



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

Αναπαράσταση Συναισθήματος και Σύνθεση Εκφράσεων Προσώπου και Χειρονομιών σε Εικονικούς Χαρακτήρες

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΑΜΑΡΥΛΛΙΣ - ΟΛΓΑ Θ. ΡΑΟΥΖΑΙΟΥ
Διπλωματούχος Ηλεκτρολόγος Μηχανικός
και Μηχανικός Υπολογιστών ΕΜΠ (2000)

Αθήνα, Σεπτέμβριος 2006



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

Αναπαράσταση Συναισθήματος και Σύνθεση Εκφράσεων Προσώπου και Χειρονομιών σε Εικονικούς Χαρακτήρες

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΑΜΑΡΥΛΛΙΣ - ΟΛΓΑ Θ. ΡΑΟΥΖΑΙΟΥ
Διπλωματούχος Ηλεκτρολόγος Μηχανικός
και Μηχανικός Υπολογιστών ΕΜΠ (2000)

Συμβουλευτική Επιτροπή : Στέφανος Κόλλιας

Ανδρέας – Γεώργιος Σταφυλοπάτης

Παναγιώτης Τσανάκας

Εγκρίθηκε από την επταμελή εξεταστική επιτροπή τη 19^η Σεπτεμβρίου 2006.

.....
Σ. Κόλλιας
Καθηγητής Ε.Μ.Π.

.....
Α. – Γ. Σταφυλοπάτης
Καθηγητής Ε.Μ.Π.

.....
Π. Τσανάκας
Καθηγητής Ε.Μ.Π.

.....
Ν. Μήτρου
Καθηγητής Ε.Μ.Π.

.....
Γ. Καραγιάννης
Καθηγητής Ε.Μ.Π.

.....
Β. Λούμος
Καθηγητής Ε.Μ.Π.

.....
Θ. Θεοχάρης
Αν. Καθηγητής ΕΚΠΑ

Αθήνα, Σεπτέμβριος 2006

.....
Αμαρυλλίς – Όλγα Θ. Ραουζαίου

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

Copyright © Αμαρυλλίς – Όλγα Θ. Ραουζαίου, 2006.

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.

Περιεχόμενα

1. Εισαγωγή.....	17
1.1. Αναδρομή.....	18
2. Διάφορες Μορφές Αναπαράστασης Συναισθήματος – οι Έξι Βασικές Εκφράσεις Προσώπου	21
2.1. Αναπαράσταση συναισθήματος.....	21
2.1.1. Τροχός Συναισθημάτων της <i>Whissel</i>	21
2.1.2. Θεωρία Συναισθημάτων του <i>Plutchik</i>	23
2.2. Συναισθημα κι Εκφράσεις Προσώπου	24
3. Τεχνικές Εμφύχωσης: FACS και MPEG-4 FBA.....	27
3.1. FACS: Σύντομη Παρουσίαση	27
3.2. MPEG-4 FBA	29
3.2.1. MPEG-4 και Πρόσωπο - FDPs και FAPs.....	29
3.2.2. MPEG-4 και ανθρώπινο σώμα	38
4. Εκφράσεις Προσώπου – Ορισμός και Ανάλυση	41
4.1. Αντιστοίχιση AUs – FAPs	41
4.2. Δημιουργία profiles-Προσδιορισμός ομάδων FAPs για κάθε έκφραση.....	43
4.2.1. Περιοχή τιμών FAPs.....	44
4.3. Ορισμός αποστάσεων.....	45
4.3.1. Αναπαράσταση των FAPs μέσω κίνησης FPs (<i>feature points</i>).....	45
4.3.2. Αποστάσεις και αλληλεπίδραση μεταξύ τους.....	46
4.4. Επαλήθευση των εκφράσεων.....	47
4.5. Σύστημα Ανάλυσης Εκφράσεων Προσώπου	50
4.5.1. Επέκταση «προφίλ» για τη δημιουργία κανόνων για είσοδο σε νευρωνικό σύστημα.....	50
4.5.2. Interface Συστήματος Ανάλυσης Εκφράσεων Προσώπου.....	52
5. Σύνθεση Εκφράσεων Προσώπου.....	53
5.1. Δημιουργία «προφίλ» βασικών εκφράσεων.....	53
5.1.1. Επαλήθευση μέσω Πειραματικών Αποτελεσμάτων.....	55
5.1.2. Προφίλ εκφράσεων για τα τεταρτημόρια του Τροχού Συναισθημάτων της <i>Whissel</i>	58
5.2. Δημιουργία «προφίλ» συναισθημάτων τα οποία ανήκουν στην ίδια κατηγορία	59
5.2.1. Επαλήθευση μέσω Πειραματικών Αποτελεσμάτων.....	60
5.3. Ενδιάμεσες εκφράσεις	62
5.3.1. Δημιουργία «προφίλ» για Ενδιάμεσες Εκφράσεις - Συναισθήματα.....	63
5.3.2. Επαλήθευση μέσω Πειραματικών Αποτελεσμάτων.....	65
5.4. Παράμετροι Σύνθεσης Εκφράσεων Προσώπου.....	68
5.4.1. Παράμετροι εκφραστικότητας για το πρόσωπο.....	69
5.5. Αλληλεπίδραση μεταξύ FAPs	71
5.6. Προφίλ των βασικών συναισθημάτων σε σχέση με τη Θεωρία Αποτίμησης (<i>Appraisal</i>)	73
5.6.1. Εκφράσεις Προσώπου των ECAs και Μοντέλο Μονάδων Επεξεργασίας (<i>component process model</i>).....	73
5.6.2. Σύνθεση Εκφράσεων Σύμφωνα με τη Θεωρία Αξιολόγησης.....	75

6. Σώμα - Χειρονομίες.....	79
6.1. Ανάλυση Χειρονομιών με Χρήση Παραμέτρων Εκφραστικότητας.....	79
6.1.1. Σύστημα Ανάλυσης Χειρονομιών	79
6.1.2. Προσδιορισμός Παραμέτρων Εκφραστικότητας.....	82
6.2. Σύνθεση Χειρονομιών	83
6.2.1. Σύνθεση χειρονομιών και εκφράσεων.....	83
6.2.2. Εκφραστικότητα Χειρονομιών.....	83
6.2.3. Μοντελοποίηση των χειρονομιών με βάση πραγματικές ακολουθίες.....	86
7. Αναπαράσταση Σώματος και Προσώπου σε Εκφραστικούς Πράκτορες - Εφαρμογές	89
7.1. Greta	89
7.2. Συνδυασμός Προσώπου και Σώματος, Ανάλυσης και Σύνθεσης	90
7.3. Η Μέθοδος Πρόκλησης Συναισθημάτων SAL	92
8. Σύνθεση Εκφράσεων και Οντολογίες.....	95
8.1. MPEG-4: FBA Node	96
8.2. Πρόσωπο	96
8.2.1. Δημιουργία Οντολογίας	97
8.2.2. Αναπαράσταση Συναισθήματος	97
8.2.3. Προφίλ Εκφράσεων Προσώπου (Βασικών και Ενδιαμέσων).....	98
8.2.4. Οντολογία Εκφράσεων Προσώπου	98
8.2.5. Παραγωγή Κανόνων στην Οντολογία των Εκφράσεων.....	100
8.2.6. Σενάρια Χρήσης Οντολογίας	101
8.3. Σώμα	103
8.3.1. Δημιουργία Οντολογίας	103
8.3.2. Παράμετροι Σώματος	103
8.3.3. Οντολογία Σώματος.....	104
8.3.4. Σενάρια Χρήσης	105
9. Συμπεράσματα-Μελλοντική Έρευνα και Περιορισμοί	107
10. Παράρτημα – Αποδόσεις όρων.....	111
11. Βιβλιογραφία	113
12. Κατάλογος Δημοσιεύσεων της συγγραφέως.....	121
13. Βιογραφικό Σημείωμα.....	127

Εικόνες

2.1	Το επίπεδο activation – evaluation	22
2.2	Τρισεδιάστατη απεικόνιση συναισθημάτων σύμφωνα με τον Plutchik.	24
2.3	Βασικές ανθρώπινες εκφράσεις: (α) λύπη, (β) θυμός, (γ) χαρά, (δ) φόβος, (ε) απέχθεια, (f) έκπληξη	25
3.1	Το FACS λαμβάνει υπόψιν τους μυς που αλλάζουν την «εμφάνιση» του προσώπου.	28
3.2	Ο ορισμός των Feature Points στο MPEG-4	30
3.3	Μονάδες κίνησης των χαρακτηριστικών του προσώπου [53]	31
3.4	Γράφημα όλων των σταδίων κωδικοποίησης της κίνησης του προσώπου	37
3.5	Γράφημα όλων των σταδίων κωδικοποίησης της κίνησης του σώματος.	39
4.1	Στιγμιότυπο Συστήματος Ανάλυσης Εκφράσεων	52
5.1	Παραδείγματα απεικόνισης «προφίλ» (α)-(γ) Θυμός, (δ)-(ε) Έκπληξη, (στ) Χαρά	56
5.2	Απεικόνιση βασικών εκφράσεων με χρήση του τρισεδιάστατου μοντέλου του Poser: (α) λύπη, (β) θυμός, (γ) χαρά, (δ) φόβος, (ε) απέχθεια και (στ) έκπληξη	57
5.3	Απεικόνιση κάποιων βασικών εκφράσεων με χρήση του MPEG-4 αποκωδικοποιητή GretaPlayer: (α) λύπη, (β) θυμός, (γ) φόβος	57
5.4	Μοντέλο προσώπου Poser: Απεικόνιση «προφίλ» των συναισθημάτων (α) φοβισμένος, (β) τρομοκρατημένος και (γ) ανήσυχος.	61
5.5	Μοντέλο προσώπου MPEG-4: Απεικόνιση «προφίλ» των συναισθημάτων (α) φοβισμένος, (β) τρομοκρατημένος και (γ) ανήσυχος.	61
5.6	Μοντέλο προσώπου MPEG-4 Greta: Απεικόνιση «προφίλ» των συναισθημάτων (α) φοβισμένος, (β) τρομοκρατημένος και (γ) ανήσυχος.	62
5.7	Μοντέλο προσώπου Poser: Απεικόνιση «προφίλ» των συναισθημάτων (α) φοβισμένος, (β) απογοητευμένος και (γ) λυπημένος.	66
5.8	Μοντέλο προσώπου MPEG-4: Απεικόνιση «προφίλ» των συναισθημάτων (α) φοβισμένος, (β) απογοητευμένος και (γ) λυπημένος.	67
5.9	Προφίλ για (α) φόβος, (β-γ) απογοήτευση, (δ) λύπη	68
5.10	Προφίλ για (α) θυμός, (β) καχυποψία, (γ) απέχθεια	68
5.11	Προφίλ για (α) φόβος, (β) ενοχή, (γ) λύπη	68
5.12	Αλληλουχία ADSR για την έκπληξη	69
5.13	Αλληλουχία ADSR για το FAP37 στην έκφραση της έκπληξης	69
5.14	Greta: Πρόβλεψη ενδιάμεσων σταδίων των εκφράσεων προσώπου σύμφωνα με τη θεωρία αξιολόγησης του Scherer για την περίπτωση του φόβου- (α) ουδέτερο (neutral), (β) ξαφνικό συμβάν (novelty-sudden), (γ) δυσαρέσκεια (unpleasant), (δ) αντιφατικό (discrepant), (ε) παρεμπόδιση στόχου (goal obstructive), (στ) χαμηλός έλεγχος -τελική έκφραση-φόβος (low control-final expression –fear). Κάθε έκφραση προκύπτει από την «πρόσθεση» των AUs των προηγούμενων εκφράσεων και της τελικής.	76
5.15	Greta: Πρόβλεψη ενδιάμεσων σταδίων των εκφράσεων προσώπου σύμφωνα με τη θεωρία αξιολόγησης του Scherer για την περίπτωση του έντονου θυμού- (α)	76

	ουδέτερο (neutral), (β) έντονη απόκριση σε ξαφνικό συμβάν (novelty- high), (γ) παρεμπόδιση στόχου (goal obstructive), (δ) υψηλός έλεγχος/ μεγάλη ένταση (control high/power high), (ε) τελική έκφραση-έντονος θυμός (final expression – hot anger).	
5.16	Xface: Πρόβλεψη ενδιάμεσων σταδίων των εκφράσεων προσώπου σύμφωνα με τη θεωρία αξιολόγησης του Scherer για την περίπτωση του φόβου- (α) ουδέτερο (neutral), (β) ξαφνικό συμβάν (novelty-sudden), (γ) δυσάρεσκεια (unpleasant), (δ) παρεμπόδιση στόχου (goal obstructive), (ε) χαμηλός έλεγχος-τελική έκφραση-φόβος (low control-final expression –fear). Κάθε έκφραση προκύπτει από την «πρόσθεση» των AUs των προηγούμενων εκφράσεων και της τελικής.	77
5.17	Xface: Πρόβλεψη ενδιάμεσων σταδίων των εκφράσεων προσώπου σύμφωνα με τη θεωρία αξιολόγησης του Scherer για την περίπτωση του έντονου θυμού- (α) ουδέτερο (neutral), (β) έντονη απόκριση σε ξαφνικό συμβάν (novelty- high), (γ) παρεμπόδιση στόχου (goal obstructive), (δ)υψηλός έλεγχος/ μεγάλη ένταση (control high/power high), (ε)τελική έκφραση-έντονος θυμός (final expression – hot anger).	78
5.18	Xface: Πρόβλεψη ενδιάμεσων σταδίων των εκφράσεων προσώπου σύμφωνα με τη θεωρία αξιολόγησης του Scherer για την περίπτωση της λύπης- (α) ουδέτερο (neutral), (β) μικρή εξοικείωση (familiarity low), (γ) παρεμπόδιση στόχου (goal obstructive), (δ)χαμηλός έλεγχος (control low), (ε)τελική έκφραση-λύπη (final expression – sadness).	78
6.1	Συνθετική δημιουργία χειρονομιών	80
6.2	Στιγμιότυπο ακολουθίας προς εξέταση	81
6.3	Έγχρωμη μάσκα	81
6.4	Βασικά Σχήματα Χεριών	83
6.5	Frames του υποκειμένου 4 από τη χειρονομία «Θεέ μου! »	86
6.6	Frames του υποκειμένου 21 από τη χειρονομία «Θεέ μου! »	87
6.7	Πειραματικά αποτελέσματα για τις έξι παραμέτρους της Ενότητας 6.1.2	87
6.8	Πειραματικά αποτελέσματα για τις είκοσι χειρονομίες για τις παραμέτρους : α)ολική ενεργοποίηση, β)χωρική έκταση, γ)ρευστότητα, δ)δύναμη/ενέργεια	87
6.9	Μέση τιμή των έξι παραμέτρων για τρία άτομα	88
6.10	(α) μέση τιμή Ολικής Ενεργοποίησης, (β) μέση τιμή Δύναμης	88
6.11	Frames της συντεθειμένης έκφρασης.	88
7.1	Ουδέτερη έκφραση του προσώπου της Greta	90
7.2	Μοντέλο σώματος H-Anim και το αντίστοιχο μοντέλο προσομοίωσης υφής δέρματος.	90
7.3	Διάγραμμα προτεινόμενου συστήματος.	92
7.4	Διαπροσωπεία SAL	94
8.1	Απεικόνιση του αντικειμένου Σχεδιασμού Κίνησης Προσώπου και Σώματος του MPEG-4 στην οντολογία.	96
8.2	Περιγραφή των ψυχολογικών μοντέλων του συναισθήματος στην οντολογία.	98
8.3	Περιγραφή των προφίλ των βασικών και των ενδιάμεσων εκφράσεων στην οντολογία.	100
8.4	Παράδειγμα δημιουργίας προφίλ Έκφρασης Προσώπου και απεικόνισης τελικού αποτελέσματος στον MPEG-4 αποκωδικοποιητή GretaPlayer.	101
8.5	Διάγραμμα της οντολογίας για το σώμα.	105
9.1	Ολοκληρωμένο σύστημα αντιγραφής συναισθηματικής κατάστασης χρήστη από ECA.	108

Πίνακες

2.1	Χαρακτηριστικά εκφράσεων	26
3.1	Ορισμός Action Units	29
3.2	Μονάδες μέτρησης των FAPs	32
3.3	Ομάδες FAPs	33
3.4	Ορισμοί FAPs	36
4.1	Αντιστοίχιση FAPs - AUs	43
4.2	Λεξιλόγιο FAPs για την περιγραφή των βασικών εκφράσεων	44
4.3	Ποσοτικοποίηση των FAPs	46
4.4	FAPs και αποστάσεις, πρωτεύουσες και μη	47
4.5	Στατιστικά δεδομένα για τιμές λεξιλογίου FAPs για τις βασικές εκφράσεις	49
4.6	Κανόνες βασιζόμενοι στα «προφίλ» συναισθημάτων για είσοδο σε νευρωνικό σύστημα	52
5.1	«Προφίλ» βασικών εκφράσεων	55
5.2	«Προφίλ» τεταρτημορίων	59
5.3	«Προφίλ» των συναισθημάτων τρόμος και ανησυχία	61
5.4	Επιλεγμένα συναισθήματα από τη μελέτη της Whissel [35]	62
5.5	«Προφίλ» του συναισθήματος ενοχή	66
5.6	Προφίλ των βασικών συναισθημάτων φοβισμένος και λυπημένος και του παραγόμενου ενδιάμεσου συναισθήματος απογοητευμένος	67
5.7	Αλληλεπίδραση μεταξύ FAPs	72
6.1	Συναισθήματα υποδηλούμενα μέσω χειρονομιών	80
6.2	Εκτελεσθείσες Χειρονομίες	81
7.1	Επίδραση παραμέτρων εκφραστικότητας σε κεφάλι, εκφράσεις προσώπου και χειρονομίες	91
8.1	Ενδιάμεσο Συναισθήμα Απογοητευμένος και τα βασικά συναισθήματα από τα οποία προέκυψε, με χρήση των μετρήσεων ενεργοποίησης και αποτίμησης	98

ΠΕΡΙΛΗΨΗ

Η παρούσα διδακτορική διατριβή ασχολείται με την αναπαράσταση συναισθήματος στην επικοινωνία ανθρώπου-μηχανής. Η διατριβή πραγματεύεται το θέμα από τη μεριά της σύνθεσης και της ανάλυσης μέσω σύνθεσης, σε συμφωνία με το διεθνές πρότυπο κωδικοποίησης MPEG-4. Συνδυάζει τα αποτελέσματα ψυχολογικών ερευνών για την αντίληψη της έκφρασης και του συναισθήματος με πειραματικά αποτελέσματα ανάλυσης εικόνων από διεθνείς βάσεις δεδομένων. Μέχρι τώρα είχαν γίνει κάποιες περιορισμένες προσπάθειες στο χώρο Επικοινωνίας Ανθρώπου-Μηχανής, οι περισσότερες από τις οποίες έχουν επικεντρωθεί στα βασικά συναισθήματα. Η παρούσα διατριβή, πέραν της συστηματοποίησης των βασικών συναισθημάτων, και της επεξεργασίας και σύνθεσης ενδιάμεσων συναισθημάτων, συμβάλλει αποφασιστικά στην αληθοφανή αναπαράσταση ενός εικονικού κόσμου με ανάλυση και χρήση μεγαλύτερης ποικιλίας εκφράσεων συναισθήματος. Εκτός της έμφασης που δίνεται στη μοντελοποίηση καθημερινών εκφράσεων διαφορετικών χρηστών, η διατριβή επεκτείνει τη θεωρία αναπαράστασης συναισθήματος και ασχολείται και με τη μοντελοποίηση χειρονομιών και την επεξεργασία των σχετικών παραμέτρων εκφραστικότητας, ενώ δίνεται η δυνατότητα σύνδεσης των ανωτέρω σε πολυδιάστατα συστήματα Επικοινωνίας Ανθρώπου-Μηχανής κι εφαρμογής τους σε εκφραστικούς εικονικούς χαρακτήρες.

ABSTRACT

The main research area of this Ph.D is emotion representation in Human-Computer Interaction (HCI) and especially emotion synthesis and analysis by synthesis in the framework of the MPEG-4 International Standard. It combines psychological results for the perception of emotion with experimental results derived from the analysis of relative databases. The emotions are modeled using Whissel's emotions wheel, which transforms the emotions into a two-dimensional space, thus facilitating their reproduction. Research in facial expression analysis and synthesis has mainly concentrated on archetypal emotions. In particular, sadness, anger, joy, fear, disgust and surprise are categories of emotions that attracted most of the attention in human computer interaction environments. The current thesis describes an approach to synthesize expressions, including intermediate ones, via the tools provided in the MPEG-4 standard based on real measurements and on universally accepted assumptions of their meaning and extends emotion representation theory by modeling gestures using expressivity parameters. The results of the synthesis process can be applied to avatars (ECAs-Embodied Conversational Agents) and to multimodal HCI systems, so as to convey the communicated messages more vividly than plain textual information or simply to make interaction more lifelike.

Πρόλογος

Η ένταξη μου στο Εργαστήριο Ψηφιακής Επεξεργασίας Εικόνας, Βίντεο και Πολυμέσων το 1999 με σκοπό την πραγματοποίηση διπλωματικής εργασίας επέδρασε πολύ περισσότερο στη ζωή μου από το αρχικά αναμενόμενο. Οι δέκα μήνες της διπλωματικής εργασίας μετατράπηκαν σε χρόνια δημιουργικής έρευνας, η οποία τελικά οδήγησε στην παρούσα διδακτορική διατριβή. Στα χρόνια αυτά προσπάθησα να συμμετέχω ενεργά στις εξελίξεις της περιοχής της Επικοινωνίας Ανθρώπου – Μηχανής, ειδικότερα στον τομέα της σύνθεσης εκφράσεων, και θέλω να ελπίζω ότι οι εργασίες που προέκυψαν από την έρευνα αυτή, καθώς και η παρούσα διατριβή θα συμβάλλουν στην πρόοδο του τομέα που πραγματεύονται.

Στη μέχρι τώρα πορεία μου αξιοσημείωτα είναι η προσοχή και το ενδιαφέρον που επέδειξε ο επιβλέπων της διατριβής, καθηγητής του ΕΜΠ κ. Στέφανος Κόλλιας, καθώς και το γεγονός ότι παραστάθηκε με εξαιρετική διάθεση, όποτε του το ζήτησα, σε οποιοδήποτε πρόβλημά μου. Τον ευχαριστώ ιδιαίτερα. Θερμές ευχαριστίες οφείλω και στα άλλα μέλη της τριμελούς επιτροπής, καθηγητές του ΕΜΠ κ. Ανδρέα Σταφυλοπάτη και κ. Παναγιώτη Τσανάνα.

Επίσης θα ήθελα να ευχαριστήσω το διδάκτορα του ΕΜΠ κ. Κώστα Καρπούζη. Με τη δική του προτροπή άρχισε η ενασχόληση με το συγκεκριμένο θέμα, το οποίο αποδείχτηκε εξαιρετικά ενδιαφέρον για εμένα. Οι γνώσεις, οι εμπειρίες, η διάθεση και το ενδιαφέρον του οδήγησαν τη συγκεκριμένη δουλειά σε αυτό το σημείο, ενώ οποιαδήποτε στιγμή ήταν έτοιμος να προσφέρει ό,τι ήταν δυνατό ώστε να ξεπεραστεί κάθε δυσκολία.

Ευχαριστίες οφείλω και στο διδάκτορα του ΕΜΠ κ. Νικόλα Τσαπατσούλη για την πολύτιμη συνεργασία του στην έρευνά μου, αλλά και στα μέλη του Εργαστηρίου Ψηφιακής Επεξεργασίας Εικόνας, Βίντεο και Πολυμέσων και ειδικότερα σε αυτούς που συνεργαστήκαμε στενότερα συνδυάζοντας την έρευνά μας με σκοπό καλύτερα αποτελέσματα. Για τη συνεργασία στις επιστημονικές εργασίες, που αποτελούν τον κορμό της παρούσας διατριβής, θα ήθελα να ευχαριστήσω τους Σπύρο Ιωάννου, Νάσο Δροσόπουλο, Μανόλη Γουάλλες και Βασίλη Τζουβάρα, Διδάκτορες ΕΜΠ, Παρασκευή Τζούβελη, Γιώργο Καρυδάκη, Λώρη Μαλατέστα και Μηνά Περτσελάκη, Υποψήφιους Διδάκτορες ΕΜΠ. Ιδιαίτερη μνεία θα ήθελα να κάνω στο φίλο και συνεργάτη Θέμη Μπαλωμένο, που χάθηκε πρόωρα.

Θα ήθελα επίσης να ευχαριστήσω την οικογένειά μου, καθώς και τους φίλους μου για τη συμμετοχή και τη συμβολή τους στις σημαντικές, αλλά και στις λιγότερο σημαντικές στιγμές της ζωής μου.

Αμαρυλλίς Ραουζαίου
Αθήνα, Σεπτέμβριος 2006

Κεφάλαιο 1

Εισαγωγή

Σήμερα, περισσότερο από ποτέ, η επικοινωνία ανθρώπου-μηχανής αποτελεί προς εξέταση θέμα για ένα ευρύ φάσμα επιστημών, φάσμα που εκτείνεται από τον τομέα της πληροφορικής μέχρι τον τομέα της ψυχολογίας. Στην επικοινωνία αυτή σημαντικό ρόλο παίζει η αναγνώριση από τον υπολογιστή της διάθεσης του χρήστη, αλλά και η δυνατότητα αναπαράστασης μίας έκφρασης ευρέως αναγνωρίσιμης στην οθόνη του. Η αναγνώριση των εκφράσεων είναι μία καθημερινή πρακτική για τον ανθρώπινο νου. Ο άνθρωπος, ον κοινωνικό, προσπαθεί να αντιληφθεί τις προθέσεις του συνομιλητή του, τις σκέψεις του, τη διάθεση απέναντί του. Κατά τη διαδικασία αυτή συνδυάζει στοιχεία προερχόμενα από την έκφραση του προσώπου, την εκφορά του λόγου, τις χρησιμοποιούμενες λέξεις, τη στάση του σώματος, είναι μία διαδικασία αυθόρμητη, συγχρόνως όμως εξαιρετικά πολύπλοκη. Ο υπολογιστής πρέπει να αποτελεί ένα εύχρηστο εργαλείο, γεγονός που καθιστά άμεση προτεραιότητα την αντίληψη της διάθεσης του χρήστη του. Αν ο χρήστης είναι ικανοποιημένος μπορεί ο υπολογιστής να συνεχίσει να λειτουργεί με τον ίδιο τρόπο, ενώ αν αυτός εμφανίζεται δυσαρεστημένος πρέπει άμεσα να προβεί σε διορθωτικές ενέργειες. Αν η διαδικασία αυτή απλά διευκολύνει τον απλό χρήστη, κρίνεται εντούτοις απαραίτητη στην περίπτωση ατόμων με ειδικές ανάγκες που δεν έχουν την ευχέρεια να εκφράσουν ξεκάθαρα τη διάθεσή τους.

Εκτός όμως από την αναγνώριση της έκφρασης, εξαιρετικά χρήσιμη στα πλαίσια αποδοτικότερης επικοινωνίας ανθρώπου-μηχανής είναι και η σύνθεση εκφράσεων. Ο χρήστης θα εργάζεται πιο ξεκούραστα αν βλέπει ένα χαμογελαστό πρόσωπο στην οθόνη του να του λέει ποιο θα είναι το επόμενο βήμα στην εγκατάσταση ενός προγράμματος ή να τον ξεναγεί στα μεγαλύτερα μυσεία του κόσμου. Το θέμα αυτό αποτελεί και το θέμα της παρούσας διδακτορικής διατριβής. Ένα θέμα που χρησιμοποιεί και συνδυάζει γνώσεις άλλων ερευνητικών πεδίων παρέχοντας ένα ολοκληρωμένο αποτέλεσμα αποτέλεσε ικανή πρόκληση ήδη από την περίοδο της διπλωματικής μου εργασίας.

Στην παρούσα διδακτορική διατριβή ακολουθείται η εξής δομή: στο παρόν κεφάλαιο ακολουθεί μία **Ενότητα** που περιλαμβάνει μία σύντομη αναδρομή της ερευνητικής περιοχής του διδακτορικού, στο **Κεφάλαιο 2** αναλύονται οι ψυχολογικές μελέτες που χρησιμοποιήθηκαν στην έρευνά μας και δίνονται στοιχεία για το συναίσθημα γενικά και ειδικότερα για τις έξι βασικές εκφράσεις, που αποτέλεσαν και τη βάση της διατριβής. Στο **Κεφάλαιο 3** παρουσιάζονται οι Τεχνικές Εμφύχωσης για το πρόσωπο αλλά και για το σώμα, ενώ το **Κεφάλαιο 4** πραγματεύεται τον ορισμό των εκφράσεων του προσώπου και την ανάλυση που οδήγησε εκεί. Αντίθετα το **Κεφάλαιο 5** ασχολείται με τη σύνθεση των εκφράσεων του προσώπου, βασικών κι ενδιαμέσων, με διάφορες μεθόδους. Το **Κεφάλαιο 6** περνάει πλέον στο θέμα του σώματος και ειδικότερα των χειρονομιών, ενώ στο **Κεφάλαιο 7** παρουσιάζεται εν συντομία ο Πράκτορας με σώμα και δυνατότητα ομιλίας (ECA - Embodied Conversational Agent) Greta, ο συνδυασμός προσώπου και σώματος και η μέθοδος πρόκλησης συναισθημάτων SAL. Το **Κεφάλαιο 8** παρουσιάζει τις οντολογίες που δημιουργήθηκαν τόσο για το πρόσωπο, όσο και για το σώμα, με χρήση των αποτελεσμάτων της παρούσας διδακτορικής διατριβής. Τέλος το **Κεφάλαιο 9** κλείνει τη διατριβή αναφερόμενο στα συμπεράσματα που εξήχθησαν, στα μελλοντικά βήματα που μπορούν να γίνουν στο συγκεκριμένο τομέα και στους περιορισμούς που υπάρχουν, τεχνικής αλλά και ηθικής φύσεως. Ακολουθούν η Βιβλιογραφία της σχετικής έρευνας, ένας κατάλογος με τις δημοσιεύσεις της συγγραφέως και ένα σύντομο βιογραφικό σημείωμα.

1.1. Αναδρομή

Ικανός αριθμός μελετών έχει πραγματοποιηθεί σχετικά με την αντίληψη της έκφρασης ή ακόμα και του συναισθήματος ενός ανθρώπου. Τα περισσότερα πειράματα πραγματοποιήθηκαν με παρεμφερείς συνθήκες [1]: ένα σύνολο από φωτογραφίες που απεικονίζουν εκφράσεις προσώπου, δίνονται για χαρακτηρισμό σε μία ομάδα ανθρώπων. Τα αποτελέσματα είναι ικανοποιητικά, κυρίως σε ό,τι αφορά σε εκφράσεις που δεν περιέχουν τις ίδιες κινήσεις χαρακτηριστικών του προσώπου [2]. Για παράδειγμα, ο *θυμός* και η *χαρά* διαχωρίζονται πολύ εύκολα, δε συμβαίνει όμως το ίδιο για το *θυμό* και το *φόβο*, ενώ διαφορές παρατηρήθηκαν ανάμεσα στην αληθινή έκφραση και σε μία αντίστοιχη υποκριτική [3]. Η δεύτερη παρουσιάζει μεγαλύτερη ασυμμετρία και εμφανίζεται είτε για πολύ λίγο είτε για υπερβολικό χρόνο.

Ένα σημαντικό ερώτημα σχετικά με τη μοντελοποίηση εκφράσεων προσώπου έγκειται στη φύση της απαιτούμενης πληροφορίας. Για παράδειγμα, πόσα χαρακτηριστικά του προσώπου είναι απαραίτητα ώστε να αποδοθεί ρεαλιστικά μία έκφραση? Το ερώτημα αυτό είναι πολύ σημαντικό και για τον καθορισμό των κατάλληλων βάσεων δεδομένων.

Σε μία έρευνά του ο Dittrich [4] διερεύνησε κατά πόσο η θέση και ο αριθμός των ορατών χαρακτηριστικών του προσώπου επιδρά στην αντίληψη της έκφρασης κατά την

κίνηση της κεφαλής. Σύμφωνα με βιολογικές μελέτες, ακόμα κι ένας μικρός αριθμός ορατών χαρακτηριστικών μπορεί να αποδώσει την έκφραση, χωρίς να είναι απαραίτητο αυτά να είναι τα μάτια, η μύτη ή το στόμα, δηλαδή ακόμα και λιγότερο σημαντικές περιοχές του προσώπου έχουν την ικανότητα να «απεικονίζουν» μία έκφραση ρεαλιστικά.

Η έρευνα σχετικά με την αντίληψη του προσώπου έχει πάρει διάφορες κατευθύνσεις. Κάποιες εργασίες ασχολούνται απλά με τον εντοπισμό της ταυτότητας του χρήστη [5], των χαρακτηριστικών του προσώπου ή της κίνησης των χειριών ([6], [7], [8]) ή ακόμα της έκφρασης του προσώπου ([9]). Άλλες εργασίες ασχολούνται με την αναγνώριση γνωστών προσώπων ([5]) ή με την αντίληψη των συναισθημάτων ([2]).

