon_2$  είναι κάθετες μεταξύ τους. Οι οξείες γωνίες που σχηματίζει η  $\varepsilon$  με τις  $\varepsilon_1, \varepsilon_2$  είναι συμπληρωματικές. Δηλαδή:  $(\widehat{\varepsilon, \varepsilon_1}) = \frac{\pi}{2} - (\widehat{\varepsilon, \varepsilon_2})$

Συνεπώς για την επιλογή της ευθείας  $\varepsilon_1$  αρκεί να ισχύει  $(\widehat{\varepsilon, \varepsilon_1}) < \frac{\pi}{4}$ .

Το κριτήριο για την επιλογή της ευθείας εκφράζεται ως:

$$\text{Εάν } \left( \varphi < \frac{\pi}{4} \text{ ή } \varphi > \frac{3\pi}{4} \right) \text{ επέλεξε την } \varepsilon_1 \text{ αλλιώς επέλεξε την } \varepsilon_2$$

Με την έκφραση αυτή ο όρος " $\varphi > \frac{3\pi}{4}$ " καλύπτει την περίπτωση όπου η γωνία  $\varphi$  που υπολογίστηκε ήταν η αμβλεία γωνία.

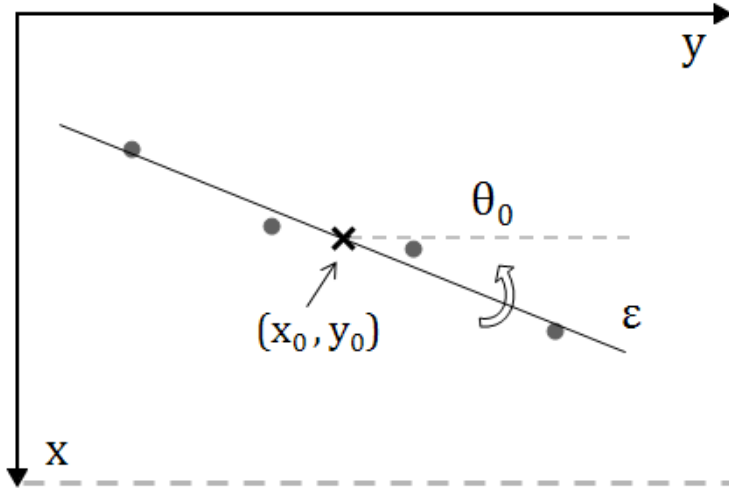
Στο σχήμα 5-7 που ακολουθεί στην επόμενη παράγραφο φαίνεται η θέση και προσανατολισμός των αξόνων  $x, y$  σε σχέση με τα pixel της εικόνας. Έχει γίνει εναλλαγή των αξόνων  $x, y$  ώστε ο  $x$  άξονας να είναι κάθετος με κατεύθυνση προς τα κάτω και ο  $y$  είναι οριζόντιος με κατεύθυνση προς τα δεξιά. Η επιλογή αυτή συμβαδίζει με τον τρόπο αρίθμησης των γραμμών και στηλών pixel μιας εικόνας, και με τις συντεταγμένες σε pixel έτσι όπως προκύπτουν από τη διαδικασία της ανίχνευσης σημείων. Στις περισσότερες περιπτώσεις τα άκρα των ματιών τείνουν να βρίσκονται σε μια ευθεία που είναι σχεδόν οριζόντια, δηλαδή τείνει να είναι παράλληλη με τον οριζόντιο άξονα  $y$  και συνεπώς κάθετη στον άξονα  $x$ . Για το λόγο αυτό, δεδομένου του τρόπου ορισμού των αξόνων, η κλασσική μέθοδος με αποκλίσεις κατά τον άξονα  $y$  δεν είναι τόσο κατάλληλη όσο η μέθοδος με αποκλίσεις κατά τον άξονα  $x$  ή με κάθετες στην ευθεία αποκλίσεις. Από τις τρεις μεθόδους που αναπτύχθηκαν στα προηγούμενα, προτιμούνται οι δύο τελευταίες, δηλαδή η χρήση των συναρτήσεων `rotate_feature_points_dx` ή `rotate_feature_points_po`.

### 5.3.2 Προσδιορισμός του κέντρου περιστροφής ( $x_0, y_0$ ) και της γωνίας στροφής $\theta_0$ .

Στο παρακάτω σχήμα φαίνεται το κέντρο περιστροφής ( $x_0, y_0$ ) και η γωνία στροφής  $\theta_0$  της ευθείας  $\varepsilon$  η οποία έχει προκύψει με μία από τις παραπάνω μεθόδους ελαχιστοποίησης της συνάρτησης τετραγώνων των αποκλίσεων.

Ως κέντρο περιστροφής επιλέχθηκε το σημείο  $(x_0, y_0) = (\bar{x}, \bar{y})$  που ανήκει στην ευθεία  $\varepsilon$  και στις τρεις μεθόδους. Η γωνία περιστροφής  $\theta_0$  ισούται με τη γωνία που σχηματίζει η ευθεία  $\varepsilon$  με τον άξονα  $y$ . Η έκφραση της  $\theta_0$  σε σχέση με τις παραμέτρους της ευθείας  $\varepsilon$  εξαρτάται από τη μέθοδο που χρησιμοποιείται.





**Σχήμα 5-7** Κέντρο και γωνία περιστροφής χαρακτηριστικών σημείων

1. Με την κλασσική MET η ευθεία που προκύπτει είναι της μορφής  $\varepsilon: y = a + bx$ . Μετασχηματίζοντας με επίλυση ως προς  $x$ :  $x = \frac{1}{b}y - \frac{a}{b}$ . Οπότε η γωνία  $\theta_0$  προκύπτει από τη σχέση  $\tan\theta_0 = \frac{1}{b} \Leftrightarrow \theta_0 = \arctan\left(\frac{1}{b}\right)$ ,  $b \neq 0$ . Αν  $b = 0$ , η ευθεία είναι παράλληλη με τον  $x$  άξονα και απαιτείται στροφή  $\frac{\pi}{2}$  ή  $-\frac{\pi}{2}$ .

Η σχέση για τη γωνία  $\theta_0$  προκύπτει και από τις τριγωνομετρικές ιδιότητες των συμπληρωματικών γωνιών.

Για την υλοποίηση σε C χρησιμοποιήθηκε η συνάρτηση `atan` της βιβλιοθήκης `math.h` η οποία επιστρέφει γωνία που ανήκει στο διάστημα  $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ .

2. Χρησιμοποιώντας την MET με αποκλίσεις παράλληλες στον  $x$  άξονα (έπειτα από εναλλαγή των  $x, y$ ) προκύπτει όπως είδαμε ευθεία παλινδρόμησης της μορφής  $\varepsilon: x = c + dy$ . Η γωνία ισούται με  $\theta_0 = \arctan(d)$ ,  $\theta_0 \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ .
3. Τέλος η MET με κάθετες στην ευθεία αποκλίσεις καταλήγει επίσης σε ευθεία της μορφής

$$\varepsilon_i: y = a_i + b_i x, \quad i = 1 \text{ ή } 2$$

$$\text{Επομένως } \theta_0 = \arctan\left(\frac{1}{b_i}\right), \quad b_i \neq 0, \quad i = 1 \text{ ή } 2, \quad \theta_0 \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$$

όπως προηγουμένως.

Και εδώ αν  $b_i = 0$ ,  $i = 1 \text{ ή } 2$ , απαιτείται στροφή  $\frac{\pi}{2}$  ή  $-\frac{\pi}{2}$ .

### 5.3.3 Οριακές θέσεις – Ειδικές περιπτώσεις ευθειών

Στις συναρτήσεις που υλοποιούν τις παραπάνω τρεις MET, συχνά εκτελείται διαίρεση με μεταβλητές που είναι δυνατόν να έχουν μηδενική τιμή. Πρέπει να εκτελούνται οι κατάλληλοι έλεγχοι στη ροή του προγράμματος ώστε να αντιμετωπίζεται η διαίρεση με το μηδέν και να μην προκαλούνται εξαιρέσεις σφαλμάτων.

Αναφέρουμε εδώ τις περιπτώσεις όπου υπάρχει πιθανή διαίρεση με το μηδέν, τη σημασία τους και πώς αντιμετωπίζονται.

#### 1. Κλασσική MET, αποκλίσεις παράλληλες στον $y$ άξονα.

**α.** Πριν τη διαίρεση για την παράμετρο  $b = \frac{S_{ts}}{S_{tt}}$  (Εξ. 5-7), ελέγχεται αν  $S_{tt} = 0$ .

Επειδή  $S_{tt} = \sum_{i=1}^n (x_i - \bar{x})^2$  είναι άθροισμα τετραγώνων, αυτό σημαίνει:

$$S_{tt} = 0 \Leftrightarrow x_i - \bar{x} = 0, i = 1, 2, \dots, n \Leftrightarrow x_i = \bar{x} \quad \forall i$$

Όλα τα σημεία έχουν την ίδια τετμημένη και βρίσκονται πάνω σε οριζόντια ευθεία, παράλληλη στον  $y$  άξονα (Σχήμα 5-7). Σε αυτήν την περίπτωση δεν απαιτείται μετασχηματισμός περιστροφής των σημείων ενδιαφέροντος του προσώπου.

**β.** Πριν τη διαίρεση  $\theta_0 = \arctan\left(\frac{1}{b}\right)$  ελέγχεται ότι η τιμή  $b \neq 0$ . Αν  $b = 0$ , η ευθεία είναι κάθετη και παράλληλη με τον  $x$  άξονα:  $y = a = \bar{y}$ . Συνεπώς απαιτείται στροφή κατά γωνία  $\frac{\pi}{2}$  ή  $-\frac{\pi}{2}$ . Το πρόσημο της γωνίας καθορίζεται με βάση τη θέση των υπόλοιπων σημείων του προσώπου. Συγκρίνουμε με βάση ένα σταθερό σημείο, όπως είναι η άκρη της μύτης (σημείο 9.3) οπότε προκύπτει η συνθήκη:

$$\text{Εάν } (9.3.y < \bar{y}) \text{ ανάθεσε } \theta_0 = \frac{\pi}{2} \text{ αλλιώς ανάθεσε } \theta_0 = -\frac{\pi}{2} \quad \text{Εξ. 5-18}$$

#### 2. Κλασσική MET, αποκλίσεις παράλληλες στον $x$ άξονα.

**α.** Πριν τη διαίρεση για την παράμετρο  $d = \frac{S_{st}}{S_{ss}}$  (Εξ. 5-9), ελέγχεται αν  $S_{ss} = 0$ .

Επειδή  $S_{ss} = \sum_{i=1}^n (y_i - \bar{y})^2$  είναι άθροισμα τετραγώνων, αυτό σημαίνει:

$$S_{ss} = 0 \Leftrightarrow y_i - \bar{y} = 0, i = 1, 2, \dots, n \Leftrightarrow y_i = \bar{y} \quad \forall i$$

Όλα τα σημεία έχουν την ίδια τεταγμένη και βρίσκονται πάνω σε κάθετη ευθεία, παράλληλη στον  $x$  άξονα (Σχήμα 5-7). Συνεπώς απαιτείται στροφή κατά γωνία  $\frac{\pi}{2}$  ή  $-\frac{\pi}{2}$ .

Χρησιμοποιείται και πάλι το κριτήριο της Εξ. 5-18.

**β.** Εδώ  $\theta_0 = \arctan(d)$  οπότε δε χρειάζεται να ελεγχθεί κάποια τιμή για διαίρεση με το μηδέν.

#### 3. MET με κάθετες στην ευθεία αποκλίσεις.

**α.** Πριν τη διαίρεση για την παράμετρο  $B = \frac{1}{2} \frac{S_{tt} - S_{ss}}{S_{ts}}$  ελέγχεται ότι  $S_{ts} \neq 0$ . Αν βρεθεί ότι  $S_{ts} = 0$ , η συνάρτηση επιστρέφει χωρίς να εκτελείται περιστροφή.

**β.** Πριν τους υπολογισμούς  $\theta_1 = \arctan\left(\frac{1}{b_1}\right)$ ,  $\theta_2 = \arctan\left(\frac{1}{b_2}\right)$  ελέγχεται ότι  $b_1, b_2 \neq 0$ .

Αν  $b_1 = 0$  ή  $b_2 = 0$  απαιτείται στροφή κατά γωνία  $\frac{\pi}{2}$  ή  $-\frac{\pi}{2}$ . Χρησιμοποιείται και πάλι το κριτήριο της Εξ. 5-18.

Παρομοίως πριν τον υπολογισμό  $d = \frac{S_{st}}{S_{ss}}$ , ελέγχεται ότι  $S_{ss} \neq 0$ . Αν  $S_{ss} = 0$ , η γωνία στροφής  $\theta_0$  πρέπει να τεθεί ίση με  $\frac{\pi}{2}$  ή  $-\frac{\pi}{2}$ . Χρησιμοποιείται και πάλι το κριτήριο της Εξ. 5-18.

### 5.3.4 Μετασχηματισμός σημείων – Περιστροφή και μεταφορές

Έχοντας καθορίσει το σημείο περιστροφής  $(x_0, y_0)$  και τη γωνία περιστροφής  $\theta_0$  σύμφωνα με την προηγούμενη παράγραφο, μπορούμε να προσδιορίσουμε τους μετασχηματισμούς με τη βοήθεια των αντίστοιχων πινάκων και τη χρήση ομογενών συντεταγμένων [73].

Επειδή ο βασικός μετασχηματισμός περιστροφής αναφέρεται σε περιστροφή γύρω από την αρχή των αξόνων, αρχικά πραγματοποιείται μεταφορά ώστε το σημείο περιστροφής  $(x_0, y_0)$  να συμπίπτει με την αρχή των αξόνων. Μεταφέρουμε τα σημεία σύμφωνα με το διάνυσμα:

$$\mathbf{d} = \begin{bmatrix} d_x \\ d_y \end{bmatrix} = \begin{bmatrix} -x_0 \\ -y_0 \end{bmatrix}$$

Ο πίνακας μεταφοράς περιγράφεται ως:

$$T(\mathbf{d}) = \begin{bmatrix} 1 & 0 & d_x \\ 0 & 1 & d_y \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -x_0 \\ 0 & 1 & -y_0 \\ 0 & 0 & 1 \end{bmatrix}$$

Ακολουθεί η περιστροφή των σημείων. Η περιστροφή κατά γωνία  $\theta = \theta_0$  γύρω από την αρχή των αξόνων περιγράφεται με τον πίνακα μετασχηματισμού:

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos\theta_0 & -\sin\theta_0 & 0 \\ \sin\theta_0 & \cos\theta_0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Η δεύτερη μεταφορά κατά διάνυσμα

$$\mathbf{d}' = \begin{bmatrix} d'_x \\ d'_y \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$$

Έχει πίνακα μεταφοράς

$$T(\mathbf{d}') = \begin{bmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{bmatrix}$$

Συνολικά λοιπόν ένα σημείο  $\mathbf{P}$  αντιστοιχίζεται σε ένα νέο σημείο  $\mathbf{P}'$  βάσει του μετασχηματισμού:

$$\mathbf{P}' = T(\mathbf{d}')R(\theta)T(\mathbf{d}) \cdot \mathbf{P}$$

$$\mathbf{P}' = \begin{bmatrix} p'_x \\ p'_y \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta_0 & -\sin\theta_0 & 0 \\ \sin\theta_0 & \cos\theta_0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -x_0 \\ 0 & 1 & -y_0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} p_x \\ p_y \\ 1 \end{bmatrix}$$

$$P' = \begin{bmatrix} p'_x \\ p'_y \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta_0 & -\sin\theta_0 & -x_0\cos\theta_0 + y_0\sin\theta_0 + x_0 \\ \sin\theta_0 & \cos\theta_0 & -x_0\sin\theta_0 - y_0\cos\theta_0 + y_0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} p_x \\ p_y \\ 1 \end{bmatrix}$$

Απ'όπου προκύπτει

$$\begin{bmatrix} p'_x \\ p'_y \end{bmatrix} = \begin{bmatrix} \cos\theta_0(p_x - x_0) - \sin\theta_0(p_y - y_0) + x_0 \\ \sin\theta_0(p_x - x_0) + \cos\theta_0(p_y - y_0) + y_0 \end{bmatrix}$$

Ο υπολογισμός των παραμέτρων  $(x_0, y_0)$ ,  $\theta_0$  καθώς και ο υπολογισμός των νέων συντεταγμένων μετά τους μετασχηματισμούς πραγματοποιείται από τις συναρτήσεις `rotate_feature_points_dy`, `rotate_feature_points_dx`, `rotate_feature_points_po`.

Οι συναρτήσεις έχουν τρία ορίσματα. Ένα παράδειγμα κλήσης της συνάρτησης : `rotate_feature_points_dx (fp_x, fp_y, 'n')`.

Τα πρώτα δύο ορίσματα `fp_x`, `fp_y` είναι διανύσματα που αποθηκεύουν τις συντεταγμένες των σημείων. Το τελευταίο όρισμα είναι χαρακτήρας που εισάγεται για να δηλώσει ο χρήστης αν θέλει ή όχι να τυπωθούν στην οθόνη πληροφορίες σχετικά με τις ευθείες που υπολογίστηκαν.

Οποιαδήποτε μέθοδος προτιμηθεί, η συνάρτηση καλείται δύο φορές. Μία για το καρέ αναφοράς που θεωρείται ουδέτερο καρέ και άλλη μία φορά για το τρέχον, υπό εξέταση καρέ.

Η επιλογή της μεθόδου για τον μετασχηματισμό των σημείων πραγματοποιείται μέσω της κλήσης της συνάρτησης `analyze_emotions` με την κατάλληλη τιμή για την παράμετρο `rotate`. Η παράμετρος `rotate` είναι η τέταρτη κατά σειρά παράμετρος, όπως φαίνεται και από τη δήλωση της συνάρτησης: `analyze_emotions (char printfp, char printfaps, char printprofiles, char rotate)`

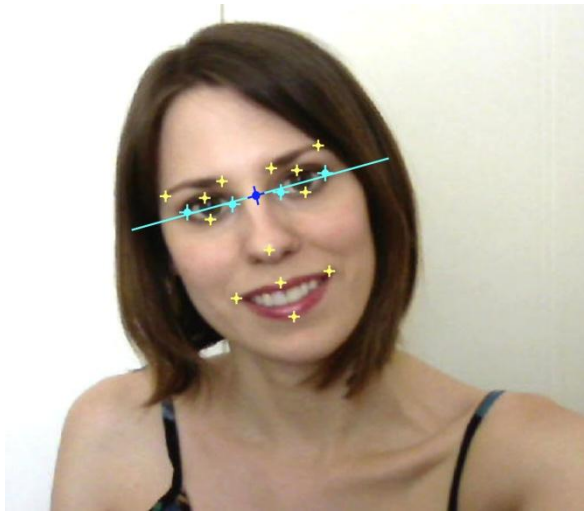
| Τιμή παραμέτρου <code>rotate</code> της συνάρτησης <code>analyze_emotions</code> | Μέθοδος περιστροφής χαρακτηριστικών σημείων  |
|--|--|
| 'x' ή 'X'  | Μέθοδος ελαχίστων τετραγώνων με αποκλίσεις κατά x-άξονα.   |
| 'y' ή 'Y'  | Μέθοδος ελαχίστων τετραγώνων με αποκλίσεις κατά y-άξονα.   |
| 'p' ή 'P'  | Μέθοδος ελαχίστων τετραγώνων με αποκλίσεις κάθετες στην ευθεία παλινδρόμησης (perpendicular offsets).      |
| 'n' ή 'N'  | Δεν εφαρμόζεται μετ/μός περιστροφής των σημείων ενδιαφέροντος.   |
| άλλος χαρακτήρας   | Εμφάνιση μηνύματος σφάλματος:<br>'Could not rotate feature points, the method you requested is not valid.' |

**Πίνακας 5-4** Παράμετρος `rotate` για την επιλογή μεθόδου περιστροφής των σημείων

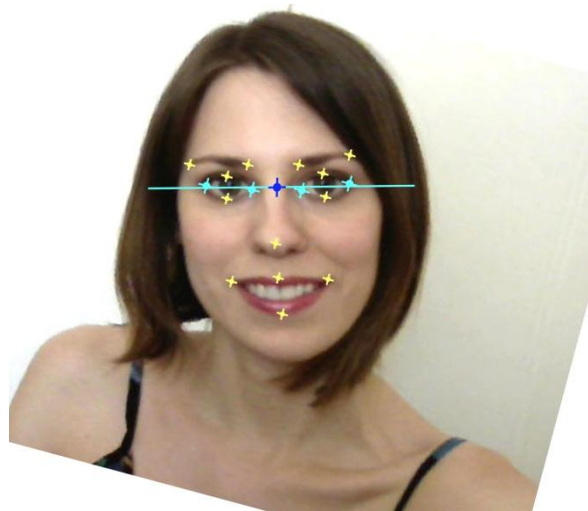
Δε διατίθεται μετασχηματισμός για να αντισταθμίσει την περιστροφή κατά άξονες  $x, y$  δεδομένης της διδιάστατης φύσης της εικόνας.

Δίνονται κάποια παραδείγματα φυσικών προσώπων πριν και μετά την περιστροφή.

**Τονίζεται ότι η εφαρμογή δεν περιστρέφει όλη την εικόνα, αλλά μόνο τα 19 χαρακτηριστικά σημεία στο πρόσωπο.** Εδώ οι περιστραμμένες εικόνες παρατίθενται για περιγραφικούς λόγους.

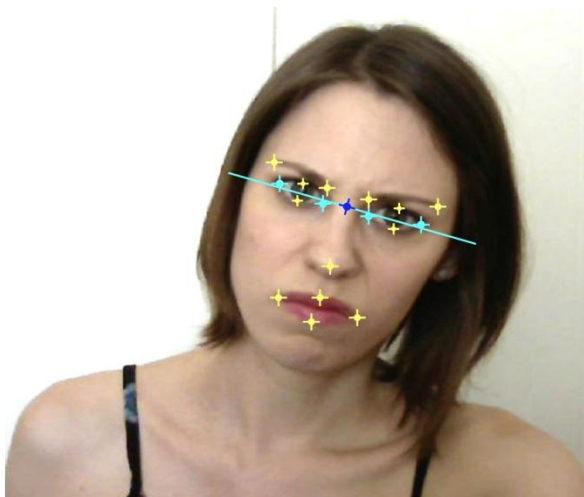


*προσανατολισμός σημείων  
πριν την περιστροφή*

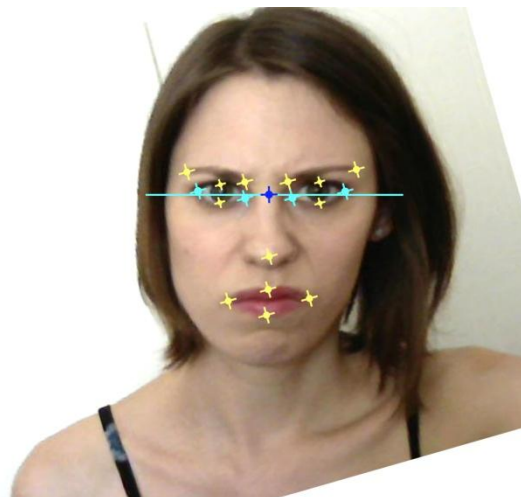


*προσανατολισμός σημείων  
μετά την περιστροφή*

**Σχήμα 5-8** Περιστροφή χαρακτηριστικών σημείων του προσώπου – έκφραση χαράς



*προσανατολισμός σημείων  
πριν την περιστροφή*



*προσανατολισμός σημείων  
μετά την περιστροφή*

**Σχήμα 5-9** Περιστροφή χαρακτηριστικών σημείων του προσώπου – έκφραση θυμού

## 5.4 Παράγωγα χαρακτηριστικά σημεία

Ο υπολογισμός κάποιων FAPs απαιτεί τη γνώση συγκεκριμένων σημείων του προσώπου που δεν παρέχονται από τη διαδικασία της ανίχνευσης. Αυτό συμβαίνει κατά τον υπολογισμό των FAPs  $F_{53} - F_{58}$ ,  $F_{19} - F_{22}$  και  $F_{33}$ ,  $F_{34}$  (δες επίσης περιγραφή των FAPs στον πίνακα 5-11). Η πλειονοψηφία των πρόσθετων σημείων στο πρόσωπο υπολογίζεται με γραμμική παρεμβολή, ως μέσος όρος κάποιων ήδη γνωστών σημείων. Εξαίρεση αποτελεί το σημείο 8.0.1 που ακολουθεί διαφορετικό τρόπο υπολογισμού για το τρέχον καρέ. Στο υπό εξέταση καρέ, όπου εικονίζεται (πιθανώς) μια συναισθηματική έκφραση, το στόμα μπορεί να έχει κάποιον ασύμμετρο μορφασμό (π.χ. στην έκφραση της αηδίας), οπότε το σημείο 8.0.1 κρίθηκε σκόπιμο να υπολογιστεί από τις συντεταγμένες που θα είχε σε σχέση με την άκρη της μύτης, που θεωρείται σταθερό σημείο, και όχι με γραμμική παρεμβολή.