Όσον αφορά στην εκφραστικότητα του σώματος, οι Wallbott και Scherer ([10]) χώρισαν τις κινήσεις σε πέντε κατηγορίες: αργή/γρήγορη, αδύναμη/ ενεργητική, κίνηση μικρού εύρους/ κίνηση μεγάλου εύρους και ευχάριστη/δυσάρεστη. Σε μεταγενέστερη εργασία, ο Wallbott ([11]) διατύπωσε την άποψη ότι εκτός από τις προαναφερθείσες, στατικές παραμέτρους, υπάρχουν και τρεις δυναμικές που μπορούν να γίνουν αντιληπτές από έναν παρατηρητή: «ποσότητα» της κίνησης, έκταση της στο χώρο και ένταση/δυναμικότητα της κίνησης.

Όσον αφορά στους συναισθηματικούς πράκτορες (emotional agents), έχουν αναπτυχθεί αρκετές ομιλούσες κεφαλές (talking heads) που έχουν τη δυνατότητα να εξωτερικεύουν συναίσθημα. Ειδικότερα, η Kshirsagar και οι συνεργάτες της [12] ανέπτυξαν έναν πράκτορα που έχει την ικανότητα να αντιδρά στο συναίσθημα του χρήστη, όπως αυτό εκφράζεται μέσω του προσώπου του. Αυτή η αντίδραση βασίζεται σε ένα υπολογιστικό μοντέλο συναισθηματικής συμπεριφοράς που ενσωματώνει ένα μοντέλο προσωπικότητας. Για παράδειγμα, το Carmen's Bright IDEAS ([13]) είναι ένα διαδραστικό έργο, κατά το οποίο οι χαρακτήρες εκφράζονται με χειρονομίες ανάλογα με τη συναισθηματική τους διάθεση και την προσωπικότητά τους, ενώ ένα μοντέλο αντιγραφής συμπεριφοράς αναπτύχθηκε από τους Marsella και Gratch ([14]). Στο μοντέλο αυτό έχει ενσωματωθεί πληροφορία σχετικά με την προσωπικότητα του πράκτορα και τον κοινωνικό του ρόλο.

Ειδικότερα για τη σύνθεση εκφράσεων σε έναν τέτοιο πράκτορα, έχουν πρόσφατα δημοσιευτεί διάφορες σχετικές μελέτες ([15], [16]).

Η Pelachaud και οι συνεργάτες της ([17]) δημιούργησαν έναν πράκτορα με σώμα (Embodied Conversational Agent-ECA), την Greta, η οποία ενσωματώνει στοιχεία επικοινωνίας και συνομιλίας, όπως αυτά προτάθηκαν από την Isabella Poggi ([18]) και χρησιμοποιεί τη γλώσσα απεικόνισης 'Affective Presentation Markup Language' (APML) για έλεγχο της συμπεριφοράς (Ενότητα 7.1).

Ειδικότερα για τη σύνθεση συγκεκριμένων εκφράσεων σε συνθετικό πρόσωπο έχουν δημοσιευτεί διάφορες σχετικές εργασίες ([19], [20], [21], [22]), ενώ οι Ekman και Friesen ([9]) και οι Pandzic και Forchheimer ([23]) ασχολήθηκαν με την εξομοίωση των κινήσεων των μυών του προσώπου σε συνθετικά πρόσωπα.

Τα ευρωπαϊκά προγράμματα HUMAINE [87] και ERMIS [86] - στα οποία συμμετέχει ή συμμετείχε και η γράφουσα- έχουν ως θέμα την αναγνώριση και απεικόνιση συναισθήματος μέσω συνδυασμού διαφόρων εκφραστικών μέσων (πρόσωπο, σώμα, ομιλία), χρησιμοποιώντας τεχνογνωσία πολλών ερευνητικών πεδίων (όραση υπολογιστών, επεξεργασία εικόνας, ψυχολογία).

Το ευρωπαϊκό πρόγραμμα NECA ([24]) ανέπτυξε διαδραστικότητα μεταξύ των ECAs μέσω διαδικτύου. Οι διάλογοι δημιουργούνται μέσω μίας αρχιτεκτονικής διαδοχικών τμημάτων που αποτελούνται από σύνθεση ομιλίας, χειρονομιών και έκφρασης προσώπου.

Για δημιουργία συμπεριφοράς που εμπεριέχει και συναισθημα, ο Chi ([25]) προσπάθησε να δημιουργήσει πιο εκφραστικές χειρονομίες με μετατροπή των ιδιοτήτων των κινηματικών και χωρικών καμπυλών των υπαρχουσών ενεργειών. Οι Buyn και Badler [26] μετέτρεψαν τις ακολουθίες δεδομένων κίνησης των χαρακτηριστικών του προσώπου στη μορφή MPEG-4 FAP. Τέλος, ο Barrientos ([27]) πρότεινε μηχανισμό εξαγωγής εκφραστικών ή συναισθηματικών παραμέτρων από την ανάλυση του γραφικού χαρακτήρα και μετά χρησιμοποίησε αυτές τις παραμέτρους για να ελέγξει τις χειρονομίες του πράκτορα μέσα από προκαθορισμένα clips.

Κεφάλαιο 2

Διάφορες Μορφές Αναπαράστασης Συναισθήματος – οι Έξι Βασικές Εκφράσεις Προσώπου

2.1. Αναπαράσταση συναισθήματος

Ένα από τα πιο σημαντικά στοιχεία στην ανάλυση των εκφράσεων είναι η εύρεση χαρακτηρισμού για τις διάφορες συναισθηματικές καταστάσεις. Δεδομένου όμως ότι ο ίδιος ο άνθρωπος χαρακτηρίζει τις εκφράσεις κατά τρόπο αρκετά σύνθετο, ένας απλός χαρακτηρισμός δεν είναι αρκετός, επιβάλλεται η εύρεση μίας συνεχούς αναπαράστασης των συναισθημάτων που θα προσεγγίζει όσο το δυνατόν περισσότερο την ανθρώπινη αντίληψη του θέματος.

Οι ψυχολόγοι έχουν εξετάσει ένα ευρύ φάσμα συναισθημάτων, αλλά πολύ λίγες από αυτές τις μελέτες έχουν αποτελέσματα αξιοποιήσιμα στο χώρο της γραφικής και της υπολογιστικής όρασης.

2.1.1. Τροχός Συναισθημάτων της Whissel

Η Whissel [35] προσέγγισε το θέμα με το επίπεδο activation-evaluation («Τροχός Συναισθημάτων»), το οποίο αποτελεί μία αναπαράσταση, απλή, αλλά συγχρόνως ικανή να «περιγράψει» ένα ευρύ φάσμα συναισθημάτων και η οποία βασίζεται σε δύο σημεία:

- *Valence*: Το κοινό χαρακτηριστικό των διαφόρων συναισθηματικών καταστάσεων αποτελεί το γεγονός ότι είναι αποτέλεσμα της επίδρασης που έχουν στο άτομο οι συνάνθρωποι του ή οι διάφορες καταστάσεις που βιώνει. Η επίδραση αυτή αποτιμάται από τη μεριά του χρήστη είτε ως θετική είτε ως αρνητική.
- *Activation level (επίπεδο δραστηριότητας)*: Ερευνητές έχουν επισημάνει το γεγονός ότι το άτομο, βιώνοντας τις διάφορες συναισθηματικές καταστάσεις, παρουσιάζει τάση συγκεκριμένων αντιδράσεων. Βασιζόμενοι στην παρατήρηση αυτή έγινε προσπάθεια να χαρακτηριστεί η έκφραση με βάση την «ενεργοποίηση» του προσώπου.

Οι άξονες στον «Τροχό Συναισθημάτων» αντικατοπτρίζουν τα δύο αυτά ζητήματα. Ο κατακόρυφος άξονας είναι ο άξονας της «ενεργοποίησης» (activation), ενώ ο οριζόντιος άξονας είναι ο άξονας της «αποτίμησης» του συναισθήματος (evaluation).

Το βασικότερο πλεονέκτημα αυτής της αναπαράστασης είναι ότι παρέχει έναν τρόπο περιγραφής συναισθηματικών καταστάσεων πιο κατανοητό από την περιγραφή με λέξεις. Τα συναισθήματα σε αυτόν τον «τροχό» είναι σημεία στο επίπεδο, οπότε η μετάβαση σε λέξεις που εκφράζουν συναισθήματα γίνεται εύκολα με χρήση διαφόρων τεχνικών, όπως ανάλυση συντελεστών, απευθείας βαθμονόμηση κ.α. [45].

Τα βασικά συναισθήματα δεν είναι ομαλά κατανεμημένα στο επίπεδο, αντιθέτως σχηματίζουν μία κυκλική διάταξη, γεγονός που, σε συνδυασμό με άλλες σχετικές ενδείξεις, οδηγεί στο συμπέρασμα ότι υπάρχει μία εγγενή κυκλική δομή στα συναισθήματα. Σε αυτό το πλαίσιο ο ορισμός του κέντρου ως αρχή των αξόνων έχει πολλαπλές εφαρμογές. Η ένταση ενός συναισθήματος μπορεί να ποσοτικοποιηθεί ως η απόσταση ενός δεδομένου σημείου από την αρχή των αξόνων, οπότε και ένα έντονο συναισθήμα μπορεί να περιγραφεί ως μία κατάσταση κατά την οποία η ένταση ενός συναισθήματος έχει περάσει ένα συγκεκριμένο όριο. Μία παρατήρηση που μπορεί να γίνει είναι ότι τα έντονα συναισθήματα είναι σαφέστερα διαχωρισμένα σε σχέση με τα λιγότερο έντονα, ίδιας υφής, συναισθήματα. Ως επέκταση της αντίληψης του τροχού μπορούμε να θεωρήσουμε τα βασικά συναισθήματα ως σημεία στην περιφέρεια του κύκλου [36] (Εικόνα 2.1).

Το επίπεδο activation-evaluation αποτελεί μία πολύ χρήσιμη αντίληψη, η οποία έχει βρει πολλές εφαρμογές στην έρευνα, στον τομέα της πληροφορικής. Εντούτοις πρέπει να επισημανθεί το γεγονός ότι με την σύμπτυξη του μεγάλου εύρους των συναισθημάτων στις δύο διαστάσεις είναι αναπόφευκτη η απώλεια πληροφορίας ενώ συχνά προκύπτουν και διαφορετικά αποτελέσματα. Για παράδειγμα ο Plutchik (Ενότητα 2.1.2) τοποθετεί το φόβο και το θυμό σε αντιδιαμετρικά σημεία, ενώ στο επίπεδο activation-evaluation της Whissel βρίσκονται σε μικρή απόσταση μεταξύ τους. Οπότε πρέπει με συνέπεια να ακολουθείται αποκλειστικά μία συγκεκριμένη αναπαράσταση.



Εικόνα 2.1: Το επίπεδο activation – evaluation

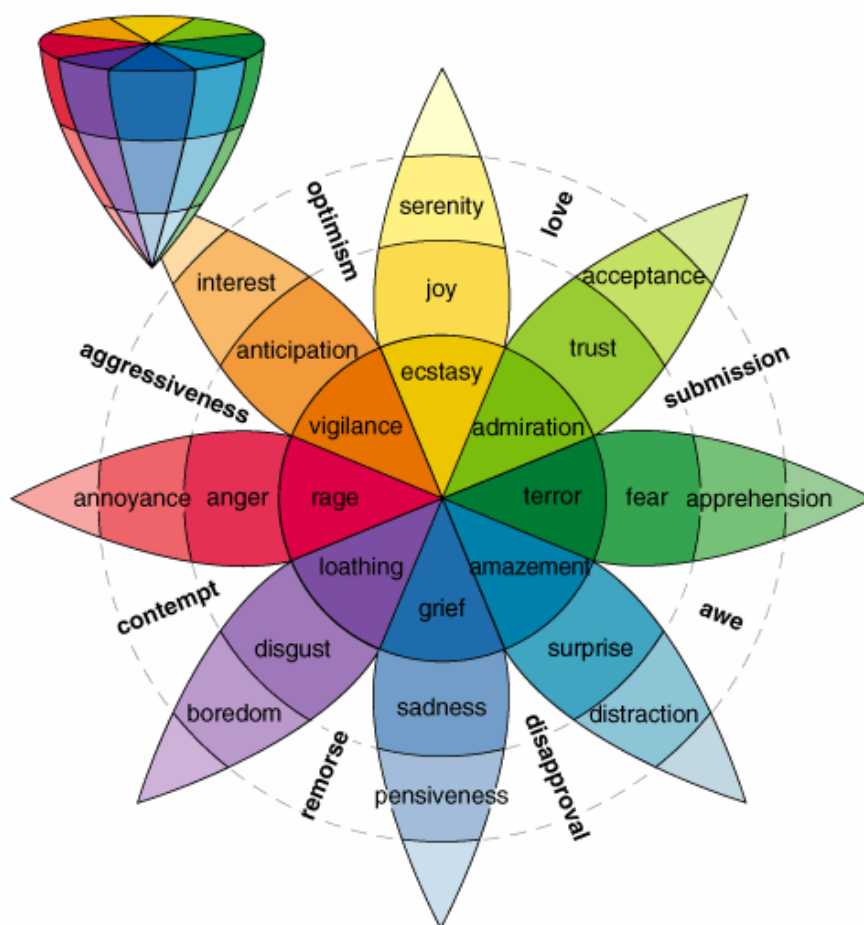
2.1.2. Θεωρία Συναισθημάτων του Plutchik

Η ψυχολογική θεωρία του Plutchik (1980) σχετικά με τα βασικά συναισθήματα, έχει δέκα αξιώματα [36]:

1. Η έννοια του συναισθήματος ισχύει σε όλα τα εξελικτικά επίπεδα σε περιπτώσεις ανθρώπων, αλλά και ζώων.
2. Τα συναισθήματα παρουσιάζουν μία εξελικτική πορεία και δημιουργούν ποικιλία εκφράσεων.
3. Τα συναισθήματα βοήθησαν τους διάφορους οργανισμούς να αντιμετωπίσουν ζητήματα επιβίωσης που δημιουργήθηκαν από το περιβάλλον.
4. Παρά τις διαφορετικές μορφές έκφρασης των συναισθημάτων στα διαφορετικά είδη, υπάρχουν ορισμένα κοινά στοιχεία ή πρότυπα προσδιορίσιμα.
5. Υπάρχει ένας μικρός αριθμός βασικών συναισθημάτων.
6. Όλα τα άλλα συναισθήματα είναι μίξεις καταστάσεων ή παράγωγες περιπτώσεις, δηλαδή προκύπτουν από τα βασικά συναισθήματα ή από συνδυασμό αυτών.
7. Τα βασικά συναισθήματα είναι υποθετικά κατασκευάσματα ή ιδανικές περιπτώσεις, των οποίων οι ιδιότητες και τα χαρακτηριστικά μπορούν να προκύψουν από διαφορετικά είδη.
8. Τα βασικά συναισθήματα μπορούν να θεωρηθούν ως αντίθετοι πόλοι συγκεκριμένων ζευγών.
9. Όλα τα συναισθήματα διαφοροποιούνται στο βαθμό ομοιότητας του ενός με το άλλο.
10. Κάθε συναίσθημα μπορεί να υπάρξει σε διαφορετικούς βαθμούς έντασης και με διαφορετικά επίπεδα διέγερσης.

Ο Plutchik θεωρεί ως βασικά συναισθήματα τα εξής: αποδοχή (acceptance), θυμός (anger), αναμονή (anticipation), απέχθεια (disgust), χαρά (joy), φόβος (fear), λύπη (sadness), έκπληξη (surprise).

Το μοντέλο τρισδιάστατης απεικόνισης των συναισθημάτων του Plutchik στην **Εικόνα 2.2** περιγράφει τις σχέσεις μεταξύ των συναισθημάτων, οι οποίες είναι ανάλογες με τα χρώματα σε έναν τροχό χρωμάτων. Η κάθετη διάσταση του κύβου αντιπροσωπεύει την ένταση και ο κύκλος αντιπροσωπεύει τους βαθμούς ομοιότητας μεταξύ των συναισθημάτων. Οι οκτώ τομείς έχουν ως σκοπό να δείξουν ότι υπάρχουν οκτώ, σύμφωνα με τον ορισμό, βασικά συναισθήματα, τα οποία αντιμετωπίζονται ως τέσσερα ζευγάρια αντιθέτων. Στο αναπτυγμένο μοντέλο της **Εικόνας 2.2**, τα συναισθήματα στα κενά διαστήματα είναι οι αρχικές δυάδες συναισθημάτων που προκύπτουν από τη μίξη δύο βασικών συναισθημάτων.



Εικόνα 2.2: Τρισδιάστατη απεικόνιση συναισθημάτων σύμφωνα με τον Plutchik

2.2. Συναισθήμα κι Εκφράσεις Προσώπου

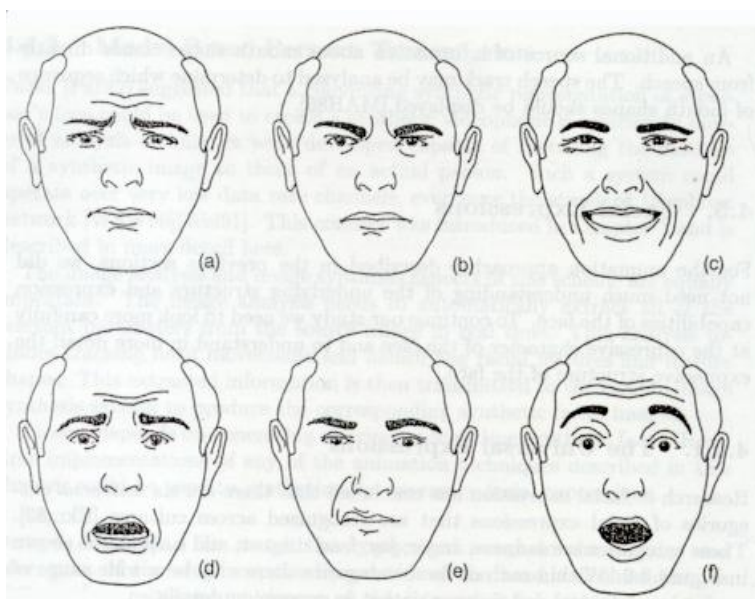
Η έκφραση κάποιου προσώπου μία δεδομένη στιγμή εκφράζεται από ένα σύνολο σημάτων. Στα σήματα αυτά περιλαμβάνονται το σχήμα και ο προσανατολισμός του προσώπου, αλλά και γενικότερα ολόκληρης της κεφαλής, και φυσικά η θέση και η κίνηση των διαφόρων χαρακτηριστικών του προσώπου (μάτια, φρύδια, στόμα κτλ). Στα σήματα αυτά συχνά περιλαμβάνονται με σημαίνοντα ρόλο το χρώμα του δέρματος του προσώπου, οι πιθανές ρυτίδες έκφρασης κτλ. Φυσικά τη σημαντικότερη συμβολή στη δημιουργία των εκφράσεων την έχουν οι μύες του προσώπου, οι οποίοι και έχουν την ικανότητα να δημιουργήσουν μεγάλη μεταβολή στην έκφραση σε πολύ σύντομο χρονικό διάστημα, συμβάλλοντας τα μέγιστα στη χωρίς λόγια επικοινωνία. Βέβαια δεν πρέπει να παραλείψουμε τις δυσκολίες που δημιουργούνται από την αλληλεπίδραση των διαφορετικών σημάτων τα οποία μπορούν είτε να αλληλοσυμπληρώνονται είτε να αλληλοαναιρούνται.

Η σχέση συναισθήματος και προσώπου είναι συναρπαστική και πολυσύνθετη. Μερικές από τις χιλιάδες εκφράσεις του προσώπου είναι συνδεδεμένες με συγκεκριμένα συναισθήματα. Έρευνες έχουν δείξει ότι άνθρωποι διαφορετικών πολιτισμών

κατηγοριοποιούν τις εκφράσεις με παρόμοιο τρόπο σε έξι βασικές κατηγορίες και στη συνέχεια «μιμούνται» τις εκφράσεις αυτές στην καθημερινή τους ζωή. Οι έξι αυτές βασικές κατηγορίες είναι: λύπη, θυμός, χαρά, φόβος, απέχθεια και έκπληξη, όπως φαίνονται στην **Εικόνα 2.3**.

Παρόλα αυτά μπορούμε να συναντήσουμε εκφράσεις οι οποίες, εκ πρώτης όψεως, δε φαίνεται να ανήκουν σε κάποια από τις έξι κατηγορίες, αυτό όμως συμβαίνει για τρεις κυρίως λόγους, ή πρόκειται για μία καθαρά υποκειμενική έκφραση ή είναι μία από τις έξι εκφράσεις αλλά με διαφορετικό όνομα ή είναι συνδυασμός δύο γνωστών εκφράσεων. Βέβαια, σε καθεμία από αυτές τις κατηγορίες, είναι δυνατόν να υπάρχει μία μεγάλη ποικιλία έντασης στις εκφράσεις και στις λεπτομέρειες τους.

Οι τρεις εκφραστικές περιοχές του προσώπου είναι τα *φρύδια*, τα *μάτια* και το *στόμα*.



Εικόνα 2.3: Βασικές ανθρώπινες εκφράσεις: (α) λύπη, (β) θυμός, (γ) χαρά, (δ) φόβος, (ε) απέχθεια, (ϝ) έκπληξη

Συνοπτικά, τα χαρακτηριστικά των εκφράσεων απεικονίζονται στον **Πίνακα 2.1**:

α/α	Έκφραση	Περιγραφή
1	Λύπη	Τα εσωτερικά σημεία των φρυδιών κάμπτονται προς τα πάνω. Τα μάτια είναι ελαφρώς κλειστά. Το στόμα είναι χαλαρωμένο.
2	Θυμός	Τα εσωτερικά φρύδια τραβιούνται μαζί προς τα κάτω. Τα μάτια είναι πολύ ανοικτά. Τα χείλια πιέζονται το ένα στο άλλο ή ανοίγουν και εκθέτουν τα δόντια.

3	Χαρά	Τα φρύδια είναι χαλαρωμένα. Το στόμα είναι ανοικτό και οι στοματικές γωνίες τραβιούνται πίσω, προς τα αυτιά.
4	Φόβος	Τα φρύδια τραβιούνται μαζί προς τα πάνω. Τα εσωτερικά σημεία των φρυδιών κάμπτονται προς τα πάνω. Τα μάτια είναι ανήσυχα, άγρυπνα.
5	Απέχθεια	Τα φρύδια και τα βλέφαρα είναι χαλαρωμένα. Το άνω χείλος αυξάνει και σουφρώνει, συχνά ασυμμετρικά.
6	Έκπληξη	Τα φρύδια κινούνται προς τα πάνω. Τα άνω βλέφαρα είναι πολύ ανοικτά, ενώ τα κάτω χαλαρά. Το σαγόνι είναι κατεβασμένο.

Πίνακας 2.1: Χαρακτηριστικά εκφράσεων

Ο συνδυασμός των παραπάνω στοιχείων, τα οποία κυρίως προέρχονται από ψυχολογικές μελέτες, αποτέλεσε αφετηρία και θεωρητική βάση της παρούσας διατριβής. Εφαρμόστηκαν με επιτυχία με τη βοήθεια των δύο γνωστότερων τεχνικών εμφύχωσης, του FACS (Facial Action Coding System) και του MPEG-4 FBA (Facial Body Animation), τα οποία παρουσιάζονται στο επόμενο κεφάλαιο, και έδωσαν τα αποτελέσματα που παρουσιάζονται στη συνέχεια της διατριβής.

Κεφάλαιο 3

Τεχνικές Εμφύχωσης: FACS και MPEG-4 FBA

3.1. FACS: Σύντομη Παρουσίαση

Το Facial Action Coding System (FACS) είναι μία ευρέως διαδεδομένη μέθοδος μέτρησης και περιγραφής της «συμπεριφοράς» του προσώπου. Οι Paul Ekman και W.V. Friesen ανέπτυξαν το συγκεκριμένο σύστημα κωδικοποίησης των κινήσεων του προσώπου το 1978, χρησιμοποιώντας γνώση ανθρώπινης ανατομίας και ορίζοντας τον τρόπο με τον οποίο η σύσπαση κάθε μυός του προσώπου (μεμονωμένα και σε συνδυασμό) αλλάζει την εμφάνιση του προσώπου [30], [33]. Εξέτασαν ακολουθίες κινούμενης εικόνας που απεικονίζουν εκφράσεις προσώπου, ώστε να καθορίσουν τις αλλαγές που συντελούνται από τις συσπάσεις των μυών, καθώς και τους συνδυασμούς αυτών και να διαχωρίσουν τους πιθανούς συνδυασμούς κινήσεων, ενώ για τον καθορισμό των AUs (Action Units) χρησιμοποιήθηκαν γνώσεις ανατομίας. Το εγχειρίδιο του FACS πρωτοεκδόθηκε το 1978.

Μονάδες μέτρησης του FACS είναι τα Action Units (AUs). Δεν επιλέχθηκαν οι μύες γιατί ορισμένα AUs συνδυάζουν την κίνηση περισσότερων του ενός μυών (*Εικόνα 3.1*) ή το αποτέλεσμα της σύσπασης ενός μυός εκφράζεται με δύο ή περισσότερα AUs.

Ο κωδικοποιητής ενός τέτοιου συστήματος αναλύει την παρατηρούμενη έκφραση σε συγκεκριμένα AUs και δίνει ως αποτέλεσμα μία λίστα των AUs που δημιουργούν την έκφραση. Επίσης είναι δυνατόν να ληφθούν υπόψιν η διάρκεια, η ένταση και η πιθανή ασυμμετρία. Τα αποτελέσματα αυτά είναι μόνο περιγραφικά και δεν παρέχουν στοιχεία σχετικά με τη σημασία της έκφρασης. Σε επόμενο στάδιο είναι δυνατόν τα στοιχεία του FACS να μετατραπούν σε χρήσιμες για την ψυχολογία έννοιες με τη βοήθεια του FACS-AID, ενός συστήματος ερμηνείας της αντίστοιχης βάσης δεδομένων διαθέσιμου στους ερευνητές.



Εικόνα 3.1: Το FACS λαμβάνει υπόψη τους μυς που αλλάζουν την «εμφάνιση» του προσώπου.

Ο Πίνακας 3.1 περιέχει τα Action Units, όπως αυτά έχουν οριστεί στο FACS.

AU	Description	Facial muscle
1	Inner Brow Raiser	<i>Frontalis, pars medialis</i>
2	Outer Brow Raiser	<i>Frontalis, pars lateralis</i>
4	Brow Lowerer	<i>Corrugator supercilii, Depressor supercilii</i>
5	Upper Lid Raiser	<i>Levator palpebrae superioris</i>
6	Cheek Raiser	<i>Orbicularis oculi, pars orbitalis</i>
7	Lid Tightener	<i>Orbicularis oculi, pars palpebralis</i>
9	Nose Wrinkler	<i>Levator labii superioris alaeque nasi</i>
10	Upper Lip Raiser	<i>Levator labii superioris</i>
11	Nasolabial Deepener	<i>Levator anguli oris (a.k.a. Caninus)</i>
12	Lip Corner Puller	<i>Zygomaticus major</i>
13	Cheek Puffer	<i>Zygomaticus minor</i>
14	Dimpler	<i>Buccinator</i>
15	Lip Corner Depressor	<i>Depressor anguli oris (a.k.a. Triangularis)</i>
16	Lower Lip Depressor	<i>Depressor labii inferioris</i>
17	Chin Raiser	<i>Mentalis</i>
18	Lip Pucker	<i>Incisivii labii superioris and Incisivii labii inferioris</i>
20	Lip stretcher	<i>Risorius w/ platysma</i>
22	Lip Funneler	<i>Orbicularis oris</i>
23	Lip Tightener	<i>Orbicularis oris</i>
24	Lip Pressor	<i>Orbicularis oris</i>
25	Lips part	<i>Depressor labii inferioris or relaxation of Mentalis, or Orbicularis oris</i>
26	Jaw Drop	<i>Masseter, relaxed Temporalis and internal Pterygoid</i>
27	Mouth Stretch	<i>Pterygoids, Digastric</i>
28	Lip Suck	<i>Orbicularis oris</i>
41	Lid droop	<i>Relaxation of Levator palpebrae superioris</i>
42	Slit	<i>Orbicularis oculi</i>
43	Eyes Closed	<i>Relaxation of Levator palpebrae superioris; Orbicularis oculi, pars palpebralis</i>
44	Squint	<i>Orbicularis oculi, pars palpebralis</i>
45	Blink	<i>Relaxation of Levator palpebrae superioris; Orbicularis oculi, pars palpebralis</i>
46	Wink	<i>Relaxation of Levator palpebrae superioris; Orbicularis oculi, pars palpebralis</i>
51	Head turn left	

52	Head turn right	
53	Head up	
54	Head down	
55	Head tilt left	
56	Head tilt right	
57	Head forward	
58	Head back	
61	Eyes turn left	
62	Eyes turn right	
63	Eyes up	
64	Eyes down	

Πίνακας 3.1: Ορισμός Action Units

3.2. MPEG-4 FBA

Στην πρώτη έκδοση του MPEG-4 αναπτύχθηκε ένας τρόπος αναπαράστασης και ανάλυσης του ανθρωπίνου προσώπου (Face animation). Το 1999 το MPEG-4 βασιζόμενο στο H-Anim, ασχολήθηκε για πρώτη φορά με το ανθρώπινο σώμα (Body Animation). Στις πρώτες επεκτάσεις του (MPEG-4 έκδοση 2 αρχές 2000), το πρότυπο απέκτησε την επίσημη διεθνή τυποποιημένη μορφή, ενώ στην τελευταία έκδοση του MPEG-4 υπάρχουν στοιχεία για την κωδικοποίηση όχι μόνο εικονικών ανθρώπων (virtual human) αλλά και ζώων και φυτών.

Στην κατεύθυνση αυτή έχουν αναπτυχθεί τα εξής δύο πλαίσια εφαρμογών:

- το FBA (Face Body Animation)
- το BBA (Bone-base Animation)

3.2.1. MPEG-4 και Πρόσωπο - FDPs και FAPs

Χρησιμοποιούμε ένα σύνολο παραμέτρων για τον προσδιορισμό του σχήματος, του μεγέθους και της υφής του προσώπου (FDPs – Facial Definition Parameters), σε συνδυασμό με ένα άλλο σύνολο παραμέτρων που χρησιμοποιείται για τον προσδιορισμό των «κινήσεων» του προσώπου είτε αυτές προέρχονται από έκφραση συναισθημάτων είτε προκύπτουν κατά την ομιλία (FAPs – Facial Animation Parameters) [29].

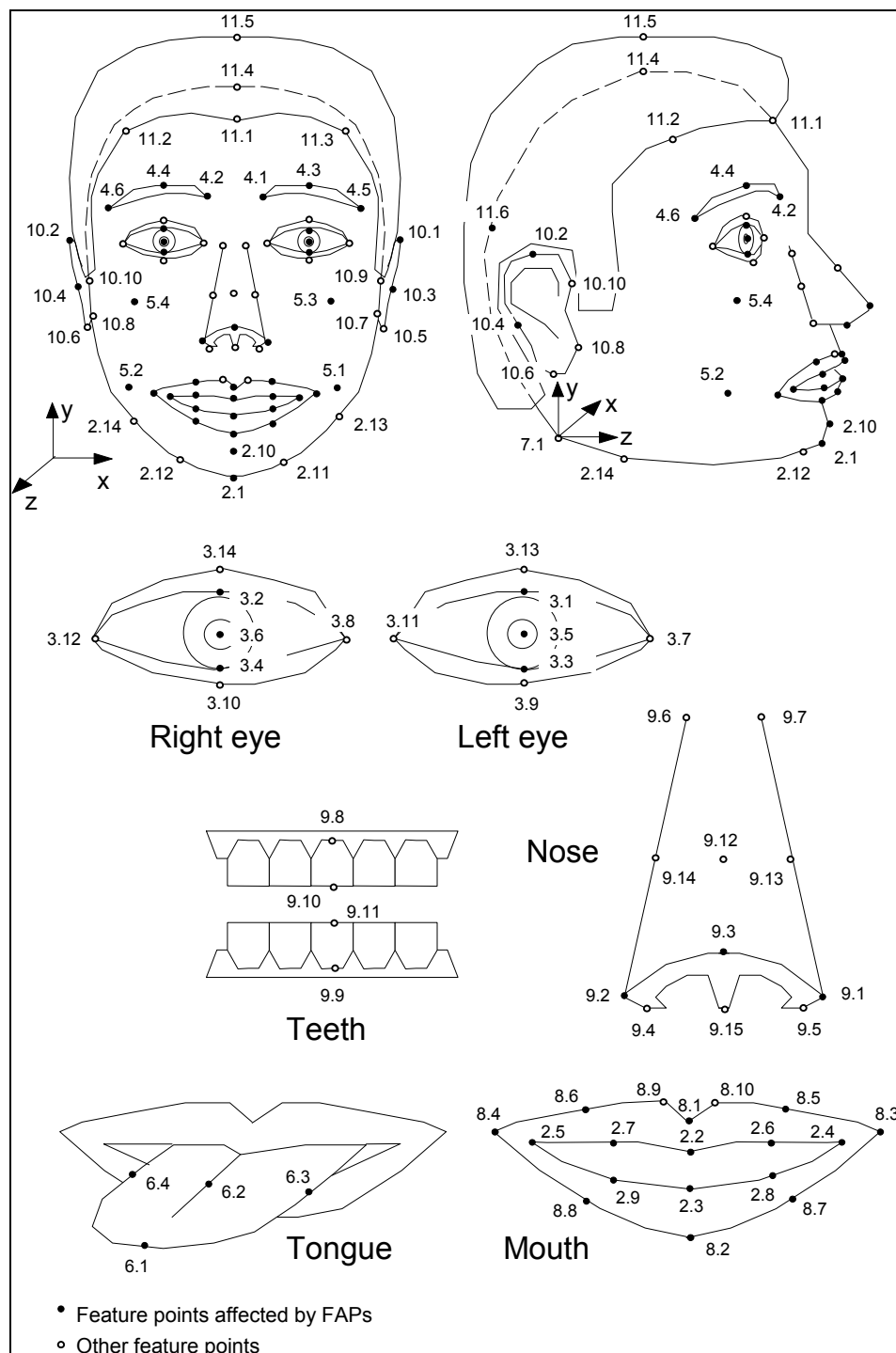
Το σύνολο των FDPs μας παρέχει τη δυνατότητα ακριβούς προσδιορισμού των κινήσεων συγκεκριμένων χαρακτηριστικών του προσώπου, ενώ από τα FAPs, με σωστή ερμηνεία, μπορούμε να προσδιορίσουμε εκφράσεις και εκφορά λόγου σε διάφορα μοντέλα προσώπων χωρίς να είναι απαραίτητες η αρχικοποίηση και η βαθμονόμηση (calibration) της κάμερας. Η μετατροπή των FAPs σε φωνήματα μπορεί να μας οδηγήσει σε ένα μοντέλο κεφαλιού που θα έχει τη δυνατότητα να μιλάει σε όλες τις γλώσσες.