Από τις συντεταγμένες των 19 χαρακτηριστικών σημείων εισόδου υπολογίζουμε 9 παράγωγα σημεία, ως εξής:

| Περιγραφή παράγωγου χαρακτηριστικού σημείου                             | Feature Point (f.p.) | Σχέσεις υπολογισμού  |                                      |
|---|----------------------|--|--------------------------------------|
| Μέσον στόματος<br>/ μέσος όρος άκρων στόματος<br>(εξωτερικό περίγραμμα) | 8.0.1                | neutral frame:   |                                      |
|   |                      | $8.0.1.x = \frac{8.3.x + 8.4.x}{2}$  | $8.0.1.y = \frac{8.3.y + 8.4.y}{2}$  |
|   |                      | current frame:   |                                      |
|   |                      | $8.0.1.x = 9.3.x + (8.0.1.x_{Neutral} - 9.3.x_{Neutral})$<br>$8.0.1.y = 9.3.y + (8.0.1.y_{Neutral} - 9.3.y_{Neutral})$ |                                      |
| Μέσος όρος αριστερής γωνίας στόματος & κέντρου πάνω χείλους             | 8.0.5                | $8.0.5.x = \frac{8.1.x + 8.3.x}{2}$  | $8.0.5.y = \frac{8.1.y + 8.3.y}{2}$  |
| Μέσος όρος δεξιάς γωνίας στόματος & κέντρου πάνω χείλους                | 8.0.6                | $8.0.6.x = \frac{8.1.x + 8.4.x}{2}$  | $8.0.6.y = \frac{8.1.y + 8.4.y}{2}$  |
| Μέσος όρος αριστερής γωνίας στόματος & κέντρου κάτω χείλους             | 8.0.7                | $8.0.7.x = \frac{8.2.x + 8.3.x}{2}$  | $8.0.7.y = \frac{8.2.y + 8.3.y}{2}$  |
| Μέσος όρος δεξιάς γωνίας στόματος & κέντρου κάτω χείλους                | 8.0.8                | $8.0.8.x = \frac{8.2.x + 8.4.x}{2}$  | $8.0.8.y = \frac{8.2.y + 8.4.y}{2}$  |
| Μέσον δεξιού ματιού<br>/ μέσος όρος άκρων δεξιού ματιού                 | 3.0.2                | $3.0.2.x = \frac{3.8.x + 3.12.x}{2}$   | $3.0.2.y = \frac{3.8.y + 3.12.y}{2}$ |
| Μέσον αριστερού ματιού / μέσος όρος άκρων αριστερού ματιού              | 3.0.1                | $3.0.1.x = \frac{3.7.x + 3.11.x}{2}$   | $3.0.1.y = \frac{3.7.y + 3.11.y}{2}$ |
| Μέσον αριστερού φρυδιού / μέσος όρος άκρων αριστερού φρυδιού            | 4.0.3                | $4.0.3.x = \frac{4.1.x + 4.5.x}{2}$  | $4.0.3.y = \frac{4.1.y + 4.5.y}{2}$  |
| Μέσον δεξιού φρυδιού / μέσος όρος άκρων αριστερού φρυδιού               | 4.0.4                | $4.0.4.x = \frac{4.2.x + 4.6.x}{2}$  | $4.0.4.y = \frac{4.2.y + 4.6.y}{2}$  |

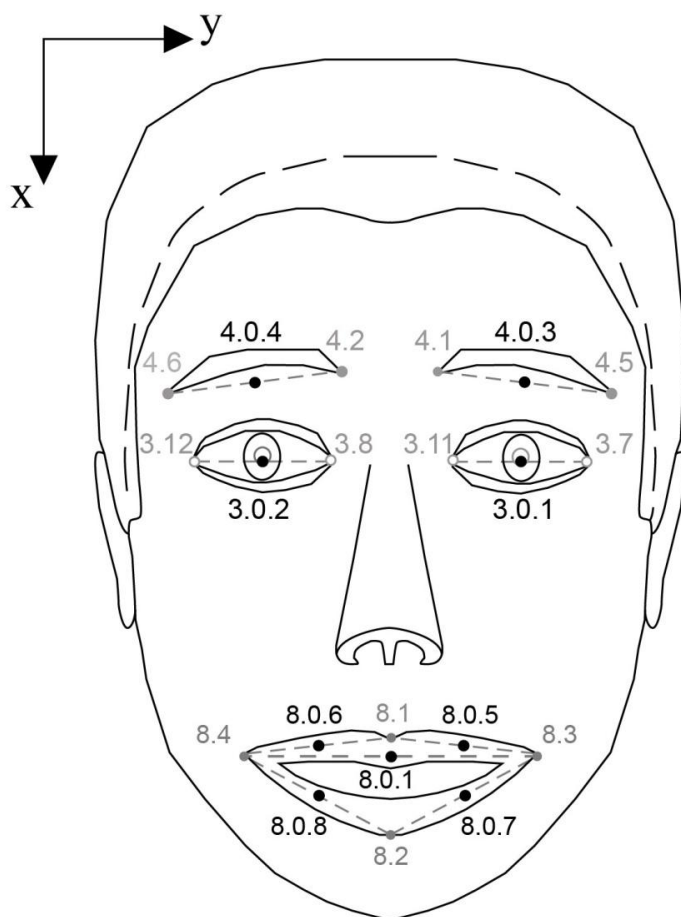
**Πίνακας 5-5** Παράγωγα χαρακτηριστικά σημεία

Ο δείκτης *Neutral* συμβολίζει το μέγεθος υπολογισμένο στην ουδέτερη κατάσταση.

Στους κωδικούς των παράγωγων σημείων το ενδιαμέσο ψηφίο 0 (π.χ. στο σημείο 8.0.1) χρησιμοποιήθηκε για να δηλώσει ότι πρόκειται για παράγωγο σημείο. Το πρώτο ψηφίο επιλέχθηκε ανάλογα με το χαρακτηριστικό του προσώπου με το οποίο σχετίζεται το παράγωγο σημείο. Το ψηφίο 8 αναφέρεται στο στόμα ενώ το ψηφίο 3 αναφέρεται στα μάτια. Το ψηφίο 4 αναφέρεται στα φρύδια. Το τρίτο ψηφίο απαριθμεί τα νέα σημεία. Για λόγους συνοχής, σημεία στο δεξί μέρος του προσώπου απαριθμούνται με ζυγά ψηφία ενώ σημεία του αριστερού μέρους έχουν μονά ψηφία. Για τα φρύδια, επιλέχθηκαν ως τρίτα ψηφία τα 3 και 4 για να υπάρχει αντιστοιχία ανάμεσα στα σημεία 4.0.3, 4.0.4 που προκύπτουν από γραμμική παρεμβολή και τα σημεία 4.3, 4.4 που βρίσκονται στο μέσον των φρυδιών σύμφωνα με το μοντέλο προσώπου του σχήματος 3-2, στη σελίδα 60. Παρομοίως για τα σημεία 8.0.5 – 8.0.8, διατηρείται η αντιστοιχία με τα σημεία 8.5 – 8.8 του ίδιου μοντέλου προσώπου.

Οι χαρακτηρισμοί «αριστερό» και «δεξί» έχουν το ίδιο σημείο αναφοράς (κατά την αντίληψη του εικονιζόμενου ατόμου) όπως και προηγουμένως .

Τα σημεία του πίνακα 5-5 δίνονται και στο παρακάτω σχήμα.



**Σχήμα 5-10** Παράγωγα σημεία ενδιαφέροντος ( Βάσει του [47] )

Με γκρι απεικονίζονται τα σημεία από τα οποία υπολογίζονται τα παράγωγα σημεία (με μαύρο χρώμα).

| Περιγραφή χαρακτηριστικού σημείου   | Κωδικός σημείου | Μεταβλητή στο πρόγραμμα |
|---|-----------------|-------------------------|
| Αριστερή γωνία (εσωτερική γωνία) δεξιού ματιού  | 3.8             | fp [ 0 ]                |
| Δεξιά γωνία (εξωτερική γωνία) δεξιού ματιού   | 3.12            | fp [ 1 ]                |
| Αριστερή γωνία (εξωτερική γωνία) αριστερού ματιού   | 3.7             | fp [ 2 ]                |
| Δεξιά γωνία (εσωτερική γωνία) αριστερού ματιού  | 3.11            | fp [ 3 ]                |
| Κέντρο ίριδας δεξιού ματιού   | 3.6             | fp [ 4 ]                |
| Κέντρο ίριδας αριστερού ματιού  | 3.5             | fp [ 5 ]                |
| Κέντρο κάτω βλεφάρου του δεξιού ματιού  | 3.4             | fp [ 6 ]                |
| Κέντρο πάνω βλεφάρου του δεξιού ματιού  | 3.2             | fp [ 7 ]                |
| Κέντρο κάτω βλεφάρου του αριστερού ματιού   | 3.3             | fp [ 8 ]                |
| Κέντρο πάνω βλεφάρου του αριστερού ματιού   | 3.1             | fp [ 9 ]                |
| Δεξιά γωνία στόματος (εξωτερικό περίγραμμα)   | 8.4             | fp [ 10 ]               |
| Αριστερή γωνία στόματος (εξωτερικό περίγραμμα)  | 8.3             | fp [ 11 ]               |
| Κάτω χείλος, κέντρο εξωτερικού περιγράμματος  | 8.2             | fp [ 12 ]               |
| Πάνω χείλος , κέντρο εξωτερικού περιγράμματος   | 8.1             | fp [ 13 ]               |
| Αριστερή (εσωτερική) γωνία δεξιού φρυδιού   | 4.2             | fp [ 14 ]               |
| Δεξιά (εξωτερική) γωνία δεξιού φρυδιού  | 4.6             | fp [ 15 ]               |
| Αριστερή (εξωτερική) γωνία αριστερού φρυδιού  | 4.5             | fp [ 16 ]               |
| Δεξιά (εσωτερική) γωνία αριστερού φρυδιού   | 4.1             | fp [ 17 ]               |
| Άκρη μύτης  | 9.3             | fp [ 18 ]               |
| Μέσον στόματος<br>/ Μέσος όρος άκρων στόματος   | 8.0.1           | fp [ 19 ]               |
| Μέσον δεξιού ματιού<br>/ Μέσος όρος άκρων δεξιού ματιού   | 3.0.2           | fp [ 20 ]               |
| Μέσον αριστερού ματιού<br>/ Μέσος όρος άκρων αριστερού ματιού                                     | 3.0.1           | fp [ 21 ]               |
| Μέσον αριστερού φρυδιού<br>/ Μέσος όρος άκρων αριστερού φρυδιού                                   | 4.0.3           | fp [ 22 ]               |
| Μέσον δεξιού φρυδιού<br>/ Μέσος όρος άκρων αριστερού φρυδιού                                      | 4.0.4           | fp [ 23 ]               |
| Μέσος όρος αριστερής γωνίας στόματος<br>και κέντρου πάνω χείλους                                  | 8.0.5           | fp [ 24 ]               |
| Μέσος όρος δεξιάς γωνίας στόματος<br>και κέντρου πάνω χείλους                                     | 8.0.6           | fp [ 25 ]               |
| Μέσος όρος αριστερής γωνίας στόματος<br>και κέντρου κάτω χείλους                                  | 8.0.7           | fp [ 26 ]               |
| Μέσος όρος δεξιάς γωνίας στόματος<br>και κέντρου κάτω χείλους                                     | 8.0.8           | fp [ 27 ]               |
| <b>Πίνακας 5-6</b> Αντιστοιχία χαρακτηριστικών σημείων με μεταβλητές αποθήκευσης του προγράμματος |                 |                         |



Η αντιστοιχία των σημείων με τη μεταβλητή αποθήκευσής τους στο πρόγραμμα, δίνεται στον προηγούμενο πίνακα (Πίνακας 5-6). Συνολικά διαθέτουμε 28 σημεία, από τα οποία υπολογίζονται τα 23 FAPs.

## 5.5 Υπολογισμός μονάδων των FAPs

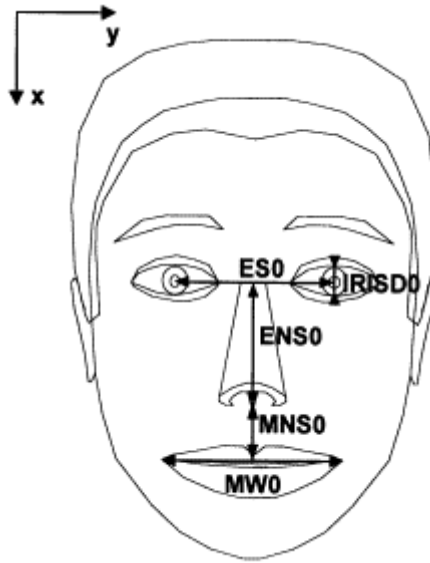
Καθώς τα FAPs μπορούν να αναφέρονται σε πρόσωπα διαφόρων διαστάσεων και αναλογιών, οι τιμές των FAPs ορίζονται σε σύστημα μονάδων FAPU (FAP Units). Οι μονάδες FAPU χρειάζονται για την κανονικοποίηση των αποτελεσμάτων κατά τη διαδικασία υπολογισμού των FAPs. Τα FAPU προκύπτουν από σταθερές οριζόντιες, κάθετες (ή εναλλακτικά ευκλείδειες) αποστάσεις σημείων στο πρόσωπο και υπολογίζονται από το καρέ που απεικονίζει το πρόσωπο σε ουδέτερη κατάσταση. Οι σχέσεις ορισμού που χρησιμοποιήθηκαν δίνονται στον παρακάτω πίνακα:

| Πίνακας 5-7 Ορισμοί Μονάδων FAPU: FAP Units |  |   |  |
|---|--|---|--|
| Ονομασία και Περιγραφή                      |  | Τρόπος υπολογισμού<br>(ουδέτερο καρέ)         | Κανονικοποιημένη μονάδα<br>FAPU με Normalization factor<br>n.f.=1024 |
| IRISD0                                      | Διάμετρος ίριδας<br>(Iris diameter)                | $\frac{(3.3.x - 3.1.x) + (3.4.x - 3.2.x)}{2}$ | IRISD = IRISD0 / n.f.  |
| ESO   | Απόσταση ματιών<br>(Eye separation)                | $3.5.y - 3.6.y$                               | ES = ESO / n.f.  |
| ENS0  | Απόσταση ματιών-μύτης<br>(Eye-nose separation)     | $9.3.x - \frac{3.5.x + 3.6.x}{2}$             | ENS = ENS0 / n.f.  |
| MNS0  | Απόσταση μύτης-στόματος<br>(Mouth-nose separation) | $\frac{8.3.x + 8.4.x}{2} - 9.3.x$             | MNS = MNS0 / n.f.  |
| MW0   | Μήκος στόματος<br>(Mouth width)                    | $8.3.y - 8.4.y$                               | MW = MW0 / n.f.  |

Αρχικά υπολογίζονται οι αποστάσεις της 1<sup>ης</sup> στήλης χρησιμοποιώντας τους τύπους της 3<sup>ης</sup> στήλης. Στη συνέχεια οι μονάδες κανονικοποιούνται μέσω της διαίρεσης με τον παράγοντα n.f.=1024 (normalization factor). Οι αποστάσεις της 1<sup>ης</sup> στήλης που χρειάζονται για τον υπολογισμό των FAPU, δίνονται και στο σχήμα 5-11.

Σημειώνεται ότι ο τρόπος υπολογισμού των FAPU που υλοποιήθηκε είναι ελαφρώς διαφορετικός από τον τρόπο υπολογισμού που παρουσιάστηκε στο κεφάλαιο 3 για το MPEG-4, στην παράγραφο 3.6.4 (πίνακας 3-3). Στη θεωρητική περίπτωση, υπάρχει συμμετρία ανάμεσα στα ομόλογα χαρακτηριστικά του προσώπου. Για παράδειγμα τα άκρα του στόματος βρίσκονται στο ίδιο ύψος (ίδια κάθετη συντ/νη). Επειδή το πρόγραμμα απευθύνεται σε φυσικά πρόσωπα, έχουν χρησιμοποιηθεί μέσοι όροι στους παραπάνω τύπους λόγω των ομόλογων χαρακτηριστικών.

Στο παρακάτω σχήμα δίνονται οι αποστάσεις που χρειάζονται για τα FAPU.



Σχήμα 5-11 Αποστάσεις για τις μονάδες των FAPs [47]

## 5.6 «Λεξιλόγιο» FAPs της εφαρμογής

Το πρότυπο MPEG-4 ορίζει 66 FAPs χαμηλού επιπέδου και 2 FAPs υψηλού επιπέδου. Από τα 66 FAPs χαμηλού επιπέδου, 23 υπολογίζονται στα πλαίσια της παρούσας εργασίας. Χρησιμοποιείται ο συμβολισμός  $F_j$  για την αναφορά σε κάποιο FAP, όπου ο δείκτης  $j$  λαμβάνει ακέραιες τιμές. Επίσης τα FAPs μπορούν να αναφέρονται με τα ονόματά τους που προκύπτουν από σύντμηση της λεκτικής περιγραφής τους. Η υλοποίηση των FAPs έγινε βάσει του άρθρου “Parameterized facial expression synthesis based on MPEG-4” των Α. Ραουζαίου και συνεργατών [29]. Υλοποιήθηκαν όσα FAPs ήταν μέρος του λεξιλογίου για τη δημιουργία των προφίλ των αρχέτυπων συναισθημάτων στο προαναφερθέν άρθρο. Εξαίρεση είναι μόνο τα FAPs  $F_{16}$ ,  $F_{18}$  και  $F_{41}$ ,  $F_{42}$ , τα οποία δεν υπολογίζονται καθώς δεν είναι διαθέσιμη πληροφορία που απαιτείται για τον υπολογισμό τους, όπως το βάθος και τα μάγουλα του εικονιζόμενου προσώπου. Λεπτομέρειες για αυτά τα FAPs δίνονται στον πίνακα 5-8 που ακολουθεί.

Από τα υπόλοιπα 41 FAPs, κάποια δεν μπορούν να υπολογιστούν γιατί παρουσιάζουν τις ίδιες αδυναμίες όπως εξηγήθηκε προηγουμένως, δηλαδή έλλειψη πληροφορίας για το βάθος και τις μεταβολές στο βάθος των σημείων, ή έλλειψη πληροφορίας λόγω της μη-ύπαρξης σημείων που αφορούν τα χαρακτηριστικά που σχετίζονται με τα συγκεκριμένα FAPs. Κάποια άλλα FAPs δεν έχουν ενδιαφέρον να υπολογιστούν στα πλαίσια της εκτίμησης συναισθημάτων, καθώς έχουν ρόλο κυρίως στη δημιουργία συνθετικών εικόνων προσώπων, και κατ'επέκταση στη σύνθεση συναισθημάτων παρά στην αναγνώριση. Παραδείγματα τέτοιων FAPs (π.χ.  $F_{14}$ ,  $F_{29}$  και  $F_{67}$ ) δίνονται παρακάτω (Πίνακας 5-8).

| Πίνακας 5-8 FAPs που δεν μπορούν να υλοποιηθούν                      |                |   |  |
|--|----------------|---|--|
| FAP No.  | FAP Name       | Περιγραφή   | Αδυναμία Υλοποίησης                                  |
| F <sub>16</sub>  | push_b_lip     | Μεταβολή βάθους του κάτω χείλους (σημείο στο μέσον) | Έλλειψη πληροφορίας βάθους για τα σημεία των χειλιών |
| F <sub>18</sub>  | depress_chin   | Ανοδική κίνηση με συμπίεση του πηγουνιού            |  |
| F <sub>41</sub>  | lift_l_cheek   | Κάθετη μετατόπιση, αριστερό μάγουλο                 | Έλλειψη πληροφορίας σημείων για τα μάγουλα           |
| F <sub>42</sub>  | lift_r_cheek   | Κάθετη μετατόπιση, δεξί μάγουλο                     |  |
| Παραδείγματα FAPs που δεν μπορούν ή/και δε χρειάζεται να υλοποιηθούν |                |   |  |
| F <sub>14</sub>  | thrust_jaw     | Μεταβολή βάθους του (κάτω) σιαγώνα                  | Έλλειψη πληροφορίας βάθους                           |
| F <sub>29</sub>  | dilate_l_pupil | Διαστολή κόρης αριστερού ματιού                     | Έλλειψη σημείων                                      |
| F <sub>67</sub>  | pull_l_ear     | Οριζόντια μετατόπιση αριστερού αφτιού               | Έλλειψη σημείων /Χρησιμότητα στη σύνθεση εικόνων     |

Στον παρακάτω πίνακα (5-9), παρουσιάζεται αναλυτικά το «λεξιλόγιο» της εφαρμογής, δηλαδή όλα τα διαθέσιμα FAPs. Εκτός από τα 23 FAPs υπολογίζονται και κάποιοι συνδυασμοί FAPs, σε αντιστοιχία με το άρθρο [29]. Οι τέσσερις συνδυασμοί δίνονται στις τελευταίες γραμμές του πίνακα.

Στον πίνακα δίνεται ο συμβολισμός  $F_j$ , η ομάδα στην οποία ανήκει το FAP καθώς και το χαρακτηριστικό του προσώπου στο οποίο αναφέρεται, το όνομα και μια σύντομη περιγραφή.

Τα FAPs F<sub>51</sub> έως F<sub>60</sub> αναφέρονται στις κινήσεις των σημείων του εξωτερικού περιγράμματος των χειλιών. Υπάρχει αντιστοιχία των FAPs F<sub>51</sub> - F<sub>60</sub> με τα FAPs F<sub>4</sub> - F<sub>13</sub>, που αναφέρονται στις ίδιες κινήσεις του εσωτερικού αντί του εξωτερικού περιγράμματος των χειλιών. Στο άρθρο [29] στο οποίο στηρίχθηκε η παρούσα υλοποίηση, οι κανόνες των μοντελοποιημένων προφίλ περιλαμβάνουν τα FAPs F<sub>4</sub> - F<sub>13</sub> και F<sub>53</sub>, F<sub>54</sub>. Δεδομένου ότι η διαδικασία ανίχνευσης παρέχει μόνο τα σημεία του εξωτερικού περιγράμματος των χειλιών, δεν υπάρχει τρόπος υπολογισμού των FAPs F<sub>4</sub> - F<sub>13</sub> με βάση τα σημεία του εσωτερικού περιγράμματος. Για τις ανάγκες της υλοποίησης, αν ο χρήστης εισάγει κανόνες που χρησιμοποιούν τα F<sub>4</sub> - F<sub>13</sub>, οι τιμές αυτών των FAPs κατά την αποτίμηση των κανόνων θα δίνονται από τις ίδιες σχέσεις με τα F<sub>51</sub> - F<sub>60</sub>. Αυτή η παραδοχή δε θεωρείται ότι προκαλεί μεγάλα σφάλματα, καθώς τα άκρα στόματος εσωτερικού ή εξωτερικού περιγράμματος των χειλιών σχεδόν ταυτίζονται (σημείο 8.4 με 2.5, όπως επίσης σημείο 8.3 με 2.4).

**Πίνακας 5-9** Λεξιλόγιο FAPs της εφαρμογής

| Group                   | FAP No.                          | FAP Name                                      | Περιγραφή  |
|-------------------------|----------------------------------|---|--|
| Jaw (2)                 | F <sub>3</sub>                   | open_jaw                                      | Άνοιγμα σιαγόνων   |
| Eyelids (3)             | F <sub>19</sub>                  | close_t_l_eyelid                              | Κλείσιμο επάνω βλεφάρου αριστερού ματιού   |
|                         | F <sub>20</sub>                  | close_t_r_eyelid                              | Κλείσιμο επάνω βλεφάρου δεξιού ματιού  |
|                         | F <sub>21</sub>                  | close_b_l_eyelid                              | Κλείσιμο κάτω βλεφάρου αριστερού ματιού  |
|                         | F <sub>22</sub>                  | close_b_r_eyelid                              | Κλείσιμο κάτω βλεφάρου δεξιού ματιού   |
| Eyebrows (4)            | F <sub>31</sub>                  | raise_l_i_eyebrow                             | Σήκωμα εσωτ. γωνίας αριστερού φρυδιού  |
|                         | F <sub>32</sub>                  | raise_r_i_eyebrow                             | Σήκωμα εσωτ. γωνίας δεξιού φρυδιού   |
|                         | F <sub>33</sub>                  | raise_l_m_eyebrow                             | Σήκωμα κέντρου αριστερού φρυδιού   |
|                         | F <sub>34</sub>                  | raise_r_m_eyebrow                             | Σήκωμα κέντρου δεξιού φρυδιού  |
|                         | F <sub>35</sub>                  | raise_l_o_eyebrow                             | Σήκωμα εξωτ. γωνίας αριστερού φρυδιού  |
|                         | F <sub>36</sub>                  | raise_r_o_eyebrow                             | Σήκωμα εξωτ. γωνίας δεξιού φρυδιού   |
|                         | F <sub>37</sub>                  | squeeze_l_eyebrow                             | Συμπίεση αριστερού φρυδιού   |
|                         | F <sub>38</sub>                  | squeeze_r_eyebrow                             | Συμπίεση δεξιού φρυδιού  |
| Outer lip positions (8) | F <sub>51</sub>                  | lower_t_midlip_o                              | Χαμήλωμα κέντρου (εσωτερικού περιγράμματος) επάνω χείλους  |
|                         | F <sub>52</sub>                  | raise_b_midlip_o                              | Σήκωμα κέντρου (εσωτερικού περιγράμματος) κάτω χείλους   |
|                         | F <sub>53</sub>                  | stretch_l_cornerlip_o                         | Τέντωμα αριστερής γωνίας στόματος (εξωτερικό περίγραμμα)   |
|                         | F <sub>54</sub>                  | stretch_r_cornerlip_o                         | Τέντωμα δεξιάς γωνίας στόματος (εξωτερικό περίγραμμα)  |
|                         | F <sub>55</sub>                  | lower_t_lip_lm_o                              | Χαμήλωμα σημείου χείλους που βρίσκεται μεταξύ αριστερής γωνίας στόματος και κέντρου πάνω χείλους (στο μέσον της απόστασης) |
|                         | F <sub>56</sub>                  | lower_t_lip_rm_o                              | Χαμήλωμα σημείου χείλους που βρίσκεται μεταξύ δεξιάς γωνίας στόματος και κέντρου πάνω χείλους (στο μέσον της απόστασης)    |
|                         | F <sub>57</sub>                  | raise_b_lip_lm_o                              | Σήκωμα σημείου χείλους που βρίσκεται μεταξύ αριστερής γωνίας στόματος και κέντρου κάτω χείλους (στο μέσον της απόστασης)   |
|                         | F <sub>58</sub>                  | raise_b_lip_rm_o                              | Σήκωμα σημείου χείλους που βρίσκεται μεταξύ δεξιάς γωνίας στόματος και κέντρου κάτω χείλους (στο μέσον της απόστασης)      |
|                         | F <sub>59</sub>                  | raise_l_cornerlip_o                           | Σήκωμα αριστερής γωνίας στόματος   |
|                         | F <sub>60</sub>                  | raise_r_cornerlip_o                           | Σήκωμα δεξιάς γωνίας στόματος  |
| Combinations – No group | F <sub>19</sub> +F <sub>21</sub> | close_t_l_eyelid + close_b_l_eyelid           | Κλείσιμο αριστερού ματιού: συνδυασμός κλείσιμου επάνω και κάτω βλεφάρου αριστερού ματιού                                   |
|                         | F <sub>20</sub> +F <sub>22</sub> | close_t_r_eyelid + close_b_r_eyelid           | Κλείσιμο δεξιού ματιού: συνδυασμός κλείσιμου επάνω και κάτω βλεφάρου δεξιού ματιού   |
|                         | F <sub>53</sub> +F <sub>54</sub> | stretch_l_cornerlip_o + stretch_r_cornerlip_o | Τέντωμα στόματος: τέντωμα δεξιάς και αριστερής γωνίας στόματος (εξωτερικό περίγραμμα)                                      |
|                         | F <sub>37</sub> +F <sub>38</sub> | squeeze_l_eyebrow + squeeze_r_eyebrow         | Σμίξιμο φρυδιών: συμπίεση δεξιού και αριστερού φρυδιού   |

Επίσης τα σημεία 8.1 με 2.2 και 8.2 με 2.3 βρίσκονται σε κοντινές αποστάσεις μεταξύ τους, σε σύγκριση με το μέγεθος ολόκληρου του προσώπου, και ακολουθούν τις ίδιες κινήσεις χωρίς η απόσταση μεταξύ τους να μεταβάλλεται σημαντικά. Εάν η διαδικασία ανίχνευσης παρείχε σημεία του εσωτερικού περιγράμματος των χειλιών, τότε εύκολα θα υπολογίζαμε τα FAPs  $F_4 - F_{13}$  από παρόμοιες σχέσεις με τα  $F_{51} - F_{60}$ . Αυτές δίνονται σε ξεχωριστό πίνακα (Πίνακας 5-13, στη σελίδα 112), και μπορούν να αποτελέσουν αντικείμενο μελλοντικής επέκτασης του κώδικα.

**Πίνακας 5-10** Αντιστοιχία FAPs  $F_4 - F_{13}$  με  $F_{51} - F_{60}$

| Group                   | FAP No.  | FAP Name              |  | FAP No.  | FAP Name            | Group                   |
|-------------------------|----------|-----------------------|--|----------|---------------------|-------------------------|
|                         |          |                       |  |          |                     |                         |
| Outer lip positions (8) | $F_{51}$ | lower_t_midlip_o      |  | $F_4$    | lower_t_midlip      | Inner lip positions (2) |
|                         | $F_{52}$ | raise_b_midlip_o      |  | $F_5$    | raise_b_midlip      |                         |
|                         | $F_{53}$ | stretch_l_cornerlip_o |  | $F_6$    | stretch_l_cornerlip |                         |
|                         | $F_{54}$ | stretch_r_cornerlip_o |  | $F_7$    | stretch_r_cornerlip |                         |
|                         | $F_{55}$ | lower_t_lip_lm_o      |  | $F_8$    | lower_t_lip_lm      |                         |
|                         | $F_{56}$ | lower_t_lip_rm_o      |  | $F_9$    | lower_t_lip_rm      |                         |
|                         | $F_{57}$ | raise_b_lip_lm_o      |  | $F_{10}$ | raise_b_lip_lm      |                         |
|                         | $F_{58}$ | raise_b_lip_rm_o      |  | $F_{11}$ | raise_b_lip_rm      |                         |
|                         | $F_{59}$ | raise_l_cornerlip_o   |  | $F_{12}$ | raise_l_cornerlip   |                         |
|                         | $F_{60}$ | raise_r_cornerlip_o   |  | $F_{13}$ | raise_r_cornerlip   |                         |

Από τα παραπάνω είναι εμφανής η σχέση μεταξύ των FAPs  $F_4 - F_{13}$  και  $F_{51} - F_{60}$ . Υπάρχει αντιστοιχία ανάμεσα στη σειρά και τα ονόματα των FAPs, με μόνη διαφορά την κατάληξη \_o (συντομογραφία του outer) που παρατίθεται στο τέλος των ονομάτων των FAPs  $F_{51} - F_{60}$ , καθώς αυτά τα FAPs αναφέρονται στο εξωτερικό περίγραμμα των χειλιών.

## 5.7 Υπολογισμός των FAPs

Στον πίνακα 5-11 δίνεται ο συμβολισμός  $F_j$ , το αντίστοιχο όνομα των FAPs, οι αποστάσεις  $D_i$  και τα χαρακτηριστικά  $f_i$  που χρειάζονται στον υπολογισμό της τιμής του  $F_j$ . Επίσης δίνεται η μονάδα FAPU για κάθε FAP του πίνακα. Οι δείκτες για τα  $D_i$  και  $f_i$  είναι οι δείκτες που χρησιμοποιούνται και στο πρόγραμμα για την αποθήκευση των αντίστοιχων τιμών.

Το πρώτο βήμα για τον υπολογισμό των FAPs είναι ο υπολογισμός των αποστάσεων  $D_i$  στο ουδέτερο και στο υπό εξέταση καρέ. Οι αποστάσεις  $D_i$  είναι είτε κάθετες, είτε οριζόντιες ή ευκλείδειες αποστάσεις ανάμεσα σε σημεία του προσώπου. Για το ουδέτερο καρέ χρησιμοποιείται ο συμβολισμός  $D_{i-NEUTRAL}$ .

**Πίνακας 5-11 Υπολογισμός των FAPs**

| Group                   | FAP No.         | FAP Name                                      | Χαρακτηριστικό περιγραφής $D_i$ | Υλοποίηση χαρακτηριστικού $f_i$    | FAP Unit | Μεταβλητή |
|-------------------------|-----------------|---|---------------------------------|------------------------------------|----------|-----------|
| Jaw (2)                 | $F_3$           | open_jaw                                      | $D_{11} = d(8.1, 8.2)$          | $f_{11} = D_{11} - D_{11-NEUTRAL}$ | MNS      | fap[11]   |
| Eyelids (3)             | $F_{19}$        | close_t_l_eyelid                              | $D_{20} = d(3.13, 3.0.1)$       | $f_{20} = D_{20-NEUTRAL} - D_{20}$ | IRISD    | fap[20]   |
|                         | $F_{20}$        | close_t_r_eyelid                              | $D_{21} = d(3.14, 3.0.2)$       | $f_{21} = D_{21-NEUTRAL} - D_{21}$ | IRISD    | fap[21]   |
|                         | $F_{21}$        | close_b_l_eyelid                              | $D_{22} = d(3.9, 3.0.1)$        | $f_{22} = D_{22-NEUTRAL} - D_{22}$ | IRISD    | fap[22]   |
|                         | $F_{22}$        | close_b_r_eyelid                              | $D_{23} = d(3.10, 3.0.2)$       | $f_{23} = D_{23-NEUTRAL} - D_{23}$ | IRISD    | fap[23]   |
| Eyebrows (4)            | $F_{31}$        | raise_l_i_eyebrow                             | $D_5 = d(4.1, 3.11)$            | $f_5 = D_5 - D_{5-NEUTRAL}$        | ENS      | fap[5]    |
|                         | $F_{32}$        | raise_r_i_eyebrow                             | $D_6 = d(4.2, 3.8)$             | $f_6 = D_6 - D_{6-NEUTRAL}$        | ENS      | fap[6]    |
|                         | $F_{33}$        | raise_l_m_eyebrow                             | $D_9 = d(4.0.3, 3.7)$           | $f_9 = D_9 - D_{9-NEUTRAL}$        | ENS      | fap[9]    |
|                         | $F_{34}$        | raise_r_m_eyebrow                             | $D_{10} = d(4.0.4, 3.12)$       | $f_{10} = D_{10} - D_{10-NEUTRAL}$ | ENS      | fap[10]   |
|                         | $F_{35}$        | raise_l_o_eyebrow                             | $D_7 = d(4.5, 3.7)$             | $f_7 = D_7 - D_{7-NEUTRAL}$        | ENS      | fap[7]    |
|                         | $F_{36}$        | raise_r_o_eyebrow                             | $D_8 = d(4.6, 3.12)$            | $f_8 = D_8 - D_{8-NEUTRAL}$        | ENS      | fap[8]    |
|                         | $F_{37}$        | squeeze_l_eyebrow                             | $D_1 = d(4.5, 3.11)$            | $f_1 = D_1 - D_{1-NEUTRAL}$        | ES       | fap[1]    |
|                         | $F_{38}$        | squeeze_r_eyebrow                             | $D_2 = d(4.6, 3.8)$             | $f_2 = D_2 - D_{2-NEUTRAL}$        | ES       | fap[2]    |
| Outer lip positions (8) | $F_{51}$        | lower_t_midlip_o                              | $D_3 = d(9.3, 8.1)$             | $f_3 = D_3 - D_{3-NEUTRAL}$        | MNS      | fap[3]    |
|                         | $F_{52}$        | raise_b_midlip_o                              | $D_4 = d(9.3, 8.2)$             | $f_4 = D_4 - D_{4-NEUTRAL}$        | MNS      | fap[4]    |
|                         | $F_{53}$        | stretch_l_cornerlip_o                         | $D_{16} =  8.3.y - 8.0.1.y $    | $f_{16} = D_{16} - D_{16-NEUTRAL}$ | MW       | fap[16]   |
|                         | $F_{54}$        | stretch_r_cornerlip_o                         | $D_{17} =  8.4.y - 8.0.1.y $    | $f_{17} = D_{17} - D_{17-NEUTRAL}$ | MW       | fap[17]   |
|                         | $F_{55}$        | lower_t_lip_lm_o                              | $D_{24} =  8.0.5.x - 9.3.x $    | $f_{24} = D_{24} - D_{24-NEUTRAL}$ | MNS      | fap[24]   |
|                         | $F_{56}$        | lower_t_lip_rm_o                              | $D_{25} =  8.0.6.x - 9.3.x $    | $f_{25} = D_{25} - D_{25-NEUTRAL}$ | MNS      | fap[25]   |
|                         | $F_{57}$        | raise_b_lip_lm_o                              | $D_{26} =  8.0.7.x - 9.3.x $    | $f_{26} = D_{26-NEUTRAL} - D_{26}$ | MNS      | fap[26]   |
|                         | $F_{58}$        | raise_b_lip_rm_o                              | $D_{27} =  8.0.8.x - 9.3.x $    | $f_{27} = D_{27-NEUTRAL} - D_{27}$ | MNS      | fap[27]   |
|                         | $F_{59}$        | raise_l_cornerlip_o                           | $D_{18} =  8.3.x - 9.3.x $      | $f_{18} = D_{18-NEUTRAL} - D_{18}$ | MNS      | fap[18]   |
|                         | $F_{60}$        | raise_r_cornerlip_o                           | $D_{19} =  8.4.x - 9.3.x $      | $f_{19} = D_{19-NEUTRAL} - D_{19}$ | MNS      | fap[19]   |
| Combinations – No group | $F_{19}+F_{21}$ | close_t_l_eyelid + close_b_l_eyelid           | $D_{12} = d(3.1, 3.3)$          | $f_{12} = D_{12} - D_{12-NEUTRAL}$ | IRISD    | fap[12]   |
|                         | $F_{20}+F_{22}$ | close_t_r_eyelid + close_b_r_eyelid           | $D_{13} = d(3.2, 3.4)$          | $f_{13} = D_{13} - D_{13-NEUTRAL}$ | IRISD    | fap[13]   |
|                         | $F_{53}+F_{54}$ | stretch_l_cornerlip_o + stretch_r_cornerlip_o | $D_{14} = d(8.3, 8.4)$          | $f_{14} = D_{14} - D_{14-NEUTRAL}$ | MW       | fap[14]   |
|                         | $F_{37}+F_{38}$ | squeeze_l_eyebrow + squeeze_r_eyebrow         | $D_{15} = d(4.5, 4.6)$          | $f_{15} = D_{15-NEUTRAL} - D_{15}$ | ES       | fap[15]   |

Με  $d(p_1, p_2)$  συμβολίζεται η ευκλείδεια απόσταση των σημείων  $p_1$  και  $p_2$ .

Με  $|x|$  συμβολίζεται η απόλυτη τιμή του αριθμού  $x$ .

Ο όρος  $D_{F-NEUTRAL}$  συμβολίζει την απόσταση  $D_i$  στην ουδέτερη κατάσταση.

Ο τρόπος υπολογισμού των  $D_i$  δίνεται στην αντίστοιχη στήλη του παραπάνω πίνακα. Η σύγκριση της κάθε απόστασης  $D_i$  με την απόσταση  $D_{i-NEUTRAL}$  στο ουδέτερο καρέ παρέχει πληροφορίες για τις κινήσεις των χαρακτηριστικών ως συνέπεια της έκφρασης του συναισθήματος στο πρόσωπο. Χρησιμοποιούμε τη διαφορά των  $D_i$ , η οποία συμβολίζεται με  $f_i$ , και η οποία υποδεικνύει ότι το αντίστοιχο FAP είναι ενεργό. Η διαφορά ανάμεσα στα  $D_i$ ,  $D_{i-NEUTRAL}$  είναι μια μεταβολή με κατευθυντικότητα. Ανάλογα με την περιγραφή του FAP, (στήλη FAP Name στον πίνακα) το  $f_i$  θα εκφράζεται είτε ως  $f_i = D_i - D_{i-NEUTRAL}$  είτε ως  $f_i = D_{i-NEUTRAL} - D_i$ . Η σειρά αφαίρεσης των δύο όρων εξαρτάται από την κατεύθυνση κίνησης που έχει οριστεί ως θετική. Αυτό βρίσκεται εύκολα από το όνομα του FAP. Το κάθε όνομα δηλώνει ότι η θετική φορά της κίνησης είναι εκείνη η κατεύθυνση που ορίζεται όταν πραγματοποιείται η αντίστοιχη κίνηση που περιγράφεται λεκτικά (π.χ. lower, raise, close, stretch). Ο μειωτέος και αφαιρετέος στη σχέση για το  $f_i$  επιλέγονται ώστε να υπάρχει αντιστοιχία του προσήμου με τη λεκτική περιγραφή.

Ο υπολογισμός του κάθε FAP ολοκληρώνεται με την κανονικοποίηση της τιμής  $f_i$  που υπολογίστηκε, διαιρώντας με την αντίστοιχη μονάδα κανονικοποίησης:

$$F_j = \frac{f_i}{MNS} = \frac{D_i - D_{i-NEUTRAL}}{MNS}$$

Ακολουθεί ένα παράδειγμα υπολογισμού ενός FAP, στην προκειμένη περίπτωση του  $F_3$ , για να γίνει κατανοητή η μέθοδος υπολογισμού.

|  |
|--|
| <b>Παράδειγμα: Υπολογισμός του FAP <math>F_3</math></b>  |
| <p><b>Βήμα 1<sup>ο</sup>:</b></p> <p>Το FAP <math>F_3</math> ποσοτικοποιεί το άνοιγμα των σιαγόνων (open_jaw). Αρχικά υπολογίζεται η απόσταση <math>D_{11-NEUTRAL}</math> από τα σημεία του καρέ αναφοράς:</p> $D_{11-NEUTRAL} = d(8.1, 8.2) _{NEUTRAL} = \sqrt{(8.1.x - 8.2.x)^2 + (8.1.y - 8.2.y)^2} \Big _{NEUTRAL}$ <p>Στη συνέχεια υπολογίζεται η ίδιο απόσταση στο εξεταζόμενο καρέ:</p> $D_{11} = d(8.1, 8.2) = \sqrt{(8.1.x - 8.2.x)^2 + (8.1.y - 8.2.y)^2}$ |
| <p><b>Βήμα 2<sup>ο</sup>:</b></p> <p>Από το όνομα και την περιγραφή του FAP <math>F_3</math> συμπεραίνουμε ότι οι θετικές τιμές θα αντιστοιχούν σε άνοιγμα των σιαγόνων, δηλαδή αύξηση της απόστασης <math>D_{11}</math>. Για να συμβαδίζει το πρόσημο με την περιγραφή του FAP, το <math>f_{11}</math> θα δίνεται από</p> $f_{11} = D_{11} - D_{11-NEUTRAL}$  |
| <p><b>Βήμα 3<sup>ο</sup>:</b></p> <p>Η μονάδα κανονικοποίησης FAPU για το <math>F_3</math> είναι η <math>MNS</math>. (Οι μονάδες FAPU έχουν υπολογιστεί εκ των προτέρων). Οπότε η τιμή για το <math>F_3</math> δίνεται εν τέλει από την εξίσωση:</p> $F_3 = \frac{f_{11}}{MNS} = \frac{D_{11} - D_{11-NEUTRAL}}{MNS}$  |

**Πίνακας 5-12** Παράδειγμα υπολογισμού ενός FAP

Στον επόμενο πίνακα παρατίθενται οι σχέσεις υπολογισμού των FAPs  $F_4 - F_{13}$  εάν τα σημεία 2.2 – 2.9 του εσωτερικού περιγράμματος των χειλιών ήταν γνωστά.