3.2.1.1 Σύνολο Παραμέτρων Προσδιορισμού Προσώπου (FDPs)

Με τα FDPs μπορούμε να προσαρμόσουμε ένα μοντέλο προσώπου σε ένα δεδομένο πρόσωπο. Το σύνολο των FDPs μπορεί να περιλαμβάνει πλέγμα 3D με πληροφορίες για

την υφή, τρισδιάστατα χαρακτηριστικά σημεία και, προαιρετικά, κάποια εικόνα της υφής και άλλα χαρακτηριστικά όπως μαλλιά, γυαλιά οράσεως, ηλικία ή φύλο. Το πλέγμα 3D χρησιμοποιείται για τον προσδιορισμό του σχήματος του προσώπου, ενώ τα τρισδιάστατα χαρακτηριστικά σημεία χρησιμοποιούνται για τον προσδιορισμό των χαρακτηριστικών του προσώπου στο συγκεκριμένο τρισδιάστατο σχήμα.

Τα Feature Points (FPs) των FDPs φαίνονται στην **Εικόνα 3.2**.



Εικόνα 3.2: Ο ορισμός των Feature Points στο MPEG-4

Τα FDPs περιέχουν τα εξής πεδία:

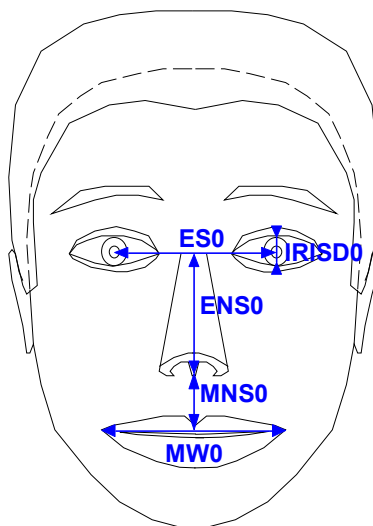
- **FeaturePointsCoord** – τα τρισδιάστατα χαρακτηριστικά σημεία που χρησιμοποιούνται για τη ζυγοστάθμιση του μοντέλου του προσώπου
- **TextureCoords** – συντεταγμένες της υφής για τα χαρακτηριστικά σημεία
- **TextureType** – πληροφορία για τον αποκωδικοποιητή σχετική με το τύπο της εικόνας της υφής.
- **FaceDefTables** – περιγραφή της συμπεριφορά των FAPs.
- **FaceSceneGraph** - περιέχει την εικόνα της υφής ή πληροφορίες για την ιεραρχία του μοντέλου.

3.2.1.2 Σύνολο Παραμέτρων Κίνησης Προσώπου (FAPs)

Τα FAPs βασίζονται στη μελέτη των ελαχίστων κινήσεων του προσώπου και συνδέονται στενά με τις κινήσεις των μυών. Αντιπροσωπεύουν ένα ολοκληρωμένο σύνολο βασικών ενεργειών του προσώπου, επιτρέποντας την απεικόνιση της πλειοψηφίας των φυσιολογικών ανθρωπίνων εκφράσεων, ενώ οι υπερβολικές τιμές μας επιτρέπουν να ορίσουμε ενέργειες αδύνατες για έναν άνθρωπο, αλλά απαραίτητες, για παράδειγμα, για τους χαρακτήρες cartoon. Όλες οι παράμετροι που εμπεριέχουν μεταφορική κίνηση εκφράζονται με τους όρους των μονάδων κίνησης των χαρακτηριστικών του προσώπου (FAPU – Facial Animation Parameter Units).

Προέκυψε η ανάγκη ορισμού των μονάδων αυτών ώστε να είναι δυνατή η εφαρμογή των FAPs σε οποιοδήποτε μοντέλο προσώπου με σταθερό τρόπο, καθώς και η παραγωγή λογικών και αναμενόμενων αποτελεσμάτων, όσον αφορά στις εκφράσεις και στην εκφορά του λόγου.

Τα FAPUs απεικονίζονται στην **Εικόνα 3.3** και αντιστοιχούν, στην ουσία, σε κλάσματα της απόστασης ορισμένων σημείων – κλειδιών του προσώπου, επιλεγμένα με τέτοιο τρόπο ώστε να δίνουν την απαιτούμενη ακρίβεια.



Εικόνα 3.3: Μονάδες κίνησης των χαρακτηριστικών του προσώπου [53]

Οι παράμετροι FAPUs που απεικονίζονται στην **Εικόνα 3.3**, δίνουν τις σχέσεις του **Πίνακα 3.2** για τις μονάδες μέτρησης του MPEG-4:

Ονομασία	Περιγραφή	Σχέσεις ορισμού μονάδων μέτρησης MPEG-4
IRISD0	Διάμετρος της ίριδας (εξ ορισμού ισούνται με την απόσταση άνω και κάτω βλεφάρου) στο ουδέτερο πρόσωπο	$IRISD = \frac{IRISD0}{1024}$
ES0	Απόσταση ματιών	$ES = \frac{ES0}{1024}$
ENS0	Απόσταση μύτης-ματιών	$ENS = \frac{ENS0}{1024}$
MNS0	Απόσταση στόματος-μύτης	$MNS = \frac{MNS0}{1024}$
MW0	Πλάτος στόματος	$MW = \frac{MW0}{1024}$
AU	Γωνιακή μονάδα	$AU = 10E - 5rad$

Πίνακας 3.2: Μονάδες μέτρησης των FAPs

Για κάθε FAP ορίζουμε αν έχει μόνο θετικές (unidirectional) ή και αρνητικές τιμές (bi-directional), προς ποια κατεύθυνση είναι οι θετικές τιμές και ποιες ακριβώς είναι οι μονάδες μέτρησης (μοίρες ή κάποιο από τα FAPUs), ενώ μόνο για την περίπτωση του σαγονιού χρησιμοποιούμε μονάδες που μετρούν την «ένταση» πίεσης του σαγονιού με κλίμακα από το 1 έως το 10.

Υπάρχουν FAPs που προσδιορίζουν την κίνηση του βολβού του ματιού, του βλεφάρου, το μέγεθος της κόρης του ματιού, την κίνηση των φρυδιών (με τρία σημεία για κάθε φρύδι), την κίνηση της μύτης και των ρουθουνιών ειδικότερα, όσο μικρή κι αν είναι αυτή. Άλλα FAPs προσδιορίζουν τις κινήσεις από τα μάγουλα και την πίεση που πιθανώς ασκείται στο σαγόني καθώς και την κίνηση του στόματος (χρησιμοποιώντας τρία σημεία και προσδιορίζοντας συγχρόνως και την πιθανή κατακόρυφη κίνηση του άνω και του κάτω χείλους ή το τράβηγμα των άκρων του στόματος, εσωτερικά αλλά και εξωτερικά). FAPs υπάρχουν ακόμη και για την κίνηση της άκρης της γλώσσας ή και για πιθανή κίνηση των αυτιών.

Υπάρχουν κάποια «ανώτερα FAPs» που περιγράφουν, χωρίς να διευκρινίζουν λεπτομέρειες, κάποια από τις έξι γνωστές εκφράσεις για το πρόσωπο που θα αποκωδικοποιηθεί. Τα FAPs αυτά (έξι, ένα για κάθε βασική έκφραση) έχουν προτεραιότητα έναντι των υπολοίπων, δηλαδή ακόμα και αν, σύμφωνα με τα υπόλοιπα FAPs, ορίζεται κάτι διαφορετικό, ο αποκωδικοποιητής θα λάβει υπόψη του μόνο τα ανώτερα FAPs.

Επίσης, έχουν οριστεί και άλλες παράμετροι που προσδιορίζουν τη στάση του κεφαλιού και την ακριβή έκφραση του προσώπου.

Το MPEG-4 προσδιορίζει 84 σημεία χαρακτηριστικών στο ουδέτερο πρόσωπο, τα οποία παρέχουν χωρική αναφορά για τον ορισμό των FAPs. Ο ορισμός του Facial Animation στο πλαίσιο του προτύπου ISO MPEG-4 [1] βασίζεται στο Σύστημα Κωδικοποίησης Ενεργειών του Προσώπου (FACS – Facial Action Coding System) (Ενότητα 3.1). Συγκεκριμένα, τα σύνολα παραμέτρων ορισμού προσώπου FDP (Facial Definition Parameter) και οι παράμετροι κίνησης προσώπου FAP (Facial Animation Parameter) σχεδιάστηκαν στο πρότυπο MPEG-4 για να επιτρέψουν τον ορισμό του σχήματος και της υφής του προσώπου, μειώνοντας την ανάγκη προσδιορισμού της τοπολογίας της υπάρχουσας γεωμετρίας μέσω των FDPs και την κίνηση των χαρακτηριστικών του προσώπου, αναπαράγοντας εκφράσεις, συναισθήματα και εκφορά λόγου μέσω των FAPs.

Ο ορισμός του viseme έχει συμπεριληφθεί στο πρότυπο για τον συγχρονισμό των κινήσεων του στόματος ως προς τα φωνήματα με τη συνολική κίνηση των χαρακτηριστικών του προσώπου (facial animation). Παρακολουθώντας τα νεύματα (gestures) του προσώπου που αντιστοιχούν σε FDP και/ή FAP κινήσεις ως προς το χρόνο, είναι δυνατή η απόδοση στοιχείων (cues) που αφορούν τις εκφράσεις και τα συναισθήματα του χρήστη.

Ποικίλα αποτελέσματα έχουν παρουσιαστεί που αφορούν στην κατάταξη βασικών εκφράσεων του προσώπου, βασιζόμενα κυρίως σε χαρακτηριστικά ή σημεία που εξάγονται από περιοχές του προσώπου γύρω από το στόμα ή τα μάτια. Αυτά τα αποτελέσματα υποδηλώνουν ότι οι εκφράσεις του προσώπου σε συνδυασμό με κάποιες χειρονομίες και την φωνή, όταν αυτή είναι διαθέσιμη, παρέχουν στοιχεία που μπορούν να χρησιμοποιηθούν για την κατανόηση της συναισθηματικής κατάστασης του ατόμου.

Τα FAPs ομαδοποιούνται ανάλογα με την λογική ορισμού τους, όπως φαίνεται στον **Πίνακα 3.3** [29].

Ομάδες FAP		Πλήθος FAP
1	Εκφράσεις	2
2	Σαγόνι, πηγούνι, εσωτερικό του κάτω χείλους, άκρες χειλιών, μέσο χειλιών	16
3	Βολβός οφθαλμού, κόρη οφθαλμού, βλέφαρο	12
4	Φρύδι	8
5	Μάγουλο	4
6	Γλώσσα	5
7	Περιστροφή κεφαλιού	3
8	Θέση εξωτερικού χείλος	10
9	Μύτη	4
10	Αυτιά	4

Πίνακας 3.3: Ομάδες FAPs

Τα FAPs, όπως αυτά έχουν οριστεί στο πλαίσιο του προτύπου MPEG-4 περιλαμβάνονται στον **Πίνακα 3.4**:

α/α	Όνομα FAP	Περιγραφή FAP	Μονάδα μέτρησης	Μετατόπιση / κλίση / περιστροφή / κλίση / περιστροφή	Κατεύθυνση Κίνησης	Ομάδα
1	Viseme	Καθορίζει την ακολουθία των visemes (κίνηση στόματος για προφορά συγκεκριμένων φωνημάτων) για το συγκεκριμένο καρέ.	-	-	-	1
2	Expression	Ένα σύνολο τιμών που καθορίζουν την έκφραση του προσώπου.	-	-	-	1
3	open_jaw	Κατακόρυφη μετατόπιση σαγονιών (δεν επηρεάζεται από το άνοιγμα του στόματος)	MNS	M	Κάτω	2
4	lower_t_midlip	Κατακόρυφη μετατόπιση του μέσου του εσωτερικού του άνω χείλους	MNS	Δ	Κάτω	2
5	raise_b_midlip	Κατακόρυφη μετατόπιση του μέσου του εσωτερικού του κάτω χείλους	MNS	Δ	Πάνω	2
6	stretch_l_cornerlip	Οριζόντια μετατόπιση της αριστερής εσωτερικής γωνίας των χειλιών	MW	Δ	Αριστερά	2
7	stretch_r_cornerlip	Οριζόντια μετατόπιση της δεξιάς εσωτερικής γωνίας των χειλιών	MW	Δ	Δεξιά	2
8	lower_t_lip_lm	Κατακόρυφη μετατόπιση του μέσου μεταξύ της αριστερής γωνίας και του μέσου του εσωτερικού του άνω χείλους	MNS	Δ	Κάτω	2
9	lower_t_lip_rm	Κατακόρυφη μετατόπιση του μέσου μεταξύ της δεξιάς γωνίας και του μέσου του εσωτερικού του άνω χείλους	MNS	Δ	Κάτω	2
10	raise_b_lip_lm	Κατακόρυφη μετατόπιση του μέσου μεταξύ της αριστερής γωνίας και του μέσου του εσωτερικού του κάτω χείλους	MNS	Δ	Πάνω	2
11	raise_b_lip_rm	Κατακόρυφη μετατόπιση του μέσου μεταξύ της δεξιάς γωνίας και του μέσου του εσωτερικού του κάτω χείλους	MNS	Δ	Πάνω	2
12	raise_l_cornerlip	Κατακόρυφη μετατόπιση της εσωτερικής αριστερής γωνίας των χειλιών	MNS	Δ	Πάνω	2
13	raise_r_cornerlip	Κατακόρυφη μετατόπιση της εσωτερικής δεξιάς γωνίας των χειλιών	MNS	Δ	Πάνω	2
14	thrust_jaw	Προώθηση του σαγονιού	MNS	M	Μπροστά	2
15	shift_jaw	Οριζόντια μετατόπιση του σαγονιού	MW	Δ	Δεξιά	2
16	push_b_lip	Προώθηση της μέσης του κάτω χείλους	MNS	Δ	Μπροστά	2
17	push_t_lip	Προώθηση της μέσης του άνω χείλους	MNS	Δ	Μπροστά	2
18	depress_chin	Ανύψωση και συμπίεση του σαγονιού (όπως στην έκφραση της λύπης).	MNS	Δ	Πάνω	2
19	close_t_l_eyelid	Κατακόρυφη μετατόπιση του άνω αριστερού	IRISD	Δ	Κάτω	3

		βλέφαρου				
20	close_t_r_eyelid	Κατακρόρυφη μετατόπιση του άνω δεξιού βλέφαρου	IRISD	Δ	Κάτω	3
21	close_b_l_eyelid	Κατακρόρυφη μετατόπιση του κάτω αριστερού βλέφαρου	IRISD	Δ	Πάνω	3
22	close_b_r_eyelid	Κατακρόρυφη μετατόπιση του κάτω δεξιού βλέφαρου	IRISD	Δ	Πάνω	3
23	yaw_l_eyeball	Οριζόντιος προσανατολισμός του βολβού του αριστερού οφθαλμού	AU	Δ	Αριστερά	3
24	yaw_r_eyeball	Οριζόντιος προσανατολισμός του βολβού του δεξιού οφθαλμού	AU	Δ	Αριστερά	3
25	pitch_l_eyeball	Κατακρόρυφος προσανατολισμός του βολβού του αριστερού οφθαλμού	AU	Δ	Κάτω	3
26	pitch_r_eyeball	Κατακρόρυφος προσανατολισμός του βολβού του δεξιού οφθαλμού	AU	Δ	Κάτω	3
27	thrust_l_eyeball	Προώθηση του βολβού του αριστερού οφθαλμού	ES	Δ	Μπροστά	3
28	thrust_r_eyeball	Προώθηση του βολβού του δεξιού οφθαλμού	ES	Δ	Μπροστά	3
29	dilate_l_pupil	Διαστολή κόρης αριστερού οφθαλμού	IRISD	Δ	Αυξητικά	3
30	dilate_r_pupil	Διαστολή κόρης δεξιού οφθαλμού	IRISD	Δ	Αυξητικά	3
31	raise_l_i_eyebrow	Κατακρόρυφη μετατόπιση του αριστερού εσωτερικού φρυδιού	ENS	Δ	Πάνω	4
32	raise_r_i_eyebrow	Κατακρόρυφη μετατόπιση του δεξιού εσωτερικού φρυδιού	ENS	Δ	Πάνω	4
33	raise_l_m_eyebrow	Κατακρόρυφη μετατόπιση του αριστερού μέσου φρυδιού	ENS	Δ	Πάνω	4
34	raise_r_m_eyebrow	Κατακρόρυφη μετατόπιση του δεξιού μέσου φρυδιού	ENS	Δ	Πάνω	4
35	raise_l_o_eyebrow	Κατακρόρυφη μετατόπιση του αριστερού εξωτερικού φρυδιού	ENS	Δ	Πάνω	4
36	raise_r_o_eyebrow	Κατακρόρυφη μετατόπιση του δεξιού εξωτερικού φρυδιού	ENS	Δ	Πάνω	4
37	squeeze_l_eyebrow	Οριζόντια μετατόπιση του αριστερού φρυδιού	ES	Δ	Δεξιά	4
38	squeeze_r_eyebrow	Οριζόντια μετατόπιση του δεξιού φρυδιού	ES	Δ	Αριστερά	4
39	puff_l_cheek	Οριζόντια μετατόπιση του αριστερού μάγουλου	ES	Δ	Αριστερά	5
40	puff_r_cheek	Οριζόντια μετατόπιση του δεξιού μάγουλου	ES	Δ	Δεξιά	5
41	lift_l_cheek	Κατακρόρυφη μετατόπιση του αριστερού μάγουλου	ENS	M	Πάνω	5
42	lift_r_cheek	Κατακρόρυφη μετατόπιση του δεξιού μάγουλου	ENS	M	Πάνω	5
43	shift_tongue_tip	Οριζόντια μετατόπιση της άκρης της γλώσσας	MW	Δ	Δεξιά	6
44	raise_tongue_tip	Κατακρόρυφη μετατόπιση της άκρης της γλώσσας	MNS	Δ	Πάνω	6
45	thrust_tongue_tip	Προώθηση της άκρης της γλώσσας	MW	Δ	Μπροστά	6
46	raise_tongue	Κατακρόρυφη μετατόπιση της γλώσσας	MNS	Δ	Πάνω	6
47	tongue_roll	Διπλωμα της γλώσσας σε σχήμα U	AU	M	Κοίλη προς τα πάνω	6
48	head_pitch	Γωνία κεφαλής από την κορυφή της σπονδυλικής στήλης	AU	Δ	Κάτω	7
49	head_yaw	Εκτροπή κεφαλής από την κορυφή της σπονδυλικής στήλης	AU	Δ	Αριστερά	7

50	head_roll	Στροφή κεφαλής από την κορυφή της σπονδυλικής στήλης	AU	Δ	Δεξιά	7
51	lower_t_midlip_o	Κατακόρυφη μετατόπιση του μέσου του εξωτερικού του άνω χείλους	MNS	Δ	Κάτω	8
52	raise_b_midlip_o	Κατακόρυφη μετατόπιση του μέσου του εξωτερικού του κάτω χείλους	MNS	Δ	Πάνω	8
53	stretch_l_cornerlip_o	Οριζόντια μετατόπιση της αριστερής εξωτερικής γωνίας των χειλιών.	MW	Δ	Αριστερά	8
54	stretch_r_cornerlip_o	Οριζόντια μετατόπιση της δεξιάς εξωτερικής γωνίας των χειλιών.	MW	Δ	Δεξιά	8
55	lower_t_lip_lm_o	Κατακόρυφη μετατόπιση του μέσου μεταξύ της αριστερής γωνίας και του μέσου του εξωτερικού του άνω χείλους	MNS	Δ	Κάτω	8
56	lower_t_lip_rm_o	Κατακόρυφη μετατόπιση του μέσου μεταξύ της δεξιάς γωνίας και του μέσου του εξωτερικού του άνω χείλους	MNS	Δ	Κάτω	8
57	raise_b_lip_lm_o	Κατακόρυφη μετατόπιση του μέσου μεταξύ της αριστερής γωνίας και του μέσου του εξωτερικού του κάτω χείλους	MNS	Δ	Πάνω	8
58	raise_b_lip_rm_o	Κατακόρυφη μετατόπιση του μέσου μεταξύ της δεξιάς γωνίας και του μέσου του εξωτερικού του κάτω χείλους	MNS	Δ	Πάνω	8
59	raise_l_cornerlip_o	Κατακόρυφη μετατόπιση της αριστερής εξωτερικής γωνίας των χειλιών	MNS	Δ	Πάνω	8
60	raise_r_cornerlip_o	Κατακόρυφη μετατόπιση της δεξιάς εξωτερικής γωνίας των χειλιών	MNS	Δ	Πάνω	8
61	stretch_r_nose	Οριζόντια μετατόπιση της αριστερής πλευράς της μύτης	ENS	Δ	Αριστερά	9
62	stretch_r_nose	Οριζόντια μετατόπιση της δεξιάς πλευράς της μύτης	ENS	Δ	Δεξιά	9
63	raise_nose	Κατακόρυφη μετατόπιση της άκρης της μύτης	ENS	Δ	Πάνω	9
64	bend_nose	Οριζόντια μετατόπιση της άκρης της μύτης	ENS	Δ	Δεξιά	9
65	raise_l_ear	Κατακόρυφη μετατόπιση του αριστερού αφτιού	ENS	Δ	Πάνω	10
66	raise_r_ear	Κατακόρυφη μετατόπιση του δεξιού αφτιού	ENS	Δ	Πάνω	10
67	pull_l_ear	Οριζόντια μετατόπιση του αριστερού αφτιού	ENS	Δ	Αριστερά	10
68	pull_r_ear	Οριζόντια μετατόπιση του δεξιού αφτιού	ENS	Δ	Δεξιά	10

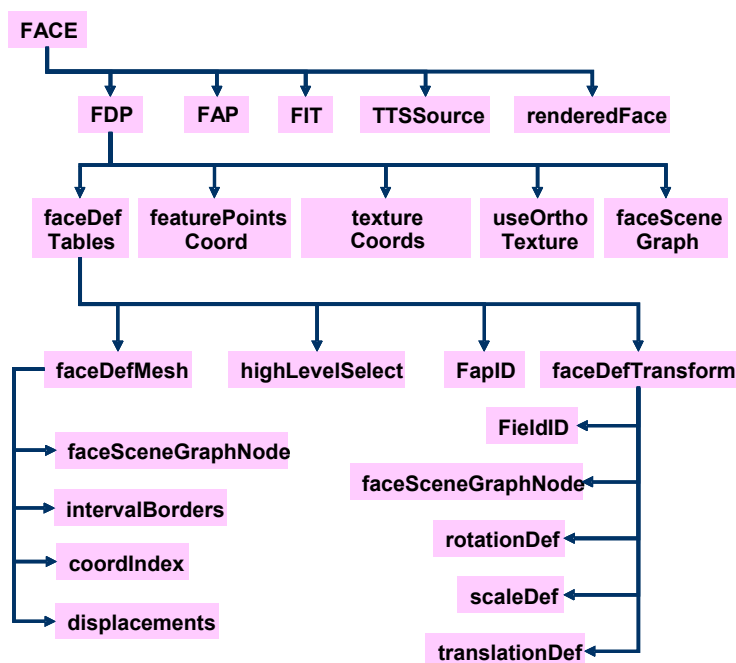
Πίνακας 3.4: Ορισμοί FAPs

3.2.1.3 Πίνακες Παρεμβολής FAPs

Όπως αναφέρθηκε προηγουμένως, ο κωδικοποιητής μπορεί να επιτρέψει στον αποκωδικοποιητή να προσεγγίσει τις τιμές κάποιων FAPs με βάση τα μεταδιδόμενα FAPs. Εναλλακτικά, ο αποκωδικοποιητής μπορεί να προσδιορίσει τους κανόνες παρεμβολής με τη χρήση πινάκων παρεμβολής FAPs (FIT – FAPs Interpolation Tables). Ένας FIT επιτρέπει ένα μικρότερο σύνολο από FAPs να σταλεί για προσδιορισμό κίνησης προσώπου. Αυτό το

μικρό σύνολο μπορεί αργότερα να χρησιμοποιηθεί για τον καθορισμό των τιμών άλλων FAPs, με χρήση λογικής πολωνυμικής συσχέτισης (rational polynomial mapping) μεταξύ των παραμέτρων. Για παράδειγμα, το FAP του εσωτερικού του πάνω χείλους μπορεί να σταλεί και στη συνέχεια να χρησιμοποιηθεί για τον προσδιορισμό του FAP του εξωτερικού του πάνω χείλους, με χρήση συνάρτησης λογικού πολυωνύμου που προσδιορίζεται στο FIT.

Σχηματικά, ο ολοκληρωμένος τρόπος κωδικοποίησης της κίνησης του προσώπου, σύμφωνα με το πρότυπο MPEG-4 δίνεται στην **Εικόνα 3.4**:



Εικόνα 3.4: Γράφημα όλων των σταδίων κωδικοποίησης της κίνησης του προσώπου

3.2.1.4 Καθορισμός του μοντέλου του προσώπου

Κάθε αποκωδικοποιητής MPEG-4 που λαμβάνει τα FAPs πρέπει να παράσχει ένα πρότυπο μοντέλο προσώπου MPEG-4 που να περιέχει και κίνηση. Συνήθως, αυτό είναι ένα πρότυπο το οποίο γνωρίζει ο αποκωδικοποιητής. Ο κωδικοποιητής δεν γνωρίζει τη μορφή του προτύπου προσώπου. Χρησιμοποιώντας παραμέτρους ορισμού του προσώπου (FDP- Face Definition Parameter), το MPEG-4 επιτρέπει στον κωδικοποιητή να διευκρινίσει με μεγάλη ακρίβεια το μοντέλο του προσώπου, του οποίου τα χαρακτηριστικά θα κινηθούν.

Για να υλοποιηθεί αυτό πρέπει να έχει καθοριστεί η στατική γεωμετρία του προτύπου προσώπου στη φυσική του κατάσταση με σκοπό να καθοριστούν οι ιδιότητες της επιφάνειας του προσώπου καθώς και οι κανόνες κίνησης, όπως αυτοί καταγράφονται στους πίνακες κίνησης προσώπου (FAT- Facial Animation Tables), οι οποίοι διευκρινίζουν πώς αυτό το μοντέλο μετασχηματίζεται από τα FAPs.

3.2.2. MPEG-4 και ανθρώπινο σώμα

Οι τελευταίες αναθεωρήσεις και προσθήκες του προτύπου MPEG-4 ασχολούνται και με την κίνηση του σώματος ορίζοντας παραμέτρους που περιγράφουν πλέον μέρη του ανθρώπινου σώματος και όχι του προσώπου. Η περιγραφή βέβαια του ανθρώπινου σώματος, κυρίως της κίνησης του, παρουσιάζει πολλές δυσκολίες λόγω των βαθμών ελευθερίας και της γεωμετρίας του σώματος.

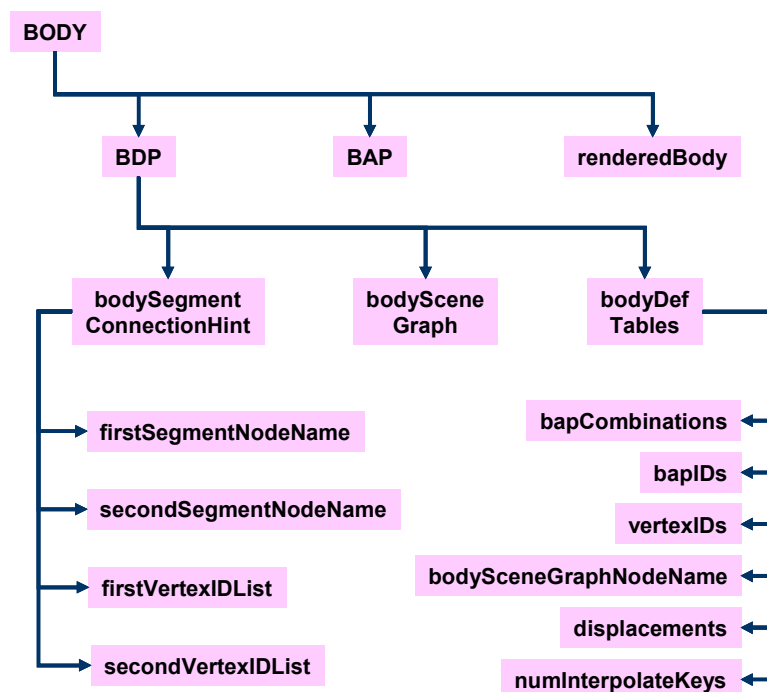
3.2.2.1. BDPs και BAPs

Στο πλαίσιο του προτύπου MPEG-4, παράμετροι έχουν οριστεί όχι μόνο για το πρόσωπο αλλά και για την κίνηση του σώματος (FBA - Face Body Animation) με τον καθορισμό των συγκεκριμένων συνόλων σημείων προσώπου και σώματος.

Η δεύτερη έκδοση του προτύπου, ακολουθώντας την ίδια διαδικασία με αυτή του ορισμού του προσώπου και της κίνησης (μέσω των FDPs και FAPs), περιγράφει την ανατομία του ανθρώπινου σώματος με ομάδες διακριτών συμβόλων, ελαττώνοντας την ανάγκη για προσδιορισμό της υπάρχουσας γεωμετρίας. Αυτά τα σύμβολα μπορούν να χρησιμοποιηθούν έτσι ώστε να ανιχνεύουν αυτόματα μετρήσεις και ενδείξεις κίνησης σε μια ακολουθία βίντεο, βοηθώντας έτσι τον υπολογισμό της πραγματικής κίνησης που διαγράφεται από το υποκείμενο, και, αν είναι απαραίτητο, την προσέγγιση της από μια συνθετική κίνηση.

Γενικά, ένα MPEG-4 σώμα είναι μια συλλογή από κόμβους. Το σύνολο των παραμέτρων ορισμού του σώματος (BDP – Body Definition Parameter) παρέχει πληροφορία για την επιφάνεια του σώματος, τις διαστάσεις και την υφή του, ενώ το σύνολο παραμέτρων κίνησης του σώματος (BAP – Body Animation Parameter) μετασχηματίζει τη στάση του σώματος. Τα BAPs περιγράφουν την τοπολογία του ανθρώπινου σκελετού, λαμβάνοντας υπόψιν τους περιορισμούς που υπεισέρχονται λόγω αρθρώσεων και ανεξάρτητων βαθμών ελευθερίας στο μοντέλο του σκελετού από διαφορετικά μέρη του σώματος. Το MPEG-4 παρέχει λοιπόν μία υψηλού επιπέδου γλώσσα για την περιγραφή της κίνησης, δεδομένου ότι κάθε BAP έχει εξαιρετικά περιγραφικό όνομα.

Ο ολοκληρωμένος τρόπος κωδικοποίησης της κίνησης του σώματος, στα πλαίσια του προτύπου MPEG-4 απεικονίζεται στην **Εικόνα 3.5**.