**Πίνακας 5-13** Υπολογισμός των FAPs  $F_4 - F_{13}$

| Group  | FAP No.  | FAP Name            | Χαρακτηριστικό περιγραφής $D_i$      | Υλοποίηση χαρακτηριστικού $f_i$    | FAP Unit |
|--|----------|---------------------|--------------------------------------|------------------------------------|----------|
| Inner lip positions (2)  | $F_4$    | lower_t_midlip      | $D_4 = d(9.3, 8.1)$                  | $f_4 = D_4 - D_{4-NEUTRAL}$        | MNS      |
|  | $F_5$    | raise_b_midlip      | $D_5 = d(9.3, 8.2)$                  | $f_5 = D_{5-NEUTRAL} - D_5$        | MNS      |
|  | $F_6$    | stretch_l_cornerlip | $D_6 = \text{abs}(2.4.y - 2.0.1.y)$  | $f_6 = D_6 - D_{6-NEUTRAL}$        | MW       |
|  | $F_7$    | stretch_r_cornerlip | $D_7 = \text{abs}(2.5.y - 2.0.1.y)$  | $f_7 = D_7 - D_{7-NEUTRAL}$        | MW       |
|  | $F_8$    | lower_t_lip_lm      | $D_8 = \text{abs}(2.6.x - 9.3.x)$    | $f_8 = D_8 - D_{8-NEUTRAL}$        | MNS      |
|  | $F_9$    | lower_t_lip_rm      | $D_9 = \text{abs}(2.7.x - 9.3.x)$    | $f_9 = D_9 - D_{9-NEUTRAL}$        | MNS      |
|  | $F_{10}$ | raise_b_lip_lm      | $D_{10} = \text{abs}(2.8.x - 9.3.x)$ | $f_{10} = D_{10-NEUTRAL} - D_{10}$ | MNS      |
|  | $F_{11}$ | raise_b_lip_rm      | $D_{11} = \text{abs}(2.9.x - 9.3.x)$ | $f_{11} = D_{11-NEUTRAL} - D_{11}$ | MNS      |
|  | $F_{12}$ | raise_l_cornerlip   | $D_{12} = \text{abs}(2.4.x - 9.3.x)$ | $f_{12} = D_{12-NEUTRAL} - D_{12}$ | MNS      |
|  | $F_{13}$ | raise_r_cornerlip   | $D_{13} = \text{abs}(2.5.x - 9.3.x)$ | $f_{13} = D_{13-NEUTRAL} - D_{13}$ | MNS      |
| <p>Με <math>d(p_1, p_2)</math> συμβολίζεται η ευκλείδεια απόσταση των σημείων <math>p_1</math> και <math>p_2</math>.<br/> Με <math>\text{abs}(x)</math> συμβολίζεται η απόλυτη τιμή του αριθμού <math>x</math>.<br/> Ο όρος <math>D_{i-NEUTRAL}</math> συμβολίζει την απόσταση <math>D_i</math> στην ουδέτερη κατάσταση.<br/> Υπάρχει αντιστοιχία ανάμεσα στα FAPs <math>F_4-F_{13}</math> και <math>F_{51}-F_{60}</math>.</p> |          |                     |                                      |                                    |          |

#### Παρατηρήσεις:

- Οι δείκτες των  $D_i$  και  $f_i$  δεν έχουν κάποια συγκεκριμένη έννοια, οπότε επιλέγησαν οι ίδιοι δείκτες με τους δείκτες των FAPs.
- Τα σημεία 2.2 – 2.9 δεν ακολουθούν τη σύμβαση των υπόλοιπων σημείων όπου τα μονά ψηφία στο τέλος δηλώνουν τη θέση στο αριστερό μέρος του προσώπου (κατά την αντίληψη του εικονιζόμενου) και τα ζυγά τη θέση στο δεξί. Στα σημεία 2.2 – 2.9 συμβαίνει ακριβώς το αντίθετο.
- Εάν δεν είναι διαθέσιμα όλα τα σημεία εκ των 2.2 – 2.9, θα πρέπει να υπολογιστούν επιπλέον σημεία, όπου αυτό είναι δυνατό, με γραμμική παρεμβολή όπως έγινε στο υποκεφάλαιο 5.4 στη σελίδα 102 (Παράγωγα χαρακτηριστικά σημεία).



## 5.8 Σύστημα ανάλυσης και εκτίμησης συναισθημάτων

Μετά τον υπολογισμό των FAPs, γίνεται αποτίμηση των κανόνων που είναι αποθηκευμένοι στο αρχείο κειμένου `emotion profiles.txt`. Στην παρούσα υποπαράγραφο περιγράφεται το θεωρητικό υπόβαθρο του συστήματος αναγνώρισης συναισθήματος που υλοποιήθηκε. Ο ευριστικός αλγόριθμος υλοποιεί βασικές αρχές της ασαφούς λογικής, όπως αυτές περιγράφηκαν στο κεφάλαιο 4 της ασαφούς λογικής.

Το σύστημα εκτίμησης του συναισθήματος βασίζεται στη μεθοδολογία που προτάθηκε το 2002 στο άρθρο των Α. Ραουζαίου και συνεργατών [29] και περιγράφεται λεπτομερώς παρακάτω.

Θεωρούμε ως είσοδο του συστήματος ένα διάνυσμα χαρακτηριστικών (feature vector)  $\mathbf{f}$  το οποίο περιέχει, στη γενική περίπτωση,  $n_f$  στοιχεία.

Από τα στοιχεία του  $\mathbf{f}$  μπορούν να υπολογιστούν οι τιμές των FAPs οι οποίες αποτελούν στοιχεία του διανύσματος εισόδου  $\mathbf{G}$ . Τα στοιχεία του  $\mathbf{G}$  αντιπροσωπεύουν τις παρατηρηθείσες τιμές των FAPs που σχετίζονται με τα χαρακτηριστικά που περιέχονται στο  $\mathbf{f}$ . Για την ακριβή αντιστοίχιση, μπορεί να σχηματίζεται ένας πίνακας αντιστοιχίας ανάλογα με την εφαρμογή.

Στο πρόγραμμα τα στοιχεία του διανύσματος  $\mathbf{f}$  δεν αποθηκεύονται σε ξεχωριστό διάνυσμα. Τα στοιχεία του διανύσματος  $\mathbf{G}$  αποθηκεύονται στο διάνυσμα `fap[FAPMAX]` του προγράμματος (Πίνακας 5-11, σελ.110). Ο πίνακας αυτός μπορεί να χρησιμοποιηθεί ως «πίνακας αντιστοιχίας».

Αν το FAP  $F_j$  συμμετέχει στο προφίλ  $P_i^{(k)}$ , δηλαδή το  $k$ -οστό προφίλ του συναισθήματος  $i$ , το εύρος τιμών του  $F_j$  μπορεί να γραφεί ως  $X_{i,j}^{(k)}$ .

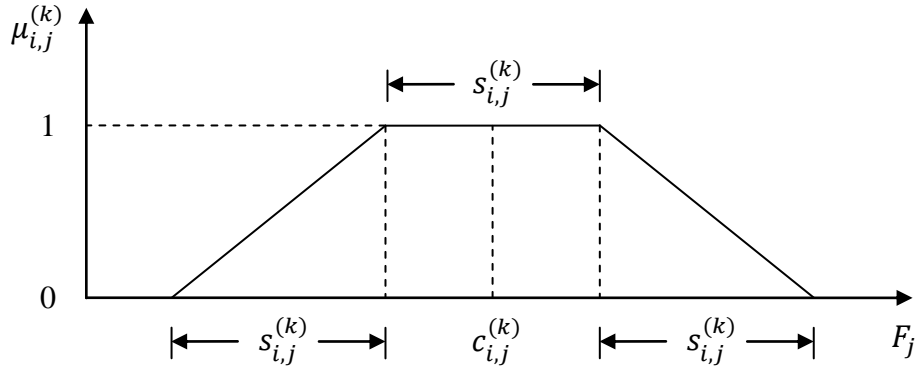
Ονομάζοντας  $c_{i,j}^{(k)}$  και  $s_{i,j}^{(k)}$  το μέσον και το μήκος του διαστήματος  $X_{i,j}^{(k)}$  αντίστοιχα, μπορούμε να ορίσουμε μια ασαφή κλάση  $A_{i,j}^{(k)}$  για το  $F_j$ , χρησιμοποιώντας την τραπεζοειδή συνάρτηση συμμετοχής (ΣΣ)  $\mu_{i,j}^{(k)}$  (δες και παράγραφο 4.3.1, στη σελίδα 74) όπως ορίζεται από τη σχέση:

$$\mu_{i,j}^{(k)}(F_j) = \begin{cases} 0, & F_j < a \\ \frac{F_j - a}{b - a}, & a \leq F_j < b \\ 1, & b \leq F_j < c \\ \frac{d - F_j}{d - c}, & c \leq F_j < d \\ 0, & F_j \geq d \end{cases}$$

Όπου οι παράμετροι της τραπεζοειδούς ΣΣ  $\mu_{i,j}^{(k)}$  ορίζονται ως:

$$a = c_{i,j}^{(k)} - \frac{3}{2}s_{i,j}^{(k)}, \quad b = c_{i,j}^{(k)} - \frac{1}{2}s_{i,j}^{(k)}, \quad c = c_{i,j}^{(k)} + \frac{1}{2}s_{i,j}^{(k)}, \quad d = c_{i,j}^{(k)} + \frac{3}{2}s_{i,j}^{(k)}$$

Η  $\mu_{i,j}^{(k)}$  φαίνεται παρακάτω στο σχήμα 5-12.



**Σχήμα 5-12** Συνάρτηση συμμετοχής  $\mu$  για την αναγνώριση συναισθημάτων

Η τραπεζοειδής συνάρτηση συμμετοχής του σχήματος προτάθηκε στο άρθρο [29].

Για τις ανάγκες της παρούσας εργασίας κρίθηκε σκόπιμο να διερευνηθεί η χρήση μιας τροποποιημένης μορφής της παραπάνω τραπεζοειδούς ΣΣ, έστω η  $\lambda_{i,j}^{(k)}$ , η οποία επιτρέπει στις πλάγιες πλευρές, για τις οποίες  $0 < \lambda_{i,j}^{(k)} < 1$ , να εκτείνονται σε διάστημα μικρότερο ή μεγαλύτερο του μήκους  $s_{i,j}^{(k)}$  ανάλογα με τις απαιτήσεις για το κάθε FAP. Με άλλα λόγια, η προκύπτουσα ΣΣ έχει σχήμα απλού, ανισοσκελούς τραπεζίου αντί για το ισοσκελές τραπέζιο της  $\mu_{i,j}^{(k)}$ .

Η συνάρτηση συμμετοχής  $\lambda_{i,j}^{(k)}$  ορίζεται με τη βοήθεια δύο επιπρόσθετων συντελεστών  $\rho_1, \rho_2$ . Κάθε συντελεστής αντιπροσωπεύει το λόγο του μήκους διαστήματος για το οποίο  $0 < \lambda_{i,j}^{(k)} < 1$  δεξιά ή αριστερά του  $c_{i,j}^{(k)}$ , προς το μήκος διαστήματος  $s_{i,j}^{(k)}$ . Με γεωμετρικούς όρους, κάθε συντελεστής  $\rho_1, \rho_2$  αντιπροσωπεύει το λόγο της προβολής μίας πλάγιας πλευράς του τραπεζίου επί της μεγάλης βάσης, προς το μήκος της μικρής βάσης.

Η  $\lambda_{i,j}^{(k)}$  ορίζεται ως:

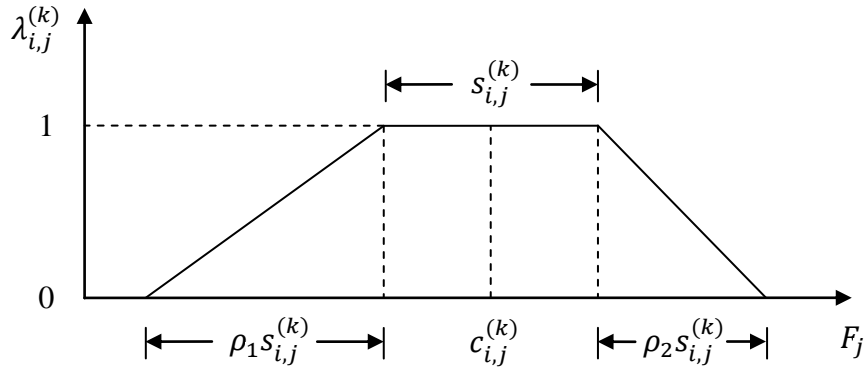
$$\lambda_{i,j}^{(k)}(F_j) = \begin{cases} 0, & F_j < a \\ \frac{F_j - a}{b - a}, & a \leq F_j < b \\ 1, & b \leq F_j < c \\ \frac{d - F_j}{d - c}, & c \leq F_j < d \\ 0, & F_j \geq d \end{cases}$$

Όπου οι παράμετροι  $a, b, c, d$  της τραπεζοειδούς ΣΣ  $\lambda_{i,j}^{(k)}$  δίνονται από τις σχέσεις:

$$\begin{aligned} a &= c_{i,j}^{(k)} - \left(\rho_1 + \frac{1}{2}\right) s_{i,j}^{(k)}, & b &= c_{i,j}^{(k)} - \frac{1}{2} s_{i,j}^{(k)}, \\ c &= c_{i,j}^{(k)} + \frac{1}{2} s_{i,j}^{(k)}, & d &= c_{i,j}^{(k)} + \left(\rho_2 + \frac{1}{2}\right) s_{i,j}^{(k)}. \end{aligned} \quad \text{Εξ. 5-19}$$

Τιμές  $0 < \rho_i < 1$ ,  $i = 1, 2$  δίνουν προβολές μικρότερου μήκους από το  $s_{i,j}^{(k)}$ , τιμές  $\rho_i = 1$ ,  $i = 1, 2$  δίνουν προβολές ίσου μήκους με το  $s_{i,j}^{(k)}$  (όπως είδαμε στην ΣΣ  $\mu_{i,j}^{(k)}$ ) και τέλος τιμές  $\rho_i > 1$ ,  $i = 1, 2$  δίνουν προβολές μεγαλύτερου μήκους από την  $s_{i,j}^{(k)}$ . Τα  $\rho_i$ ,  $i = 1, 2$  δεν έχει νόημα να λάβουν αρνητικές τιμές.

Η συνάρτηση συμμετοχής  $\lambda_{i,j}^{(k)}$  φαίνεται στο εξής σχήμα:

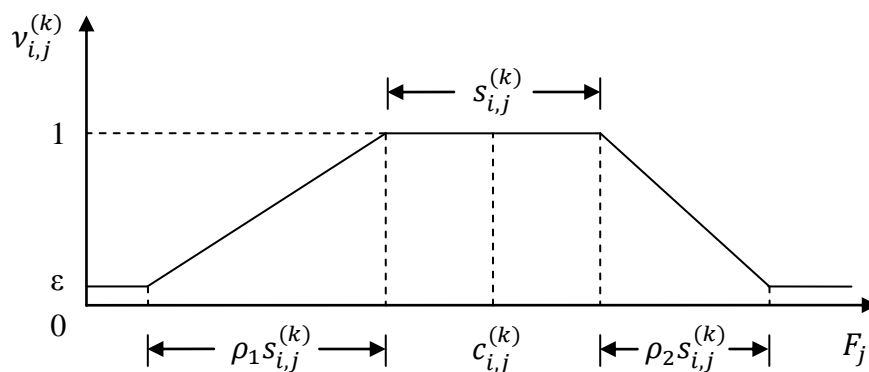


**Σχήμα 5-13** Συνάρτηση συμμετοχής  $\lambda$  για την αναγνώριση συναισθημάτων

Εάν θέλουμε η συνάρτηση συμμετοχής να μην έχει μηδενικές τιμές, δηλαδή όλα τα  $F_j$  να έχουν έναν ελάχιστο βαθμό συμμετοχής  $\varepsilon$ , μπορούμε να τροποποιήσουμε τη ΣΣ ως εξής:

$$v_{i,j}^{(k)}(F_j) = \begin{cases} \varepsilon & , F_j < a \\ \varepsilon + \frac{F_j - a}{b - a}(1 - \varepsilon) & , a \leq F_j < b \\ 1 & , b \leq F_j < c \\ 1 - \frac{F_j - c}{d - c}(1 - \varepsilon) & , c \leq F_j < d \\ \varepsilon & , F_j \geq d \end{cases} \quad \text{Εξ. 5-20}$$

Όπου οι παράμετροι  $a, b, c, d$  της τραπεζοειδούς ΣΣ  $v_{i,j}^{(k)}$  δίνονται και πάλι από τις σχέσεις Εξ. 5-19. Η ΣΣ  $v_{i,j}^{(k)}$  φαίνεται στο σχήμα που ακολουθεί:



**Σχήμα 5-14** Συνάρτηση συμμετοχής  $v$  για την αναγνώριση συναισθημάτων

Η παράμετρος  $\varepsilon$  επιλέγεται με τιμές  $0 \leq \varepsilon \leq 1$ . Συνήθεις τιμές κυμαίνονται στο διάστημα  $0 \leq \varepsilon < 0,10$ . Εάν επιλεγεί  $\varepsilon = 0$ , η  $\Sigma \Sigma v_{i,j}^{(k)}$  ταυτίζεται με την  $\lambda_{i,j}^{(k)}$  του σχήματος 5-13.

Η συνάρτηση συμμετοχής  $v_{i,j}^{(k)}$  υλοποιήθηκε στο πρόγραμμα μέσω της συνάρτησης  $\text{mf}(x, x1, x2, r1, r2, e)$  που επιτρέπει στο χρήστη να ορίζει τις παραμέτρους της συνάρτησης  $v_{i,j}^{(k)}$ . Οι παράμετροι  $x1, x2$  αντιπροσωπεύουν τα άκρα του διαστήματος  $X_{i,j}^{(k)}$  ενώ οι παράμετροι  $r1, r2$  αντιπροσωπεύουν τους συντελεστές  $\rho_1, \rho_2$  έτσι όπως δίνονται στα παραπάνω σχήματα. Τέλος η παράμετρος  $e$  είναι η παράμετρος  $\varepsilon$  της  $\Sigma \Sigma v_{i,j}^{(k)}$ . Μπορεί να τεθεί ίση με μηδέν,  $\varepsilon = 0$ , αν θέλουμε να χρησιμοποιήσουμε τη  $\Sigma \Sigma \lambda_{i,j}^{(k)}$  αντί της  $v_{i,j}^{(k)}$ . Η μεταβλητή  $x$  δέχεται την τιμή του FAP  $F_j$ .

Στη συνέχεια ορίζουμε ως  $\Delta_{i,j}^{(k)}$  το σύνολο των ασαφών κλάσεων  $A_{i,j}^{(k)}$  που αντιστοιχούν στο προφίλ  $P_i^{(k)}$ . Η βεβαιότητα ότι η παρατηρηθείσα μέσω του διανύσματος  $\mathbf{G}$  συναισθηματική κατάσταση του προσώπου αντιστοιχεί στο προφίλ  $P_i^{(k)}$  και το συναίσθημα  $i$ , συμβολίζονται με  $p_i^{(k)}$  και  $b_i$  αντίστοιχα. Οι βεβαιότητες αυτές υπολογίζονται από τις παρακάτω σχέσεις:

$$p_i^{(k)} = \prod_{A_{i,j}^{(k)} \in \Delta_{i,j}^{(k)}} r_{i,j}^{(k)} \quad \text{Εξ. 5-21}$$

όπου το  $r_{i,j}^{(k)}$  δίνεται από τη σχέση

$$r_{i,j}^{(k)} = \max\{g_i \cap A_{i,j}^{(k)}\} \quad \text{Εξ. 5-22}$$

και εκφράζει τη συσχέτιση του  $i$ -οστού στοιχείου του διανύσματος εισόδου με την κλάση  $A_{i,j}^{(k)}$ . Τα  $g_i$  της Εξ. 5-22 είναι στοιχεία του ασαφοποιημένου διανύσματος εισόδου  $\mathbf{g} = A'(\mathbf{G}) = \{g_1, g_2, \dots\}$ . Το διάνυσμα  $\mathbf{g}$  προκύπτει από διαδικασία ασαφοποίησης ασαφών σημείων (singleton fuzzification procedure).

Η βεβαιότητα  $b_i$  δίνεται από:

$$b_i = \max_k(p_i^{(k)}) \quad \text{Εξ. 5-23}$$

Όπου και πάλι τα  $p_i^{(k)}$  δίνονται από την εξίσωση Εξ. 5-21.

Η τελική εκτίμηση του συναισθήματος  $q$  γίνεται βάσει της σχέσης:

$$q = \arg \max_i b_i \quad \text{Εξ. 5-24}$$

Παρατηρώντας την εξίσωση Εξ. 5-21 είναι εμφανές ότι τα διάφορα προφίλ συναισθημάτων αντιστοιχούν σε ασαφείς τομές ασαφών συνόλων και βρίσκονται με την εφαρμογή τριγωνικής νόρμας γινομένου t-norm της γενικής μορφής  $t(a, b) = a \cdot b$ .

Παρομοίως η βεβαιότητα  $b_i$  ότι το παρατηρηθέν διάνυσμα εισόδου αντιστοιχεί σε ένα συγκεκριμένο συναίσθημα προκύπτει από την ασαφή ένωση κάποιων ασαφών συνόλων. Η εξίσωση Εξ. 5-23 αντιπροσωπεύει μια τριγωνική s-norm μεγίστου, της γενικής μορφής  $u(a, b) = \max(a, b)$ .

Η συσχέτιση  $r_{i,j}^{(k)}$  υπολογίζεται από τη συνάρτηση  $\text{mf}(x, x1, x2, r1, r2, e)$ . Η μέγιστη τιμή της  $r_{i,j}^{(k)}$  είναι 1, συνεπώς η μέγιστη βεβαιότητα  $p_i^{(k)}$  είναι και αυτή ίση με 1. Η συνάρτηση `analyze_emotions` υπολογίζει το γινόμενο της Εξ. 5-21 για τη βεβαιότητα  $p_i^{(k)}$  με έναν επαναληπτικό βρόχο. Μετά από κάθε υπολογισμό της βεβαιότητας  $p_i^{(k)}$  για ένα προφίλ  $k$ , ενημερώνεται η τρέχουσα εκτίμηση για το συναίσθημα βάσει των σχέσεων Εξ. 5-23 και Εξ. 5-24.

Όπως αναφέρθηκε και προηγουμένως, η συνάρτηση  $\text{mf}(x, x1, x2, r1, r2, e)$  υλοποιεί τη ΣΣ  $v_{i,j}^{(k)}$  (Σχήμα 5-14, Εξ. 5-20 στη σελίδα 115). Ο λόγος που εισήχθη ο παράμετρος  $\varepsilon$  είναι για να μη μηδενίζεται η βεβαιότητα  $p_i^{(k)}$  (Εξ. 5-21) όταν ένα μεμονωμένο  $r_{i,j}^{(k)}$  είναι μηδενικό. Χρειάζεται όμως προσοχή στην ερμηνεία των αποτελεσμάτων. Η εισαγωγή της παραμέτρου  $\varepsilon$  ευνοεί τους κανόνες που έχουν λίγα FAPs. Για παράδειγμα, έστω δύο κανόνες για δύο προφίλ  $P_i^{(k1)}$  και  $P_i^{(k2)}$  οι οποίοι δεν ικανοποιούνται, δηλαδή όλες οι τιμές των υπολογισθέντων FAPs βρίσκονται εκτός των διαστημάτων των κανόνων και των ασαφών κλάσεων. Αν τα προφίλ διατυπώθηκαν με τη χρήση  $n_1, n_2$  FAPs αντίστοιχα, με  $n_1 < n_2$ , τότε για τις βεβαιότητες ισχύει  $p_i^{(k1)} > p_i^{(k2)}$  επειδή  $\varepsilon^{n_1} > \varepsilon^{n_2}$  ( $\varepsilon < 1$ ). Παρότι δηλαδή κανένα προφίλ δεν αντιπροσωπεύει το εικονιζόμενο συναίσθημα, το αριθμητικό αποτέλεσμα της βεβαιότητας δεν έχει την ίδια τιμή, αλλά επηρεάζεται από το πλήθος των FAPs. Το πλήθος των FAPs καθορίζει ουσιαστικά την τάξη μεγέθους της βεβαιότητας. Στην παραπάνω περίπτωση η βεβαιότητα θα έχει ούτως ή άλλως μικρές τιμές (για μικρό  $\varepsilon$ ) οπότε οποιαδήποτε απόφαση για το συναίσθημα συνοδεύεται από μικρό βαθμό εμπιστοσύνης.