Εικόνα 3.5: Γράφημα όλων των σταδίων κωδικοποίησης της κίνησης του σώματος

3.2.2.2. Κίνηση βασιζόμενη στα οστά BBA (Bone Based Animation)

Το MPEG-4 BBA προσφέρει ένα τυποποιημένο σχήμα ανταλλαγής που επεκτείνει το MPEG-4 FBA [52]. Στο BBA ο σκελετός είναι μια ιεραρχική δομή που αποτελείται από οστά. Σε αυτήν την ιεραρχία κάθε οστό έχει έναν γονέα και μπορεί να έχει ως παιδιά άλλα οστά, μυς ή τρισδιάστατα αντικείμενα. Για τη μετακίνηση κάθε οστού πρέπει να καθορίσουμε την επίδραση αυτής της μετακίνησης στο δέρμα του μοντέλου μας, καθώς και την επίδραση αυτής της κίνησης στα παιδιά του και το αντίστροφο.

Στα διαδοχικά πλαίσια BBA η περιστροφή αντιπροσωπεύεται ως γωνίες Euler. Περιέχει όλα τα κινούμενα πλαίσια ή δεδομένα στα χρονικά βασικά πλαίσια. Στη δεύτερη περίπτωση ο αποκωδικοποιητής θα υπολογίσει τα ενδιάμεσα πλαίσια από τη χρονική παρεμβολή (γραμμική παρεμβολή για τη μετάφραση και για την κλίμακα και σφαιρική γραμμική παρεμβολή για την περιστροφή και τον προσανατολισμό της κλίμακας). Οι αναπαραστάσεις, οι οποίες είναι βασισμένες στα οστά, οφείλονται στην σύνθεση και στην ανάλυση των χειρονομιών των χεριών και άλλων κινήσεων του σώματος.

Ο συνδυασμός των κινήσεων του προσώπου και του σώματος στα πλαίσια του προτύπου MPEG-4 μπορεί να έχει πολλαπλές εφαρμογές, ενώ συγχρόνως η χρήση παραμέτρων καλά ορισμένων πιστοποιεί καλύτερο αποτέλεσμα και πιο εύχρηστο στη διεθνή επιστημονική κοινότητα.

Κεφάλαιο 4

Εκφράσεις Προσώπου – Ορισμός και Ανάλυση

4.1. Αντιστοίχιση AUs – FAPs

Στα προηγούμενα κεφάλαια αναφέρθηκαν δύο διαφορετικές ομάδες παραμέτρων καθορισμού της «συμπεριφοράς» του προσώπου. Αυτές οι δύο ομάδες παρόλο που έχουν οριστεί σαφώς κι εντελώς ανεξάρτητα η μία από την άλλη έχουν ως κοινή αφετηρία ανατομικές μελέτες και βασίζονται στις κινήσεις των μυών του προσώπου, δηλαδή οι παράμετροι του προτύπου MPEG-4 (FAPs) είναι στενά συνδεδεμένες με τα Action Units (AUs). Μία προσέγγιση αντιστοίχισης αποτυπώνεται στον **Πίνακα 4.1**.

Για την αντιστοίχιση αυτή, πέραν του ορισμού τους, ελήφθησαν υπόψιν στοιχεία από ψυχολογικές μελέτες [39], [40], [32] σχετικές με το σχεδιασμό κίνησης του προσώπου κατά τη δημιουργία εκφράσεων, καθώς και από πειραματικό υλικό γνωστών βάσεων δεδομένων, όπως η στατική βάση του Ekman και η δυναμική βάση του MediaLab.

AU Description	FAPs
AU1 Inner Brow Raiser	<i>raise_l_i_eyebrow (F₃₁) + raise_r_i_eyebrow (F₃₂)</i>
AU2 Outer Brow Raiser	<i>raise_l_o_eyebrow(F₃₅) + raise_r_o_eyebrow (F₃₆)</i>
AU4 Brow Lowerer	<i>raise_l_o_eyebrow(F₃₅) + raise_r_o_eyebrow(F₃₆) + raise_l_m_eyebrow(F₃₃) + raise_r_m_eyebrow(F₃₄) + raise_l_i_eyebrow(F₃₁) + raise_r_i_eyebrow(F₃₂) + squeeze_l_eyebrow (F₃₇) + squeeze_r_eyebrow (F₃₈)</i> +
AU5 Upper Lid Raiser	<i>close_t_l_eyelid (F₁₉) + close_t_r_eyelid (F₂₀)</i>
AU6 Cheek Raiser	<i>lift_l_cheek (F₄₁) + lift_r_cheek (F₄₂)</i>
AU7 Lid Tightener	<i>close_b_l_eyelid (F₂₁) + close_b_r_eyelid (F₂₂)</i>
AU9	<i>lower_t_midlip(F₄) + raise_nose(F₆₃) + stretch_l_nose(F₆₁) + stretch_r_nose(F₆₂)</i>

Nose Wrinkler	
AU10 Upper Lip Raiser	$raise_nose (F_{63}) (+ stretch_l_nose (F_{61}) + stretch_r_nose (F_{62})) + lower_t_midlip (F_4)$
AU11 Nasolabial Deepener	$stretch_l_nose (F_{61}) + stretch_r_nose (F_{62})$
AU12 Lip Corner Puller	$push_t_lip (F_{17}) + push_b_lip (F_{16}) [+ lower_t_midlip (F_4) + raise_b_midlip (F_5)]$
AU13 Cheek Puffer	$raise_l_cornerlip (F_{12}) + raise_r_cornerlip (F_{13}) + puff_l_cheek (F_{39}) + puff_r_cheek (F_{40}) + lift_l_cheek (F_{41}) + lift_r_cheek (F_{42})$
AU14 Dimpler	$stretch_l_cornerlip (F_6) + stretch_r_cornerlip (F_7) + raise_l_cornerlip_o (F_{59}) + raise_r_cornerlip_o (F_{60}) + stretch_l_cornerlip_o (F_{53}) + stretch_r_cornerlip_o (F_{54})$
AU15 Lip Corner Depressor	$raise_l_cornerlip (F_{12}) + raise_r_cornerlip (F_{13})$
AU16 Lower Lip Depressor	$raise_b_midlip (F_5) + raise_b_lip_lm (F_{10}) + raise_b_lip_rm (F_{11}) + raise_b_lip_lm_o (F_{57}) + raise_b_lip_rm_o (F_{58}) + push_b_lip (F_{16})$
AU17 Chin Raiser	$depress_chin (F_{18})$
AU18 Lip Puckerer	$push_b_lip (F_{16}) + push_t_lip (F_{17}) + stretch_l_cornerlip (F_6) + stretch_r_cornerlip (F_7) + stretch_l_cornerlip_o (F_{53}) + stretch_r_cornerlip_o (F_{54}) + [lower_t_lip_lm (F_8) + raise_b_lip_lm (F_{10}) + lower_t_lip_lm_o (F_{55}) + raise_b_lip_lm_o (F_{57}) + lower_t_lip_rm (F_9) + raise_b_lip_rm (F_{11}) + lower_t_lip_rm_o (F_{56}) + raise_b_lip_rm_o (F_{58})]$
AU20 Lip stretcher	$raise_b_midlip (F_5) + raise_l_cornerlip (F_{12}) + raise_r_cornerlip (F_{13}) + stretch_l_cornerlip (F_6) + stretch_r_cornerlip (F_7) + lower_t_lip_lm (F_8) + raise_b_lip_lm (F_{10}) + lower_t_lip_lm_o (F_{55}) + raise_b_lip_lm_o (F_{57}) + raise_l_cornerlip_o (F_{59}) + lower_t_lip_rm (F_9) + raise_b_lip_rm (F_{11}) + lower_t_lip_rm_o (F_{56}) + raise_b_lip_rm_o (F_{58}) + raise_r_cornerlip_o (F_{60})$
AU22 Lip Funneler	$push_b_lip (F_{16}) + push_t_lip (F_{17}) + stretch_l_cornerlip (F_6) + stretch_r_cornerlip (F_7) + stretch_l_cornerlip_o (F_{53}) + stretch_r_cornerlip_o (F_{54}) + [lower_t_lip_lm (F_8) + raise_b_lip_lm (F_{10}) + lower_t_lip_lm_o (F_{55}) + raise_b_lip_lm_o (F_{57}) + lower_t_lip_rm (F_9) + raise_b_lip_rm (F_{11}) + lower_t_lip_rm_o (F_{56}) + raise_b_lip_rm_o (F_{58})]$
AU23 Lip Tightener	$push_b_lip (F_{16}) + push_t_lip (F_{17}) + stretch_l_cornerlip (F_6) + stretch_r_cornerlip (F_7) + [lower_t_lip_lm (F_8) + raise_b_lip_lm (F_{10}) + lower_t_lip_lm_o (F_{55}) + raise_b_lip_lm_o (F_{57}) + lower_t_lip_rm (F_9) + raise_b_lip_rm (F_{11}) + lower_t_lip_rm_o (F_{56}) + raise_b_lip_rm_o (F_{58})]$
AU24 Lip Pressor	$lower_t_midlip (F_4) + raise_b_midlip (F_5) + lower_t_lip_lm (F_8) + raise_b_lip_lm (F_{10}) + lower_t_lip_rm (F_9) + raise_b_lip_rm (F_{11})$
AU25 Lips part	$lower_t_midlip (F_4) + raise_b_midlip (F_5) + lower_t_lip_lm (F_8) + raise_b_lip_lm (F_{10}) + lower_t_lip_rm (F_9) + raise_b_lip_rm (F_{11})$
AU26 Jaw Drop	$raise_b_midlip (F_5) + raise_b_lip_lm (F_{10}) + raise_b_lip_rm (F_{11}) + thrust_jaw (F_{14})$
AU27 Mouth Stretch	$stretch_l_cornerlip (F_6) + stretch_r_cornerlip (F_7) + stretch_l_cornerlip_o (F_{53}) + stretch_r_cornerlip_o (F_{54}) + lower_t_lip_lm (F_8) + raise_b_lip_lm (F_{10}) + lower_t_lip_lm_o (F_{55}) + raise_b_lip_lm_o (F_{57}) + lower_t_lip_rm (F_9) + raise_b_lip_rm (F_{11}) + lower_t_lip_rm_o (F_{56}) + raise_b_lip_rm_o (F_{58})$
AU28 Lip Suck	$push_b_lip (F_{16}) + push_t_lip (F_{17}) + lower_t_lip_lm (F_8) + raise_b_lip_lm (F_{10}) + lower_t_lip_lm_o (F_{55}) + raise_b_lip_lm_o (F_{57}) + lower_t_lip_rm (F_9) + raise_b_lip_rm (F_{11}) + lower_t_lip_rm_o (F_{56}) + raise_b_lip_rm_o (F_{58})$
AU41 Lid droop	$close_t_l_eyelid (F_{19}) + close_t_r_eyelid (F_{20})$
AU42 Slit	$close_t_l_eyelid (F_{19}) + close_t_r_eyelid (F_{20}) + close_b_l_eyelid (F_{21}) + close_b_r_eyelid (F_{22})$

	(F_{22})
AU43 Eyes Closed	$close_t_l_eyelid(F_{19}) + close_t_r_eyelid(F_{20}) + close_b_l_eyelid(F_{21}) + close_b_r_eyelid(F_{22})$
AU44 Squint	$close_t_l_eyelid(F_{19}) + close_t_r_eyelid(F_{20}) + close_b_l_eyelid(F_{21}) + close_b_r_eyelid(F_{22}) + squeeze_l_eyebrow(F_{37}) + squeeze_r_eyebrow(F_{38})$
AU45 Blink	$close_t_l_eyelid(F_{19}) + close_t_r_eyelid(F_{20}) + close_b_l_eyelid(F_{21}) + close_b_r_eyelid(F_{22})$
AU46 Wink	$\{close_t_l_eyelid(F_{19}) + close_b_l_eyelid(F_{21})\} OR \{close_t_r_eyelid(F_{20}) + close_b_r_eyelid(F_{22})\}$
AU51 Head turn left	$head_yaw(F_{49})$
AU52 Head turn right	$head_yaw(F_{49})$
AU53 Head up	$head_pitch(F_{48})$
AU54 Head down	$head_pitch(F_{48})$
AU55 Head tilt left	$head_roll(F_{50})$
AU56 Head tilt right	$head_roll(F_{50})$
AU57 Head forward	$head_pitch(F_{48})$
AU58 Head back	$head_pitch(F_{48})$
AU61 Eyes turn left	$yaw_l_eyeball(F_{23}) + yaw_r_eyeball(F_{24})$
AU62 Eyes turn right	$yaw_l_eyeball(F_{23}) + yaw_r_eyeball(F_{24})$
AU63 Eyes up	$pitch_l_eyeball(F_{25}) + pitch_r_eyeball(F_{26})$
AU64 Eyes down	$pitch_l_eyeball(F_{25}) + pitch_r_eyeball(F_{26})$

Πίνακας 4.1: Αντιστοίχιση FAPs - AUs

4.2. Δημιουργία profiles-Προσδιορισμός ομάδων FAPs για κάθε έκφραση

Η περιγραφή των βασικών εκφράσεων με βάση την κίνηση των μυών και τα AUs υπήρξε η αφετηρία για την περιγραφή των εκφράσεων αυτών μέσω FAPs.

Στον **Πίνακα 4.2** αποτυπώνεται το πρώτο βήμα περιγραφής των βασικών εκφράσεων και μερικών παραλλαγών τους, με τη βοήθεια των παραμέτρων του MPEG-4 (FAPs). Ο **Πίνακας 4.2** απεικονίζει το σύνολο των FAPs (λεξιλόγιο) που χρησιμοποιείται για κάθε βασική έκφραση και όχι κάποιο συγκεκριμένο «προφίλ», οπότε πιθανή απεικόνιση δε θα οδηγούσε με βεβαιότητα στη συγκεκριμένη έκφραση. Ο ορισμός «προφίλ» για κάθε έκφραση, δηλαδή το σύνολο των FAPs, συνοδευόμενα από κατάλληλο διάστημα τιμών, των οποίων η συνδυασμένη απεικόνιση έχει ως αποτέλεσμα τη συγκεκριμένη έκφραση, δίνεται αναλυτικά στο Κεφάλαιο της Σύνθεσης Εκφράσεων Προσώπου (Κεφάλαιο 5).

Χαρά	<i>open_jaw</i> (F ₃), <i>lower_t_midlip</i> (F ₄), <i>raise_b_midlip</i> (F ₅), <i>stretch_l_cornerlip</i> (F ₆), <i>stretch_r_cornerlip</i> (F ₇), <i>raise_l_cornerlip</i> (F ₁₂), <i>raise_r_cornerlip</i> (F ₁₃), <i>close_t_l_eyelid</i> (F ₁₉), <i>close_t_r_eyelid</i> (F ₂₀), <i>close_b_l_eyelid</i> (F ₂₁), <i>close_b_r_eyelid</i> (F ₂₂), <i>raise_l_m_eyebrow</i> (F ₃₃), <i>raise_r_m_eyebrow</i> (F ₃₄), <i>lift_l_cheek</i> (F ₄₁), <i>lift_r_cheek</i> (F ₄₂), <i>stretch_l_cornerlip_o</i> (F ₅₃), <i>stretch_r_cornerlip_o</i> (F ₅₄)
Λύπη	<i>close_t_l_eyelid</i> (F ₁₉), <i>close_t_r_eyelid</i> (F ₂₀), <i>close_b_l_eyelid</i> (F ₂₁), <i>close_b_r_eyelid</i> (F ₂₂), <i>raise_l_i_eyebrow</i> (F ₃₁), <i>raise_r_i_eyebrow</i> (F ₃₂), <i>raise_l_m_eyebrow</i> (F ₃₃), <i>raise_r_m_eyebrow</i> (F ₃₄), <i>raise_l_o_eyebrow</i> (F ₃₅), <i>raise_r_o_eyebrow</i> (F ₃₆)
Θυμός	<i>lower_t_midlip</i> (F ₄), <i>raise_b_midlip</i> (F ₅), <i>push_b_lip</i> (F ₁₆), <i>depress_chin</i> (F ₁₈), <i>close_t_l_eyelid</i> (F ₁₉), <i>close_t_r_eyelid</i> (F ₂₀), <i>close_b_l_eyelid</i> (F ₂₁), <i>close_b_r_eyelid</i> (F ₂₂), <i>raise_l_i_eyebrow</i> (F ₃₁), <i>raise_r_i_eyebrow</i> (F ₃₂), <i>raise_l_m_eyebrow</i> (F ₃₃), <i>raise_r_m_eyebrow</i> (F ₃₄), <i>raise_l_o_eyebrow</i> (F ₃₅), <i>raise_r_o_eyebrow</i> (F ₃₆), <i>squeeze_l_eyebrow</i> (F ₃₇), <i>squeeze_r_eyebrow</i> (F ₃₈)
Φόβος	<i>open_jaw</i> (F ₃), <i>lower_t_midlip</i> (F ₄), <i>raise_b_midlip</i> (F ₅), <i>lower_t_lip_lm</i> (F ₈), <i>lower_t_lip_rm</i> (F ₉), <i>raise_b_lip_lm</i> (F ₁₀), <i>raise_b_lip_rm</i> (F ₁₁), <i>close_t_l_eyelid</i> (F ₁₉), <i>close_t_r_eyelid</i> (F ₂₀), <i>close_b_l_eyelid</i> (F ₂₁), <i>close_b_r_eyelid</i> (F ₂₂), <i>raise_l_i_eyebrow</i> (F ₃₁), <i>raise_r_i_eyebrow</i> (F ₃₂), <i>raise_l_m_eyebrow</i> (F ₃₃), <i>raise_r_m_eyebrow</i> (F ₃₄), <i>raise_l_o_eyebrow</i> (F ₃₅), <i>raise_r_o_eyebrow</i> (F ₃₆), <i>squeeze_l_eyebrow</i> (F ₃₇), <i>squeeze_r_eyebrow</i> (F ₃₈)
Απέχθεια	<i>open_jaw</i> (F ₃), <i>lower_t_midlip</i> (F ₄), <i>raise_b_midlip</i> (F ₅), <i>lower_t_lip_lm</i> (F ₈), <i>lower_t_lip_rm</i> (F ₉), <i>raise_b_lip_lm</i> (F ₁₀), <i>raise_b_lip_rm</i> (F ₁₁), <i>close_t_l_eyelid</i> (F ₁₉), <i>close_t_r_eyelid</i> (F ₂₀), <i>close_b_l_eyelid</i> (F ₂₁), <i>close_b_r_eyelid</i> (F ₂₂), <i>raise_l_m_eyebrow</i> (F ₃₃), <i>raise_r_m_eyebrow</i> (F ₃₄), <i>lower_t_lip_lm_o</i> (F ₅₅), <i>lower_t_lip_rm_o</i> (F ₅₆), <i>raise_b_lip_lm_o</i> (F ₅₇), <i>raise_b_lip_rm_o</i> (F ₅₈), <i>raise_l_cornerlip_o</i> (F ₅₉), <i>raise_r_cornerlip_o</i> (F ₆₀)
Έκπληξη	<i>open_jaw</i> (F ₃), <i>raise_b_midlip</i> (F ₅), <i>stretch_l_cornerlip</i> (F ₆), <i>stretch_r_cornerlip</i> (F ₇), <i>raise_b_lip_lm</i> (F ₁₀), <i>raise_b_lip_rm</i> (F ₁₁), <i>close_t_l_eyelid</i> (F ₁₉), <i>close_t_r_eyelid</i> (F ₂₀), <i>close_b_l_eyelid</i> (F ₂₁), <i>close_b_r_eyelid</i> (F ₂₂), <i>raise_l_i_eyebrow</i> (F ₃₁), <i>raise_r_i_eyebrow</i> (F ₃₂), <i>raise_l_m_eyebrow</i> (F ₃₃), <i>raise_r_m_eyebrow</i> (F ₃₄), <i>raise_l_o_eyebrow</i> (F ₃₅), <i>raise_r_o_eyebrow</i> (F ₃₆), <i>squeeze_l_eyebrow</i> (F ₃₇), <i>squeeze_r_eyebrow</i> (F ₃₈), <i>stretch_l_cornerlip_o</i> (F ₅₃), <i>stretch_r_cornerlip_o</i> (F ₅₄)

Πίνακας 4.2: Λεξιλόγιο FAPs για την περιγραφή των βασικών εκφράσεων

4.2.1. Περιοχή τιμών FAPs

Σημαντικό ρόλο τόσο στη σύνθεση όσο και στην ανάλυση των εκφράσεων έχει η περιοχή τιμών της κάθε παραμέτρου. Από τη μεριά της σύνθεσης υπάρχει σχετική μελέτη [35], η οποία αναφέρεται στον ορισμό της περιοχής τιμών των FAPs. Παρολ' αυτά οι συγκεκριμένες περιοχές τιμών είναι ιδιαίτερα ευρείες με αποτέλεσμα να μην μπορούν να χρησιμοποιηθούν στην περιοχή της ανάλυσης. Με σκοπό, λοιπόν την παραγωγή πιο περιορισμένων περιοχών τιμών αναλύθηκαν δύο γνωστές βάσεις δεδομένων, η στατική του Ekman [32] και η δυναμική του Medialab [41]. Τα δεδομένα και των δύο βάσεων απεικονίζουν ακραίες περιπτώσεις εκφράσεων, αλλά είναι ιδιαίτερα χρήσιμα στον καθορισμό ορίων των αντίστοιχων περιοχών τιμών [42]. Η διαδικασία όμως ορισμού τιμών έχει ως προϋπόθεση την «ποσοτικοποίηση» των FAPs μέσω της κίνησης των σημείων του προσώπου (facial points) και ως αποτέλεσμα την επαλήθευση και συμπλήρωση του λεξιλογίου των εκφράσεων καθώς και τον ορισμό «προφίλ» για κάθε βασική έκφραση.

Τα προφίλ, όπως αυτά δημιουργήθηκαν από τα δεδομένα της ανάλυσης, επαληθεύτηκαν με διαδικασία σύνθεσης και χρησιμοποιήθηκαν σε εφαρμογές ανάλυσης αλλά και σύνθεσης, παρατίθενται στο Κεφάλαιο της σύνθεσης (Κεφάλαιο 5).

4.3. Ορισμός αποστάσεων

4.3.1. Αναπαράσταση των FAPs μέσω κίνησης FPs (feature points)

Παρά το γεγονός ότι τα FAPs προσφέρονται για σχεδιασμό κίνησης, είναι ακατάλληλα για την ανάλυση εκφράσεων προσώπου από βίντεο ή μεμονωμένες εικόνες κυρίως λόγω της απουσίας σαφούς ποσοτικού ορισμού των περισσότερων FAPs, οπότε η μέτρηση τους προϋποθέτει την περιγραφή τους μέσω εντοπίσιμων σημείων της περιοχής του προσώπου.

Στον **Πίνακα 4.3** καταγράφεται η αναπαράσταση των FAPs με τη βοήθεια ευκλείδειων αποστάσεων μεταξύ χαρακτηριστικών σημείων του προσώπου. Η πρώτη στήλη περιλαμβάνει τις ονομασίες των FAPs, όπως αυτές έχουν οριστεί στα πλαίσια του προτύπου MPEG-4, στη δεύτερη στήλη δίνονται οι αντίστοιχες αποστάσεις $s(x,y)$, όπου τα x και y είναι σημεία του προσώπου (Εικόνα 3.2). Στην τρίτη στήλη ορίζεται η κανονικοποίηση για κάθε απόσταση, όπου $D_{i-NEUTRAL}$ είναι η απόσταση D_i όταν το πρόσωπο έχει «ουδέτερη» έκφραση, και στην τέταρτη αναφέρονται οι μονάδες μέτρησης.

Από τον πίνακα φαίνεται ότι η αντιστοιχία FAPs – αποστάσεων δεν είναι πάντα ένα προς ένα (π.χ. features $f_{12}f_{15}$), ενώ ταυτόχρονα δεν περιλαμβάνεται το σύνολο των FAPs.

Όνομα FAP	Ορισμός απόστασης	Ορισμός χαρακτηριστικών	Μονάδες μέτρησης
<i>squeeze_l_eyebrow</i> (F_{37})	$D_1=s(4.5,3.11)$	$f_{1=} D_{1-NEUTRAL} - D_1$	ES
<i>squeeze_r_eyebrow</i> (F_{38})	$D_2=s(4.6,3.8)$	$f_{2=} D_{2-NEUTRAL} - D_2$	ES
<i>lower_t_midlip</i> (F_4)	$D_3=s(9.3,8.1)$	$f_{3=} D_3 - D_{3-NEUTRAL}$	MNS
<i>raise_b_midlip</i> (F_5)	$D_4=s(9.3,8.2)$	$f_{4=} D_{4-NEUTRAL} - D_4$	MNS
<i>raise_l_l_eyebrow</i> (F_{31})	$D_5=s(4.1,3.11)$	$f_{5=} D_5 - D_{5-NEUTRAL}$	ENS
<i>raise_r_l_eyebrow</i> (F_{32})	$D_6=s(4.2,3.8)$	$f_{6=} D_6 - D_{6-NEUTRAL}$	ENS
<i>raise_l_o_eyebrow</i> (F_{35})	$D_7=s(4.5,3.7)$	$f_{7=} D_7 - D_{7-NEUTRAL}$	ENS
<i>raise_r_o_eyebrow</i> (F_{36})	$D_8=s(4.6,3.12)$	$f_{8=} D_8 - D_{8-NEUTRAL}$	ENS
<i>raise_l_m_eyebrow</i> (F_{33})	$D_9=s(4.3,3.7)$	$f_{9=} D_9 - D_{9-NEUTRAL}$	ENS
<i>raise_r_m_eyebrow</i> (F_{34})	$D_{10}=s(4.4,3.12)$	$f_{10=} D_{10} - D_{10-NEUTRAL}$	ENS
<i>open_jaw</i> (F_3)	$D_{11}=s(8.1,8.2)$	$f_{11=} D_{11} - D_{11-NEUTRAL}$	MNS
<i>close_t_l_eyelid</i> (F_{19}) – <i>close_b_l_eyelid</i> (F_{21})	$D_{12}=s(3.1,3.3)$	$f_{12=} D_{12} - D_{12-NEUTRAL}$	IRISD
<i>close_t_r_eyelid</i> (F_{20}) – <i>close_b_r_eyelid</i> (F_{22})	$D_{13}=s(3.2,3.4)$	$f_{13=} D_{13} - D_{13-NEUTRAL}$	IRISD

$stretch_l_cornerlip$ (F_6) $(stretch_l_cornerlip_0)(F_{53}) -$ $stretch_r_cornerlip$ (F_7) $(stretch_r_cornerlip_0)(F_{54})$	$D_{14}=s(8.4,8.3)$	$f_{14}= D_{14} - D_{14-NEUTRAL}$	MW
$squeeze_l_eyebrow$ (F_{37}) AND $squeeze_r_eyebrow$ (F_{38})	$D_{15}=s(4.6,4.5)$	$f_{15}= D_{15-NEUTRAL} - D_{15}$	ES

Πίνακας 4.3: Ποσοτικοποίηση των FAPs

4.3.2. Αποστάσεις και αλληλεπίδραση μεταξύ τους

Ο Πίνακας 4.3 της Ενότητας 4.3.1 παρέχει μία πρώτη αντιστοίχιση FAPs και αποστάσεων. Παρόλο που η αντιστοίχιση αυτή ήδη παρέχει έναν ικανοποιητικό τρόπο υπολογισμού των FAPs με τη βοήθεια κάποιων χαρακτηριστικών σημείων του προσώπου, θα ήταν πιο παραστατική αν λάμβανε υπόψιν τις άλλες μεταβολές που λαμβάνουν χώρα ακούσια στο πρόσωπο ακολουθώντας την κίνηση των μυών. Στην κατεύθυνση αυτή γίνεται μία προσπάθεια συνδυασμού των διαφόρων αποστάσεων με διαφορετικά βέβαια βάρη η κάθε μία, για τον υπολογισμό κάθε FAP. Για κάθε FAP ορίζεται μία απόσταση η οποία είναι η βασική και κάποιες άλλες, με δικό της βάρος η κάθε μία, που παίρνουν, κατά μεγαλύτερο ή μικρότερο βαθμό, μέρος στην κίνηση που περιγράφεται από το εξεταζόμενο FAP (Πίνακας 4.4).

FAP	Πρωτεύουσα Απόσταση	Άλλες Αποστάσεις	Καταστάσεις	Σχόλια
$squeeze_l_eyebrow$ (F_{37})	D1	D5, D7, D9, D15	low, medium, high	
$squeeze_r_eyebrow$ (F_{38})	D2	D6, D8, D10, D15	low, medium, high	
$lower_t_midlip$ (F_4)	D3	D11	low, medium	το άνω χείλος δεν έχει πολλά περιθώρια κίνησης, οπότε η κατάσταση 'high' αγνοήθηκε
$raise_b_midlip$ (F_5)	D4	D11	very low, low, high	μία κατάσταση medium είναι παρόμοια με τη neutral
$raise_l_i_eyebrow$ (F_{31})	D5	D1, D7, D9, D15	low, medium, high	
$raise_r_i_eyebrow$ (F_{32})	D6	D2, D8, D10,	low, medium,	

		D15	high	
<i>raise_l_o_eyebrow</i> (F_{35})	D7	D1, D5, D9, D15	low, medium, high	
<i>raise_r_o_eyebrow</i> (F_{36})	D8	D2, D6, D10, D15	low, medium, high	
<i>raise_l_m_eyebrow</i> (F_{33})	D9	D1, D5, D7, D15	low, medium, high	
<i>raise_r_m_eyebrow</i> (F_{34})	D10	D2, D6, D8, D15	low, medium, high	
<i>open_jaw</i> (F_3)	D11	D4	low, medium, high	
<i>close_t_l_eyelid</i> (F_{19}) – <i>close_b_l_eyelid</i> (F_{21})	D12	-	low, high	η κατάσταση medium ενός κλειστού ματιού δεν μπορεί να ανιχνευθεί
<i>close_t_r_eyelid</i> (F_{20}) – <i>close_b_r_eyelid</i> (F_{22})	D13	-	low, high	η κατάσταση medium ενός κλειστού ματιού δεν μπορεί να ανιχνευθεί
<i>stretch_l_cornerlip</i> (F_6) (<i>stretch_l_cornerlip_o</i>) (F_{53}) – <i>stretch_r_cornerlip</i> (F_7) (<i>stretch_r_cornerlip_o</i>) (F_{54})	D14	D3, D4, D11, D16, D17		
<i>squeeze_l_eyebrow</i> (F_{37}) AND <i>squeeze_r_eyebrow</i> (F_{38}) <i>wrinkles_between_eyebrows</i>	D15	D1, D2, D5, D6, D7, D8, D10	low, medium, high	
<i>raise_left_outer_cornerlip</i>	D16	D14	low, medium, high	
<i>raise_right_outer_cornerlip</i>	D17	D14	low, medium, high	

Πίνακας 4.4: FAPs και αποστάσεις, πρωτεύουσες και μη

4.4. Επαλήθευση των εκφράσεων

Από την ανάλυση των εικόνων των βάσεων του Ekman και του Medialab, με τη βοήθεια των αποστάσεων που ορίστηκαν προηγουμένως, προέκυψαν στατιστικά δεδομένα που περιέχονται στον **Πίνακα 4.5**. Η μέση τιμή αποτελεί την τυπική τιμή που μπορεί να χρησιμοποιηθεί στην απεικόνιση μιας συγκεκριμένης έκφρασης, ενώ η τυπική απόκλιση χρησιμοποιείται για πιθανές παραλλαγές της.