Όμως το ίδιο πρόβλημα μπορεί να προκύψει ακόμα και όταν σε κάποιο προφίλ υπάρχουν FAPs που βρίσκονται εντός του διαστήματος της ασαφούς κλάσης. Για παράδειγμα, έστω ότι οι βεβαιότητες δύο προφίλ προκύπτουν από τους παρακάτω υπολογισμούς (χρησιμοποιείται  $\varepsilon = 0,1$ ):

$$p_i^{(k1)} = 0,1 \cdot 0,1 \cdot 0,1 \cdot 0,1 \cdot 0,1 = 10^{-5} \text{ (5 FAPs)}$$

$$p_i^{(k2)} = 0,1 \cdot 0,1 \cdot 0,1 \cdot 0,2 \cdot 0,3 \cdot 0,4 \cdot 0,5 \cdot 0,6 = 0,72 \cdot 10^{-5} \text{ (8 FAPs)}$$

Λόγω της εξίσωσης Εξ. 5-23 θα επιλεγεί το προφίλ  $p_i^{(k1)}$  παρότι κανένα FAP δεν εμπίπτει στα δοσμένα διαστήματα τιμών. Είναι προφανές ότι εδώ θα έπρεπε να επιλέγεται το  $p_i^{(k2)}$  έναντι του  $p_i^{(k1)}$ .

Για να αντιμετωπιστούν οι παραπάνω περιπτώσεις, αλλά και για να παρέχεται μια εικόνα για το πλήθος των FAPs που εμπίπτουν στα διαστήματα τιμών των κανόνων, εισάγουμε έναν «μετρητή επιτυχιών» (hit counter). Ο μετρητής αυξάνει κάθε φορά που η συνάρτηση  $\text{mf}(x, x1, x2, r1, r2, e)$  για τη συσχέτιση  $r_{i,j}^{(k)}$  επιστρέφει αποτέλεσμα μεγαλύτερο από την παράμετρο  $\varepsilon$ . Μετά την αποτίμηση κάθε κανόνα, ο μετρητής θα κρατά το πλήθος των «επιτυχημένων» FAPs, που βρέθηκε δηλαδή ότι εμπίπτουν στα διαστήματα τιμών του κανόνα.

Η τιμή του μετρητή (hit counter) χρησιμοποιείται και στο κριτήριο για την ενημέρωση της τρέχουσας εκτίμησης για το συναίσθημα. Πρέπει η υπολογισθείσα βεβαιότητα του υπό εξέταση προφίλ να είναι μεγαλύτερη από την αποθηκευμένη τιμή της τρέχουσας εκτίμησης, και ταυτόχρονα θα πρέπει ο μετρητής να είναι μη-μηδενικός. Μόνο τότε θα πραγματοποιείται ενημέρωση της τρέχουσας εκτίμησης. Η πληροφορία του μετρητή μπορεί να εκτυπώνεται ως κλάσμα και ως ποσοστό (Παράδειγμα, hits: 4 out of 12 = 33.3%). Περισσότερες πληροφορίες για την εμφάνιση των αποτελεσμάτων και τις επιλογές εκτύπωσης δίνονται στην παράγραφο 5.10.

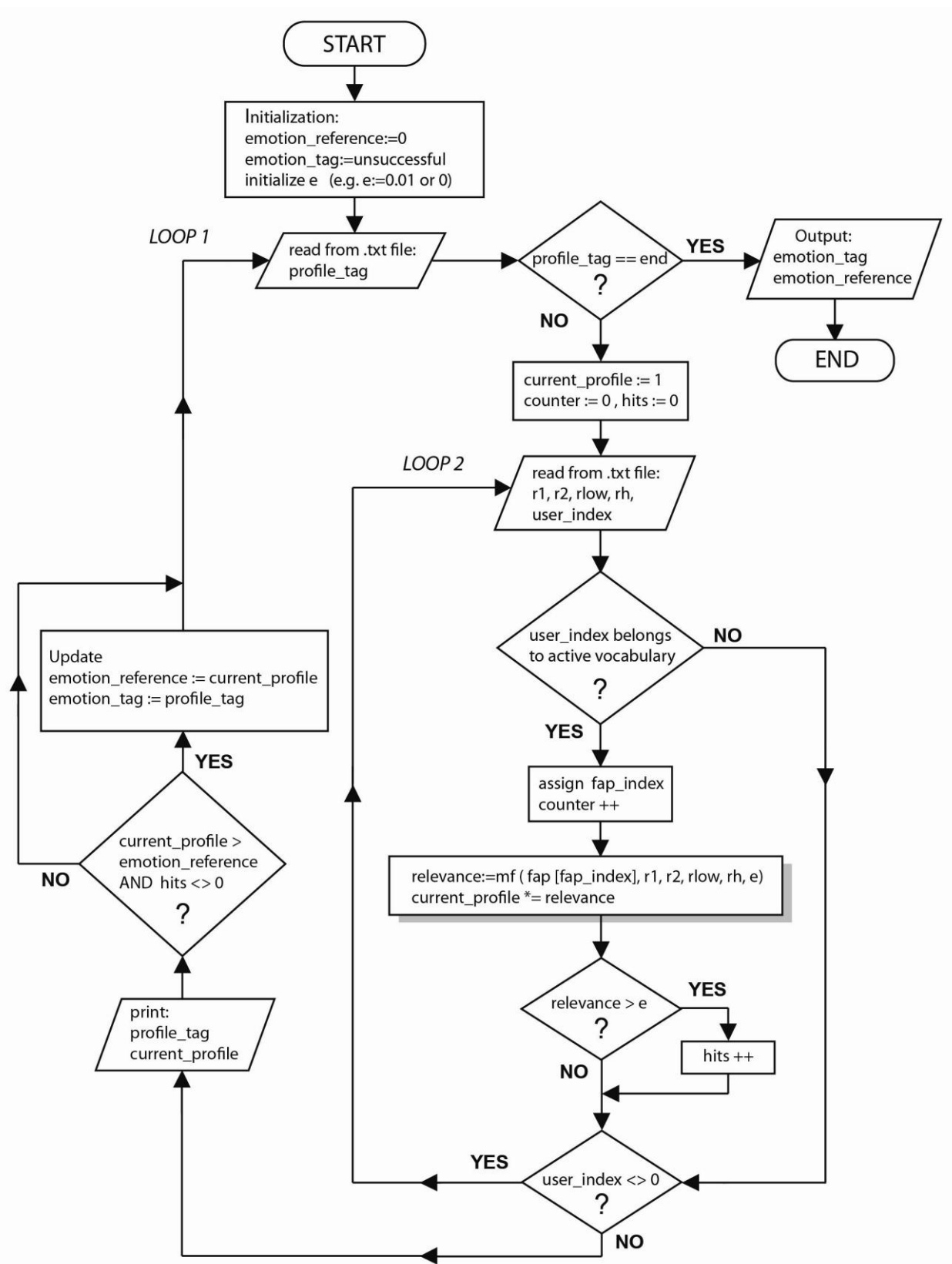
Αξίζει να σημειωθεί ότι στο σύστημα που περιγράφηκε στο παρόν υποκεφάλαιο, δεν έχει διατυπωθεί εκ των προτέρων υπόθεση για το πλήθος των συναισθημάτων που μπορούν να αναγνωριστούν. Το πλήθος των συναισθημάτων εξαρτάται καθαρά από τα διαθέσιμα προφίλ που έχουν μοντελοποιηθεί. Κατά συνέπεια το παραπάνω σύστημα μπορεί να χρησιμοποιηθεί τόσο για την ανίχνευση των αρχέτυπων συναισθημάτων όσο και την αναγνώριση ενδιάμεσων προφίλ.

Η υλοποίηση ακολουθεί αυτό το ευέλικτο μοντέλο. Οι πληροφορίες για τα προφίλ συναισθημάτων διαβάζονται από ένα αρχείο .txt, το `emotion_profiles.txt` (δες επίσης παράγραφο 5.2 για τις εισόδους) και δεν υπάρχει καμία υπόθεση εκ των προτέρων για το πλήθος των προφίλ ή για το πλήθος των FAPs που συμμετέχουν σε κάθε προφίλ. Κάθε προφίλ αποτιμάται ανεξάρτητα από την ύπαρξη ή όχι άλλων προφίλ. Αυτό επιτρέπει στο χρήστη να ελέγχει εξολοκλήρου τα διαθέσιμα προφίλ, αρκεί να έχουν μοντελοποιηθεί με διαστήματα τιμών των FAPs. Ο μόνος περιορισμός έγκειται στο λεξιλόγιο των FAPs (vocabulary) δηλαδή το πλήθος των διαθέσιμων FAPs που υπολογίζονται από τη συνάρτηση `analyze_emotions`, και παρουσιάστηκαν αναλυτικά στον πίνακα 5-11.

## 5.9 Αποτίμηση Κανόνων

Το θεωρητικό υπόβαθρο για την εκτίμηση του συναισθήματος του εικονιζόμενου προσώπου παρουσιάστηκε λεπτομερώς στις προηγούμενες παραγράφους. Παρακάτω δίνεται σχηματικά το διάγραμμα ροής της υπολογιστικής υλοποίησης της αποτίμησης κανόνων καθώς πραγματοποιείται η ανάγνωση πληροφοριών από το αρχείο κειμένου.

Η βασική δομή αυτού του τμήματος του προγράμματος καθορίζεται από δύο επαναληπτικούς βρόχους. Ο εξωτερικός βρόχος (loop 2) επαναλαμβάνεται για κάθε μοντελοποιημένο προφίλ/κανόνα που βρίσκεται στο αρχείο κειμένου. Ο εσωτερικός βρόχος (loop 1) διαβάζει επαναληπτικά το σταθερό φορμάτ της μορφής "`fap[%d]-[%f %f]-[%f %f]`" που αντιστοιχεί σε όλες τις πληροφορίες σχετικά με ένα FAP του εκάστοτε κανόνα. Η ανάγνωση συνεχίζεται μέχρι να διαβαστεί η ετικέτα "end" (ή "End") που σηματοδοτεί το τέλος των κανόνων. Εάν διαβαστεί κάποιο FAP από τα  $F_{16}$ ,  $F_{18}$ ,  $F_{41}$  και  $F_{42}$  που δεν είναι δυνατόν να υπολογιστεί στα πλαίσια της εφαρμογής αλλά ανήκει στο λεξιλόγιο των κανόνων που μοντελοποιήθηκαν στα πλαίσια του άρθρου [29], τότε το πρόγραμμα αγνοεί αυτήν την ανάγνωση και συνεχίζει με το επόμενο FAP αν αυτό υπάρχει (δες διάγραμμα ροής που ακολουθεί). Εάν διαβαστεί κάποιος άλλος δείκτης, εκτυπώνεται ενημερωτικό μήνυμα για το FAP που δεν αναγνωρίστηκε και το προφίλ στο οποίο άνηκε, και η πληροφορία αγνοείται από το πρόγραμμα προχωρώντας στο επόμενο FAP, αν αυτό υπάρχει.



Σχήμα 5-15 Διάγραμμα ροής της αποτίμησης κανόνων

## 5.10 Επιλογές εκτύπωσης και εμφάνιση αποτελεσμάτων

### 5.10.1 Παράμετροι συναρτήσεων που ελέγχουν την εκτύπωση

Κατά την κλήση της συνάρτησης εκτίμησης συναισθημάτων `analyze_emotions` από το κυρίως πρόγραμμα, ο χρήστης μπορεί να επιλέξει πόσες πληροφορίες θέλει να εκτυπωθούν, περνώντας στην `analyze_emotions` τις κατάλληλες παραμέτρους.

Η `analyze_emotions` ορίζεται ως:

```
int analyze_emotions (char printfp, char printfaps, char printprofiles,
char rotate)
```

Οι παράμετροι `printfp`, `printfaps` και `printprofiles` ελέγχουν την εκτύπωση των πληροφοριών που υπολογίζονται με το πρόγραμμα. Η παράμετρος `rotate` χρησιμοποιείται για να επιλεγεί μία από τις τρεις μεθόδους για το μετασχηματισμό περιστροφής των σημείων (δες επίσης παράγραφο 5.3 στη σελίδα 89). Στην πράξη η συνάρτηση `analyze_emotions` καλεί άλλες συναρτήσεις για την εκτύπωση των αποτελεσμάτων, τις `print_frames`, `print_faps_basic`, `print_faps_details`. Οι παράμετροι `printfp`, `printfaps` και `printprofiles` χρησιμοποιούνται σε συνθήκες κατάλληλων δομών ελέγχου για την κλήση αυτών των συναρτήσεων. Η παράμετρος `printfp` περνά επίσης σαν παράμετρος στις συναρτήσεις για το μετασχηματισμό περιστροφής, οι οποίες εκτυπώνουν τις πληροφορίες για την ευθεία παλινδρόμησης ανάλογα με την τιμή της `printfp`. Για παράδειγμα, η συνάρτηση που χρησιμοποιεί τη μέθοδο ελαχίστων τετραγώνων με x-αποκλίσεις, ορίζεται ως:

```
int rotate_feature_points_dx (double fp_x[], double fp_y[], char printr)
```

Όταν καλείται, στη θέση της παραμέτρου `printr` εισάγεται η παράμετρος `printfp`.

Για τις παραμέτρους `printfp` και `printfaps` ο χρήστης έχει τρεις επιλογές. Μπορεί να επιλέξει τη λεπτομερή εκτύπωση πληροφοριών, την εκτύπωση των βασικών πληροφοριών ή να μην εκτυπώνεται καμία πληροφορία. Η παράμετρος `printfp` ελέγχει την εκτύπωση των πληροφοριών που σχετίζονται με τα σημεία ενδιαφέροντος στο πρόσωπο και η παράμετρος `printfaps` ελέγχει την εκτύπωση πληροφοριών των FAPs.

Για την παράμετρο `printprofiles` ο χρήστης έχει δύο επιλογές, να εκτυπώνονται ή όχι πληροφορίες που υπολογίζονται κατά την αποτίμηση του κάθε επιμέρους κανόνα.

Για την παράμετρο `rotate` ο χρήστης έχει τρεις επιλογές για να επιλέξει μία από τις τρεις μεθόδους ελαχίστων τετραγώνων. Οι επιλογές συζητήθηκαν στην παράγραφο 5.3.4 (μετ/μός περιστροφής) στη σελίδα 99, Πίνακας 5-4.



Οι πληροφορίες σχετικά με την εκτύπωση των αποτελεσμάτων συνοψίζονται στον ακόλουθο πίνακα:

| Τιμή παραμέτρου                | Παράμετροι συνάρτησης <code>analyze_emotions</code> για την εκτύπωση πληροφοριών   |   |   |
|--------------------------------|--|---|---|
|                                | <code>printfp</code>   | <code>printfaps</code>  | <code>printprofiles</code>  |
| 'd' ή 'D' (details)            | Εκτύπωση συντ/νων των σημείων του ουδέτερου και υπό εξέταση καρέ πριν και μετά το μετασχηματισμό περιστροφής. Εκτύπωση ευθειών παλινδρόμησης μετ/μού περιστροφής. (Εφόσον έχει επιλεγεί μέθοδος περιστροφής) | Λεπτομερής εκτύπωση πληροφοριών: ομάδα στην οποία ανήκει το FAP – $F_j$ – όνομα FAP – αντίστοιχη μεταβλητή αποθήκευσης – τιμή FAP. Υπολογισμένες τιμές των συνδυασμών FAPs. | Εκτύπωση πληροφοριών για κάθε προφίλ του .txt που εξετάζεται: Ετικέτα προφίλ, τιμή βεβαιότητας του προφίλ και τιμή του μετρητή «επιτυχιών» (κλασματική τιμή και ποσοστό). |
| 'y' ή 'Y' (yes)                | Εκτύπωση συντ/νων σημείων του ουδέτερου και υπό εξέταση καρέ μετά το μετασχηματισμό περιστροφής. (Εφόσον έχει επιλεγεί μέθοδος περιστροφής)  | Βασική εκτύπωση: ομάδα FAPs, κωδικός $F_j$ , όνομα και τιμή FAP.  |   |
| άλλος χαρακτήρας               | Καμία εκτύπωση συντ/νων.   | Καμία εκτύπωση τιμών των FAPs   | Δεν εκτυπώνονται πληροφορίες για τα επιμέρους εξεταζόμενα προφίλ  |
| ανεξάρτητο από τις παραμέτρους | Σε κάθε περίπτωση, εκτύπωση της τελικής εκτίμησης για το συναίσθημα: ετικέτα του προφίλ και τιμή της βεβαιότητας.  |   |   |

**Πίνακας 5-14** Επιλογές εκτύπωσης

### 5.10.2 Παραδείγματα εκτύπωσης πληροφοριών

Στους παρακάτω πίνακες δίνονται παραδείγματα εκτύπωσης δεδομένων από το πρόγραμμα για διάφορες τιμές των παραμέτρων `printfp`, `printfaps` και `printprofiles`.

| Λεπτομερής εκτύπωση συντεταγμένων των σημείων για <code>printfp='d' ή 'D'</code>  |  |
|---|--|
| Coordinates before rotation:<br>Neutral frame:<br>( <code>fp_x_neutr</code> , <code>fp_y_neutr</code> )<br>0 - ( 281.78 , 237.25 )<br>1 - ( 282.27 , 180.84 ) | Current frame:<br>( <code>fp_x</code> , <code>fp_y</code> ):<br>0 - ( 273.93 , 241.66 )<br>1 - ( 273.93 , 187.71 ) |

|                          |                          |
|--------------------------|--------------------------|
| 2 - ( 281.78 , 362.81 )  | 2 - ( 271.97 , 365.26 )  |
| 3 - ( 285.70 , 307.38 )  | 3 - ( 276.88 , 311.80 )  |
| 4 - ( 277.86 , 206.35 )  | 4 - ( 270.99 , 212.72 )  |
| 5 - ( 277.86 , 335.34 )  | 5 - ( 270.50 , 339.26 )  |
| 6 - ( 289.14 , 206.84 )  | 6 - ( 283.25 , 213.21 )  |
| 7 - ( 270.99 , 206.35 )  | 7 - ( 261.18 , 211.74 )  |
| 8 - ( 290.12 , 335.34 )  | 8 - ( 282.76 , 339.75 )  |
| 9 - ( 268.54 , 334.85 )  | 9 - ( 259.22 , 339.75 )  |
| 10 - ( 410.77 , 224.99 ) | 10 - ( 411.26 , 236.27 ) |
| 11 - ( 410.77 , 327.00 ) | 11 - ( 410.28 , 326.51 ) |
| 12 - ( 427.94 , 275.50 ) | 12 - ( 440.69 , 278.45 ) |
| 13 - ( 407.83 , 275.01 ) | 13 - ( 398.51 , 279.92 ) |
| 14 - ( 255.79 , 243.13 ) | 14 - ( 236.17 , 246.07 ) |
| 15 - ( 253.82 , 155.34 ) | 15 - ( 237.15 , 162.21 ) |
| 16 - ( 259.22 , 393.21 ) | 16 - ( 246.47 , 396.65 ) |
| 17 - ( 255.79 , 307.38 ) | 17 - ( 236.66 , 307.87 ) |
| 18 - ( 366.14 , 273.05 ) | 18 - ( 353.88 , 278.45 ) |

Rotation of neutral frame:

Least Squares, y-offsets:  
 Linear equation:  $y = -2193.290371 + 8.715085 x$   
 Point of rotation:  $(x_0, y_0) = (282.884196, 272.069482)$   
 Angle  $\theta_0 = 0.114244$

Rotation of current frame:

Least Squares, y-offsets:  
 Linear equation:  $y = 1849.101060 - 5.735296 x$   
 Point of rotation:  $(x_0, y_0) = (274.178474, 276.606267)$   
 Angle  $\theta_0 = -0.172624$

Coordinates after rotation:

Neutral frame:

|                          | ( fp_x_neutr , fp_y_neutr ) |
|--------------------------|-----------------------------|
| 0 - ( 285.76 , 237.80 )  |                             |
| 1 - ( 292.67 , 182.55 )  |                             |
| 2 - ( 271.44 , 360.91 )  |                             |
| 3 - ( 281.66 , 307.01 )  |                             |
| 4 - ( 285.38 , 207.06 )  |                             |
| 5 - ( 270.68 , 333.54 )  |                             |
| 6 - ( 296.53 , 208.82 )  |                             |
| 7 - ( 278.56 , 206.28 )  |                             |
| 8 - ( 282.86 , 334.92 )  |                             |
| 9 - ( 261.48 , 332.00 )  |                             |
| 10 - ( 415.31 , 240.39 ) |                             |
| 11 - ( 403.68 , 340.41 ) |                             |
| 12 - ( 426.60 , 291.86 ) |                             |
| 13 - ( 406.68 , 289.11 ) |                             |
| 14 - ( 259.26 , 240.63 ) |                             |
| 15 - ( 267.32 , 154.33 ) |                             |
| 16 - ( 245.56 , 388.17 ) |                             |
| 17 - ( 251.94 , 303.62 ) |                             |
| 18 - ( 365.49 , 282.46 ) |                             |

Current frame:

|                          | ( fp_x , fp_y ): |
|--------------------------|------------------|
| 0 - ( 267.93 , 243.25 )  |                  |
| 1 - ( 258.67 , 191.70 )  |                  |
| 2 - ( 287.23 , 361.70 )  |                  |
| 3 - ( 282.88 , 309.78 )  |                  |
| 4 - ( 260.06 , 216.10 )  |                  |
| 5 - ( 281.32 , 337.11 )  |                  |
| 6 - ( 272.23 , 214.49 )  |                  |
| 7 - ( 250.23 , 216.82 )  |                  |
| 8 - ( 293.48 , 335.50 )  |                  |
| 9 - ( 270.29 , 339.48 )  |                  |
| 10 - ( 402.30 , 214.86 ) |                  |
| 11 - ( 416.83 , 301.27 ) |                  |
| 12 - ( 438.53 , 250.19 ) |                  |
| 13 - ( 397.23 , 258.73 ) |                  |
| 14 - ( 231.49 , 253.86 ) |                  |
| 15 - ( 218.05 , 173.55 ) |                  |
| 16 - ( 267.50 , 396.01 ) |                  |
| 17 - ( 242.59 , 312.84 ) |                  |
| 18 - ( 353.01 , 264.88 ) |                  |