Όνομα FAP	Κατάσταση	Θυμός	Λύπη	Χαρά	Απέχθεια	Φόβος	Έκπληξη
open_jaw (F ₃)	Μέση τιμή	*	*	-	-	291	885
	Τυπική απόκλιση	*	*	-	-	189	316
lower_t_midlip (F ₄)	Μέση τιμή	73	*	-271	-234	-	*
	Τυπική απόκλιση	51	*	110	109	-	*
raise_b_midlip (F ₅)	Μέση τιμή	*	*	-	-177	218	543
	Τυπική απόκλιση	*	*	-	108	135	203
stretch_l_cornerlip (F ₆), stretch_l_cornerlip_o (F ₅₃), stretch_r_cornerlip (F ₇), stretch_r_cornerlip_o (F ₅₄)	Μέση τιμή	*	*	234	*	*	-82
	Τυπική απόκλιση	*	*	98	*	*	39
lower_t_lip_lm (F ₈)	Μέση τιμή	*	*	*	-	*	*
	Τυπική απόκλιση	*	*	*	-	*	*
lower_t_lip_rm (F ₉)	Μέση τιμή	*	*	*	-	*	*
	Τυπική απόκλιση	*	*	*	-	*	*
raise_b_lip_lm (F ₁₀)	Μέση τιμή	*	*	*	-	*	*
	Τυπική απόκλιση	*	*	*	-	*	*
raise_b_lip_rm (F ₁₁)	Μέση τιμή	*	*	*	-	*	*
	Τυπική απόκλιση	*	*	*	-	*	*
close_t_l_eyelid (F ₁₉), close_b_l_eyelid (F ₂₁)	Μέση τιμή	-	-153	-254	-203	244	254
	Τυπική απόκλιση	-	112	133	148	126	83
close_t_r_eyelid (F ₂₀), close_b_r_eyelid (F ₂₂)	Μέση τιμή	-	-161	-242	-211	249	252
	Τυπική απόκλιση		109	122	145	128	81
raise_l_i_eyebrow (F ₃₁)	Μέση τιμή	-83	85	*	*	104	224
	Τυπική απόκλιση	48	55	*	*	69	103
raise_r_i_eyebrow (F ₃₂)	Μέση τιμή	-85	80	*	*	111	211
	Τυπική απόκλιση	51	54	*	*	72	97
raise_l_m_eyebrow (F ₃₃)	Μέση τιμή	-149	-	24	-80	72	144
	Τυπική απόκλιση	40	-	22	53	58	64
raise_r_m_eyebrow (F ₃₄)	Μέση τιμή	-144	-	25	-82	75	142
	Τυπική απόκλιση	39	-	22	54	60	62
raise_l_o_eyebrow (F ₃₅)	Μέση τιμή	-66	-	*	*	-	54
	Τυπική απόκλιση	35	-	*	*	-	31
raise_r_o_eyebrow (F ₃₆)	Μέση τιμή	-70	-	*	*	-	55

	Τυπική απόκλιση	38		*	*	-	31
squeeze_l_eyebrow (F ₃₇)	Μέση τιμή	57	*	*	*	-	-
	Τυπική απόκλιση	28	*	*	*	-	-
squeeze_r_eyebrow (F ₃₈)	Μέση τιμή	58	*	*	*	-	-
	Τυπική απόκλιση	31	*	*	*	-	-
lift_l_cheek (F ₄₁)	Μέση τιμή	*	*	-	*	*	*
	Τυπική απόκλιση	*	*	-	*	*	*
lift_r_cheek (F ₄₂)	Μέση τιμή	*	*	-	*	*	*
	Τυπική απόκλιση	*	*	-	*	*	*
stretch_l_cornerlip_o (F ₅₃)	Μέση τιμή	*	*	-	*	*	-
	Τυπική απόκλιση	*	*	-	*	*	-
stretch_r_cornerlip_o (F ₅₄)	Μέση τιμή	*	*	*	*	*	-
	Τυπική απόκλιση	*	*	*	*	*	
lower_t_lip_lm_o (F ₅₅)	Μέση τιμή	*	*	*	-	*	*
	Τυπική απόκλιση	*	*	*	-	*	*
lower_t_lip_rm_o (F ₅₆)	Μέση τιμή	*	*	*	-	*	*
	Τυπική απόκλιση	*	*	*	-	*	*
raise_b_lip_lm_o (F ₅₇)	Μέση τιμή	*	*	*	-	*	*
	Τυπική απόκλιση	*	*	*	-	*	*
raise_b_lip_rm_o (F ₅₈)	Μέση τιμή	*	*	*	-	*	*
	Τυπική απόκλιση	*	*	*	-	*	*
raise_l_cornerlip_o (F ₅₉)	Μέση τιμή	*	*	*	-	*	*
	Τυπική απόκλιση	*	*	*	-	*	*
raise_r_cornerlip_o (F ₆₀)	Μέση τιμή	*	*	*	-	*	*
	Τυπική απόκλιση	*	*	*	-	*	*

Πίνακας 4.5: Στατιστικά δεδομένα για τιμές λεξιλογίου FAPs για τις βασικές εκφράσεις

Οι μονάδες μέτρησης που χρησιμοποιούνται είναι οι ορισμένες από το πρότυπο όπως αυτές παρουσιάστηκαν στην **Ενότητα 3.2** [29]. Το σύμβολο (*) υποδηλώνει απουσία του αντίστοιχου FAP στο «λεξιλόγιο» της συγκεκριμένης έκφρασης, ενώ το σύμβολο (-) υποδηλώνει ότι παρόλο που το αντίστοιχο FAP περιλαμβάνεται στο «λεξιλόγιο», η τιμή του δεν έχει επαληθευθεί μέσω της στατιστικής ανάλυσης. Ο εντοπισμός των σημείων του προσώπου που συμμετέχουν στην «ποσοτικοποίηση» των FAPs βασίστηκε στην εργασία [46], με χειρωνακτική υποβοήθηση όπου κρινόταν απαραίτητο.

4.5. Σύστημα Ανάλυσης Εκφράσεων Προσώπου

4.5.1. Επέκταση «προφίλ» για τη δημιουργία κανόνων για είσοδο σε νευρωνικό σύστημα

Στο πλαίσιο του συστήματος Ανάλυσης Εκφράσεων Προσώπου, τα προφίλ που δημιουργήθηκαν για τη σύνθεση (Κεφάλαιο 5), μετατράπηκαν σε κανόνες με μορφή κατάλληλη για είσοδο σε νευρωνικό. Στον **Πίνακα 4.6** καταγράφεται το σύνολο των κανόνων αυτών στη μορφή που χρησιμοποιήθηκε στο νευρωνικό του συστήματος Ανάλυσης. Οι κανόνες περιέχουν τα FAPs στη μορφή FI_J, όπου ο δείκτης I είναι ο αύξων αριθμός του κάθε FAP σύμφωνα με τον ορισμό του MPEG-4 και ο δείκτης J δίνει πληροφορία σχετικά με το διάστημα τιμών του αντίστοιχου FAP στο συγκεκριμένο κανόνα και παίρνει τιμές **High**, **Medium**, **Low** και σπανιότερα **VeryLow**. Το αποτέλεσμα σε κάθε κανόνα, το οποίο άλλωστε αποτελεί και την έξοδο του νευρωνικού, είναι το τεταρτημόριο του Τροχού Συναισθημάτων της Whissel, στο οποίο ανήκει κάθε φορά η προς εξέταση έκφραση.

Rule	FAP	Quadrant
1	F3_H+F4_L+F5_VL+[F53+F54]_H+[F19+F21]_H+[F20+F22]_H	(+,+)
2	F3_M+F4_L+F5_L+[F53+F54]_H+[F19+F21]_H+[F20+F22]_H	(+,+)
3	F3_M+F4_L+F5_H+[F53+F54]_H+[F19+F21]_H+[F20+F22]_H	(+,+)
4	F3_H+F4_L+F5_L+[F53+F54]_H+[F19+F21]_H+[F20+F22]_H	(+,+)
5	F3_L+F4_M+F5_H+[F53+F54]_M+[F19+F21]_H+[F20+F22]_H +F31_M+F32_M+F33_M+F34_M+F37_M+F38_M+F59_H+F60_H	(+,+)
6	F3_H+F4_L+F5_VL+[F53+F54]_M+[F19+F21]_H+[F20+F22]_H	(+,+)
7	F3_L+F4_L+F5_H+[F53+F54]_H+[F19+F21]_H+[F20+F22]_H +[F37+F38]_M+F59_H+F60_H	(+,+)
8	F3_H+F5_VL+[F53+F54]_L+[F19+F21]_L+[F20+F22]_L+F31_H +F32_H+F33_H+F34_H+F35_H+F36_H+F37_L+F38_L+[F37 +F38]_L	(+,+)
9	F3_H+F5_VL+[F53+F54]_M+[F19+F21]_L+[F20+F22]_L+F31_H +F32_H+F33_H+F34_H+F35_H+F36_H+F37_L+F38_L	(+,+)
10	F3_M+F5_L+[F53+F54]_L+[F19+F21]_L+[F20+F22]_L+F31_H +F32_H+F33_H+F34_H+F35_H+F36_H	(+,+)
11	F3_H+F5_VL+[F53+F54]_L+[F19+F21]_L+[F20+F22]_L+F31_H +F32_H+F33_H+F34_H+F35_M+F36_M	(+,+)
12	F3_H+F5_VL+[F53+F54]_L+[F19+F21]_L+[F20+F22]_L+F31_M +F32_M+F33_H+F34_H+F35_M+F36_M	(+,+)
13	F3_L+F4_M+F5_H+F31_L+F32_L+F33_L+F34_L+F37_H+F38_H +[F37+F38]_H+F59_M+F60_M	(-,+)
14	F3_L+F4_M+F5_L+F31_L+F32_L+F33_L+F34_L+F37_M+F38_M +[F37+F38]_H	(-,+)
15	F3_L+F4_M+F5_H+F31_M+F32_M+F33_L+F34_L+F37_H+F38_H	(-,+)

	8_H+[F37+F38]_H	
16	F3_L+F4_M+F5_L+F31_L+F32_L+F33_L+F34_L+F37_H+F38_H+[F37+F38]_H+F59_M+F60_M	(-,+)
17	F3_H+F4_L+F5_VL+[F53+F54]_L+F31_M+F32_M+F33_L+F34_L+F35_L+F36_L+F37_H+F38_H+[F37+F38]_H+F59_L+F60_L	(-,+)
18	F3_H+F4_M+F5_VL+[F53+F54]_L+F31_L+F32_L+F33_L+F34_L+F35_L+F36_L+F37_H+F38_H+[F37+F38]_H	(-,+)
19	F3_H+F4_M+F5_VL+[F53+F54]_L+F31_L+F32_L+F33_L+F34_L+[F37+F38]_H+F59_L+F60_L	(-,+)
20	F3_H+F4_L+F5_VL+[F53+F54]_L+F31_M+F32_M+F33_M+F34_M+F37_H+F38_H+[F37+F38]_H+F59_M+F60_M	(-,+)
21	F3_M+F4_L+F5_L+[F53+F54]_L+F31_L+F32_L+F33_L+F34_L+F37_H+F38_H+[F37+F38]_H+F59_M+F60_M	(-,+)
22	F3_L+F4_M+F5_H+F31_M+F32_M+F33_L+F34_L+F37_H+F38_H+[F37+F38]_M	(-,+)
23	F3_L+[F53+F54]_M+F5_H+[F19+F21]_L+[F20+F22]_L+F31_M+F32_M+F33_M+F34_M	(-,+)
24	F3_M+[F19+F21]_L+[F20+F22]_L+F31_H+F32_H+F37_M+F38_M	(-,+)
25	F3_M+F4_M+F5_H+[F19+F21]_L+[F20+F22]_L+F33_M+F34_M+F35_H+F36_H	(-,+)
26	F3_M+F5_L+[F19+F21]_H+[F20+F22]_H+F31_H+F32_H+F33_M+F34_M+F35_M+F36_M+[F37+F38]_H	(-,-)
27	F3_M+F4_L+F5_L+[F53+F54]_L+[F19+F21]_H+[F20+F22]_H+F31_L+F32_L+F33_L+F34_L+[F37+F38]_M	(-,-)
28	F3_M+F4_L+F5_H+[F53+F54]_L+[F19+F21]_H+[F20+F22]_H+F31_L+F32_L+F33_L+F34_L+F35_L+F36_L+F37_M+F38_M+[F37+F38]_L	(-,-)
29	F3_L+F4_L+F5_H+[F53+F54]_L+[F19+F21]_H+[F20+F22]_H+F31_M+F32_M+F33_M+F34_M+[F37+F38]_M	(-,-)
30	F3_L+F4_L+F5_H+[F53+F54]_L+[F19+F21]_H+[F20+F22]_H+F31_L+F32_L+F33_L+F34_L+[F37+F38]_H	(-,-)
31	F3_L+F4_L+F5_H+[F53+F54]_M+[F19+F21]_H+[F20+F22]_H+F31_L+F32_L+F33_L+F34_L+[F37+F38]_M	(-,-)
32	F3_L+F4_M+F5_H+[F53+F54]_L+[F19+F21]_H+[F20+F22]_H+F31_L+F32_L+F33_L+F34_L+[F37+F38]_H+F59_H	(-,-)
33	F3_M+F4_L+F5_H+[F53+F54]_M+[F19+F21]_H+[F20+F22]_H+F31_M+F32_M+F33_M+F34_M+F35_M+F36_M+[F37+F38]_H+F60_H	(-,-)
34	F3_L+F4_L+[F53+F54]_L+F31_M+F32_M+F33_M+F34_M+F35_M+F36_M+F37_M+F38_M+[F37+F38]_M	(-,-)
35	F3_H+F4_L+F5_VL+[F53+F54]_L+[F19+F21]_H+[F20+F22]_H+F31_L+F32_L+F33_L+F34_L+F35_L+F36_L+F37_H+F38_H+[F37+F38]_H+F59_H	(-,-)
36	F3_L+F4_M+F5_H+[F53+F54]_L+[F19+F21]_H+[F20+F22]_H+F31_L+F32_L+F33_L+F34_L+[F37+F38]_H+F60_H	(-,-)
37	F3_M+F4_L+F5_H+[F53+F54]_M+[F19+F21]_H+[F20+F22]_H	(-,-)

	+F31_M+F32_M+F33_M+F34_M+F35_M+F36_M+[F37+F38]_H+F59_H	
38	F3_L+F4_L+F5_H+[F53+F54]_M+[F19+F21]_H+[F20+F22]_H+F31_L+F32_L+F33_L+F34_L+ [F37+F38]_H+F59_H+F60_H	(-, -)
39	F3_L+F4_L+F5_H+[F53+F54]_L+[F19+F21]_H+[F20+F22]_H+F31_M+F32_M+F33_M+F34_M+ [F37+F38]_H+F59_H+F60_H	(-, -)
40	F3_H+F4_L+F5_VL+[F53+F54]_L+[F19+F21]_H+[F20+F22]_H+F31_L+F32_L+F33_L+F34_L+F35_L+F36_L+F37_H+F38_H+[F37+F38]_H+F60_H	(-, -)
41	F3_L+F4_M+F31_M+F32_M+F33_M+F34_M+F35_M+F36_M+F37_M+F38_M+[F37+F38]_M+ F59_M+F60_M	Neutral

Πίνακας 4.6: Κανόνες βασιζόμενοι στα «προφίλ» συναισθημάτων για είσοδο σε νευρωνικό σύστημα

Επίσης πρέπει να σημειωθεί ότι στο πλαίσιο της διαμόρφωσης κανόνων για το σύστημα ανάλυσης εκφράσεων, δημιουργήθηκε η ανάγκη, η οποία και ικανοποιήθηκε, ορισμού κανόνων, οι οποίοι δε θα λάμβαναν υπόψιν το στόμα, για τις περιπτώσεις κατά τις οποίες ο προς εξέταση άνθρωπος μιλάει και η κίνηση του στόματος δεν παίζει ρόλο στην έκφραση του συναισθήματος.

4.5.2. Interface Συστήματος Ανάλυσης Εκφράσεων Προσώπου

Το σύστημα Ανάλυσης Εκφράσεων Προσώπου [54] δημιουργήθηκε με τη βοήθεια της Matlab και χρησιμοποίησε ως είσοδο τους κανόνες του **Πίνακα 4.6** και τις αποστάσεις, όπως αυτές ορίστηκαν στον **Πίνακα 4.3**. Η αρχική μορφή των κανόνων έδινε ως αποτέλεσμα συναισθημα, όμως σε επόμενη έκδοση αυτών και μετά από παραίνεση ειδικών, οι κανόνες αυτοί προσαρμόστηκαν, ώστε να έχουν ως αποτέλεσμα κάποιο από τα τεταρτημόρια (ή την αρχή των αξόνων για neutral) του Τροχού Συναισθημάτων της Whissel. Ένα στιγμιότυπο του συστήματος φαίνεται στην **Εικόνα 4.1**.



Εικόνα 4.1: Στιγμιότυπο Συστήματος Ανάλυσης Εκφράσεων

Κεφάλαιο 5

Σύνθεση Εκφράσεων Προσώπου

5.1. Δημιουργία «προφίλ» βασικών εκφράσεων

Όπως διατυπώθηκε ήδη στο κεφάλαιο της Ανάλυσης, ως «προφίλ» έχουμε ορίσει ένα σύνολο από FAPs, υποσύνολο του «λεξιλογίου» της συγκεκριμένης έκφρασης (Πίνακας 4.2), συνοδευόμενο από τις αντίστοιχες περιοχές τιμών, το οποίο, απεικονιζόμενο, οδηγεί στην αντίστοιχη έκφραση. Στη δημιουργία των «προφίλ» ακολουθούμε διαδικασία τριών βημάτων:

1. ορισμός συνόλου FAPs που συμμετέχουν στη διαμόρφωση μίας βασικής έκφρασης, μετασχηματίζοντας σε FAPs την, προτεινόμενη από ψυχολογικές μελέτες, μορφοποίηση του προσώπου.
2. χρήση των αντίστοιχων περιοχών τιμών του Πίνακα 4.5.
3. απεικόνιση των «προφίλ» με σκοπό την επαλήθευση της καταλληλότητας των απεικονίσεων που προκύπτουν.

Η περιοχή τιμών για τα FAPs υπολογίστηκε ως εξής:

Έστω m_{ij} και σ_{ij} η μέση τιμή και η τυπική απόκλιση αντίστοιχα του FAP F_j για τη βασική έκφραση i (όπου $i = \{1 \Rightarrow \text{Θυμός}, 2 \Rightarrow \text{Λύπη}, 3 \Rightarrow \text{Χαρά}, 4 \Rightarrow \text{Απέχθεια}, 5 \Rightarrow \text{Φόβος}, 6 \Rightarrow \text{Εκπληξη}\}$), όπως αυτές υπολογίστηκαν στον Πίνακα 4.5. Το διάστημα τιμών X_{ij} του FAP F_j για τη βασική έκφραση i ορίζεται ως εξής:

$$X_{ij} = [m_{ij} - \sigma_{ij}, m_{ij} + \sigma_{ij}] \quad (1)$$

για τα FAPs των δύο κατευθύνσεων και

$$X_{ij} = [\max(0, m_{ij} - \sigma_{ij}), m_{ij} + \sigma_{ij}] \text{ ή } X_{ij} = [m_{ij} - \sigma_{ij}, \min(0, m_{ij} + \sigma_{ij})] \quad (2)$$

για της μίας κατεύθυνσης [29].

Ο Πίνακας 5.1 περιέχει κάποιες εκδοχές «προφίλ» βασικών εκφράσεων, τα οποία δημιουργήθηκαν με τη μέθοδο που περιγράφηκε.

Προφίλ Εκφράσεων	FAPs και Διαστήματα Τιμών
Θυμός	$F_4 \in [22, 124]$, $F_{31} \in [-131, -25]$, $F_{32} \in [-136, -34]$, $F_{33} \in [-189, -109]$, $F_{34} \in [-183, -105]$,

$(P_{\Theta}^{(0)})$	$F_{35} \in [-101,-31], F_{36} \in [-108,-32], F_{37} \in [29,85], F_{38} \in [27,89]$
$P_{\Theta}^{(1)}$	$F_{19} \in [-330,-200], F_{20} \in [-335,-205], F_{21} \in [200,330], F_{22} \in [205,335], F_{31} \in [-200,-80], F_{32} \in [-194,-74], F_{33} \in [-190,-70], F_{34} \in [-190,-70]$
$P_{\Theta}^{(2)}$	$F_{19} \in [-330,-200], F_{20} \in [-335,-205], F_{21} \in [200,330], F_{22} \in [205,335], F_{31} \in [-200,-80], F_{32} \in [-194,-74], F_{33} \in [70,190], F_{34} \in [70,190]$
$P_{\Theta}^{(3)}$	$F_{16} \in [45,155], F_{18} \in [45,155], F_{19} \in [-330,-200], F_{20} \in [-330,-200], F_{31} \in [-200,-80], F_{32} \in [-194,-74], F_{33} \in [-190,-70], F_{34} \in [-190,-70], F_{37} \in [65,135], F_{38} \in [65,135]$
$P_{\Theta}^{(4)}$	$F_{16} \in [-355,-245], F_{18} \in [145,255], F_{19} \in [-330,-200], F_{20} \in [-330,-200], F_{31} \in [-200,-80], F_{32} \in [-194,-74], F_{33} \in [-190,-70], F_{34} \in [-190,-70], F_{37} \in [65,135], F_{38} \in [65,135]$
Δύπη $(P_{\Lambda}^{(0)})$	$F_{19} \in [-265,-41], F_{20} \in [-270,-52], F_{21} \in [-265,-41], F_{22} \in [-270,-52], F_{31} \in [30,140], F_{32} \in [26,134]$
Χαρά $(P_X^{(0)})$	$F_4 \in [-381,-161], F_6 \in [136,332], F_7 \in [136,332], F_{19} \in [-387,-121], F_{20} \in [-364,-120], F_{21} \in [-387,-121], F_{22} \in [-364,-120], F_{33} \in [2,46], F_{34} \in [3,47], F_{53} \in [136,332], F_{54} \in [136,332]$
$P_X^{(1)}$	$F_6 \in [160,240], F_7 \in [160,240], F_{12} \in [260,340], F_{13} \in [260,340], F_{19} \in [-449,-325], F_{20} \in [-426,-302], F_{21} \in [325,449], F_{22} \in [302,426], F_{33} \in [70,130], F_{34} \in [70,130], F_{41} \in [130,170], F_{42} \in [130,170], F_{53} \in [160,240], F_{54} \in [160,240]$
$P_X^{(2)}$	$F_6 \in [160,240], F_7 \in [160,240], F_{12} \in [260,340], F_{13} \in [260,340], F_{19} \in [-449,-325], F_{20} \in [-426,-302], F_{21} \in [-312,-188], F_{22} \in [-289,-165], F_{33} \in [70,130], F_{34} \in [70,130], F_{41} \in [130,170], F_{42} \in [130,170], F_{53} \in [160,240], F_{54} \in [160,240]$
$P_X^{(3)}$	$F_6 \in [160,240], F_7 \in [160,240], F_{12} \in [260,340], F_{13} \in [260,340], F_{19} \in [-449,-325], F_{20} \in [-426,-302], F_{21} \in [61,185], F_{22} \in [38,162], F_{33} \in [70,130], F_{34} \in [70,130], F_{41} \in [130,170], F_{42} \in [130,170], F_{53} \in [160,240], F_{54} \in [160,240]$
Απέχθεια $(P_A^{(0)})$	$F_4 \in [-343,-125], F_5 \in [-285,-69], F_{19} \in [-351,-55], F_{20} \in [-356,-66], F_{21} \in [-351,-55], F_{22} \in [-356,-66], F_{33} \in [-123,-27], F_{34} \in [-126,-28]$
Φόβος $(P_{\Phi}^{(0)})$	$F_3 \in [102,480], F_5 \in [83,353], F_{19} \in [118,370], F_{20} \in [121,377], F_{21} \in [118,370], F_{22} \in [121,377], F_{31} \in [35,173], F_{32} \in [39,183], F_{33} \in [14,130], F_{34} \in [15,135]$
$P_{\Phi}^{(1)}$	$F_3 \in [400,560], F_5 \in [333,373], F_{19} \in [-400,-340], F_{20} \in [-407,-347], F_{21} \in [-400,-340], F_{22} \in [-407,-347]$
$P_{\Phi}^{(2)}$	$F_3 \in [400,560], F_5 \in [307,399], F_{19} \in [-530,-470], F_{20} \in [-523,-463], F_{21} \in [-530,-470], F_{22} \in [-523,-463], F_{31} \in [460,540], F_{32} \in [460,540], F_{33} \in [460,540], F_{34} \in [460,540], F_{35} \in [460,540], F_{36} \in [460,540]$
$P_{\Phi}^{(3)}$	$F_3 \in [400,560], F_5 \in [-240,-160], F_{19} \in [-630,-570], F_{20} \in [-630,-570], F_{21} \in [-630,-570], F_{22} \in [-630,-570], F_{31} \in [460,540], F_{32} \in [460,540], F_{37} \in [60,140], F_{38} \in [60,140]$
$P_{\Phi}^{(4)}$	$F_3 \in [400,560], F_5 \in [-240,-160], F_{19} \in [-630,-570], F_{20} \in [-630,-570], F_{21} \in [-630,-570], F_{22} \in [-630,-570], F_{31} \in [460,540], F_{32} \in [460,540], F_{33} \in [360,440], F_{34} \in [360,440], F_{35} \in [260,340], F_{36} \in [260,340], F_{37} \in [60,140], F_{38} \in [60,140]$
$P_{\Phi}^{(5)}$	$F_3 \in [400,560], F_5 \in [-240,-160], F_{19} \in [-630,-570], F_{20} \in [-630,-570], F_{21} \in [-630,-570], F_{22} \in [-630,-570], F_{31} \in [460,540], F_{32} \in [460,540], F_{33} \in [360,440], F_{34} \in [360,440], F_{35} \in [260,340], F_{36} \in [260,340], F_{37} \in 0, F_{38} \in 0$
$P_{\Phi}^{(6)}$	$F_3 \in [400,560], F_5 \in [-240,-160], F_8 \in [-120,-80], F_9 \in [-120,-80], F_{10} \in [-120,-80], F_{11} \in [-120,-80], F_{19} \in [-630,-570], F_{20} \in [-630,-570], F_{21} \in [-630,-570], F_{22} \in [-630,-570],$

	$F_{31} \in [460,540], F_{32} \in [460,540], F_{33} \in [360,440], F_{34} \in [360,440], F_{35} \in [260,340], F_{36} \in [260,340], F_{37} \in 0, F_{38} \in 0$
$P_{\Phi}^{(7)}$	$F_3 \in [400,560], F_5 \in [-240,-160], F_{19} \in [-630,-570], F_{20} \in [-630,-570], F_{21} \in [-630,-570], F_{22} \in [-630,-570], F_{31} \in [360,440], F_{32} \in [360,440], F_{33} \in [260,340], F_{34} \in [260,340], F_{35} \in [160,240], F_{36} \in [160,240]$
$P_{\Phi}^{(8)}$	$F_3 \in [400,560], F_5 \in [-240,-160], F_{19} \in [-630,-570], F_{20} \in [-630,-570], F_{21} \in [-630,-570], F_{22} \in [-630,-570], F_{31} \in [260,340], F_{32} \in [260,340], F_{33} \in [160,240], F_{34} \in [160,240], F_{35} \in [60,140], F_{36} \in [60,140]$
$P_{\Phi}^{(9)}$	$F_3 \in [400,560], F_5 \in [307,399], F_{19} \in [-630,-570], F_{20} \in [-623,-563], F_{21} \in [-630,-570], F_{22} \in [-623,-563], F_{31} \in [460,540], F_{32} \in [460,540], F_{33} \in [460,540], F_{34} \in [460,540], F_{35} \in [460,540], F_{36} \in [460,540]$
Έκπληξη ($P_E^{(0)}$)	$F_3 \in [569,1201], F_5 \in [340,746], F_6 \in [-121,-43], F_7 \in [-121,-43], F_{19} \in [170,337], F_{20} \in [171,333], F_{21} \in [170,337], F_{22} \in [171,333], F_{31} \in [121,327], F_{32} \in [114,308], F_{33} \in [80,208], F_{34} \in [80,204], F_{35} \in [23,85], F_{36} \in [23,85], F_{53} \in [-121,-43], F_{54} \in [-121,-43]$
$P_E^{(1)}$	$F_3 \in [1150,1252], F_5 \in [-792,-700], F_6 \in [-141,-101], F_7 \in [-141,-101], F_{10} \in [-530,-470], F_{11} \in [-530,-470], F_{19} \in [-350,-324], F_{20} \in [-346,-320], F_{21} \in [-350,-324], F_{22} \in [-346,-320], F_{31} \in [314,340], F_{32} \in [295,321], F_{33} \in [195,221], F_{34} \in [191,217], F_{35} \in [72,98], F_{36} \in [73,99], F_{53} \in [-141,-101], F_{54} \in [-141,-101]$
$P_E^{(2)}$	$F_3 \in [834,936], F_5 \in [-589,-497], F_6 \in [-102,-62], F_7 \in [-102,-62], F_{10} \in [-380,-320], F_{11} \in [-380,-320], F_{19} \in [-267,-241], F_{20} \in [-265,-239], F_{21} \in [-267,-241], F_{22} \in [-265,-239], F_{31} \in [211,237], F_{32} \in [198,224], F_{33} \in [131,157], F_{34} \in [129,155], F_{35} \in [41,67], F_{36} \in [42,68]$
$P_E^{(3)}$	$F_3 \in [523,615], F_5 \in [-386,-294], F_6 \in [-63,-23], F_7 \in [-63,-23], F_{10} \in [-230,-170], F_{11} \in [-230,-170], F_{19} \in [-158,-184], F_{20} \in [-158,-184], F_{21} \in [-158,-184], F_{22} \in [-158,-184], F_{31} \in [108,134], F_{32} \in [101,127], F_{33} \in [67,93], F_{34} \in [67,93], F_{35} \in [10,36], F_{36} \in [11,37]$

Πίνακας 5.1: «Προφίλ» βασικών εκφράσεων

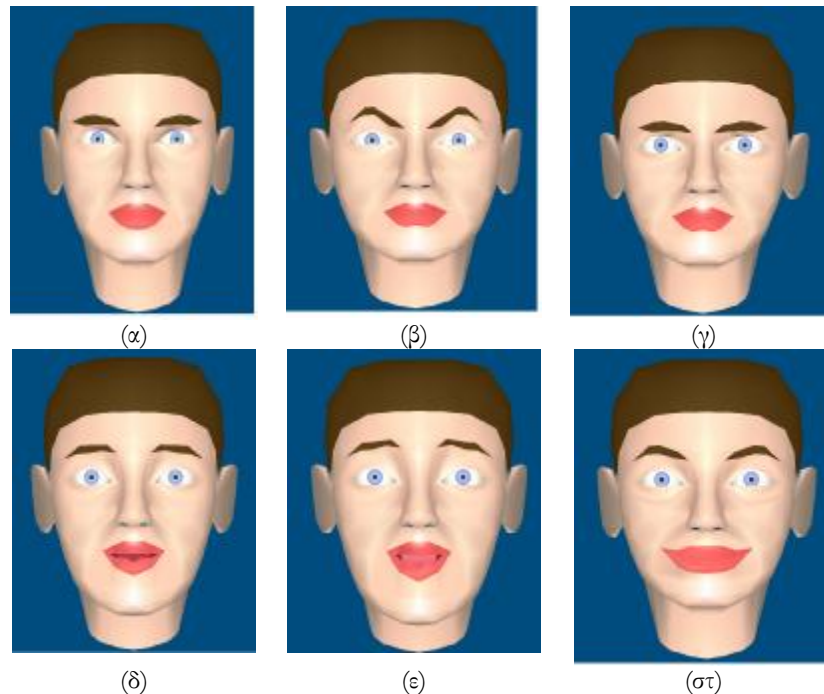
5.1.1.Επαλήθευση μέσω Πειραματικών Αποτελεσμάτων

Στα πειράματα που πραγματοποιήθηκαν στην παρούσα διατριβή για τη δημιουργία και απεικόνιση των «προφίλ» των εκφράσεων, αρχικά χρησιμοποιήθηκε το μοντέλο προσώπου που αναπτύχθηκε στο πλαίσιο του Ευρωπαϊκού Προγράμματος *ACTS MoMuSys* [49], το οποίο διατίθεται στην ιστοσελίδα <http://www.iso.ch/ittf> και στη συνέχεια χρησιμοποιήθηκε ο πιο ολοκληρωμένος (περιλαμβάνει και σώμα) MPEG-4 αποκωδικοποιητής GretaPlayer [90], ο οποίος περιγράφεται στην **Ενότητα 7.1**, αφού αποτελεί, εκτός από MPEG-4 αποκωδικοποιητή, και Πράκτορα με σώμα και δυνατότητα ομιλίας (ECA - Embodied Conversational Agent).

Για την επαλήθευση των αποτελεσμάτων, σε ορισμένες περιπτώσεις χρησιμοποιήθηκε και το τρισδιάστατο μοντέλο του υπολογιστικού πακέτου *Poser*, έκδοση 6 της Curious Labs. Το μοντέλο του *Poser* έχει ξεχωριστά μέρη για κάθε κινούμενο κομμάτι του προσώπου, κινείται από τα κουμπιά ελέγχου του λογισμικού κι έχει συνδέσμους που κινούνται ρεαλιστικά, όπως στον άνθρωπο. Το Poser μιμείται τις πραγματικές κινήσεις του προσώπου

προσθέτοντας συνδέσμους σε κάθε κομμάτι, παρέχοντας τη δυνατότητα στο χρήστη να ελέγξει το μοντέλο βασιζόμενος σε αυτές τις παραμέτρους. Μπορούμε να ελέγξουμε τα μάτια, τα φρύδια και το στόμα συμπληρώνοντας απλά τις κατάλληλες παραμέτρους. Για τη διαδικασία αυτή είναι απαραίτητη μια αντιστοίχιση των FAPs και των παραμέτρων του Poser, την οποία και πραγματοποιήσαμε βασιζόμενοι κυρίως στην εμπειρία αφού τις περισσότερες φορές πρόκειται για ευδιάκριτη αντιστοίχιση.

Η **Εικόνα 5.1** δείχνει μερικά παραδείγματα απεικόνισης βασικών εκφράσεων που δημιουργήθηκαν με το μοντέλο του Ευρωπαϊκού Προγράμματος *ACTS MoMuSys* [49]. Η **Εικόνα 5.1(α)** απεικονίζει ένα «προφίλ» της βασικής έκφρασης *θυμός*, ενώ οι **Εικόνες 5.1(β)** και **(γ)** απεικονίζουν διαφορετικά «προφίλ» της ίδιας έκφρασης. Οι διαφορές μεταξύ τους οφείλονται στις τιμές των FAPs. Τέτοιου είδους διαφορές διακρίνονται και στις **Εικόνες 5.1(δ)** και **(ε)**, οι οποίες απεικονίζουν το ίδιο «προφίλ» της έκφρασης *έκπληξη*. Η **Εικόνα 5.1(στ)** απεικονίζει ένα «προφίλ» της έκφρασης *χαρά*.

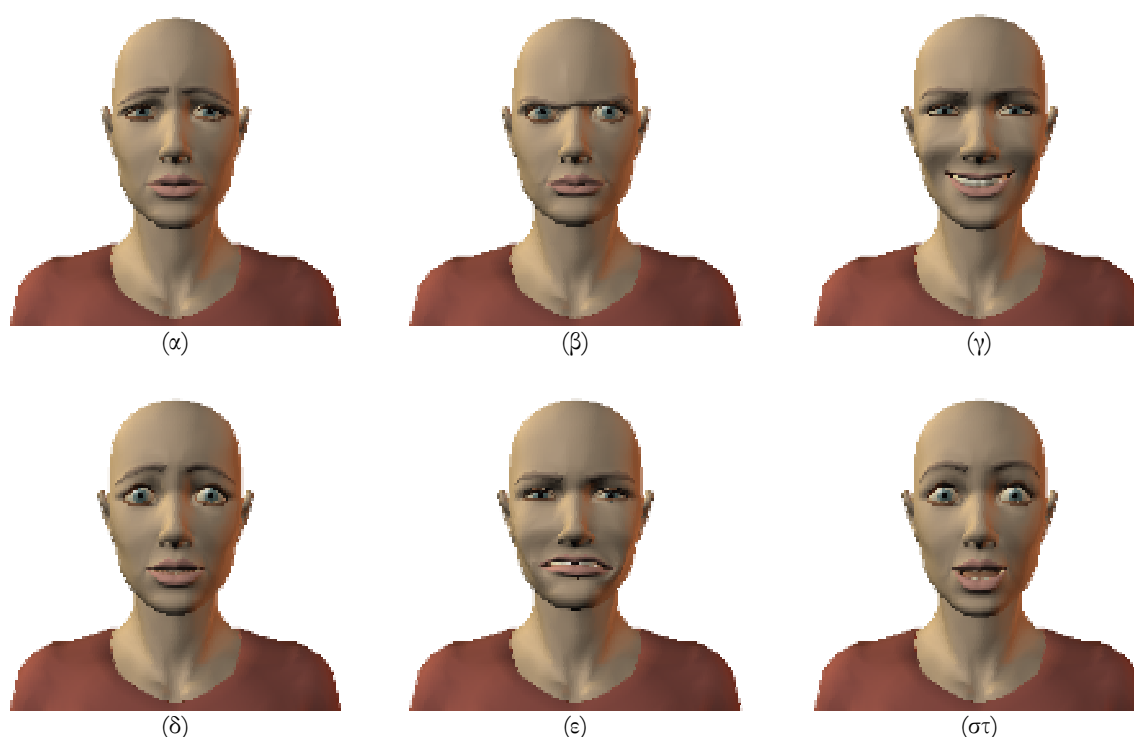


Εικόνα 5.1: Παραδείγματα απεικόνισης «προφίλ» (α)-(γ) *Θυμός*, (δ)-(ε) *Έκπληξη*, (στ) *Χαρά*

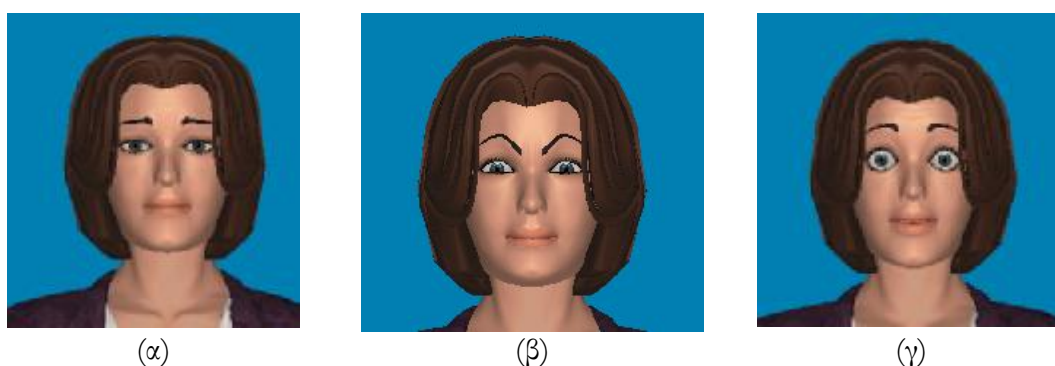
Η ομάδα αποτελεσμάτων της **Εικόνας 5.2** αποτελείται από απεικονίσεις βασικών εκφράσεων κι έγινε με τη βοήθεια του λογισμικού Poser. Η τρισδιάστατη φύση του συγκεκριμένου μοντέλου δίνει πιο ρεαλιστικά αποτελέσματα στην απεικόνιση των εκφράσεων σε σχέση με το, συμβατό με το πρότυπο MPEG-4, μοντέλο, διαπίστωση που προκύπτει από τη σύγκριση των **Εικόνων 5.1(ε)-(στ)** για τις εκφράσεις *έκπληξη* και *χαρά* με τις **Εικόνες 5.2(στ)** και **5.2(γ)** αντίστοιχα. Πάντως και στις δύο περιπτώσεις τα αποτελέσματα είναι ικανοποιητικά.

Στην **Εικόνα 5.3** φαίνονται και κάποια παραδείγματα βασικών εκφράσεων, των οποίων η απεικόνιση έγινε στο, συμβατό με MPEG-4, αποκωδικοποιητή GretaPlayer.

Στην περίπτωση του Poser οι χρήστες έχουν τη δυνατότητα να ελέγξουν την μορφοποίηση περιοχών του πολυγωνικού μοντέλου, χωρίς να περιορίζονται σε συγκεκριμένα διανύσματα, οπότε τα αποτελέσματα είναι πιο ρεαλιστικά. Στην περίπτωση των **Εικόνων 5.1 και 5.3** ο αποκωδικοποιητής χρησιμοποιεί μόνο τα FAPs που του δίνονται, οπότε το τελικό αποτέλεσμα εξαρτάται από την προκαθορισμένη αντιστοίχιση μεταξύ των παραμέτρων κίνησης του προσώπου και του μοντέλου, το οποίο αποτελείται από περιορισμένο αριθμό πολυγώνων.



Εικόνα 5.2: Απεικόνιση βασικών εκφράσεων με χρήση του τρισδιάστατου μοντέλου του Poser: (α) λύπη, (β) θυμός, (γ) χαρά, (δ) φόβος, (ε) απέχθεια και (στ) έκπληξη



Εικόνα 5.3: Απεικόνιση κάποιων βασικών εκφράσεων με χρήση του MPEG-4 αποκωδικοποιητή GretaPlayer: (α) λύπη, (β) θυμός, (γ) φόβος

5.1.2. Προφίλ εκφράσεων για τα τεταρτημόρια του Τροχού Συναισθημάτων της Whissel

Τα προφίλ που δημιουργήθηκαν σύμφωνα με τη διαδικασία που περιγράφεται στην προηγούμενη ενότητα, έχουν ως αποτέλεσμα μία από τις έξι βασικές εκφράσεις. Όμως, ύστερα από παρέμβαση ψυχολόγων, ακολούθηθηκε η ίδια διαδικασία, αλλά για προφίλ που έχουν ως αποτέλεσμα τα τεταρτημόρια του Τροχού Συναισθημάτων της Whissel. Πλέον οι διάφοροι συνδυασμοί των FAPs με τα αντίστοιχα διαστήματα τιμών δεν οδηγούν σε συμπέρασμα σχετικά με την έκφραση που αναλύεται-ή αντίστοιχα που συντίθεται, αλλά δίνουν στοιχεία για την ενεργοποίηση (activation) και αποτίμηση (evaluation) του συναισθήματος του χρήστη, ενώ απαλείφθηκαν τα προφίλ που οδηγούσαν σε αντικρουόμενα αποτελέσματα. Τα προφίλ, όπως αυτά διαμορφώθηκαν ώστε να δίνουν τεταρτημόριο, καταγράφονται στον **Πίνακα 5.2**:

Τεταρτημόριο	FAPs και Διαστήματα Τιμών
(+, +)	$F_4 \in [-381, -161]$, $F_6 \in [136, 332]$, $F_7 \in [136, 332]$, $F_{19} \in [-387, -121]$, $F_{20} \in [-364, -120]$, $F_{21} \in [-387, -121]$, $F_{22} \in [-364, -120]$, $F_{33} \in [2, 46]$, $F_{34} \in [3, 47]$, $F_{53} \in [136, 332]$, $F_{54} \in [136, 332]$
(+, +)	$F_6 \in [160, 240]$, $F_7 \in [160, 240]$, $F_{12} \in [260, 340]$, $F_{13} \in [260, 340]$, $F_{19} \in [-449, -325]$, $F_{20} \in [-426, -302]$, $F_{21} \in [325, 449]$, $F_{22} \in [302, 426]$, $F_{33} \in [70, 130]$, $F_{34} \in [70, 130]$, $F_{41} \in [130, 170]$, $F_{42} \in [130, 170]$, $F_{53} \in [160, 240]$, $F_{54} \in [160, 240]$
(+, +)	$F_6 \in [160, 240]$, $F_7 \in [160, 240]$, $F_{12} \in [260, 340]$, $F_{13} \in [260, 340]$, $F_{19} \in [-449, -325]$, $F_{20} \in [-426, -302]$, $F_{21} \in [-312, -188]$, $F_{22} \in [-289, -165]$, $F_{33} \in [70, 130]$, $F_{34} \in [70, 130]$, $F_{41} \in [130, 170]$, $F_{42} \in [130, 170]$, $F_{53} \in [160, 240]$, $F_{54} \in [160, 240]$
(+, +)	$F_3 \in [569, 1201]$, $F_5 \in [340, 746]$, $F_6 \in [-121, -43]$, $F_7 \in [-121, -43]$, $F_{19} \in [170, 337]$, $F_{20} \in [171, 333]$, $F_{21} \in [170, 337]$, $F_{22} \in [171, 333]$, $F_{31} \in [121, 327]$, $F_{32} \in [114, 308]$, $F_{33} \in [80, 208]$, $F_{34} \in [80, 204]$, $F_{35} \in [23, 85]$, $F_{36} \in [23, 85]$, $F_{53} \in [-121, -43]$, $F_{54} \in [-121, -43]$
(+, +)	$F_3 \in [1150, 1252]$, $F_5 \in [-792, -700]$, $F_6 \in [-141, -101]$, $F_7 \in [-141, -101]$, $F_{10} \in [-530, -470]$, $F_{11} \in [-530, -470]$, $F_{19} \in [-350, -324]$, $F_{20} \in [-346, -320]$, $F_{21} \in [-350, -324]$, $F_{22} \in [-346, -320]$, $F_{31} \in [314, 340]$, $F_{32} \in [295, 321]$, $F_{33} \in [195, 221]$, $F_{34} \in [191, 217]$, $F_{35} \in [72, 98]$, $F_{36} \in [73, 99]$, $F_{53} \in [-141, -101]$, $F_{54} \in [-141, -101]$
(+, +)	$F_3 \in [834, 936]$, $F_5 \in [-589, -497]$, $F_6 \in [-102, -62]$, $F_7 \in [-102, -62]$, $F_{10} \in [-380, -320]$, $F_{11} \in [-380, -320]$, $F_{19} \in [-267, -241]$, $F_{20} \in [-265, -239]$, $F_{21} \in [-267, -241]$, $F_{22} \in [-265, -239]$, $F_{31} \in [211, 237]$, $F_{32} \in [198, 224]$, $F_{33} \in [131, 157]$, $F_{34} \in [129, 155]$, $F_{35} \in [41, 67]$, $F_{36} \in [42, 68]$
(+, +)	$F_3 \in [523, 615]$, $F_5 \in [-386, -294]$, $F_6 \in [-63, -23]$, $F_7 \in [-63, -23]$, $F_{10} \in [-230, -170]$, $F_{11} \in [-230, -170]$, $F_{19} \in [-158, -184]$, $F_{20} \in [-158, -184]$, $F_{21} \in [-158, -184]$, $F_{22} \in [-158, -184]$, $F_{31} \in [108, 134]$, $F_{32} \in [101, 127]$, $F_{33} \in [67, 93]$, $F_{34} \in [67, 93]$, $F_{35} \in [10, 36]$, $F_{36} \in [11, 37]$
(-, +)	$F_4 \in [22, 124]$, $F_{31} \in [-131, -25]$, $F_{32} \in [-136, -34]$, $F_{33} \in [-189, -109]$, $F_{34} \in [-183, -105]$, $F_{35} \in [-101, -31]$, $F_{36} \in [-108, -32]$, $F_{37} \in [29, 85]$, $F_{38} \in [27, 89]$
(-, +)	$F_{19} \in [-330, -200]$, $F_{20} \in [-335, -205]$, $F_{21} \in [200, 330]$, $F_{22} \in [205, 335]$, $F_{31} \in [-200, -80]$, $F_{32} \in [-194, -74]$, $F_{33} \in [-190, -70]$, $F_{34} \in [-190, -70]$

(-, +)	$F_{19} \in [-330, -200]$, $F_{20} \in [-335, -205]$, $F_{21} \in [200, 330]$, $F_{22} \in [205, 335]$, $F_{31} \in [-200, -80]$, $F_{32} \in [-194, -74]$, $F_{33} \in [70, 190]$, $F_{34} \in [70, 190]$
(-, +)	$F_{16} \in [45, 155]$, $F_{18} \in [45, 155]$, $F_{19} \in [-330, -200]$, $F_{20} \in [-330, -200]$, $F_{31} \in [-200, -80]$, $F_{32} \in [-194, -74]$, $F_{33} \in [-190, -70]$, $F_{34} \in [-190, -70]$, $F_{37} \in [65, 135]$, $F_{38} \in [65, 135]$
(-, +)	$F_{16} \in [-355, -245]$, $F_{18} \in [145, 255]$, $F_{19} \in [-330, -200]$, $F_{20} \in [-330, -200]$, $F_{31} \in [-200, -80]$, $F_{32} \in [-194, -74]$, $F_{33} \in [-190, -70]$, $F_{34} \in [-190, -70]$, $F_{37} \in [65, 135]$, $F_{38} \in [65, 135]$
(-, +)	$F_3 \in [102, 480]$, $F_5 \in [83, 353]$, $F_{19} \in [118, 370]$, $F_{20} \in [121, 377]$, $F_{21} \in [118, 370]$, $F_{22} \in [121, 377]$, $F_{31} \in [35, 173]$, $F_{32} \in [39, 183]$, $F_{33} \in [14, 130]$, $F_{34} \in [15, 135]$
(-, +)	$F_3 \in [400, 560]$, $F_5 \in [307, 399]$, $F_{19} \in [-530, -470]$, $F_{20} \in [-523, -463]$, $F_{21} \in [-530, -470]$, $F_{22} \in [-523, -463]$, $F_{31} \in [460, 540]$, $F_{32} \in [460, 540]$, $F_{33} \in [460, 540]$, $F_{34} \in [460, 540]$, $F_{35} \in [460, 540]$, $F_{36} \in [460, 540]$
(-, +)	$F_3 \in [400, 560]$, $F_5 \in [-240, -160]$, $F_{19} \in [-630, -570]$, $F_{20} \in [-630, -570]$, $F_{21} \in [-630, -570]$, $F_{22} \in [-630, -570]$, $F_{31} \in [460, 540]$, $F_{32} \in [460, 540]$, $F_{37} \in [60, 140]$, $F_{38} \in [60, 140]$
(-, +)	$F_3 \in [400, 560]$, $F_5 \in [-240, -160]$, $F_{19} \in [-630, -570]$, $F_{20} \in [-630, -570]$, $F_{21} \in [-630, -570]$, $F_{22} \in [-630, -570]$, $F_{31} \in [460, 540]$, $F_{32} \in [460, 540]$, $F_{33} \in [360, 440]$, $F_{34} \in [360, 440]$, $F_{35} \in [260, 340]$, $F_{36} \in [260, 340]$, $F_{37} \in [60, 140]$, $F_{38} \in [60, 140]$
(-, +)	$F_3 \in [400, 560]$, $F_5 \in [-240, -160]$, $F_8 \in [-120, -80]$, $F_9 \in [-120, -80]$, $F_{10} \in [-120, -80]$, $F_{11} \in [-120, -80]$, $F_{19} \in [-630, -570]$, $F_{20} \in [-630, -570]$, $F_{21} \in [-630, -570]$, $F_{22} \in [-630, -570]$, $F_{31} \in [460, 540]$, $F_{32} \in [460, 540]$, $F_{33} \in [360, 440]$, $F_{34} \in [360, 440]$, $F_{35} \in [260, 340]$, $F_{36} \in [260, 340]$, $F_{37} \in 0$, $F_{38} \in 0$
(-, +)	$F_3 \in [400, 560]$, $F_5 \in [-240, -160]$, $F_{19} \in [-630, -570]$, $F_{20} \in [-630, -570]$, $F_{21} \in [-630, -570]$, $F_{22} \in [-630, -570]$, $F_{31} \in [260, 340]$, $F_{32} \in [260, 340]$, $F_{33} \in [160, 240]$, $F_{34} \in [160, 240]$, $F_{35} \in [60, 140]$, $F_{36} \in [60, 140]$
(-, +)	$F_3 \in [400, 560]$, $F_5 \in [307, 399]$, $F_{19} \in [-630, -570]$, $F_{20} \in [-623, -563]$, $F_{21} \in [-630, -570]$, $F_{22} \in [-623, -563]$, $F_{31} \in [460, 540]$, $F_{32} \in [460, 540]$, $F_{33} \in [460, 540]$, $F_{34} \in [460, 540]$, $F_{35} \in [460, 540]$, $F_{36} \in [460, 540]$
(-, -)	$F_{19} \in [-265, -41]$, $F_{20} \in [-270, -52]$, $F_{21} \in [-265, -41]$, $F_{22} \in [-270, -52]$, $F_{31} \in [30, 140]$, $F_{32} \in [26, 134]$
(-, -)	$F_4 \in [-343, -125]$, $F_5 \in [-285, -69]$, $F_{19} \in [-351, -55]$, $F_{20} \in [-356, -66]$, $F_{21} \in [-351, -55]$, $F_{22} \in [-356, -66]$, $F_{33} \in [-123, -27]$, $F_{34} \in [-126, -28]$

Πίνακας 5.2: «Προφίλ» τεταρτημορίων

5.2. Δημιουργία «προφίλ» συναισθημάτων τα οποία ανήκουν στην ίδια κατηγορία

Γενικά είναι δυνατόν να οριστούν έξι γενικές κατηγορίες, η κάθε μία από τις οποίες χαρακτηρίζεται από ένα βασικό συναισθημα. Στο πλαίσιο κάθε κατηγορίας οι ενδιάμεσες εκφράσεις περιγράφονται από διαφορετικές εντάσεις, οπότε παρουσιάζουν μικρές διαφοροποιήσεις στην τελική έκφραση. Από την πλευρά της σύνθεσης, η απεικόνιση των συναισθημάτων που ανήκουν στην ίδια κατηγορία περιλαμβάνει τα ίδια FAPs, αλλά με διαφορετικά διαστήματα τιμών. Για παράδειγμα, στην κατηγορία *ρόβος* περιλαμβάνονται η *ανησυχία* και ο *τρόμος* [40], συναισθήματα τα οποία μπορούν να απεικονιστούν με μείωση ή αύξηση, αντίστοιχα, των τιμών των συγκεκριμένων FAPs. Στην περίπτωση των «προφίλ» των

εκφράσεων, αυτή η διακύμανση στις τιμές συμβάλλει και στη διάκριση εκφράσεων με μικρές διαφοροποιήσεις, οπότε οδηγούμαστε σε πιο αληθοφανή απεικόνιση.

Έστω $P_i^{(k)}$ το k -ιστό «προφίλ» του συναισθήματος i και $X_{i,j}^{(k)}$ το διάστημα τιμών του FAP F_j που συμμετέχει στο $P_i^{(k)}$. Αν \mathcal{A} , I είναι συναισθήματα που ανήκουν στην ίδια κατηγορία, όπου \mathcal{A} το βασικό συναισθήμα και I το ενδιάμεσο, ισχύουν οι παρακάτω κανόνες:

Κανόνας 1:	Τα $P_{\mathcal{A}}^{(k)}$ και $P_I^{(k)}$ χρησιμοποιούν τα ίδια FAPs
Κανόνας 2:	Το διάστημα τιμών $X_{I,j}^{(k)}$ υπολογίζεται από τη σχέση: $X_{I,j}^{(k)} = \frac{a_I}{a_{\mathcal{A}}} X_{\mathcal{A},j}^{(k)}$.
Κανόνας 3:	$a_{\mathcal{A}}$ και a_I είναι οι τιμές της παραμέτρου ενεργοποίησης για τα συναισθήματα \mathcal{A} και I σύμφωνα με τη Whissel [35].

5.2.1. Επαλήθευση μέσω Πειραματικών Αποτελεσμάτων

Τα συναισθήματα *φοβισμένος*, *τρομοκρατημένος* και *ανήσυχος* θεωρείται ότι ανήκουν στην κατηγορία *φόβος* [40], του οποίου η απεικόνιση βασίζεται στο *φοβισμένος*. Στον **Πίνακα 5.3** φαίνονται τα «προφίλ» για τα συναισθήματα *τρομοκρατημένος* και *ανήσυχος*, τα οποία προέκυψαν από το «προφίλ» του *φοβισμένου* (συγκεκριμένα του $P_{\Phi}^{(8)}$ - Πίνακας 5.1). Το διάστημα τιμών $X_{T,j}^{(8)}$ του FAP F_j , το οποίο ανήκει στο όγδοο «προφίλ» του συναισθήματος *τρομοκρατημένος*, υπολογίζεται από τη σχέση $X_{T,j}^{(8)} = \frac{6.3}{4.9} X_{\Phi,j}^{(8)}$, όπου $X_{\Phi,j}^{(8)}$ είναι το διάστημα τιμών του FAP F_j το οποίο ανήκει στο όγδοο «προφίλ» του συναισθήματος *φοβισμένος*. Παρομοίως $X_{A,j}^{(8)} = \frac{3.9}{4.9} X_{\Phi,j}^{(8)}$ είναι το διάστημα τιμών του FAP F_j το οποίο ανήκει στο όγδοο «προφίλ» του συναισθήματος *ανήσυχος*. Στις **Εικόνες 5.4(α)-(γ)**, **5.5(α)-(γ)** και **5.6(α)-(γ)** φαίνεται η απεικόνιση των «προφίλ» *φοβισμένος*, *τρομοκρατημένος* και *ανήσυχος* αντιστοίχως. Οι τιμές των FAPs που χρησιμοποιήσαμε είναι οι ενδιάμεσες τιμές των αντιστοίχων διαστημάτων.

Συναισθήμα	Ενεργοποίηση	«Προφίλ»
<i>Φοβισμένος</i>	4.9	$F_3 \in [400,560]$, $F_5 \in [-240,-160]$, $F_{19} \in [-630,-570]$, $F_{20} \in [-630,-570]$, $F_{21} \in [-630,-570]$, $F_{22} \in [-630,-570]$, $F_{31} \in [260,340]$, $F_{32} \in [260,340]$, $F_{33} \in [160,240]$, $F_{34} \in [160,240]$, $F_{35} \in [60,140]$, $F_{36} \in [60,140]$
<i>Τρομοκρατημένος</i>	6.3	$F_3 \in [520,730]$, $F_5 \in [-310,-210]$, $F_{19} \in [-820,-740]$, $F_{20} \in [-820,-740]$, $F_{21} \in [-820,-740]$, $F_{22} \in [-820,-740]$, $F_{31} \in [340,440]$, $F_{32} \in [340,440]$, $F_{33} \in [210,310]$,

		$F_{34} \in [210, 310]$, $F_{35} \in [80, 180]$, $F_{36} \in [80, 180]$
<i>Ανήσυχος</i>	3.9	$F_3 \in [320, 450]$, $F_5 \in [-190, -130]$, $F_{19} \in [-500, -450]$, $F_{20} \in [-500, -450]$, $F_{21} \in [-500, -450]$, $F_{22} \in [-500, -450]$, $F_{31} \in [210, 270]$, $F_{32} \in [210, 270]$, $F_{33} \in [130, 190]$, $F_{34} \in [130, 190]$, $F_{35} \in [50, 110]$, $F_{36} \in [50, 110]$

Πίνακας 5.3: «Προφίλ» των συναισθημάτων *τρόμος* και *ανήσυχία*



(α)



(β)

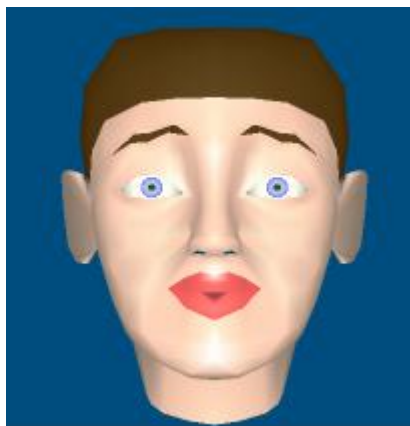


(γ)

Εικόνα 5.4: Μοντέλο προσώπου Poser: Απεικόνιση «προφίλ» των συναισθημάτων (α) *φοβισμένος*, (β) *τρομοκρατημένος* και (γ) *ανήσυχος*



(α)

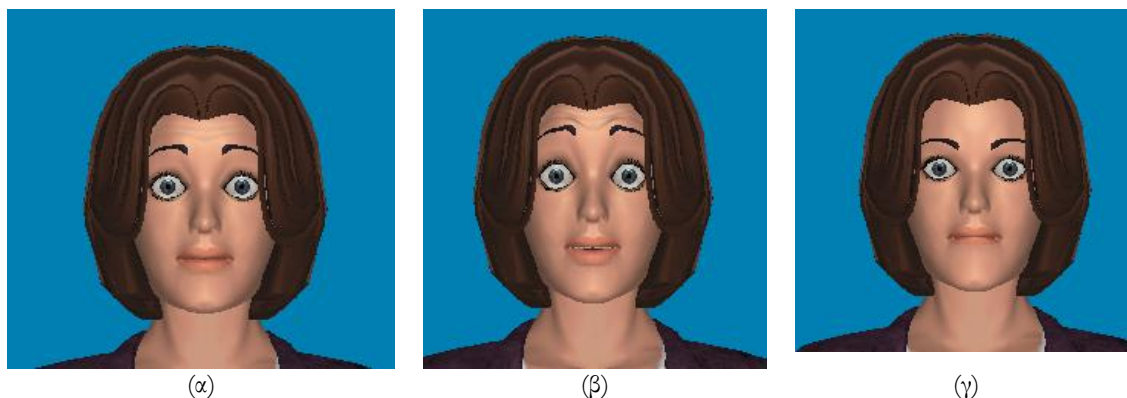


(β)



(γ)

Εικόνα 5.5: Μοντέλο προσώπου MPEG-4: Απεικόνιση «προφίλ» των συναισθημάτων (α) *φοβισμένος*, (β) *τρομοκρατημένος* και (γ) *ανήσυχος*



Εικόνα 5.6: Μοντέλο προσώπου MPEG-4 Greta: Απεικόνιση «προφίλ» των συναισθημάτων (α) φοβισμένος, (β) τρομοκρατημένος και (γ) ανήσυχος

5.3. Ενδιάμεσες εκφράσεις

Το μεγαλύτερο μέρος των μελετών που έχουν εκπονήσει ερευνητές στον τομέα της υπολογιστικής όρασης και της γραφικής ασχολείται με τις βασικές εκφράσεις και όχι με το σύνολο των βασικών κι ενδιάμεσων εκφράσεων [37]. Στην έρευνά μας χρησιμοποιήσαμε αναπαραστάσεις συναισθήματος [42], όπως αυτή της Whissel (Κεφάλαιο 2), οι οποίες μπορούν να περιγράψουν εύρος εκφράσεων.

Στην πρώτη στήλη του **Πίνακα 5.4** αναγράφονται τιμές για το ποσό ενεργοποίησης (activation) ορισμένων συναισθημάτων, για παράδειγμα ο όρος «έκπληκτος» (περίπου 3) δείχνει έντονη ενεργοποίηση, σε αντίθεση με το «λυπημένος» (περίπου -2) που δείχνει χαμηλή ενεργοποίηση. Η αποτίμηση (evaluation) των διαφόρων συναισθημάτων, ουσιαστικά δηλαδή πόσο θετικό είναι ένα συναίσθημα, αναγράφεται στη δεύτερη στήλη του **Πίνακα 5.4**, όπου ο «θυμωμένος» (-3) αντιπροσωπεύει το αρνητικό άκρο και ο «ενθουσιασμένος» (2.9) το θετικό άκρο [35].

Η τελευταία στήλη του **Πίνακα 5.4** περιέχει τη γωνιακή απόσταση των συναισθημάτων, όπως αυτή έχει οριστεί από τον Plutchik [36].

	<i>Ενεργοποίηση</i>	<i>Αποτίμηση</i>	<i>Γωνία</i>
Τρομοκρατημένος	6.3	3.4	75.7
Φοβισμένος	4.9	3.4	70.3
Ανήσυχος	3.9	2.9	126
Θυμωμένος	4.2	2.7	212
Έκπληκτος	3.3	3.8	146.7
Λυπημένος	3.8	2.4	108.5
Απογοητευμένος	5	3.2	181.3
Καχύποπτος	4.2	6.4	102.3
Ενθουσιασμένος	4	1.1	318.6

Πίνακας 5.4: Επιλεγμένα συναισθήματα από τη μελέτη της Whissel [35]

Για τη δημιουργία των «προφίλ» των ενδιάμεσων εκφράσεων διακρίνουμε δύο περιπτώσεις:

1. Συναισθήματα παρεμφερή κάποιας βασικής έκφρασης, τα οποία μπορεί να διαφέρουν αυτής στην ένταση της μυϊκής κίνησης (Ενότητα 5.2).
2. Συναισθήματα που δεν τοποθετούνται κοντά σε κάποια από τις βασικές εκφράσεις.

Για τη δημιουργία προφίλ ενδιάμεσων εκφράσεων ακολουθούμε τα βήματα που περιγράφονται στη συνέχεια:

1. Θεωρούμε ως προϋπάρχουσα γνώση είτε την *παράμετρο ενεργοποίησης* είτε τη *γωνιακή μέτρηση* του *Plutchik*. Αυτή η γνώση συνδυάζεται με τα «προφίλ» των βασικών εκφράσεων και με τη βοήθεια κανόνων δημιουργούμε τα «προφίλ» των ενδιάμεσων εκφράσεων.
2. Απεικονίζουμε τα «προφίλ» που δημιουργήσαμε, ώστε να τα ελέγξουμε και να τα διορθώσουμε για να προσεγγίζουν κατά το δυνατόν περισσότερο τη ζητούμενη έκφραση.

5.3.1. Δημιουργία «προφίλ» για Ενδιάμεσες Εκφράσεις - Συναισθήματα

Η δημιουργία «προφίλ» συναισθημάτων τα οποία δεν ανήκουν ευκρινώς σε κάποια κατηγορία δεν είναι προφανής διαδικασία. Εκτός από τον ορισμό του διαστήματος τιμών για κάθε FAP, είναι αναγκαίο να οριστεί και το λεξιλόγιο των FAPs, δηλαδή ποια FAPs θα συμμετέχουν στην απεικόνιση της συγκεκριμένης έκφρασης. Για την κατηγορία αυτή των συναισθημάτων χρησιμοποιούμε τόσο τη γωνιακή απόσταση του *Plutchik* (Πίνακας 5.4), όσο και την παράμετρο ενεργοποίησης. Έστω I ένα συναισθηματικό μεταξύ των συναισθημάτων A_1 και A_2 - τα οποία είναι τα πιο κοντινά στο επιλεγμένο συναισθηματικό - σύμφωνα με τη γωνιακή μέτρηση. Έστω επίσης V_{A_1} και V_{A_2} τα λεξιλόγια (σύνολα FAPs - Πίνακας 4.2), τα οποία αντιστοιχούν στα A_1 και A_2 αντιστοίχως. Το λεξιλόγιο V_I του συναισθηματικού I προκύπτει από την ένωση των λεξιλογίων V_{A_1} και V_{A_2} , δηλαδή $V_I = V_{A_1} \cup V_{A_2}$.