**Πίνακας 5-15** Λεπτομερής εκτύπωση συντεταγμένων σημείων ενδιαφέροντος.

| Βασική εκτύπωση συντεταγμένων των σημείων για printfp='y' ή 'Y' |                          |
|---|--------------------------|
| Coordinates after rotation:                                     |                          |
| Neutral frame:  | Current frame:           |
| ( fp_x_neutr , fp_y_neutr )                                     | ( fp_x , fp_y ):         |
| 0 - ( 339.04 , 223.02 )   | 0 - ( 326.24 , 217.72 )  |
| 1 - ( 334.41 , 155.35 )   | 1 - ( 317.79 , 146.94 )  |
| 2 - ( 334.11 , 371.64 )   | 2 - ( 317.86 , 372.51 )  |
| 3 - ( 339.78 , 304.92 )   | 3 - ( 326.10 , 301.57 )  |
| 4 - ( 332.05 , 183.80 )   | 4 - ( 321.09 , 179.37 )  |
| 5 - ( 332.55 , 343.69 )   | 5 - ( 320.42 , 341.67 )  |
| 6 - ( 344.80 , 183.27 )   | 6 - ( 334.33 , 179.62 )  |
| 7 - ( 320.77 , 184.33 )   | 7 - ( 302.46 , 179.01 )  |
| 8 - ( 344.32 , 342.67 )   | 8 - ( 333.19 , 340.93 )  |
| 9 - ( 319.79 , 342.75 )   | 9 - ( 299.85 , 339.80 )  |
| 10 - ( 487.60 , 207.84 )  | 10 - ( 498.08 , 210.24 ) |
| 11 - ( 487.98 , 330.45 )  | 11 - ( 497.93 , 320.08 ) |
| 12 - ( 509.38 , 270.55 )  | 12 - ( 558.86 , 263.88 ) |
| 13 - ( 472.59 , 271.16 )  | 13 - ( 460.71 , 265.91 ) |
| 14 - ( 309.65 , 235.86 )  | 14 - ( 285.33 , 227.72 ) |
| 15 - ( 304.42 , 132.88 )  | 15 - ( 278.54 , 121.67 ) |
| 16 - ( 305.77 , 405.57 )  | 16 - ( 280.45 , 404.65 ) |
| 17 - ( 306.41 , 297.67 )  | 17 - ( 282.10 , 293.86 ) |
| 18 - ( 405.41 , 274.31 )  | 18 - ( 391.98 , 268.51 ) |

**Πίνακας 5-16** Βασική εκτύπωση συντεταγμένων σημείων ενδιαφέροντος.

| Λεπτομερής εκτύπωση υπολογισμένων FAPs για printfap='d' ή 'D' |         |
|---|---------|
| Values of FAPs:   |         |
| Group 2 - Jaw, chin, inner lip positions                      |         |
| FAP3 - open_jaw - fap[11] =                                   | 513.39  |
| Group 3 - Eyeballs, pupils, eyelids                           |         |
| FAP19 - close_t_l_eyelid - fap[20] =                          | -1.60   |
| FAP20 - close_t_r_eyelid - fap[21] =                          | -90.49  |
| FAP21 - close_b_l_eyelid - fap[22] =                          | -106.59 |
| FAP22 - close_b_r_eyelid - fap[23] =                          | -102.55 |
| Group 4 - Eyebrows  |         |
| FAP31 - raise_l_i_eyebrow - fap[5] =                          | 122.87  |
| FAP32 - raise_r_i_eyebrow - fap[6] =                          | 132.45  |
| FAP33 - raise_l_m_eyebrow - fap[9] =                          | 67.81   |
| FAP34 - raise_r_m_eyebrow - fap[10] =                         | 87.36   |
| FAP35 - raise_l_o_eyebrow - fap[7] =                          | 23.34   |
| FAP36 - raise_r_o_eyebrow - fap[8] =                          | 76.64   |
| FAP37 - squeeze_l_eyebrow - fap[1] =                          | 9.96    |
| FAP38 - squeeze_r_eyebrow - fap[2] =                          | -1.86   |
| Group 8 - Outer lip positions                                 |         |
| FAP51 (approx.FAP4) - lower_t_midlip_o - fap[3] =             | 67.95   |
| FAP52 (approx.FAP5) - raise_b_midlip_o - fap[4] =             | -580.32 |
| FAP53 (approx.FAP6) - stretch_l_cornerlip_o - fap[16] =       | -219.28 |

|                      |                         |           |   |         |
|----------------------|-------------------------|-----------|---|---------|
| FAP54 (approx.FAP7)  | - stretch_r_cornerlip_o | - fap[17] | = | 80.80   |
| FAP55 (approx.FAP8)  | - lower_t_lip_lm_o      | - fap[24] | = | 333.43  |
| FAP56 (approx.FAP9)  | - lower_t_lip_rm_o      | - fap[25] | = | 29.01   |
| FAP57 (approx.FAP10) | - raise_b_lip_lm_o      | - fap[26] | = | -582.16 |
| FAP58 (approx.FAP11) | - raise_b_lip_rm_o      | - fap[27] | = | -277.74 |
| FAP59 (approx.FAP12) | - raise_l_cornerlip_o   | - fap[18] | = | -136.55 |
| FAP60 (approx.FAP13) | - raise_r_cornerlip_o   | - fap[19] | = | -167.85 |

#### Combinations of FAPs

|               |           |   |         |
|---------------|-----------|---|---------|
| FAP19 + FAP21 | - fap[12] | = | -49.78  |
| FAP20 + FAP22 | - fap[13] | = | -97.00  |
| FAP53 + FAP54 | - fap[14] | = | -132.99 |
| FAP37 + FAP38 | - fap[15] | = | 56.36   |

**Πίνακας 5-17** Λεπτομερής εκτύπωση υπολογισμένων FAPs.

*Ο πίνακας δείχνει το πλήρες λεξιλόγιο FAPs της εφαρμογής.*

#### Βασική εκτύπωση υπολογισμένων FAPs για printprofiles = 'y' ή 'Y'

##### Values of FAPs:

##### Jaw - Group 2

|      |            |   |        |
|------|------------|---|--------|
| FAP3 | - open_jaw | = | 762.98 |
|------|------------|---|--------|

##### Eyelids - Group 3

|       |                    |   |         |
|-------|--------------------|---|---------|
| FAP19 | - close_t_l_eyelid | = | -192.98 |
| FAP20 | - close_t_r_eyelid | = | -133.30 |
| FAP21 | - close_b_l_eyelid | = | -138.57 |
| FAP22 | - close_b_r_eyelid | = | -109.80 |

##### Eyebrows - Group 4

|       |                     |   |        |
|-------|---------------------|---|--------|
| FAP31 | - raise_l_i_eyebrow | = | 147.32 |
| FAP32 | - raise_r_i_eyebrow | = | 140.64 |
| FAP33 | - raise_l_m_eyebrow | = | 124.93 |
| FAP34 | - raise_r_m_eyebrow | = | 76.24  |
| FAP35 | - raise_l_o_eyebrow | = | 71.54  |
| FAP36 | - raise_r_o_eyebrow | = | 129.00 |
| FAP37 | - squeeze_l_eyebrow | = | -41.57 |
| FAP38 | - squeeze_r_eyebrow | = | -68.44 |

##### Outer lip positions - Group 8

|       |                         |   |         |
|-------|-------------------------|---|---------|
| FAP51 | - lower_t_midlip_o      | = | 18.88   |
| FAP52 | - raise_b_midlip_o      | = | -781.93 |
| FAP53 | - stretch_l_cornerlip_o | = | -38.24  |
| FAP54 | - stretch_r_cornerlip_o | = | -68.44  |
| FAP55 | - lower_t_lip_lm_o      | = | 154.86  |
| FAP56 | - lower_t_lip_rm_o      | = | 158.19  |
| FAP57 | - raise_b_lip_lm_o      | = | -536.26 |
| FAP58 | - raise_b_lip_rm_o      | = | -539.58 |
| FAP59 | - raise_l_cornerlip_o   | = | -619.04 |
| FAP60 | - raise_r_cornerlip_o   | = | -492.62 |

**Πίνακας 5-18** Βασική εκτύπωση υπολογισμένων FAPs.

| Εκτύπωση πληροφοριών αποτίμησης κανόνων για printfap = 'd'/'D'/'y'/'Y'  |           |   |                |   |       |         |   |       |   |
|---|-----------|---|----------------|---|-------|---------|---|-------|---|
| Evaluation of rules for profiles:   |           |   |                |   |       |         |   |       |   |
| Profile tag :   | Anger0    | - | belief = 0.000 | - | hits: | 4 of 9  | = | 44.4  | % |
| Profile tag :   | Anger1    | - | belief = 0.000 | - | hits: | 1 of 8  | = | 12.5  | % |
| Profile tag :   | Anger2    | - | belief = 0.000 | - | hits: | 3 of 8  | = | 37.5  | % |
| Profile tag :   | Anger3    | - | belief = 0.000 | - | hits: | 3 of 8  | = | 37.5  | % |
| Profile tag :   | Anger4    | - | belief = 0.000 | - | hits: | 3 of 8  | = | 37.5  | % |
| Profile tag :   | Sadness0  | - | belief = 0.707 | - | hits: | 6 of 6  | = | 100.0 | % |
| Profile tag :   | Fear0     | - | belief = 0.000 | - | hits: | 9 of 10 | = | 90.0  | % |
| Profile tag :   | Fear1     | - | belief = 0.000 | - | hits: | 1 of 6  | = | 16.7  | % |
| Profile tag :   | Fear2     | - | belief = 0.000 | - | hits: | 1 of 12 | = | 8.3   | % |
| Profile tag :   | Fear3     | - | belief = 0.000 | - | hits: | 3 of 10 | = | 30.0  | % |
| Profile tag :   | Fear4     | - | belief = 0.000 | - | hits: | 3 of 14 | = | 21.4  | % |
| Profile tag :   | Fear5     | - | belief = 0.000 | - | hits: | 3 of 14 | = | 21.4  | % |
| Profile tag :   | Fear6     | - | belief = 0.000 | - | hits: | 3 of 18 | = | 16.7  | % |
| Profile tag :   | Fear7     | - | belief = 0.000 | - | hits: | 1 of 12 | = | 8.3   | % |
| Profile tag :   | Fear8     | - | belief = 0.000 | - | hits: | 4 of 12 | = | 33.3  | % |
| Profile tag :   | Fear9     | - | belief = 0.000 | - | hits: | 1 of 12 | = | 8.3   | % |
| Profile tag :   | Joy0      | - | belief = 0.000 | - | hits: | 7 of 11 | = | 63.6  | % |
| Profile tag :   | Joy1      | - | belief = 0.000 | - | hits: | 4 of 12 | = | 33.3  | % |
| Profile tag :   | Joy2      | - | belief = 0.000 | - | hits: | 6 of 12 | = | 50.0  | % |
| Profile tag :   | Joy3      | - | belief = 0.000 | - | hits: | 4 of 12 | = | 33.3  | % |
| Profile tag :   | Disgust0  | - | belief = 0.000 | - | hits: | 6 of 8  | = | 75.0  | % |
| Profile tag :   | Surprise0 | - | belief = 0.000 | - | hits: | 7 of 16 | = | 43.8  | % |
| Profile tag :   | Surprise1 | - | belief = 0.000 | - | hits: | 2 of 18 | = | 11.1  | % |
| Profile tag :   | Surprise2 | - | belief = 0.000 | - | hits: | 4 of 16 | = | 25.0  | % |
| Profile tag :   | Surprise3 | - | belief = 0.000 | - | hits: | 7 of 16 | = | 43.8  | % |
| Estimated emotion: Sadness0 with belief = 0.707   |           |   |                |   |       |         |   |       |   |
| <p>Η τελευταία γραμμή για την τελική εκτίμηση του συναισθήματος τυπώνεται σε κάθε περίπτωση, ανεξάρτητα από την τιμή των παραμέτρων printfp, printfaps και printfprofiles.</p> <p>Οι κανόνες που χρησιμοποιήθηκαν σε αυτό το παράδειγμα βασίζονται στα προφίλ του άρθρου [29]</p> |           |   |                |   |       |         |   |       |   |

**Πίνακας 5-19** Εκτύπωση πληροφοριών αποτίμησης κανόνων.

## 5.11 Χρήση προγράμματος

Στην παρούσα παράγραφο συζητείται με ποιον τρόπο μπορεί ο χρήστης να χρησιμοποιήσει το πρόγραμμα για διάφορες εισόδους. Επίσης καταγράφονται όλες οι επιλογές που βρίσκονται στη διάθεση του χρήστη.

Υπενθυμίζουμε κάποιες βασικές **υποθέσεις** που θεωρούνται δεδομένες για την αποτελεσματική και αξιόπιστη λειτουργία του συστήματος:

- Το σύστημα δεν ελέγχει τη νοηματική ορθότητα των κανόνων ή των συντεταγμένων για να εξασφαλίσει την εγκυρότητα των αποτελεσμάτων. Η εννοιολογική/νοηματική συνοχή των δεδομένων εισόδου εναπόκειται στο χρήστη.
- Δεν υπάρχει πρόβλεψη για μεταβολές βάθους ανάμεσα στα καρέ. Το εικονιζόμενο άτομο θεωρείται ότι δεν απομακρύνεται ούτε πλησιάζει την κάμερα.
- Δεν υπάρχει πρόβλεψη για στροφή κατά x ή y άξονα. Το εικονιζόμενο άτομο θεωρείται ότι διατηρεί τον ίδιο προσανατολισμό ως προς το xy επίπεδο.
- Το σύστημα λειτουργεί μόνο με εκφράσεις του προσώπου, δεν εκμεταλλεύεται άλλες σωματικές κινήσεις, οι οποίες μάλιστα δυσχεραίνουν τη διαδικασία.
- Θεωρείται ότι τα καρέ απεικονίζουν ένα άτομο τη φορά, που κοιτά την κάμερα και δεν στρέφει το βλέμμα ή το κεφάλι αλλού.

Η ακόλουθη λίστα συνοψίζει όλα τα σημεία στα οποία ο χρήστης μπορεί να επέμβει ώστε να αλλάζουν τα δεδομένα εισόδου και να επιλέγονται οι επιθυμητές παράμετροι για το πρόγραμμα:

- Αρχείο με τις συντεταγμένες των σημείων ενδιαφέροντος.

Ο χρήστης μπορεί να τροποποιήσει το περιεχόμενο του αρχείου `framesrealtime.txt` ή να αλλάξει το αρχείο από όπου διαβάζονται τα σημεία ενδιαφέροντος. Για να διαβαστούν οι συντεταγμένες σωστά, θα πρέπει το ουδέτερο καρέ να βρίσκεται στην πρώτη γραμμή του αρχείου, χωρίς να προηγούνται τυχόν κενές γραμμές γιατί θα προκληθεί σφάλμα. Το τρέχον καρέ πρέπει να βρίσκεται στην τελευταία γραμμή του αρχείου. Στην αρχή της πρώτης γραμμής πρέπει να υπάρχει μια ετικέτα, π.χ. `frame $x$`  όπου  $x$  είναι συνήθως κάποιος αριθμός. Στην αρχή της τελευταίας γραμμής πρέπει να υπάρχει ετικέτα που πρέπει να περιλαμβάνει υποχρεωτικά το γράμμα `f` ώστε να αναγνωρίζεται σωστά. Οι συντεταγμένες χωρίζονται με στηλοθέτες ή χαρακτήρες κενού. Αν μια γραμμή περιέχει περισσότερες συντεταγμένες από τις προβλεπόμενες 38, οι υπόλοιπες μετά τις πρώτες 38 απλά αγνοούνται.

- Αρχείο που περιέχει τους κανόνες για τα μοντελοποιημένα προφίλ.

Ο χρήστης μπορεί να αλλάξει ελεύθερα το πλήθος των κανόνων που εξετάζονται ή/και τις παραμέτρους για τα FAPs που συμμετέχουν σε κάθε κανόνα. Δεν υπάρχει ελάχιστο ούτε μέγιστο πλήθος FAPs που πρέπει να συμμετέχουν στον κανόνα. Είναι ευθύνη του χρήστη να επαληθεύσει ότι τα αριθμητικά δεδομένα που εισάγονται καθώς και οι δείκτες των FAPs είναι σωστοί για κάθε κανόνα.

Ως προς το φορμάτ του αρχείου, υπενθυμίζεται ότι πρέπει να υπάρχει μια αλφαριθμητική συμβολοσειρά στην αρχή του κάθε προφίλ (χωρίς κενά) η οποία θα λειτουργεί ως ετικέτα για το μοντελοποιημένο προφίλ. Και πάλι είναι ευθύνη του χρήστη να εξασφαλίσει ότι η ετικέτα είναι σωστή και έχει νόημα σε σχέση με το προφίλ που ακολουθεί. Το τέλος του κάθε κανόνα πρέπει να συμβολίζεται με μια συμβολοσειρά της μορφής: `fap[0]-[0 0]-[0 0]`. Είναι σημαντικό να υπάρχει το μηδέν στο δείκτη για το FAP (`fap[0]`), καθώς αυτό θα χρησιμοποιηθεί ως κριτήριο για να αναγνωριστεί η λήξη του προφίλ. Επίσης είναι απαραίτητο να υπάρχει η ετικέτα `End` στο τέλος του αρχείου, ώστε να αναγνωρίζεται το τέλος της ακολουθίας κανόνων και να τερματίζει ο βρόχος ανάγνωσης. Οι πληροφορίες των FAPs διαβάζονται σε προκαθορισμένα φορμάτ

ανάγνωσης του τύπου " `fap[%d]-[%f %f]-[%f %f]`". Το κενό μπροστά από το γράμμα `f` στο " `fap[%d]`" είναι εσκεμμένο και είναι σημαντικό να υπάρχει ένας χαρακτήρας κενού για να γίνεται σωστά η ανάγνωση των πληροφοριών. Τα προφίλ μπορούν να διαχωρίζονται με οσοδήποτε κενές γραμμές χωρίς πρόβλημα.

- Επιλογές εκτύπωσης.

Όπως παρουσιάστηκε στο 5.10, σελ. 120 ο χρήστης μπορεί να περάσει τους κατάλληλους χαρακτήρες ως παραμέτρους στη συνάρτηση `analyze_emotions` ώστε να επιλέγει πόσες και ποιές λεπτομέρειες θα τυπώνονται ανάλογα με τις προτιμήσεις ή ανάγκες του.

- Επιλογή μεθόδου περιστροφής των σημείων ενδιαφέροντος

Ο χρήστης επιλέγει τη μέθοδο περιστροφής της προτίμησής του ή εκείνη που θεωρείται ότι αποδίδει καλύτερα δοσμένων των συντ/νων των σημείων. Λεπτομέρειες παρουσιάστηκαν ήδη στο 5.3.4 σελ. 100.

- Επιλογή παραμέτρου  $\varepsilon$  (στο πρόγραμμα `e`) της συνάρτησης συμμετοχής (`mf`).

Συνήθεις τιμές για την παράμετρο είναι οι τιμές  $\varepsilon = 0.01 - 0.05$ . Εάν επιθυμείται η συνάρτηση συμμετοχής να είναι τραπεζοειδής, επιλέγεται  $\varepsilon = 0$ . Ο ρόλος του  $\varepsilon$  αναπτύχθηκε στο 5.8, σελ. 115-118.

## 5.12 Περιοχές τροποποίησης και επέκτασης του προγράμματος.

Το σύστημα εκτίμησης συναισθήματος που υλοποιήθηκε μπορεί να τροποποιηθεί ή να επεκταθεί σχετικά εύκολα ανάλογα με τις ανάγκες της εφαρμογής. Στο υποκεφάλαιο αυτό συζητούνται κάποια σημεία που μπορούν να αποτελέσουν αντικείμενο μιας μελλοντικής επέκτασης.

- Πρόσθετα σημεία ενδιαφέροντος στο πρόσωπο.

Εάν η διαδικασία ανίχνευσης σημείων στο πρόσωπο παρέχει περισσότερα σημεία από τα 19 σημεία που παρουσιάστηκαν παραπάνω, τότε το πρόγραμμα μπορεί να επεκταθεί καταλλήλως ώστε να διαβάζει τα επιπλέον σημεία, να τα αποθηκεύει σε μεγαλύτερα διανύσματα και να τα επεξεργάζεται καταλλήλως. Περισσότερα σημεία θα επέτρεπαν πιθανώς και τον υπολογισμό περισσότερων FAPs, οπότε η επέκταση του λεξιλογίου των FAPs θα μπορούσε να κάνει την εκτίμηση του προφίλ συναισθήματος πιο αποδοτική. Επίσης μπορεί να γίνουν αλλαγές και στον τρόπο υπολογισμού των παράγωγων σημείων ενδιαφέροντος. Στην παρούσα εργασία τα σημεία αυτά υπολογίστηκαν με γραμμική παρεμβολή (παράγραφος 5.4).

- Προεπεξεργασία σημείων ενδιαφέροντος.

Με τον όρο προεπεξεργασία εννοούνται όλοι οι υπολογισμοί και οποιαδήποτε επεξεργασία γίνεται στα σημεία πριν χρησιμοποιηθούν για τον υπολογισμό των FAPs. Στην παρούσα εργασία η προεπεξεργασία των σημείων συνίσταται από το μετασχηματισμό περιστροφής των σημείων, για λόγους που εξηγήθηκαν στην παράγραφο 5.3. Αν κριθεί σκόπιμο, θα μπορούσε να εισαχθεί μια διαφορετική επεξεργασία των σημείων, π.χ. αλλαγή κλίμακας για να αντισταθμίσει τυχόν απομάκρυνση του προσώπου από την κάμερα.

- Διαφορετικός τρόπος υπολογισμού των FAPs.

Τα FAPs στην παρούσα εργασία υπολογίστηκαν σύμφωνα με τον ορισμό τους [47] και με τις προτάσεις του άρθρου [29]. Εάν κριθεί σκόπιμο, θα μπορούσαν να εισαχθούν τροποποιήσεις στον υπολογισμό τους. Κάποια FAPs που υπολογίζονται με κάθετες ή οριζόντιες αποστάσεις (δηλαδή διαφορές ανάμεσα σε  $x$  ή  $y$  συντεταγμένες σημείων) θα μπορούσαν να υπολογίζονται με ευκλείδειες αποστάσεις, ή το αντίστροφο. Επίσης αν κριθεί ωφέλιμο, θα μπορούσε να τροποποιηθεί και ο υπολογισμός των μονάδων FAPs.

- Επέκταση λεξιλογίου FAPs.

Ανάλογα με τη μοντελοποίηση των κανόνων που έχει προηγηθεί, μπορεί να απαιτούνται περισσότερα ή λιγότερα FAPs στο λεξιλόγιο της εφαρμογής. Ο υπολογισμός περισσότερων FAPs ώστε να εμπλουτιστεί το λεξιλόγιο προϋποθέτει κατά βάση τη διαθεσιμότητα περισσότερων σημείων ενδιαφέροντος στο εικονιζόμενο πρόσωπο.

- Χρήση διαφορετικής συνάρτησης συμμετοχής στην αποτίμηση των κανόνων.

Η συνάρτηση συμμετοχής (ΣΣ) που χρησιμοποιήθηκε αποτελεί παραλλαγή τραπεζοειδούς συνάρτησης συμμετοχής. Θα μπορούσε να χρησιμοποιηθεί και κάποια άλλη μορφή για τη ΣΣ, π.χ. καμπανοειδής ΣΣ, ΣΣ Gauss, ή κάποια άλλη μορφή που αναμένεται να έχει ικανοποιητικά αποτελέσματα. Στην περίπτωση αυτή θα χρειαστεί πιθανώς να προσαρμοστεί και το φορμάτ του αρχείου κειμένου που κρατά τις πληροφορίες των κανόνων καθώς και το τμήμα του κώδικα που πραγματοποιεί την ανάγνωση αυτών των δεδομένων.