Όπως αναφέρθηκε παραπάνω, ο ορισμός του λεξιλογίου δεν αρκεί για την απεικόνιση μιας έκφρασης, η δημιουργία «προφίλ» είναι απαραίτητη, γεγονός που θέτει διάφορα ζητήματα στην περίπτωση που χρησιμοποιούνται διαφορετικά FAPs στα διάφορα «προφίλ». Στο πλαίσιο της παρούσας διατριβής διατηρούμε, κατά τη δημιουργία «προφίλ», τα FAPs που είναι κοινά στα δύο συναισθήματα, ενώ τα FAPs που συναντώνται μόνο στη μία έκφραση συνδυάζονται με την αντίστοιχη ουδέτερη έκφραση λαμβάνοντας το μέσο όρο. Το ίδιο ισχύει και στην περίπτωση αλληλοαποκλειόμενων FAPs, όπου ο υπολογισμός του μέσου όρου των τιμών συνήθως ευνοεί την πιο ακραία από τις δύο συνδυαζόμενες εκφράσεις, ενώ τα FAPs με αντίθετο διάστημα τιμών δεν λαμβάνονται υπόψη. Στην πράξη, αυτή η προσέγγιση δίνει ικανοποιητικά αποτελέσματα, όπως άλλωστε γίνεται φανερό από τα αντίστοιχα αποτελέσματα.

Ο συνδυασμός διαφορετικών, πιθανώς αντίθετων ή αποκλειόμενων, FAPs είναι δυνατόν να χρησιμεύσει στη δημιουργία σαφούς κατηγοριοποίησης των συναισθημάτων, η οποία θα

προσεγγίζει τη σημασιολογική και θα λαμβάνει υπόψιν τα κοινά ή κοντινά FAPs που χρησιμοποιούνται στην απεικόνιση των συναισθημάτων.

Παρακάτω περιγράφεται η διαδικασία η οποία ακολουθείται για τη δημιουργία «προφίλ» ενδιάμεσων εκφράσεων από το συνδυασμό «προφίλ» βασικών συναισθημάτων:

Έστω $P_{A_1}^{(k)}$ το k -ιοστό προφίλ του συναισθήματος A_1 και $P_{A_2}^{(l)}$ το l -ιοστό «προφίλ» του συναισθήματος A_2 . Επιπλέον, $\omega = \tan^{-1}\left(\frac{a}{e}\right)$, ω_{A_1} , ω_{A_2} και ω_I , $\omega_{A_1} < \omega_I < \omega_{A_2}$ είναι οι τιμές των *γωνιακών* παραμέτρων για τα συναισθήματα A_1 , A_2 και I , σύμφωνα με τον Plutchik [36], a_{A_1} , a_{A_2} και a_I είναι οι τιμές της παραμέτρου ενεργοποίησης (*activation*) και e_{A_1} , e_{A_2} and e_I είναι οι τιμές της παραμέτρου αποτίμησης (*evaluation*) για τα συναισθήματα A_1 , A_2 και I , σύμφωνα με τη Whissel [35]. Για τη δημιουργία του «προφίλ» $P_I^{(m)}$ του ενδιάμεσου συναισθήματος I εφαρμόζονται οι παρακάτω κανόνες:

Κανόνας 1:	Το $P_I^{(m)}$ περιέχει τα FAPs που περιλαμβάνονται στο $P_{A_1}^{(k)}$ ή στο $P_{A_2}^{(l)}$.
Κανόνας 2:	<p>Αν F_j είναι ένα FAP που περιέχεται συγχρόνως στα $P_{A_1}^{(k)}$ και $P_{A_2}^{(l)}$ με το ίδιο πρόσημο (κατεύθυνση κίνησης), τότε το διάστημα τιμών $X_{I,j}^{(k)}$ υπολογίζεται ως σταθμητός συνδυασμός των $X_{A_1,j}^{(k)}$ και $X_{A_2,j}^{(l)}$ (όπου $X_{A_1,j}^{(k)}$ και $X_{A_2,j}^{(l)}$ είναι τα διαστήματα τιμών του FAP F_j στα $P_{A_1}^{(k)}$ και $P_{A_2}^{(l)}$ αντίστοιχα) με την εξής διαδικασία: (α) υπολογίζονται οι αποκλίσεις των τιμών $t(X_{A_1,j}^{(k)}) = \frac{a_I}{a_{A_1}} X_{A_1,j}^{(k)}$ και $t(X_{A_2,j}^{(l)}) = \frac{a_I}{a_{A_2}} X_{A_2,j}^{(l)}$ των $X_{A_1,j}^{(k)}$ και $X_{A_2,j}^{(l)}$,</p> <p>(β) τα κέντρα $c_{A_1,j}^{(k)}$ και $c_{A_2,j}^{(l)}$ των $t(X_{A_1,j}^{(k)})$ και $t(X_{A_2,j}^{(l)})$ παραμένουν τα ίδια με αυτά των $X_{A_1,j}^{(k)}$ και $X_{A_2,j}^{(l)}$ (γ) τα πλάτη $s_{A_1,j}^{(k)}$ και $s_{A_2,j}^{(l)}$ των $t(X_{A_1,j}^{(k)})$ και $t(X_{A_2,j}^{(l)})$ υπολογίζονται από τη σχέση $s_{Ai,j}^{(k)} = \frac{1}{3} t(X_{A_i,j}^{(k)})$,</p> <p>(δ) το πλάτος του $X_{I,j}^{(m)}$ είναι $s_{I,j}^{(m)} = \frac{\omega_I - \omega_{A_1}}{\omega_{A_2} - \omega_{A_1}} s_{A_1,j}^{(k)} + \frac{\omega_{A_2} - \omega_I}{\omega_{A_2} - \omega_{A_1}} s_{A_2,j}^{(l)}$ και το κέντρο του $c_{I,j}^{(m)} = \frac{\omega_I - \omega_{A_1}}{\omega_{A_2} - \omega_{A_1}} c_{A_1,j}^{(k)} + \frac{\omega_{A_2} - \omega_I}{\omega_{A_2} - \omega_{A_1}} c_{A_2,j}^{(l)}$</p>
Κανόνας 3:	<p>Αν το F_j περιέχεται συγχρόνως στα $P_{A_1}^{(k)}$ και $P_{A_2}^{(l)}$, αλλά με αντίθετο πρόσημο (δηλαδή αντίθετη κατεύθυνση κίνησης), τότε η απόκλιση τιμών $X_{I,j}^{(k)}$ υπολογίζεται από τη σχέση $X_{I,j}^{(m)} = \frac{a_I}{a_{A_1}} X_{A_1,j}^{(k)} \cap \frac{a_I}{a_{A_2}} X_{A_2,j}^{(l)}$. Στην</p>

	περίπτωση που η $X_{I,j}^{(k)}$ έχει απαλειφθεί (η πιθανότερη περίπτωση), το F_j αποκλείεται από το «προφίλ».
Κανόνας 4:	Αν το F_j περιέχεται αποκλειστικά σε ένα από τα $P_{A_1}^{(k)}$ και $P_{A_2}^{(l)}$ τότε το διάστημα τιμών $X_{I,j}^{(k)}$ θα συνδυαστεί με την αντίστοιχη ουδέτερη έκφραση, π.χ.: $X_{I,j}^{(m)} = \frac{a_I}{2 * a_{A_1}} X_{A_1,j}^{(k)}$ και $X_{I,j}^{(m)} = \frac{a_I}{2 * a_{A_2}} X_{A_2,j}^{(l)}$

Τα «προφίλ» που δημιουργούνται σύμφωνα με τους παραπάνω κανόνες πρέπει πρώτα να απεικονιστούν, ώστε να επαληθευτούν και να γίνουν οι απαραίτητες διορθώσεις που θα τα κάνουν πιο αληθοφανή.

5.3.2. Επαλήθευση μέσω Πειραματικών Αποτελεσμάτων

Στην παρούσα ενότητα περιγράφεται η διαδικασία δημιουργίας «προφίλ» του συναισθήματος *απογοητευμένος*, ως ειδική περίπτωση δημιουργίας προφίλ ενδιάμεσων συναισθημάτων. Σύμφωνα με τις *γωνιακές* μετρήσεις του Plutchik (Πίνακας 5.4), το συναισθήμα *απογοητευμένος* (γωνιακή μέτρηση 102.3 μοίρες) βρίσκεται ανάμεσα στα βασικά συναισθήματα *φοβισμένος* (γωνιακή μέτρηση 70.3 μοίρες) και *λυπημένος* (γωνιακή μέτρηση 108.5 μοίρες), ευρισκόμενο πιο κοντά στο δεύτερο. Σύμφωνα με όσα περιγράφηκαν σε προηγούμενες ενότητες, το λεξιλόγιο V_{EV} του συναισθήματος *απογοητευμένος* προκύπτει από την ένωση των λεξιλογίων V_{Φ} και V_{Λ} , δηλαδή από τη σχέση $V_{EV} = V_{\Phi} \cup V_{\Lambda}$, όπου V_{Φ} και V_{Λ} είναι τα λεξιλόγια που αντιστοιχούν στα συναισθήματα *φοβισμένος* και *χαρούμενος* αντίστοιχα. Στον Πίνακα 5.5 φαίνεται το «προφίλ» για το συναισθήμα *απογοητευμένος*, όπως αυτό προέκυψε από το συνδυασμό των «προφίλ» του *φοβισμένου* (ειδικότερα του $P_{\Phi}^{(8)}$) και του *λυπημένου* (ειδικότερα του $P_{\Lambda}^{(0)}$). Τα FAPs $F_3, F_5, F_{33}-F_{36}$ περιλαμβάνονται μόνο στο $P_{\Phi}^{(8)}$, οπότε τα αντίστοιχα διαστήματα τιμών στο προκύπτον «προφίλ» του ενόχου $P_{EV}^{(m)}$ (*m*-ιστό «προφίλ») υπολογίζονται από το συνδυασμό των διαστημάτων τιμών του $P_{\Phi}^{(8)}$ με το πρόσωπο ουδέτερης έκφρασης, σύμφωνα με τον Κανόνα 4, για παράδειγμα $X_{EV,3}^{(m)} = \frac{4}{2 * 4.9} X_{\Phi,4}^{(8)}$. Τα FAPs $F_{19}-F_{22}, F_{31}, F_{32}$ περιλαμβάνονται και στα δύο «προφίλ» $P_{\Phi}^{(8)}$ και $P_{\Lambda}^{(0)}$ με την ίδια κατεύθυνση κίνησης, οπότε ακολουθείται ο Κανόνας 2. Για παράδειγμα το διάστημα τιμών $X_{EV,19}^{(m)}$ για το FAP F_{29} υπολογίζεται ως εξής:

$$t(X_{\Phi,19}^{(8)}) = \frac{4}{4.9} X_{\Phi,19}^{(8)} \Rightarrow [-510, -460], c_{\Phi,19}^{(8)} = -485, s_{\Phi,19}^{(8)} = 50,$$

$$t(X_{\Lambda,19}^{(0)}) = \frac{4}{3.9} X_{\Lambda,19}^{(0)} \Rightarrow [-270, -42], c_{\Lambda,19}^{(0)} = -156, s_{\Lambda,19}^{(9)} = 228,$$

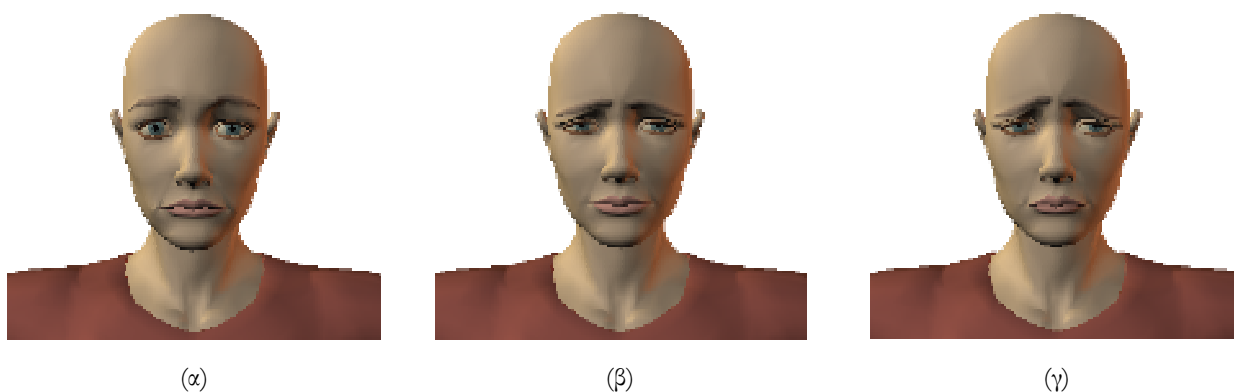
$$\alpha\phi\acute{o}\nu \ \omega_{\Phi} = 70.3^{\circ}, \ \omega_{\Lambda} = 108.5^{\circ}, \ \omega_{E\nu} = 102.3^{\circ},$$

$$c_{E\nu,19}^{(m)} = \frac{102.3 - 70.3}{108.5 - 70.3} * (-156) + \frac{108.5 - 102.3}{108.5 - 70.3} * (-485) = -209$$

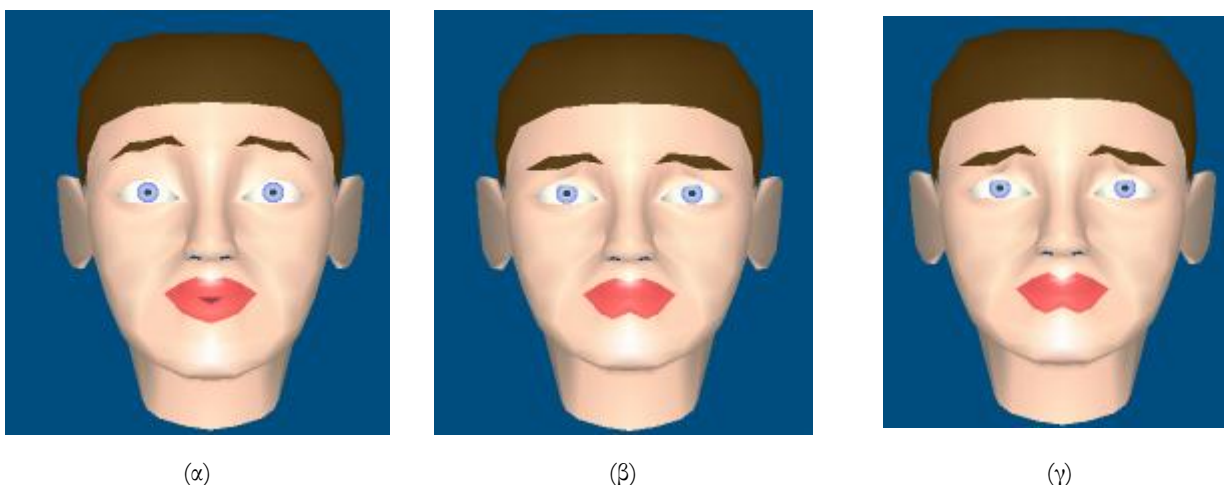
$$s_{E\nu,19}^{(m)} = \frac{102.3 - 70.3}{108.5 - 70.3} * 228 + \frac{108.5 - 102.3}{108.5 - 70.3} * 50 = 199, \text{ και } X_{E\nu,19}^{(m)} \text{ αντιστοιχεί στο διάστημα τιμών } [-110, -310].$$

Συναίσθημα	Ενεργοποίηση	Γωνιακή Μέτρηση	«Προφίλ»
Φοβισμένος	4.9	70.3	$F_3 \in [400, 560], F_5 \in [-240, -160], F_{19} \in [-630, -570], F_{20} \in [-630, -570], F_{21} \in [-630, -570], F_{22} \in [-630, -570], F_{31} \in [260, 340], F_{32} \in [260, 340], F_{33} \in [160, 240], F_{34} \in [160, 240], F_{35} \in [60, 140], F_{36} \in [60, 140]$
Απογοητευμένος	4	102.3	$F_3 \in [160, 230], F_5 \in [-100, -65], F_{19} \in [-110, -310], F_{20} \in [-120, -315], F_{21} \in [-110, -310], F_{22} \in [-120, -315], F_{31} \in [61, 167], F_{32} \in [57, 160], F_{33} \in [65, 100], F_{34} \in [65, 100], F_{35} \in [25, 60], F_{36} \in [25, 60]$
Λυπημένος	3.9	108.5	$F_{19} \in [-265, -41], F_{20} \in [-270, -52], F_{21} \in [-265, -41], F_{22} \in [-270, -52], F_{31} \in [30, 140], F_{32} \in [26, 134]$

Πίνακας 5.5: «Προφίλ» του συναισθήματος *ενοχή*



Εικόνα 5.7: Μοντέλο προσώπου Poser: Απεικόνιση «προφίλ» των συναισθημάτων (α) *φοβισμένος*, (β) *απογοητευμένος* και (γ) *λυπημένος*



Εικόνα 5.8: Μοντέλο προσώπου MPEG-4: Απεικόνιση «προφίλ» των συναισθημάτων (α) φοβισμένος, (β) απογοητευμένος και (γ) λυπημένος

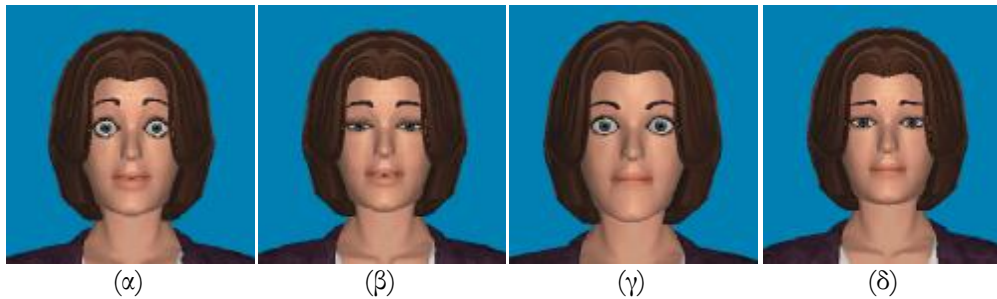
Στον Πίνακα 5.6 φαίνεται το ίδιο παράδειγμα, με αναγραφή όμως όλων των εμπλεκόμενων FAPs.

Φοβισμένος:											
F ₃	F ₅	F ₁₉	F ₂₀	F ₂₁	F ₂₂	F ₃₁	F ₃₂	F ₃₃	F ₃₄	F ₃₅	F ₃₆
400	-240	-630	-630	-630	-630	260	260	160	160	60	60
560	-160	-570	-570	-570	-570	340	340	240	240	140	140
Απογοητευμένος:											
F ₃	F ₅	F ₁₉	F ₂₀	F ₂₁	F ₂₂	F ₃₁	F ₃₂	F ₃₃	F ₃₄	F ₃₅	F ₃₆
160	-100	-110	-120	-110	-120	61	57	65	65	25	25
230	-65	-310	-315	-310	-315	167	160	100	100	60	60
Λυπημένος:											
F ₃	F ₅	F ₁₉	F ₂₀	F ₂₁	F ₂₂	F ₃₁	F ₃₂	F ₃₃	F ₃₄	F ₃₅	F ₃₆
0	0	-265	-270	-265	-270	30	26	0	0	0	0
0	0	-41	-52	-41	-52	140	134	0	0	0	0

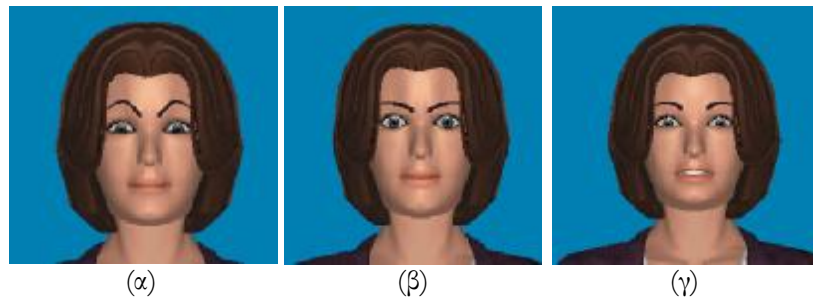
Πίνακας 5.6: Προφίλ των βασικών συναισθημάτων φοβισμένος και λυπημένος και του παραγόμενου ενδιάμεσου συναισθήματος απογοητευμένος

Με βάση τη διαδικασία που περιγράφηκε παραπάνω, η απογοήτευση (Εικόνες 5.9β, 5.9γ) και η ενοχή (Εικόνες 5.11β) δημιουργήθηκαν με χρήση των εκφράσεων φόβος (Εικόνες 5.9α, 5.11α) και λύπη (Εικόνες 5.9γ, 5.11γ), ενώ για την καχυποψία (Εικόνα 5.10β) χρησιμοποιήθηκαν οι εκφράσεις θυμός (Εικόνα 5.10α) και απέχθεια (Εικόνα 5.10γ). Από τις

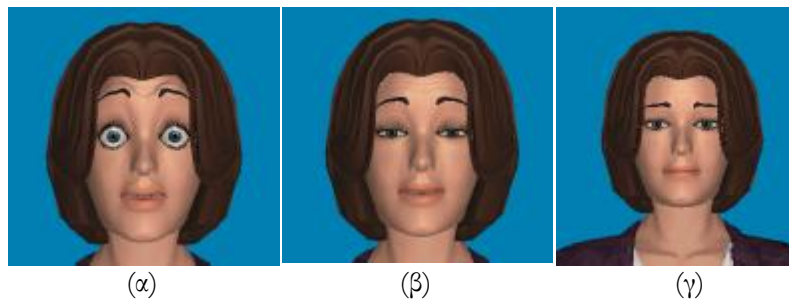
προκύπτουσες εκφράσεις, η έκφραση της **Εικόνας 5.9β** εγκρίνεται, ενώ η **5.9γ** απορρίπτεται, μετά από εισήγηση των ειδικών.



Εικόνα 5.9: Προφίλ για (α) φόβος, (β-γ) απογοήτευση, (δ) λύπη



Εικόνα 5.10: Προφίλ για (α) θυμός, (β) καχυποψία, (γ) απέχθεια



Εικόνα 5.11: Προφίλ για (α) φόβος, (β) ενοχή, (γ) λύπη

5.4. Παράμετροι Σύνθεσης Εκφράσεων Προσώπου

Στην παρούσα ενότητα περιγράφεται ο τρόπος υπολογισμού και αληθοφανούς απεικόνισης των εκφράσεων του προσώπου, γενικά, αλλά και ειδικότερα μέσω ενός MPEG-4 αποκωδικοποιητή [60].

Έπειτα από ανάλυση εκφράσεων προέκυψε ότι μία έκφραση προσδιορίζεται από τις χρονικές παραμέτρους και από τις μεταβολές όλων των εμπλεκόμενων FAPs. Γνωρίζοντας το χρόνο έναρξης και τη διάρκεια της κάθε έκφρασης, το επόμενο στάδιο περιλαμβάνει τον υπολογισμό της εξέλιξης της έντασης της έκφρασης. Ως ένταση μίας έκφρασης νοείται το

εύρος των κινήσεων των χαρακτηριστικών του προσώπου, το οποίο είναι χρονικά μεταβαλλόμενο.

Κάθε έκφραση χαρακτηρίζεται από τέσσερις χρονικές παραμέτρους:

«επίθεση» (A-attack): είναι ο χρόνος που χρειάζεται, ώστε, με αφετηρία ένα ανέκφραστο πρόσωπο, η έκφραση να φτάσει στη μέγιστη έντασή της.

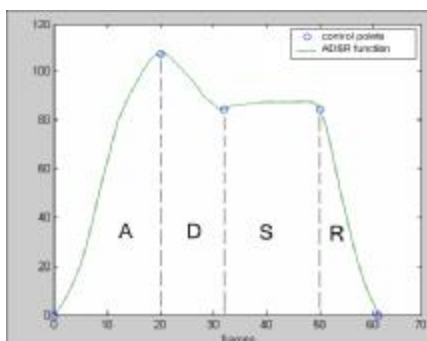
«παρακμή» (D-decay): ο χρόνος κατά τον οποίο η ένταση της έκφρασης ελαφρώς μειώνεται, μέχρι να αποκτήσει μία σταθερή τιμή.

«συντήρηση» (S-sustain): ο χρόνος κατά τον οποίο η έκφραση διατηρείται, συνήθως αντιπροσωπεύει την περισσότερη ορατή φάση μίας έκφρασης.

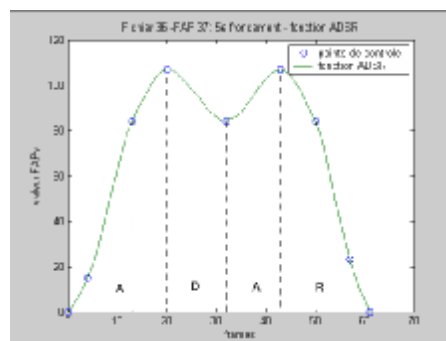
«απεμπλοκή» (R-release): ο χρόνος που μεσολαβεί από τη στιγμή της μέγιστης έντασης της έκφρασης μέχρι το πρόσωπο να ξαναγίνει ανέκφραστο.

Οι παραπάνω παράμετροι διαφέρουν από έκφραση σε έκφραση. Για παράδειγμα η λύπη χαρακτηρίζεται από μακρό χρόνο απεμπλοκής (η έκφραση χρειάζεται περισσότερο χρόνο για να «εξαφανιστεί»), ενώ η έκπληξη έχει μικρό χρόνο επίθεσης.

Για κάθε έκφραση που αναλύουμε είναι δυνατό να δημιουργήσουμε μία αλληλουχία ADSR, ώστε να μοντελοποιήσουμε τη χρονική εξέλιξη του κάθε FAP.



Εικόνα 5.12: Αλληλουχία ADSR για την έκπληξη



Εικόνα 5.13: Αλληλουχία ADSR για το FAP37 στην έκφραση της έκπληξης

5.4.1. Παράμετροι εκφραστικότητας για το πρόσωπο

Γίνεται μία προσπάθεια μέτρησης της εκφραστικότητας μίας έκφρασης με χρήση έξι χαρακτηριστικών που περιγράφονται στη συνέχεια [61]:

Γενική Ενεργοποίηση: μέτρηση δραστηριότητας (μέτρηση μετακίνησης) με διάφορα εκφραστικά μέσα, κατά τη διάρκεια μίας συνομιλίας (π.χ. ταυτόχρονη χρήση έκφρασης προσώπου και χειρονομιών κατά τη διάρκεια συνομιλίας).

Χωρική έκταση: εύρος της κίνησης (π.χ. ο χώρος που χρησιμοποιείται από το σώμα ή το πλάτος της ανύψωσης του φρυδιού)

Χρόνος: διάρκεια της κίνησης (π.χ. γρήγορη ή διαρκής ενέργεια)

Ρευστότητα: ομαλότητα και συνοχή της συνολικής κίνησης (π.χ. ομαλή, «αέρινη» έναντι ξαφνικής / σπασμωδικής)

Δύναμη/Ενέργεια: δυναμικές ιδιότητες της κίνησης (π.χ. αδύναμη / χαλαρή κίνηση έναντι πομπώδους / έντονης)

Επαναληπτικότητα: τάση προς ρυθμική επανάληψη συγκεκριμένης κίνησης κάποιου εκφραστικού μέσου.

Η γενική ενεργοποίηση παίρνει διακριτές κυμαινόμενες τιμές στο διάστημα $[0, 1]$, ενώ οι άλλες παράμετροι παίρνουν επίσης κυμαινόμενες τιμές στο διάστημα $[-1, 1]$, όπου η τιμή 0 αντιστοιχεί σε αυθόρμητες ενέργειες του πράκτορα, χωρίς έλεγχο εκφραστικότητας.

Αναλυτικότερα:

Χωρική έκταση

Αυτή η παράμετρος καθορίζει το εύρος της φυσικής μετατόπισης του κάθε FAP που εμπλέκεται στη διαδικασία της έκφρασης. Στην ουσία αντιστοιχεί σε έναν παράγοντα που μεταβάλλει την καμπύλη ADSR κάθε εμπλεκόμενου FAP κατά μήκος του y άξονα.

Χρόνος

Η παράμετρος αυτή συνδέεται με τα χρονικά χαρακτηριστικά της έκφρασης του προσώπου. Στις περισσότερες περιπτώσεις μεταβάλλει την καμπύλη ADSR κατά το x άξονα. Προκειμένου να επιτευχθεί συμβατότητα μεταξύ των εκφράσεων του προσώπου και του λόγου, η μεταβολή της καμπύλης συνοδεύεται από την κατάλληλη μετατόπιση.

Ρευστότητα

Η παράμετρος αυτή καθορίζει τη σύσπαση των μυών του προσώπου. Όσο η κίνηση γίνεται πιο απότομη, αυξάνεται και η ταχύτητα σύσπασης των μυών, δηλαδή παρατηρείται αύξηση της κλίσης των φάσεων A και R της καμπύλης ADSR. Η αντίστροφη διαδικασία μπορεί να χρησιμοποιηθεί για την προσομοίωση ομαλότερων κινήσεων.

Δύναμη

Η παράμετρος αυτή πραγματοποιεί μία μεταβολή κατά μήκος του άξονα y της καμπύλης ADSR για όλες τις εμπλεκόμενες εκφράσεις του προσώπου, ενώ πρέπει να σημειωθεί ότι επηρεάζει και το σχήμα των χειλιών. Σε αρχικό στάδιο χρησιμοποιούνται τρεις διαφορετικές διαβαθμίσεις μυϊκής πίεσης των χειλιών (φυσιολογική, έντονη και ανεπαίσθητη πίεση). Πίεση στα χείλη μπορεί να εμφανιστεί κατά τις εκφράσεις του φόβου και του θυμού, αλλά και κατά την εκφορά ενός χειλικού συμφώνου (π.χ. β) ή όταν αυξάνεται το πλάτος του στόματος, οπότε λεπταίνουν τα χείλη.

Επαναληπτικότητα

Η παράμετρος αυτή δίνει στοιχεία σχετικά με το πόσο συχνά επαναλαμβάνεται μία έκφραση του προσώπου. Μεγαλύτερη τιμή της παραμέτρου μαρτυρά μεγαλύτερο αριθμό επαναλήψεων (πχ. κινήσεις κεφαλής ή ανασήκωμα φρυδιών κατά τη διάρκεια συνομιλίας).

5.4.1.1. Σύνδεση των FAPs με τις παραμέτρους εκφραστικότητας

Προκειμένου να επιτύχουμε αληθοφανή αποτελέσματα, η εφαρμογή των προφίλ που έχουν δημιουργηθεί για τη σύνθεση εκφράσεων πρέπει να συνδυάζει και τις παραμέτρους εκφραστικότητας. Ειδικότερα:

πώς διαφοροποιείται το εύρος των εμπλεκόμενων FAPs ανάλογα με την έκφραση που θέλουμε να επιτύχουμε

πώς συνδυάζονται τα διάφορα FAPs, δηλαδή κατά ποιον τρόπο η ενεργοποίηση ενός FAP επιδρά στην ενεργοποίηση οποιουδήποτε άλλου.

μετάβαση από τη μία έκφραση προσώπου στην άλλη- δηλαδή πώς μεταβάλλεται η τιμή που λαμβάνει ένα FAP κατά την απευθείας μετάβαση μεταξύ δύο διαδοχικών εκφράσεων προσώπου.

Οι παράμετροι που αναλύθηκαν στην παρούσα ενότητα, θεωρούνται απαραίτητες και στο σχεδιασμό συναισθημάτων στον MPEG-4 αποκωδικοποιητή Greta (Ενότητα 7.1).

5.5. Αλληλεπίδραση μεταξύ FAPs

Στους αποκωδικοποιητές MPEG-4 που έχουν υλοποιηθεί μέχρι τώρα, το πρόσωπο έχει δημιουργηθεί σύμφωνα με τις αρχές της ανθρώπινης ανατομίας, όμως το κάθε FAP, ορθώς ακολουθώντας τον ορισμό του προτύπου, ενεργοποιείται χωριστά. Θεωρητικά λοιπόν μπορεί ο χρήστης να ορίσει ανύψωση του εσωτερικού άκρου του φρυδιού και συγχρόνως να αφήσει εντελώς ακίνητο το μέσο του ιδίου φρυδιού. Το μοντέλο του MPEG-4 αποτελείται από πολύγωνα και για την εφαρμογή των FAPs λαμβάνει υπόψιν τη στατική γεωμετρία του προσώπου στη φυσική κατάσταση, ενώ στηρίζει τον τρόπο κίνησης των χαρακτηριστικών στα FAT (Κεφάλαιο 3). Παρ' όλα αυτά κρίνεται απαραίτητο να δημιουργηθεί μία αντιστοιχία ανάμεσα στα FAPs, δηλαδή δεδομένης της χρήσης συγκεκριμένου FAP να είναι γνωστό ποια άλλα ενεργοποιούνται. Από ανθρωπολογικές μελέτες [63], [94] αλλά και από επαλήθευση των συνθετικών αποτελεσμάτων που προέκυψαν από εφαρμογή των προφίλ (Πίνακας 5.1), προκύπτει ο Πίνακας 5.7.