- Αποτίμηση κανόνων που μοντελοποιούν τα προφίλ συναισθημάτων.

Το κριτήριο που χρησιμοποιήθηκε για την τελική επιλογή του συναισθήματος είναι η μεγιστοποίηση της βεβαιότητας όπως εκφράστηκε με την εξίσωση  $p_i^{(k)} = \prod_{A_{i,j}^{(k)} \in \Delta_{i,j}^{(k)}} r_{i,j}^{(k)}$ , ως γινόμενο των συσχετίσεων  $r_{i,j}^{(k)}$ .

Στο υποκεφάλαιο 5.8 συζητήθηκε επίσης η χρήση του μετρητή επιτυχιών για να είναι διαθέσιμη η πληροφορία του πλήθους των FAPs που εμπίπτουν στα δοσμένα εύρη τιμών κάθε κανόνα. Αυτή η πληροφορία θα μπορούσε να χρησιμοποιηθεί για την έκφραση εναλλακτικού κριτηρίου που συνδυάζει με κάποιο τρόπο τη βεβαιότητα με το μετρητή επιτυχιών. Επίσης η αύξηση του μετρητή θα μπορούσε να δεχτεί κάποιο άλλο κριτήριο. Στην παρούσα εφαρμογή ο μετρητής για να αυξηθεί, αρκεί η συνάρτηση  $mf$  να επιστρέψει τιμή μεγαλύτερη του  $\varepsilon$ . Η αύξηση του μετρητή μπορεί να γίνει και με πιο αυστηρά κριτήρια, π.χ. ο μετρητής αυξάνει όταν  $mf(x, x1, x2, r1, r2, e) > \varepsilon + (1 - \varepsilon) \cdot a\%$  με  $a\%$  να λαμβάνει τιμές  $a\% = 10\%, 20\%$  και ούτω καθ'εξής. Το κριτήριο αυτό υλοποιεί ουσιαστικά μια  $\alpha$ -τομή (δες 4.3.1, Εξ. 4-2, on page 76).

Η αποτίμηση κανόνων είναι ένα πεδίο που επιδέχεται πολλές διαφορετικές υλοποιήσεις.



## ΚΕΦΑΛΑΙΟ 6: Συμπεράσματα και μελλοντικές επεκτάσεις

Σε αυτήν τη διπλωματική εργασία παρουσιάστηκε ένα ολοκληρωμένο πρόγραμμα γραμμένο στη γλώσσα προγραμματισμού C, το οποίο αποσκοπεί στην εκτίμηση του συναισθήματος του εικονιζόμενου προσώπου. Η μέθοδος που υλοποιήθηκε υπολογίζει τις παραμέτρους FAPs του MPEG-4 και αποτιμά κανόνες που έχουν γραφεί ως διαστήματα τιμών FAPs. Η αποτίμηση γίνεται εφαρμόζοντας βασικές αρχές της ασαφούς λογικής: ασαφή διαστήματα, συναρτήσεις συμμετοχής και τριγωνικές νόρμες, δηλαδή ασαφείς τομές και ενώσεις. Οι κανόνες είναι ανεξάρτητοι από την υλοποίηση γεγονός που εξασφαλίζει την ευελιξία της και την άμεση επεκτασιμότητα των προφίλ που εξετάζονται.

### 6.1 Αξιολόγηση της μεθόδου

Παρακάτω παρατίθενται τα συμπεράσματα σχετικά με τα προτερήματα και τα μειονεκτήματα της υλοποίησης και του γενικότερου συστήματος στο οποίο ανήκει.

#### 6.1.1 Πλεονεκτήματα

Η μέθοδος που υλοποιήθηκε στα πλαίσια της παρούσας διπλωματικής χαρακτηρίζεται από κάποια προτερήματα που δικαιολογούν το ενδιαφέρον για χρήση της σε εφαρμογές εκτίμησης συναισθημάτων, ιδιαίτερα εφαρμογές πραγματικού χρόνου.

- Ένα σημαντικό προτέρημα είναι πρωτίστως η **χαμηλή πολυπλοκότητα χώρου και χρόνου**. Δεν απαιτείται η αποθήκευση πολλών δεδομένων, ούτε πληθώρα περίπλοκων υπολογισμών. Η χαμηλή πολυπλοκότητα εξασφαλίζει γρήγορη απόκριση σε πραγματικές εφαρμογές. Αυτό καθιστά την υλοποίηση κατάλληλη για εφαρμογές πραγματικού χρόνου.
- **Ευελιξία**. Ο διαχωρισμός των κανόνων από το υπόλοιπο πρόγραμμα έπαιξε καθοριστικό ρόλο στην ανάπτυξη μια εύελικτης εφαρμογής που μπορεί να χρησιμοποιηθεί για μεταβλητό πλήθος κανόνων, μεταβλητής έκτασης και ποικίλης σημειολογίας. Το πρόγραμμα δεν αξιολογεί τους κανόνες εννοιολογικά, και εναπόκειται στο χρήστη να εξασφαλίσει την περιγραφή των επιθυμητών συναισθημάτων με κατάλληλα προφίλ. Ο μόνος περιορισμός που τίθεται είναι στη μορφοποίηση των προφίλ, η οποία σχολιάστηκε εκτενώς στο προηγούμενο κεφάλαιο, και το πεπερασμένο λεξιλόγιο FAPs της εφαρμογής.

- Η πλήρης αποσύνδεση του προγράμματος από τη μέθοδο που χρησιμοποιείται για την ανίχνευση των σημείων εισόδου επιτρέπει τη χρήση για **σημεία χειροκίνητης ή αυτόματης ανίχνευσης**.
- Η υλοποίηση επιχειρεί να δώσει κάποιες λύσεις για τις **ασυμμετρίες και ατέλειες** που υπάρχουν εγγενή σε όλα τα ανθρώπινα πρόσωπα. Αυτή η μέριμνα είναι απαραίτητη για πρακτικές, ρεαλιστικές εφαρμογές και θεωρείται ότι είναι ένα βήμα προς τη σωστή κατεύθυνση.
- Η ενσωμάτωση παραμέτρων από το πρότυπο MPEG-4 επιτρέπουν ενδεχομένως τη μελλοντική **διασύνδεση με άλλα συστήματα** συμβατά με το διεθνές αυτό πρότυπο.
- Χαμηλά ποσοστά επιτυχία για κάποιο σύνολο κανόνων δεν σημαίνουν την πλήρη αστοχία της μεθόδου και δεν αποκλείουν την **καλύτερη συμπεριφορά με άλλα δεδομένα** για τις εισόδους ή/και τους κανόνες.

### 6.1.2 Μειονεκτήματα

Κάποια από τα μειονεκτήματα οφείλονται στην προγραμματιστική υλοποίηση, ενώ κάποια άλλα μειονεκτήματα της μεθόδου εντοπίζονται στο γενικότερο σύστημα εκτίμησης συναισθημάτων και ενδέχεται να είναι εκτός του πεδίου δράσης της παρούσας εργασίας. Παρολαυτά είναι υπαρκτά προβλήματα που επηρεάζουν τα αποτελέσματα της εφαρμογής.

- Ένα βασικό πρόβλημα είναι η **ποιότητα των δεδομένων εισόδου**. Η εφαρμογή δε θέτει περιορισμούς για τη μέθοδο ανίχνευσης σημείων, μόνο για το πλήθος τους και τη μορφοποίηση εισόδου. Η επιτυχία της εκτίμησης του συναισθήματος και η αξιοπιστία του αποτελέσματος είναι άρρηκτα δεμένα με την ακρίβεια και εγκυρότητα των δεδομένων εισόδου. **Σφάλματα στα δεδομένα της ανίχνευσης μεταφέρονται** σε όλα τα επίπεδα της υλοποίησης και επηρεάζουν δυσμενώς το αποτέλεσμα.
- Ένα εγγενές πρόβλημα της μεθόδου είναι ότι οι **κινήσεις** που επιχειρείται να ανιχνευτούν πολλές φορές τείνουν να είναι συγκρίσιμες με τα σφάλματα αβεβαιότητας της ανίχνευσης. Για παράδειγμα τα FAPs που σχετίζονται με τις κινήσεις των βλεφάρων συμμετέχουν σε πολλά προφίλ συναισθημάτων γιατί τα μάτια είναι μία από τις εκφραστικές περιοχές του προσώπου. Οι κινήσεις των βλεφάρων όμως –ειδικά του κάτω βλεφάρου– είναι **μικρές συγκριτικά** με τις υπόλοιπες διαστάσεις του προσώπου, και **δύσκολα ανιχνεύσιμες** σε χαμηλής ανάλυσης εικόνες.
- Η διαδικασία της ανίχνευσης είναι **ευαίσθητη σε συνθήκες περιβάλλοντος** όπως οι συνθήκες φωτισμού, η θέση του χρήστη σε σχέση με την κάμερα, οι μεταβολές στην απόσταση και τη σχετική θέση κάμερας-χρήστη.
- Στην προτεινόμενη μέθοδο **δεν αξιοποιούνται** καθόλου οι λοιπές **κινήσεις του κεφαλιού** που συνοδεύουν μια συναισθηματική έκφραση. Για παράδειγμα στο συναίσθημα της οργής το κεφάλι μπορεί να γείρει προς τα μπροστά. Απεναντίας, αυτές οι κινήσεις **προκαλούν πρόβλημα** στην εκτίμηση του συναισθήματος. Η

συναισθηματική εμπειρία συνοδεύεται όμως από κινήσεις του σώματος και χειρονομίες, που είναι εκτός της εμβέλειας της παρούσας εφαρμογής.

- Χρησιμοποιείται **μόνο η πληροφορία της έκφρασης του προσώπου**, και μάλιστα αυτή συμπυκνώνεται στις **συντεταγμένες 19 σημείων**. Με τον τρόπο αυτό χάνουμε την πληροφορία της αλλαγής του μυϊκού τόνου, των μεταβολών στα περιγράμματα των χαρακτηριστικών. Για παράδειγμα το αίσθημα της απέχθειας συνοδεύεται συχνά από ασύμμετρο τράβηγμα του πάνω χείλους, που μπορεί εύκολα να ερμηνευθεί από άνθρωπο, χάνει όμως την περιεκτικότητα σε πληροφορία όταν η κατάσταση των χειλιών συμπυκνώνεται στις συντεταγμένες τεσσάρων μόνο σημείων.
- **Περιορισμένο, πεπερασμένο λεξιλόγιο FAPs**. Το πρόγραμμα υπολογίζει λεξιλόγιο έκτασης 23 FAPs. Αν χρησιμοποιηθούν κανόνες με άλλα FAPs, οι πληροφορίες για τα μη-υλοποιήσιμα FAPs αγνοούνται. Αυτό μοιραία μπορεί να συντελέσει στην ανακρίβεια και αναποτελεσματικότητα της εφαρμογής.
- **Αδυναμία υπολογισμού μεταβολών βάθους** χαρακτηριστικών όπως το πηγούνι ή τα χείλη, που όμως πιθανώς χρειάζονται για τα προφίλ συναισθημάτων όπως η χαρά, ο θυμός (δες παράρτημα). Επίσης **δεν υπάρχει ελευθερία στροφικής κίνησης** του προσώπου καθώς δεν είναι δυνατόν να αντισταθμιστούν όλες οι πιθανές περιστροφές του προσώπου.
- **Μη-αντιπροσωπευτικοί κανόνες** ή κανόνες που προέκυψαν από data set με μεγάλες διαφορές **καθιστούν την εφαρμογή αναποτελεσματική**.

## 6.2 Μελλοντικές εφαρμογές και επεκτάσεις

### 6.2.1 Προτεινόμενες εναλλακτικές εφαρμογές

Η ανεξαρτητοποίηση των μοντελοποιημένων προφίλ παρέχει τη δυνατότητα της χρήσης του προγράμματος για διαφορετικές εφαρμογές, όχι απαραίτητα στο πεδίο της εκτίμησης συναισθήματος. Ανάλογα με τους κανόνες που παρέχονται στο πρόγραμμα, είναι εφικτό να γίνει **ανίχνευση συγκεκριμένων κινήσεων** που δεν έχουν απαραίτητα συναισθηματική ερμηνεία, όπως για παράδειγμα το σήκωμα των φρυδιών, το κλείσιμο των ματιών, το άνοιγμα του στόματος. Τέτοιες λειτουργίες έχουν ενδιαφέρον για διαδραστικές εφαρμογές επικοινωνίας ανθρώπου-μηχανής. Εναλλακτικά μπορούν να παρέχονται κανόνες που αποσκοπούν στην **ποσοτική και ποιοτική μελέτη των FAPs** ή ενδεχομένως τη συσχέτιση FAPs με AUs (Action Units του συστήματος FACS).

Μια ακόμη ενδιαφέρουσα λειτουργία είναι και η **μελέτη της συμπεριφοράς των FAPs** σε πραγματικά δεδομένα εικόνων. Για παράδειγμα μπορεί να μελετηθεί κατά πόσον παρατηρούνται συγκεκριμένοι συνδυασμοί προσήμων ή αν ικανοποιούνται κάποια άλλα κριτήρια συσχέτισης των FAPs. Αυτή η χρήση έχει νόημα και για την **αξιολόγηση υπαρκτών κανόνων**, με σκοπό τη διόρθωση ή και προσαρμογή τους.

### 6.2.2 Προτεινόμενες επεκτάσεις

Προτείνεται να διερευνηθεί η **ανάπτυξη** μιας περισσότερο **εύρωστης μεθόδου** είτε για την ανίχνευση είτε για τη μετέπειτα επεξεργασία των δεδομένων εισόδου. Είναι επιθυμητό να αναπτυχθεί μια μέθοδος που είναι λιγότερο ευαίσθητη σε σφάλματα της ανίχνευσης.

Μία ακόμη εναλλακτική είναι η ανάπτυξη αλγορίθμου που χρησιμοποιεί **πληροφορίες από ακολουθία εικόνων** και δε συγκρίνει μόνο δύο καρέ της ακολουθίας βίντεο.

Επίσης προτείνεται η **αξιοποίηση της πλήρους συναισθηματικής έκφρασης των ανθρώπων**. Θα ήταν επιθυμητό να αναπτυχθεί μια μέθοδος που θα συλλέγει **συνδυασμό πληροφοριών** όπως η **ομιλία**, οι **κινήσεις του κεφαλιού**, του **σώματος** και οι **χειρονομίες**. Μελλοντικές εργασίες θα πρέπει επίσης να **απομακρυνθούν από τον περιορισμό των έξι αρχέτυπων συναισθημάτων** που τείνει να γίνει στερεότυπο. Η αναγνώριση συναισθηματικών/διανοητικών καταστάσεων όπως η επίδειξη ενδιαφέροντος (interest), η απορία (puzzlement), η απογοήτευση (frustration), η κούραση (fatigue), η σύγχυση (confusion), το άγχος (stress) και η ανία (boredom) έχουν πιθανώς μεγαλύτερη εννοιολογική χρησιμότητα στην επικοινωνία ανθρώπου-μηχανής.

Η τελευταία προτεινόμενη επέκταση σχετίζεται με τη χρήση **τρισδιάστατης αντί για διδιάστατης πληροφορίας**. Έχουν γίνει κάποιες εργασίες [74] που χρησιμοποιούν τεχνικές τρισδιάστατης σάρωσης του προσώπου με ενθαρρυντικά αποτελέσματα στον τομέα της φυσιολογικής αναγνώρισης προσώπων. Οι τεχνικές αυτές ή ακόμη τεχνικές με τρισδιάστατη ανίχνευση περισσότερων χαρακτηριστικών σημείων, θα μπορούσαν ίσως να επεκταθούν και στην ανάλυση της έκφρασης με απώτερο σκοπό την αναγνώριση του συναισθήματος.

## Παράρτημα



## Κανόνες για τα αρχέτυπα συναισθήματα [29]

| ANGER            |                 |                   |              |              |              |              |              |
|------------------|-----------------|-------------------|--------------|--------------|--------------|--------------|--------------|
| Vocabulary       |                 |                   | Profiles     |              |              |              |              |
| Group            | FAP No.         | FAP Name          | $P_A^{(0)}$  | $P_A^{(1)}$  | $P_A^{(2)}$  | $P_A^{(3)}$  | $P_A^{(4)}$  |
| Midlip, chin (2) | F <sub>4</sub>  | lower_t_midlip    | [22, 124]    | -            | -            | -            | -            |
|                  | F <sub>16</sub> | push_b_lip        | -            | -            | -            | [45, 155]    | [-355, -245] |
|                  | F <sub>18</sub> | depress_chin      | -            | -            | -            | [45, 155]    | [145, 255]   |
| Eyelids (3)      | F <sub>19</sub> | close_t_l_eyelid  | -            | [-330, -200] | [-330, -200] | [-330, -200] | [-330, -200] |
|                  | F <sub>20</sub> | close_t_r_eyelid  | -            | [-335, -205] | [-335, -205] | [-330, -200] | [-330, -200] |
|                  | F <sub>21</sub> | close_b_l_eyelid  | -            | [200, 330]   | [200, 330]   | -            | -            |
|                  | F <sub>22</sub> | close_b_r_eyelid  | -            | [205, 335]   | [205, 335]   | -            | -            |
| Eyebrows (4)     | F <sub>31</sub> | raise_l_i_eyebrow | [-131, -25]  | [-200, -80]  | [-200, -80]  | [-200, -80]  | [-200, -80]  |
|                  | F <sub>32</sub> | raise_r_i_eyebrow | [-136, -34]  | [-194, -74]  | [-194, -74]  | [-194, -74]  | [-194, -74]  |
|                  | F <sub>33</sub> | raise_l_m_eyebrow | [-189, -109] | [-190, -70]  | [70, 190]    | [-190, -70]  | [-190, -70]  |
|                  | F <sub>34</sub> | raise_r_m_eyebrow | [-183, -105] | [-190, -70]  | [70, 190]    | [-190, -70]  | [-190, -70]  |
|                  | F <sub>35</sub> | raise_l_o_eyebrow | [-101, -31]  | -            | -            | -            | -            |
|                  | F <sub>36</sub> | raise_r_o_eyebrow | [-108, -32]  | -            | -            | -            | -            |
|                  | F <sub>37</sub> | squeeze_l_eyebrow | [29, 85]     | -            | -            | [65, 135]    | [65, 135]    |
|                  | F <sub>38</sub> | squeeze_r_eyebrow | [27, 89]     | -            | -            | [65, 135]    | [65, 135]    |

## SADNESS

| Vocabulary   |                 |                   | Profiles    |
|--------------|-----------------|-------------------|-------------|
| Group        | FAP No.         | FAP Name          | $P_S^{(0)}$ |
| Eyelids (3)  | F <sub>19</sub> | close_t_l_eyelid  | [-265, -41] |
|              | F <sub>20</sub> | close_t_r_eyelid  | [-270, -52] |
|              | F <sub>21</sub> | close_b_l_eyelid  | [-265, -41] |
|              | F <sub>22</sub> | close_b_r_eyelid  | [-270, -52] |
| Eyebrows (4) | F <sub>31</sub> | raise_l_i_eyebrow | [30, 140]   |
|              | F <sub>32</sub> | raise_r_i_eyebrow | [26, 134]   |

## JOY

| Vocabulary              |                 |                       | Profiles     |              |              |              |
|-------------------------|-----------------|-----------------------|--------------|--------------|--------------|--------------|
| Group                   | FAP No.         | FAP Name              | $P_J^{(0)}$  | $P_J^{(1)}$  | $P_J^{(2)}$  | $P_J^{(3)}$  |
| Midlip, cornerlips (2)  | F <sub>4</sub>  | lower_t_midlip        | [-381, -161] | -            | -            | -            |
|                         | F <sub>6</sub>  | stretch_l_cornerlip   | [136, 332]   | [160, 240]   | [160, 240]   | [160, 240]   |
|                         | F <sub>7</sub>  | stretch_r_cornerlip   | [136, 332]   | [160, 240]   | [160, 240]   | [160, 240]   |
|                         | F <sub>12</sub> | raise_l_cornerlip     | -            | [260, 340]   | [260, 340]   | [260, 340]   |
|                         | F <sub>13</sub> | raise_r_cornerlip     | -            | [260, 340]   | [260, 340]   | [260, 340]   |
| Eyelids (3)             | F <sub>19</sub> | close_t_l_eyelid      | [-387, -121] | [-449, -325] | [-449, -325] | [-449, -325] |
|                         | F <sub>20</sub> | close_t_r_eyelid      | [-364, -120] | [-426, -302] | [-426, -302] | [-426, -302] |
|                         | F <sub>21</sub> | close_b_l_eyelid      | [-387, -121] | [325, 449]   | [-312, -188] | [61, 185]    |
|                         | F <sub>22</sub> | close_b_r_eyelid      | [-364, -120] | [302, 426]   | [-289, -165] | [38, 162]    |
| Eyebrows (4)            | F <sub>33</sub> | raise_l_m_eyebrow     | [2, 46]      | [70, 130]    | [70, 130]    | [70, 130]    |
|                         | F <sub>34</sub> | raise_r_m_eyebrow     | [3, 47]      | [70, 130]    | [70, 130]    | [70, 130]    |
| Cheeks (5)              | F <sub>41</sub> | lift_l_cheek          | -            | [130, 170]   | [130, 170]   | [130, 170]   |
|                         | F <sub>42</sub> | lift_r_cheek          | -            | [130, 170]   | [130, 170]   | [130, 170]   |
| Outer lip positions (8) | F <sub>53</sub> | stretch_l_cornerlip_o | [136, 332]   | [160, 240]   | [160, 240]   | [160, 240]   |
|                         | F <sub>54</sub> | stretch_r_cornerlip_o | [136, 332]   | [160, 240]   | [160, 240]   | [160, 240]   |

## DISGUST

| Vocabulary   |                 |                   | Profiles     |
|--------------|-----------------|-------------------|--------------|
| Group        | FAP No.         | FAP Name          | $P_D^{(0)}$  |
| Midlip (2)   | F <sub>4</sub>  | lower_t_midlip    | [-343, -125] |
|              | F <sub>5</sub>  | raise_b_midlip    | [-285, -69]  |
| Eyelids (3)  | F <sub>19</sub> | close_t_l_eyelid  | [-351, -55]  |
|              | F <sub>20</sub> | close_t_r_eyelid  | [-356, -66]  |
|              | F <sub>21</sub> | close_b_l_eyelid  | [-351, -55]  |
|              | F <sub>22</sub> | close_b_r_eyelid  | [-356, -66]  |
| Eyebrows (4) | F <sub>33</sub> | raise_l_m_eyebrow | [-123, -27]  |
|              | F <sub>34</sub> | raise_r_m_eyebrow | [-126, -28]  |



## FEAR (I)

| Vocabulary                                  |                 |                   | Profiles    |              |              |              |              |
|---|-----------------|-------------------|-------------|--------------|--------------|--------------|--------------|
| Group                                       | FAP No.         | FAP Name          | $P_F^{(0)}$ | $P_F^{(1)}$  | $P_F^{(2)}$  | $P_F^{(3)}$  | $P_F^{(4)}$  |
| Jaw, midlip, inner lowerlip, cornerlips (2) | F <sub>3</sub>  | open_jaw          | [102, 480]  | [400, 560]   | [400, 560]   | [400, 560]   | [400, 560]   |
|   | F <sub>5</sub>  | raise_b_midlip    | [83, 353]   | [333, 373]   | [307, 399]   | [-240, -160] | [-240, -160] |
|   | F <sub>8</sub>  | lower_t_lip_lm    | -           | -            | -            | -            | -            |
|   | F <sub>9</sub>  | lower_t_lip_rm    | -           | -            | -            | -            | -            |
|   | F <sub>10</sub> | raise_b_lip_lm    | -           | -            | -            | -            | -            |
|   | F <sub>11</sub> | raise_b_lip_rm    | -           | -            | -            | -            | -            |
| Eyelids (3)                                 | F <sub>19</sub> | close_t_l_eyelid  | [118, 370]  | [-400, -340] | [-530, -470] | [-630, -570] | [-630, -570] |
|   | F <sub>20</sub> | close_t_r_eyelid  | [121, 377]  | [-407, -347] | [-523, -463] | [-630, -570] | [-630, -570] |
|   | F <sub>21</sub> | close_b_l_eyelid  | [118, 370]  | [-400, -340] | [-530, -470] | [-630, -570] | [-630, -570] |
|   | F <sub>22</sub> | close_b_r_eyelid  | [121, 377]  | [-407, -347] | [-523, -463] | [-630, -570] | [-630, -570] |
| Eyebrows (4)                                | F <sub>31</sub> | raise_l_i_eyebrow | [35, 173]   | -            | [460, 540]   | [460, 540]   | [460, 540]   |
|   | F <sub>32</sub> | raise_r_i_eyebrow | [39, 183]   | -            | [460, 540]   | [460, 540]   | [460, 540]   |
|   | F <sub>33</sub> | raise_l_m_eyebrow | [14, 130]   | -            | [460, 540]   | -            | [360, 440]   |
|   | F <sub>34</sub> | raise_r_m_eyebrow | [15, 135]   | -            | [460, 540]   | -            | [360, 440]   |
|   | F <sub>35</sub> | raise_l_o_eyebrow | -           | -            | [460, 540]   | -            | [260, 340]   |
|   | F <sub>36</sub> | raise_r_o_eyebrow | -           | -            | [460, 540]   | -            | [260, 340]   |
|   | F <sub>37</sub> | squeeze_l_eyebrow | -           | -            | -            | [60, 140]    | [60, 140]    |
|   | F <sub>38</sub> | squeeze_r_eyebrow | -           | -            | -            | [60, 140]    | [60, 140]    |

## FEAR (II)

| Vocabulary                                  |                 |                   | Profiles     |              |              |              |              |
|---|-----------------|-------------------|--------------|--------------|--------------|--------------|--------------|
| Group                                       | FAP No.         | FAP Name          | $P_F^{(5)}$  | $P_F^{(6)}$  | $P_F^{(7)}$  | $P_F^{(8)}$  | $P_F^{(9)}$  |
| Jaw, midlip, inner lowerlip, cornerlips (2) | F <sub>3</sub>  | open_jaw          | [400, 560]   | [400, 560]   | [400, 560]   | [400, 560]   | [400, 560]   |
|   | F <sub>5</sub>  | raise_b_midlip    | [-240, -160] | [-240, -160] | [-240, -160] | [-240, -160] | [307, 399]   |
|   | F <sub>8</sub>  | lower_t_lip_lm    | -            | [-120, -80]  | -            | -            | -            |
|   | F <sub>9</sub>  | lower_t_lip_rm    | -            | [-120, -80]  | -            | -            | -            |
|   | F <sub>10</sub> | raise_b_lip_lm    | -            | [-120, -80]  | -            | -            | -            |
|   | F <sub>11</sub> | raise_b_lip_rm    | -            | [-120, -80]  | -            | -            | -            |
| Eyelids (3)                                 | F <sub>19</sub> | close_t_l_eyelid  | [-630, -570] | [-630, -570] | [-630, -570] | [-630, -570] | [-630, -570] |
|   | F <sub>20</sub> | close_t_r_eyelid  | [-630, -570] | [-630, -570] | [-630, -570] | [-630, -570] | [-623, -563] |
|   | F <sub>21</sub> | close_b_l_eyelid  | [-630, -570] | [-630, -570] | [-630, -570] | [-630, -570] | [-630, -570] |
|   | F <sub>22</sub> | close_b_r_eyelid  | [-630, -570] | [-630, -570] | [-630, -570] | [-630, -570] | [-623, -563] |
| Eyebrows (4)                                | F <sub>31</sub> | raise_l_i_eyebrow | [460, 540]   | [460, 540]   | [360, 440]   | [260, 340]   | [460, 540]   |
|   | F <sub>32</sub> | raise_r_i_eyebrow | [460, 540]   | [460, 540]   | [360, 440]   | [260, 340]   | [460, 540]   |
|   | F <sub>33</sub> | raise_l_m_eyebrow | [360, 440]   | [360, 440]   | [260, 340]   | [160, 240]   | [460, 540]   |
|   | F <sub>34</sub> | raise_r_m_eyebrow | [360, 440]   | [360, 440]   | [260, 340]   | [160, 240]   | [460, 540]   |
|   | F <sub>35</sub> | raise_l_o_eyebrow | [260, 340]   | [260, 340]   | [160, 240]   | [60, 140]    | [460, 540]   |
|   | F <sub>36</sub> | raise_r_o_eyebrow | [260, 340]   | [260, 340]   | [160, 240]   | [60, 140]    | [460, 540]   |
|   | F <sub>37</sub> | squeeze_l_eyebrow | ∈0           | ∈0           | -            | -            | -            |
|   | F <sub>38</sub> | squeeze_r_eyebrow | ∈0           | ∈0           | -            | -            | -            |

# SURPRISE

| Vocabulary  |                 |                       | Profiles       |                |                |                |
|---|-----------------|-----------------------|----------------|----------------|----------------|----------------|
| Group   | FAP No.         | FAP Name              | $P_{Su}^{(0)}$ | $P_{Su}^{(1)}$ | $P_{Su}^{(2)}$ | $P_{Su}^{(3)}$ |
| Jaw, midlip,<br>inner lowerlip,<br>cornerlips (2) | F <sub>3</sub>  | open_jaw              | [569, 1201]    | [1150, 1252]   | [834, 936]     | [523, 615]     |
|   | F <sub>5</sub>  | raise_b_midlip        | [340, 746]     | [-792, -700]   | [-589, -497]   | [-386, -294]   |
|   | F <sub>6</sub>  | stretch_l_cornerlip   | [-121, -43]    | [-141, -101]   | [-102, -62]    | [-63, -23]     |
|   | F <sub>7</sub>  | stretch_r_cornerlip   | [-121, -43]    | [-141, -101]   | [-102, -62]    | [-63, -23]     |
|   | F <sub>10</sub> | raise_b_lip_lm        | -              | [-530, -470]   | [-380, -320]   | [-230, -170]   |
|   | F <sub>11</sub> | raise_b_lip_rm        | -              | [-530, -470]   | [-380, -320]   | [-230, -170]   |
| Eyelids<br>(3)                                    | F <sub>19</sub> | close_t_l_eyelid      | [170, 337]     | [-350, -324]   | [-267, -241]   | [-184, -158]   |
|   | F <sub>20</sub> | close_t_r_eyelid      | [171, 333]     | [-346, -320]   | [-265, -239]   | [-184, -158]   |
|   | F <sub>21</sub> | close_b_l_eyelid      | [170, 337]     | [-350, -324]   | [-267, -241]   | [-184, -158]   |
|   | F <sub>22</sub> | close_b_r_eyelid      | [171, 333]     | [-346, -320]   | [-265, -239]   | [-184, -158]   |
| Eyebrows (4)                                      | F <sub>31</sub> | raise_l_i_eyebrow     | [121, 327]     | [314, 340]     | [211, 237]     | [108, 134]     |
|   | F <sub>32</sub> | raise_r_i_eyebrow     | [114, 308]     | [295, 321]     | [198, 224]     | [101, 127]     |
|   | F <sub>33</sub> | raise_l_m_eyebrow     | [80, 208]      | [195, 221]     | [131, 157]     | [67, 93]       |
|   | F <sub>34</sub> | raise_r_m_eyebrow     | [80, 204]      | [191, 217]     | [129, 155]     | [67, 93]       |
|   | F <sub>35</sub> | raise_l_o_eyebrow     | [23, 85]       | [72, 98]       | [41, 67]       | [10, 36]       |
|   | F <sub>36</sub> | raise_r_o_eyebrow     | [23, 85]       | [73, 99]       | [42, 68]       | [11, 37]       |
| Outer lip<br>positions<br>(8)                     | F <sub>53</sub> | stretch_l_cornerlip_o | [-121, -43]    | [-141, -101]   | -              | -              |
|   | F <sub>54</sub> | stretch_r_cornerlip_o | [-121, -43]    | [-141, -101]   | -              | -              |



# Αναφορές

- [1] Ronald de Sousa. (2012, March) The Stanford Encyclopedia of Philosophy. [Online]. <http://plato.stanford.edu/archives/spr2012/entries/emotion/>
- [2] Anne M. Kleinginna and Paul R. Kleinginna Jr. (1981) A Categorized List of Emotion Definitions, with Suggestions for a Consensual Definition. PDF. [Online]. <http://148.202.18.157/sitios/catedrasnacionales/material/2010b/sanz/kleinginna1981.pdf>
- [3] Richard J. Gerrig and Philip J. Zimbardo. (2012) American Psychological Association. [Online]. <http://www.apa.org/research/action/glossary.aspx>
- [4] (2012, May) Wikipedia. [Online]. [http://en.wikipedia.org/wiki/Four\\_temperaments](http://en.wikipedia.org/wiki/Four_temperaments)
- [5] (2012, June) Wikipedia. [Online]. <http://en.wikipedia.org/wiki/Emotion>
- [6] Classics in the History of Psychology. [Online]. <http://psychclassics.yorku.ca/James/emotion.htm>
- [7] (2012, April) Wikipedia. [Online]. [http://en.wikipedia.org/wiki/Cannon-Bard\\_theory](http://en.wikipedia.org/wiki/Cannon-Bard_theory)
- [8] (2012, June) Sparknotes. [Online]. <http://www.sparknotes.com/psychology/psych101/emotion/section1.html>
- [9] (2012) CliffsNotes. [Online]. [http://www.cliffsnotes.com/study\\_guide/Early-Theories-of-Emotion.topicArticleId-25438,articleId-25365.html](http://www.cliffsnotes.com/study_guide/Early-Theories-of-Emotion.topicArticleId-25438,articleId-25365.html)
- [10] (2011, March) Psychwiki. [Online]. [http://www.psychwiki.com/wiki/The\\_Schachter-Singer\\_Theory\\_of\\_Emotion](http://www.psychwiki.com/wiki/The_Schachter-Singer_Theory_of_Emotion)
- [11] Mark Holah. Psychology Holah.co.uk. [Online]. <http://www.holah.karoo.net/schachterstudy.htm>
- [12] (2012, April) Wikipedia. [Online]. [http://en.wikipedia.org/wiki/Appraisal\\_theory](http://en.wikipedia.org/wiki/Appraisal_theory)
- [13] Joshua Ian Davis, Ann Senghas, and Kevin N. Ochsner, "How does facial feedback modulate emotional experience?," *Journal of Research in Personality*, pp. 822-829, 2009.
- [14] Fritz Strack, Leonard L. Martin, and Sabine Stepper, "Inhibiting and Facilitating Conditions of the Human Smile: A Nonobtrusive Test of the Facial Feedback Hypothesis," *Journal of Personality and Social Psychology*, vol. 54, no. 5, pp. 768-777, 1988.
- [15] Andreas Hennenlotter et al., "The Link between Facial Feedback and Neural Activity within Central Circuitries of Emotion - New Insights from Botulinum Toxin-Induced Denervation of Frown Muscles," *Cerebral Cortex*, vol. 19, no. 3, pp. 537-542, June 2008.
- [16] David A. Havas, Arthur M. Glenberg, Karol A. Gutowski, Mark J. Lucarelli, and Richard J. Davidson, "Cosmetic Use of Botulinum Toxin-A Affects Processing of Emotional Language," *Psychological Science*, vol. 21, no. 7, pp. 895-900, 2010.
- [17] Charles R. Darwin, "The expression of the emotions in man and animals," in *The expression of the emotions in man and animals*. London: John Murray, 1872, ch. XIV, p. 366.
- [18] Paul Ekman, "Facial Expression and Emotion," *American Psychiatrist*, vol. 48, no. 4, pp. 384-392, April 1993.
- [19] Paul Ekman, "Universal Facial Expressions of Emotion," *California Mental Health Research Digest*, vol. 8, no. 4, pp. 151-158, Autumn 1970.
- [20] (2012, July) Wikipedia. [Online]. [http://en.wikipedia.org/wiki/Paul\\_Ekman](http://en.wikipedia.org/wiki/Paul_Ekman)
- [21] Joseph C. Hager and Paul Ekman, "The Inner and Outer Meanings of Facial Expressions," in *Social Psychophysiology: A Sourcebook*, John T. Cacioppo and R. E. Petty, Eds. New York: The Guilford Press, 1983, ch. 10.
- [22] (2002) DataFace. [Online]. <http://face-and-emotion.com/dataface/facs/manual/TOC.html>
- [23] (2012, June) Wikipedia. [Online].

- [http://en.wikipedia.org/wiki/Facial\\_Action\\_Coding\\_System](http://en.wikipedia.org/wiki/Facial_Action_Coding_System)
- [24] Gregory Johnson. (2009, June) Internet Encyclopedia of Philosophy. [Online]. <http://www.iep.utm.edu/emotion/>
- [25] (2012, July) Wikipedia. [Online]. [http://en.wikipedia.org/wiki/Robert\\_Plutchik](http://en.wikipedia.org/wiki/Robert_Plutchik)
- [26] Markus Drews. (2007, February) Robert Plutchik's Psychoevolutionary Theory of Basic Emotions. PDF. [Online]. <http://www.markusdrews.de/Plutchiks.Emotionstheorie.PLAkat.pdf>
- [27] Robert Plutchik, "The Nature of Emotions," *American Scientist*, vol. 89, pp. 344-350, July-August 2001.
- [28] 6seconds.org. (2011) Plutchik's emotion wheel. gif. [Online]. <http://www.6seconds.org/wp-content/uploads/2011/06/Plutchik.gif>
- [29] A. Raouzaïou, N. Tsapatsoulis, K. Karpouzis, and S. Kollias, "Parametrized facial expression synthesis based on MPEG-4," *EURASIP Journal on Applied Signal Processing*, vol. 2002, no. 1, January 2002.
- [30] Αμαρυλλίς Ραουζαίου, Αναπαράσταση Συναισθήματος και Σύνθεση Εκφράσεων Προσώπου και Χειρονομιών σε Εικονικούς Χαρακτήρες, 2006, Διδακτορική διατριβή.
- [31] A. Raouzaïou, K. Karpouzis, and Stefanos D. Kollias. (2003) Online Gaming and Emotion Representation. PDF. [Online]. <http://www.image.ece.ntua.gr/papers/277.pdf>
- [32] Rosalind W. Picard, "Affective Computing," MIT Media Laboratory, Cambridge, MA, Technical Report 321, 1995.
- [33] M. S. Bartlett et al., "Fully Automatic Facial Action Recognition in Spontaneous Behavior," in *Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition (AFGR '06)*, 2006, pp. 223-230.
- [34] Maja Pantic and Ioannis Patras, "Dynamics of facial expression: Recognition of Facial Actions and Their Temporal Segments From Face Profile Image Sequences," *IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics*, vol. 32, no. 2, pp. 433-499, April 2006.
- [35] Yongmian Zhang and Qiang Ji, "Active and Dynamic Information Fusion for Facial Expression Understanding from Image Sequences," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 5, May 2005.
- [36] Michel F. Valstar, Hatice Gunes, and Maja Pantic, "How to Distinguish Posed from Spontaneous Smiles using Geometric Features," in *International Conference Multimodal Interfaces*, Nagoya, Aichi, Japan, 2007, pp. 38-45.
- [37] M. Yeasin, B. Bullot, and R. Sharma, "Recognition of Facial Expressions and Measurement of Levels of Interest from Video," *Transactions on Multimedia*, vol. 8, no. 3, pp. 500-507, June 2006.
- [38] Jeffrey Cohn, L. I. Reed, Zara Ambadar, Jing Xiao, and Tsuyoshi Moriyama, "Automatic analysis and recognition of brow actions and head motion in spontaneous facial behavior," in *Proceedings of the IEEE Conference on Systems, Man, and Cybernetics*, 2004, pp. 610-616.
- [39] Zhihong Zeng, Maja Pantic, Glenn I. Roisman, and Thomas S. Huang, "A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions," *IEEE Transactions Pattern Analysis and Machine Intelligence*, vol. 31, no. 1, pp. 39-58, January 2009.
- [40] F. Nasoz and C. L. Lisetti, "MAUI avatars: Mirroring the user's sensed emotions via expressive multi-ethnic facial avatars," *Journal of Visual Languages & Computing*, vol. 17, no. 5, pp. 430-444, October 2006.
- [41] L. Maat and M. Pantic, "Gaze-X: Adaptive Affective Multimodal Interface for Single-User Office Scenarios," in *Proc. Eighth ACM Int'l Conf. Multimodal Interfaces (ICMI '06)*, 2006, pp. 171-178.

- [42] Ashish Kapoor, Winslow Burleson, and Rosalind W. Picard, "Automatic prediction of frustration," *International Journal of Human-Computer Studies*, vol. 65, pp. 724-736, 2007.
- [43] Beckman Institute. (2001) Multimodal Human Computer Interaction: Toward a Proactive Computer. [Online]. <http://itr.beckman.uiuc.edu/>
- [44] (2002, March) The MPEG Home Page. [Online]. <http://mpeg.chiariglione.org/standards/mpeg-4/mpeg-4.htm>
- [45] (2012, July) Wikipedia. [Online]. <http://en.wikipedia.org/wiki/MPEG-4>
- [46] (2012, July) ISO/IEC. [Online]. <http://www.itscj.ipsj.or.jp/sc29/29w42911.htm#MPEG-4>
- [47] A. Murat Tekalp and Jorn Ostermann, "Face and 2-D mesh animation in MPEG-4," *Signal Processing: Image Communication*, vol. 15, pp. 387-421, 1999.
- [48] Visage Technologies. [Online]. [www.visagetechologies.com/MPEG-4FBAOverview.pdf](http://www.visagetechologies.com/MPEG-4FBAOverview.pdf)
- [49] (2008, March) Ελληνική Ελεύθερη Εγκυκλοπαίδεια. [Online]. [http://www.livepedia.gr/index.php/Ασαφής\\_Λογική](http://www.livepedia.gr/index.php/Ασαφής_Λογική)
- [50] Lotfi A. Zadeh, Eugene M. Izhikevich, and Serguei A. Mokhov. (2007) Scholarpedia. [Online]. [www.scholarpedia.org/article/Fuzzy\\_logic](http://www.scholarpedia.org/article/Fuzzy_logic)
- [51] Steven D. Kaehler. Seattle Robotics Society. [Online]. [http://www.seattlerobotics.org/encoder/mar98/fuz/fl\\_part1.html](http://www.seattlerobotics.org/encoder/mar98/fuz/fl_part1.html)
- [52] Lotfi A. Zadeh. (1965) The Berkeley Initiative in Soft Computing. [Online]. <http://www-bisc.cs.berkeley.edu/Zadeh-1965.pdf>
- [53] W. Pedrycz and F. Gomide, *An Introduction to Fuzzy Sets. Analysis and Design*. Cambridge, MA: MIT Press, 1998.
- [54] J. Mendel, *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*. New Jersey: Prentice Hall, 2000.
- [55] J. A. Goguen, *L-fuzzy sets*, *J. Math. Anal. Appl.* 18., 1967.
- [56] W. Zhang, "Bipolar Fuzzy Sets," in *Proceedings of IEEE World Congress on Computational Science*, Anchorage, AK, 1998, pp. 835-840.
- [57] S. Benferhat, D. Dubois, S. Kaci, and H. Prade, "Bipolar possibility theory in preference modeling: Representation, fusion and optimal solutions.," *Information Fusion, an International Journal on Multi-Sensor, Multi-Source Information Fusion*, 2005.
- [58] K. T. Atanassov, "Intuitionistic fuzzy sets," *Fuzzy Sets and Systems*, vol. 20, no. 1, pp. 87-96, August 1986.
- [59] E. P. Klement, R. Mesiar, and E. Pap, "Triangular Norms," *Trends in Logic*, vol. 8, 2000.
- [60] V. Novak, I. Perfilieva, and J. Mockor, *Mathematical principles of fuzzy logic*. Boston/Dordrecht: Kluwer, 1999.
- [61] Lotfi A. Zadeh, Fuzzy sets as a basis for a theory of possibility, 1978.
- [62] D. Dubois and H. Prade, *Possibility Theory*. New York: Plenum Press, 1988.
- [63] Milan Mares. (2006) Scholarpedia. [Online]. [http://www.scholarpedia.org/article/Fuzzy\\_sets](http://www.scholarpedia.org/article/Fuzzy_sets)
- [64] Σταύρος Βολογιαννίδης. (2009) T.E.I. Σεργίων. [Online]. [www.teiser.gr/icd/staff/vologian/files/intel.pdf](http://www.teiser.gr/icd/staff/vologian/files/intel.pdf)
- [65] Bryan Klingenberg. Calvin College. [Online]. <http://www.calvin.edu/~pribeiro/othrlnks/Fuzzy/fuzzysets.htm>
- [66] J. Dombi. (1988) Membership function as an evaluation. PDF.
- [67] Mirko Navara. (2007) Scholarpedia. [Online]. [http://www.scholarpedia.org/article/Triangular\\_norms\\_and\\_conorms](http://www.scholarpedia.org/article/Triangular_norms_and_conorms)
- [68] Milan Mares, *Computations Over Fuzzy Quantities*. Boca Raton: CRC--Press, 1994.
- [69] Imperial College London. MMI Facial Expression Database. [Online]. <http://www.mmifacedb.com/>

- [70] Eric W. Weisstein. Wolfram MathWorld. [Online].  
<http://mathworld.wolfram.com/LeastSquaresFittingPerpendicularOffsets.html>
- [71] Eric W. Weisstein. Wolfram MathWorld. [Online].  
<http://mathworld.wolfram.com/LeastSquaresFitting.html>
- [72] John D. Cook. (2008, October) The Endeavour. [Online].  
<http://www.johndcook.com/blog/2008/10/20/comparing-two-ways-to-fit-a-line-to-data/>
- [73] Mohamed Al-Dabbagh. (2009) Orthogonal Distance Least-Squares Curve-Fitting (Web Research). [Online]. <http://cf.net16.net/odlsf>
- [74] Θ. Θεοχάρης and Α. Μπεμ, "Μετασχηματισμοί  $2\Delta/3\Delta$  και Συστήματα Συντεταγμένων," in *Γραφικά Αρχές και Αλγόριθμοι*. Αθήνα: Εκδόσεις Συμμετρία, 1999, pp. 76-83.
- [75] Mark Williams. (2007, May) Technology Review (Published by MIT). [Online].  
<http://www.technologyreview.com/news/407976/better-face-recognition-software/>