FAPs	Υποχρεωτικώς Ενεργοποιούμενα FAPs	Προαιρετικώς Ενεργοποιούμενα FAPs
Περιοχή ματιών-φρυδιών		
close_t_l_eyelid (F ₁₉)		close_b_l_eyelid (F ₂₁)
close_b_l_eyelid (F ₂₁)		close_t_l_eyelid (F ₁₉)
close_t_r_eyelid (F ₂₀)		close_b_r_eyelid (F ₂₂)
close_b_r_eyelid (F ₂₂)		close_t_r_eyelid (F ₂₀)
raise_l_i_eyebrow (F ₃₁)	raise_l_m_eyebrow (F ₃₃)	raise_l_o_eyebrow (F ₃₅)
raise_r_i_eyebrow (F ₃₂)	raise_r_m_eyebrow (F ₃₄)	raise_r_o_eyebrow (F ₃₆)
raise_l_m_eyebrow (F ₃₃)	raise_l_i_eyebrow (F ₃₁)	raise_l_o_eyebrow (F ₃₅)
raise_r_m_eyebrow (F ₃₄)	raise_r_i_eyebrow (F ₃₂)	raise_r_o_eyebrow (F ₃₆)
raise_l_o_eyebrow (F ₃₅)	raise_l_m_eyebrow (F ₃₃)	raise_l_i_eyebrow (F ₃₁)
raise_r_o_eyebrow (F ₃₆)	raise_r_m_eyebrow (F ₃₄)	raise_r_i_eyebrow (F ₃₂)
squeeze_l_eyebrow (F ₃₇)	raise_l_i_eyebrow (F ₃₂), raise_l_m_eyebrow (F ₃₄)	raise_l_o_eyebrow (F ₃₆)
squeeze_r_eyebrow (F ₃₈)	raise_r_i_eyebrow (F ₃₂), raise_r_m_eyebrow (F ₃₄)	raise_r_o_eyebrow (F ₃₆)
Περιοχή στόματος		
open_jaw (F ₃)		raise_b_midlip (F ₅) raise_b_lip_lm (F ₁₀) raise_b_lip_rm (F ₁₁) raise_b_lip_lm_o (F ₅₇) raise_b_lip_rm_o (F ₅₈)
lower_t_midlip (F ₄)	lower_t_lip_lm (F ₈) lower_t_lip_rm (F ₉)	
raise_b_midlip (F ₅)	raise_b_lip_lm (F ₁₀) raise_b_lip_rm (F ₁₁)	
stretch_l_cornerlip (F ₆)	stretch_l_cornerlip_o (F ₅₃)	stretch_r_cornerlip (F ₇) stretch_r_cornerlip_o (F ₅₄) lift_l_cheek (F ₄₁)
stretch_r_cornerlip (F ₇)	stretch_r_cornerlip_o (F ₅₄)	stretch_l_cornerlip (F ₆) stretch_l_cornerlip_o (F ₅₃) lift_r_cheek (F ₄₂)
lower_t_lip_lm (F ₈)	lower_t_midlip (F ₄)	lower_t_lip_rm (F ₉)
lower_t_lip_rm (F ₉)	lower_t_midlip (F ₄)	lower_t_lip_lm (F ₈)
raise_b_lip_lm (F ₁₀)	raise_b_midlip (F ₅)	raise_b_lip_rm (F ₁₁)
raise_b_lip_rm (F ₁₁)	raise_b_midlip (F ₅)	raise_b_lip_lm (F ₁₀)
lift_l_cheek (F ₄₁)	stretch_l_cornerlip (F ₆)	close_b_l_eyelid (F ₂₁)
lift_r_cheek (F ₄₂)	stretch_r_cornerlip (F ₇)	close_b_r_eyelid (F ₂₂)
stretch_l_cornerlip_o (F ₅₃)	stretch_l_cornerlip (F ₆)	lift_l_cheek (F ₄₁)
stretch_r_cornerlip_o (F ₅₄)	stretch_r_cornerlip (F ₇)	lift_r_cheek (F ₄₂)
lower_t_lip_lm_o (F ₅₅)	lower_t_midlip (F ₄)	lower_t_lip_rm_o (F ₅₆)
lower_t_lip_rm_o (F ₅₆)	lower_t_midlip (F ₄)	lower_t_lip_lm_o (F ₅₅)
raise_b_lip_lm_o (F ₅₇)	raise_b_midlip (F ₅)	raise_b_lip_rm_o (F ₅₈)
raise_b_lip_rm_o (F ₅₈)	raise_b_midlip (F ₅)	raise_b_lip_lm_o (F ₅₇)
raise_l_cornerlip_o (F ₅₉)	lift_l_cheek (F ₄₁)	
raise_r_cornerlip_o (F ₆₀)	lift_r_cheek (F ₄₂)	

Πίνακας 5.7: Αλληλεπίδραση μεταξύ FAPs

5.6. Προφίλ των βασικών συναισθημάτων σε σχέση με τη Θεωρία Αποτίμησης (Appraisal)

5.6.1. Εκφράσεις Προσώπου των ECAs και Μοντέλο Μονάδων Επεξεργασίας (component process model)

Όταν δουλεύεις με Πράκτορες με σώμα και δυνατότητα ομιλίας (ECAs), είναι απαραίτητη η προσέγγιση και η κατανόηση της διαδικασίας εξαγωγής και σχηματισμού της τελικής έκφρασης από τον άνθρωπο, διαδικασία που πρέπει να προηγηθεί της διαδικασίας μοντελοποίησης της ίδιας έκφρασης για έναν εικονικό πράκτορα. Αν και κάποιος θα περίμενε η διαδικασία αυτή να μην έχει σχέση με τη μοντελοποίηση της τελικής έκφρασης, η οποία θα μπορούσε να στηρίζεται σε γενικά πρότυπα, σχετικές έρευνες οδήγησαν στο συμπέρασμα ότι η μοντελοποίηση σχετίζεται με αυτό που θέλουμε να κάνει ο πράκτορας. Για παράδειγμα η προσέγγιση που λαμβάνει υπόψη τις αποστάσεις (dimensional approach) για τη μοντελοποίηση εκφράσεων, ταιριάζει πιο πολύ στην περίπτωση της αναγνώρισης εκφράσεων, π.χ. ανίχνευση θυμού. Αν θέλουμε όμως να έχουμε πιο αληθοφανή διαδραστικότητα και να μην επηρεάζεται κάθε φορά η έκφραση από την αντίστοιχη ενέργεια, είναι απαραίτητη η γνώση της χρονικής εξέλιξης μίας έκφρασης και πώς αυτή εντάσσεται στο γενικότερο περιβάλλον.

Ο Scherer ανέπτυξε ένα μοντέλο αξιολόγησης συναισθήματος, σύμφωνα με το οποίο τα συναισθήματα αντιμετωπίζονται ως αποτέλεσμα μίας ακολουθίας από ενδιάμεσα στάδια [70], [71]. Το μοντέλο του Scherer παρέχει προβλέψεις για τις εκφράσεις που μεσολαβούν μέχρι την τελική, ενώ προσπαθεί να προβλέψει και την τελική, βασιζόμενο σε ελέγχους αποτίμησης (appraisal checks) διαφορετικά ορισμένων συνιστωσών. Προκειμένου οι προβλέψεις αυτές να μπορούν να χρησιμοποιηθούν από έναν πράκτορα σε ένα δυναμικό περιβάλλον, χρησιμοποιήθηκε η προαναφερθείσα μεθοδολογία και συνδυάστηκε με τα μοντέλα των εκφράσεων που προέκυψαν από άλλα μοντέλα.

Σύμφωνα με διάφορες γνωστικές θεωρίες για το συναίσθημα, τα συναισθήματα συνδέονται στενά με την αντίστοιχη κάθε φορά κατάσταση (βιωμένη ή φανταστική) του πράκτορα. Ειδικότερα, τα συναισθήματα συνδέονται με τη νοητική διαδικασία που κατατάσσει τα στοιχεία μίας κατάστασης σε θετικά και αρνητικά. Αυτή η διαδικασία γενικά χαρακτηρίζεται ως «αποτίμηση» (appraisal). [64], [67], [68]. Η αποτίμηση δεν είναι υποχρεωτικά συνειδητή, οπότε η διαδικασία αποτίμησης μπορεί να προκύψει και ασυνείδητα, ακολουθώντας διαφορετικές μεθόδους παρουσίασης των ερεθισμάτων [69].

Η αποτίμηση (appraisal) είναι η ακολουθία των ενδιάμεσων σταδίων ανάλυσης των ερεθισμάτων (Stimulus Evaluation Checks (SECs)), η οποία αντιπροσωπεύει το μικρότερο δυνατό σύνολο των, απαραίτητων για τη διαφοροποίηση των συναισθηματικών καταστάσεων, κριτηρίων.

Τα ενδιάμεσα στάδια SECs μπορούν να ομαδοποιηθούν ως εξής: 1) Ανίχνευση σχετικότητας: περιλαμβάνει έλεγχο καινοτομίας, εγγενούς ευχαρίστησης και σχετικότητας στόχου 2) Αξιολόγηση επίπτωσης: περιλαμβάνει αιτιώδη έλεγχο απόδοσης, απόκλισης από την υπάρχουσα προσδοκία και έλεγχο επείγουσας ανάγκης 3) Δυνατότητα αντιμετώπισης κατάστασης (μπορεί το περιστατικό να ελεγχθεί?) 4) Κανονιστική αξιολόγηση σπουδαιότητας: περιλαμβάνει εσωτερικό και εξωτερικό έλεγχο προτύπων. Μια σημαντική υπόθεση της θεωρίας SEC του Scherer είναι ότι η ακολουθία των ελέγχων και των ομάδων είναι καθορισμένη. Εντούτοις, αυτό δεν αποκλείει την παράλληλη επεξεργασία καθώς, θεωρητικά, όλα τα SECs επεξεργάζονται ταυτόχρονα.

Μοντελοποιώντας τις μεμονωμένες διαφορές (συμπεριλαμβανομένης και της διαφοροποίησης της προσωπικότητας) στη γένεση των συναισθηματικών διαδικασιών, η θεωρία αξιολόγησης είναι πιθανώς ο καλύτερος τρόπος να αποδοθεί ρεαλιστικά ένα συναισθηματικό ακόμη και όταν αυτό δεν είναι εύκολο ([66]).

Οι απεικονίσεις των διαφόρων συναισθηματικών καταστάσεων που χρησιμοποιούν αυτό το μοντέλο ερμηνεύονται από την αξιολόγηση της πρότερης κατάστασης. Η αξιολόγηση, με τη σειρά της, προέκυψε από τη διαφορετική φύση (διαφορές προσωπικότητας και χρονικής στιγμής) των συναισθηματικών αποκρίσεων, συνδυάζοντας τις, ως επί το πλείστον αντικρουόμενες, αυτόματες/ασυναισθητες αξιολογήσεις των συναισθημάτων με τις πιο συνειδητές, διαδικασία που πολλές φορές οδηγεί σε παράλογες πτυχές συναισθημάτων [67]. Για τα πειραματικά αποτελέσματα, εκτός από την Greta, χρησιμοποιήθηκε το εργαλείο XFACE (εργαλείο σύνθεσης τρισδιάστατης έκφρασης προσώπου, σε πραγματικό χρόνο, το οποίο περιλαμβάνει τις κινήσεις της κεφαλής και των ματιών) των Wehrle et al. [72], [77]. Μέχρι σήμερα, η, σχετική με τη θεωρία του Scherer, ανάλυση πραγματοποιούνταν ως επί το πλείστον σε στατικές εικόνες. Πλέον η διαδικασία της σύνθεσης βίντεο θέτει το πρόβλημα της χρονικής εξέλιξης των εκφράσεων. Εξετάζονται διάφορες μέθοδοι μετάβασης μεταξύ των ενδιάμεσων σταδίων της κάθε έκφρασης.

Η θεωρία αξιολόγησης υποστηρίζει μία *αθροιστική επίδραση* των ενδιάμεσων σταδίων της τελικής έκφρασης του συναισθήματος. Η επίδραση αυτή χρειάζεται εμπειρική έρευνα προκειμένου να καθορίσει την κατάλληλη μέθοδο δημιουργίας της έκφρασης. Διακρίνονται δύο τρόποι απεικόνισης της εξελικτικής διαδικασίας που καταλήγει στη ζητούμενη έκφραση. Η πρώτη μέθοδος απεικονίζει την τελική έκφραση χειριζόμενη αθροιστικά τα ενδιάμεσα στάδια (η τελική έκφραση προκύπτει από την πρόσθεση των προηγούμενων), ενώ η δεύτερη στηρίζεται στη διαδοχική μετατροπή των ενδιάμεσων σταδίων, μέχρι να προκύψει η τελική έκφραση. Και οι δύο μέθοδοι χρειάζονται εμπειρικό έλεγχο της φυσικότητας-αληθοφάνειας της προκύπτουσας έκφρασης. Στην έρευνά μας, η πρώτη μέθοδος χρησιμοποιήθηκε στην περίπτωση του φόβου και η δεύτερη μέθοδος στην περίπτωση του έντονου θυμού.

Δείγματα των αποτελεσμάτων και των δύο μεθόδων απεικονίζονται στις **Εικόνες 5.14, 5.15, 5.16, 5.17 και 5.18**.

Additive animation (πρώτη μέθοδος)

Στην περίπτωση της πρώτης μεθόδου (φόβος – *Εικόνες 5.14, 5.16*), η τελική έκφραση προκύπτει από την άθροιση, κάθε φορά, των τιμών των AUs της συγκεκριμένης έκφρασης με τις τιμές των AUs του προηγούμενου σταδίου.

Η μέθοδος αυτή παρουσιάζει προβλήματα στις περιπτώσεις κατά τις οποίες οι διαδοχικές εκφράσεις παρουσιάζουν αντικρουόμενες απεικονίσεις. Για παράδειγμα, στην περίπτωση του έντονου θυμού, κατά το πρώτο ενδιάμεσο στάδιο ([65]) μεταξύ άλλων ανασηκώνονται τα φρύδια, ενώ το επόμενο στάδιο απαιτεί χαμηλωμένα φρύδια. Αυτή η αντίφαση έχει ως αποτέλεσμα την προβληματική απεικόνιση του συναισθήματος, οπότε στην περίπτωση αυτή προτιμάται η δεύτερη μέθοδος απεικόνισης των εκφράσεων.

Sequential animation (δεύτερη μέθοδος)

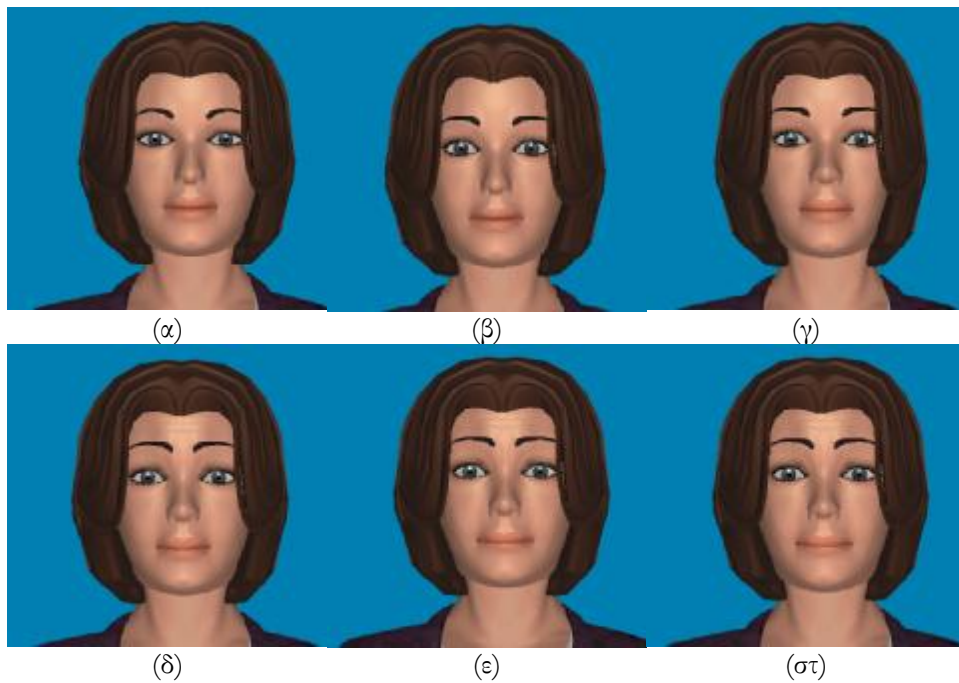
Με τη μέθοδο αυτή, όλα τα ενδιάμεσα στάδια απεικονίζονται διαδοχικά. Αυτό πραγματοποιείται είτε με την παρεμβολή της ουδέτερης έκφρασης μεταξύ των διαδοχικών σταδίων είτε με τη μετάβαση από το ένα στάδιο στο άλλο θεωρώντας ως βάση τις κοινές μεταβολές. Η πρώτη προσέγγιση έχει λιγότερο αληθοφανή αποτελέσματα σε σχέση με τη δεύτερη, της οποίας το αποτέλεσμα είναι πιο φιλικό στο μάτι, αλλά δεν αποτελεί πραγματική προσομοίωση της ανθρώπινης διαδικασίας παραγωγής μίας έκφρασης.

5.6.2. Σύνθεση Εκφράσεων Σύμφωνα με τη Θεωρία Αξιολόγησης

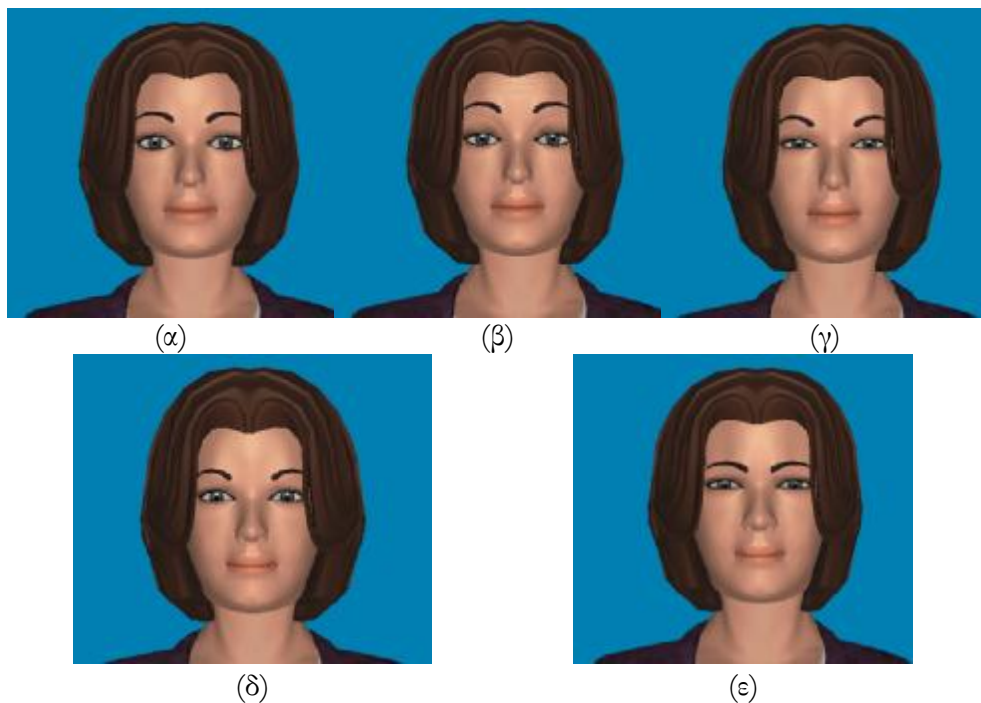
Με βάση τις προβλέψεις της θεωρίας του Scherer, χρησιμοποιήθηκε ο MPEG-4 αποκωδικοποιητής GretaPlayer. Η απεικόνιση βασίστηκε στην αντιστοίχιση των Action Units του Ekman με τα FAPs του MPEG-4 (*Πίνακας 4.1*).

Για τα ενδιάμεσα στάδια των εκφράσεων του έντονου θυμού, του φόβου και της λύπης, δημιουργήσαμε τις **Εικόνες 5.16, 5.17 και 5.18**, χρησιμοποιώντας την ευελιξία και την εκφραστικότητα του αποκωδικοποιητή MPEG-4 Xface [77].

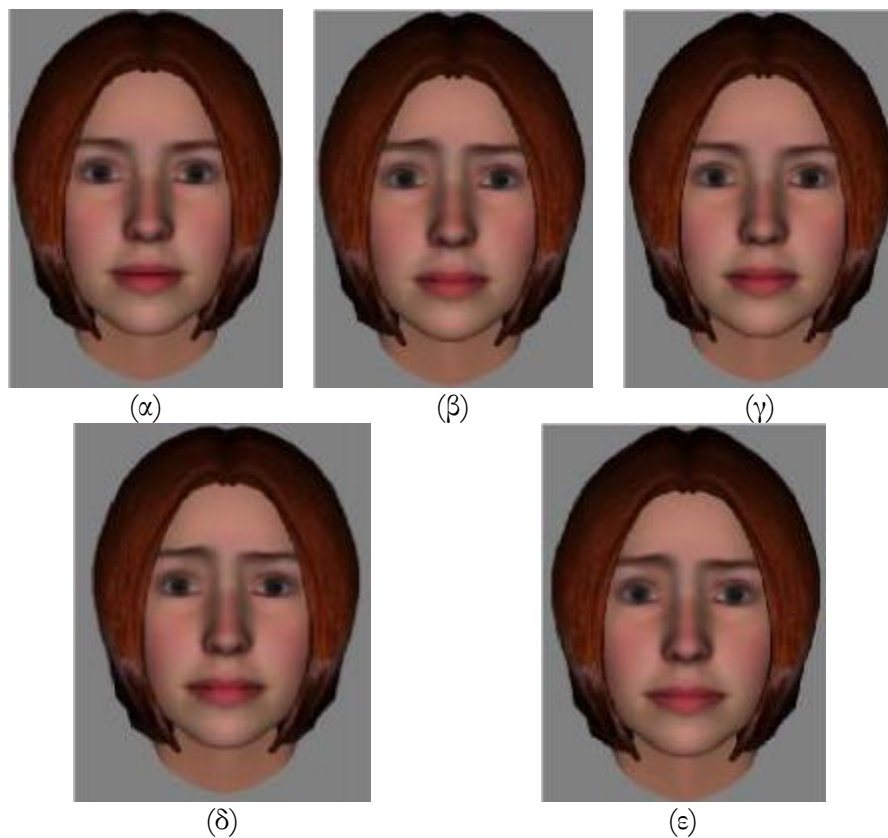
Από τη διαδικασία αυτή προκύπτουν ζητήματα σχετικά με τη χρονική εξέλιξη. Η αθροιστική μέθοδος προσδιορισμού της τελικής έκφρασης περιλαμβάνει είτε διαδοχική είτε αθροιστική διαχείριση των ενδιάμεσων σταδίων. Στο παρόν διδακτορικό υιοθετήθηκαν και οι δύο μέθοδοι, η πρώτη για την απεικόνιση του έντονου θυμού και η δεύτερη για την απεικόνιση του φόβου και της λύπης. Και στις τρεις περιπτώσεις παραμένουν ανοικτά τα ζητήματα του συγχρονισμού, του χρόνου έναρξης κάθε ενδιάμεσου σταδίου και της ταχύτητας μετάβασης μεταξύ αυτών. Μέρος των αποτελεσμάτων απεικονίζεται στις **Εικόνες 5.16, 5.17 και 5.18**.



Εικόνα 5.14: Greta: Πρόβλεψη ενδιάμεσων σταδίων των εκφράσεων προσώπου σύμφωνα με τη θεωρία αξιολόγησης του Scherer για την περίπτωση του φόβου- (α) ουδέτερο (*neutral*), (β) ξαφνικό συμβάν (*novelty-sudden*), (γ) δυσαρέσκεια (*unpleasant*), (δ) αντιφατικό (*discrepant*), (ε) παρεμπόδιση στόχου (*goal obstructive*), (στ) χαμηλός έλεγχος -τελική έκφραση-φόβος (*low control-final expression –fear*). Κάθε έκφραση προκύπτει από την «πρόσθεση» των AUs των προηγούμενων εκφράσεων και της τελικής.



Εικόνα 5.15: Greta: Πρόβλεψη ενδιάμεσων σταδίων των εκφράσεων προσώπου σύμφωνα με τη θεωρία αξιολόγησης του Scherer για την περίπτωση του έντονου θυμού- (α) ουδέτερο (*neutral*), (β) έντονη απόκριση σε ξαφνικό συμβάν (*novelty- high*), (γ) παρεμπόδιση στόχου (*goal obstructive*), (δ) υψηλός έλεγχος/ μεγάλη ένταση (*control high/power high*), (ε) τελική έκφραση-έντονος θυμός (*final expression – hot anger*).



Εικόνα 5.16: *Xface*: Πρόβλεψη ενδιάμεσων σταδίων των εκφράσεων προσώπου σύμφωνα με τη θεωρία αξιολόγησης του Scherer για την περίπτωση του φόβου- (α) ουδέτερο (*neutral*), (β) ξαφνικό συμβάν (*novelty-sudden*), (γ) δυσαρέσκεια (*unpleasant*), (δ) παρεμπόδιση στόχου (*goal obstructive*), (ε) χαμηλός έλεγχος-τελική έκφραση-φόβος (*low control-final expression –fear*). Κάθε έκφραση προκύπτει από την «πρόσθεση» των *AUs* των προηγούμενων εκφράσεων και της τελικής.





(δ)



(ε)

Εικόνα 5.17: *Xface*: Πρόβλεψη ενδιάμεσων σταδίων των εκφράσεων προσώπου σύμφωνα με τη θεωρία αξιολόγησης του Scherer για την περίπτωση του έντονου θυμού- (α) ουδέτερο (*neutral*), (β) έντονη απόκριση σε ξαφνικό συμβάν (*novelty- high*), (γ) παρεμπόδιση στόχου (*goal obstructive*), (δ) υψηλός έλεγχος/ μεγάλη ένταση (*control high/ power high*), (ε) τελική έκφραση-έντονος θυμός (*final expression – hot anger*).



(α)



(β)



(γ)



(δ)



(ε)

Εικόνα 5.18: *Xface*: Πρόβλεψη ενδιάμεσων σταδίων των εκφράσεων προσώπου σύμφωνα με τη θεωρία αξιολόγησης του Scherer για την περίπτωση της λύπης- (α) ουδέτερο (*neutral*), (β) μικρή εξοικείωση (*familiarity low*), (γ) παρεμπόδιση στόχου (*goal obstructive*), (δ) χαμηλός έλεγχος (*control low*), (ε) τελική έκφραση-λύπη (*final expression – sadness*).

Κεφάλαιο 6

Σώμα - Χειρονομίες

6.1. Ανάλυση Χειρονομιών με Χρήση Παραμέτρων Εκφραστικότητας

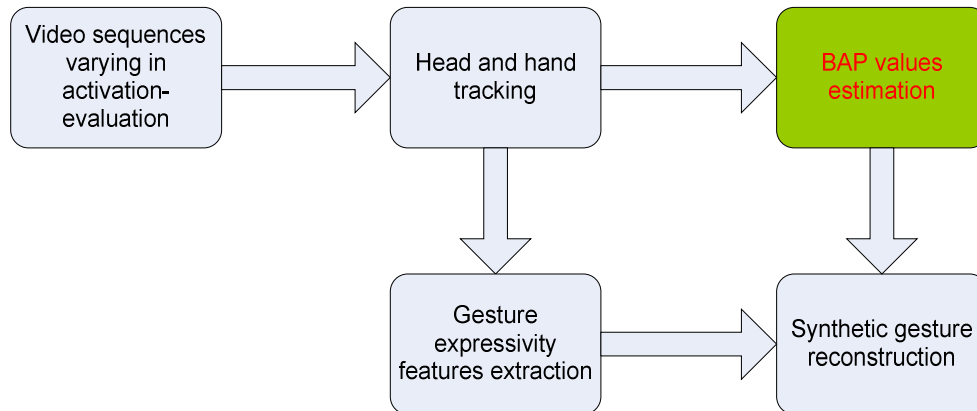
6.1.1. Σύστημα Ανάλυσης Χειρονομιών

Τόσο η ανάλυση όσο και η σύνθεση χειρονομιών αποτελούν ένα πολύ σημαντικό κομμάτι της Επικοινωνίας Ανθρώπου-Μηχανής (HCI). Μερικές φορές, μία απλή χειρονομία, όπως η τοποθέτηση των χεριών στα αυτιά, μπορεί να περάσει το μήνυμα ότι το άτομο κουράστηκε από αυτά που ακούει, και αυτό εκφράζεται καλύτερα με μία χειρονομία παρά με οποιαδήποτε φράση. Για να επωφεληθούμε από τη χρήση των χειρονομιών στην Επικοινωνία Ανθρώπου-Μηχανής, κρίνεται απαραίτητη η εύρεση τρόπων ερμηνείας χειρονομιών από υπολογιστή.

Δεδομένου ότι η επεξεργασία της οπτικής πληροφορίας παρέχει ικανοποιητικά στοιχεία προκειμένου να εξάγουμε συμπεράσματα για τις καταστάσεις ενός κινούμενου αντικειμένου στο χρόνο, οι τεχνικές που βασίζονται στην οπτική πληροφορία παρέχουν τουλάχιστον επαρκείς εναλλακτικές λύσεις για τη σύλληψη και ερμηνεία της κίνησης των χεριών. Συγχρόνως, οι εφαρμογές μπορούν να ωφεληθούν από το γεγονός ότι τα οπτικά συστήματα μπορούν να είναι πολύ αποδοτικά χωρίς να επηρεάζουν τη φυσικότητα της αλληλεπίδρασης με το χρήστη. Η ανάλυση των χειρονομιών είναι μία σύνθετη διαδικασία που μπορεί να περιλαμβάνει μοντελοποίηση και ανάλυση κίνησης, αναγνώριση προτύπων και επεξεργασία ψυχολinguιστικών μελετών.

Το σύστημα που δημιουργήθηκε χρησιμοποιεί ως είσοδο ακολουθίες εικόνων και εντοπίζει το κεφάλι και τα χέρια του ηθοποιού, οπότε, με αυτό τον τρόπο, μας δίνεται η δυνατότητα να κάνουμε μία πρώτη εκτίμηση των τιμών των παραμέτρων BAP (Body Animation Parameters) του MPEG-4 για κάθε χειρονομία και να εξάγουμε σημαντικές παραμέτρους εκφραστικότητας. Το σύνολο των αποτελεσμάτων χρησιμοποιούνται στη συνθετική και, όσο το δυνατόν, ρεαλιστική δημιουργία κάθε χειρονομίας.

Το υποσύστημα της δημιουργίας συνθετικών χειρονομιών φαίνεται στην **Εικόνα 6.1**:



Εικόνα 6.1: Συνθετική δημιουργία χειρονομιών

Στο συγκεκριμένο σύστημα χρησιμοποιήθηκαν ως είσοδο πραγματικά βίντεο επτά φοιτητών του Εργαστηρίου Επεξεργασίας Εικόνων, ο καθένας από τους οποίους έκανε επτά διαφορετικές χειρονομίες. Η κάθε χειρονομία εκτελέστηκε αρκετές φορές, δηλώνοντας κάθε φορά μία διαφορετική συναισθηματική κατάσταση. Οι χειρονομίες του πειράματος είναι οι παρακάτω:

- Εξήγησε (explain)
- Θεέ μου! (και τα δύο χέρια πάνω από το κεφάλι) (oh my god (both hands over head))
- Παράτα με (leave me alone)
- Ανάταση χεριού(προσέλκυση προσοχής) (raise hand (draw attention))
- Βαριεστημένος (χέρι κάτω από το πηγούνι) (bored (one hand under chin))
- Χαιρετισμός (wave)
- Χειροκρότημα (clap)

Η κάθε μία από τις παραπάνω χειρονομίες δήλωνε κάθε φορά μία από τις συναισθηματικές καταστάσεις που αναφέρονται στον **Πίνακα 6.1**:

συναισθημα	Τεταρτημόριο του τροχού της Whissel
ουδέτερο	(0,0)
χαρά, έκπληξη	+, +
θυμός	+, -
λύπη	-, -

Πίνακας 6.1: Συναισθήματα υποδηλούμενα μέσω χειρονομιών

Βέβαια μερικά συναισθήματα δεν εκφράστηκαν μέσω συγκεκριμένων χειρονομιών, αφού δεν είχε ιδιαίτερο νόημα, για παράδειγμα, η έκφραση χαράς μέσω της χειρονομίας

που υποδηλώνει βαρεμάρα. Δηλαδή συνολικά οι εκφράσεις που εκτελέστηκαν είναι αυτές που αναγράφονται στον **Πίνακα 6.2**.

Χειρονομία	Τεταρτημόριο του τροχού της Whissel
εξήγησε-explain	(0,0), (+, +), (-, +), (-, -)
Θεέ μου!-oh my god	(+, +), (-, +)
παράτα με-leave me alone	(-, +), (-, -)
ανάταση χεριού -raise hand	(0,0), (+, +), (-, -)
βαριεστημένος -bored	(-, -)
χαιρετισμός-wave	(0,0), (+, +), (-, +), (-, -)
χειροκρότημα-clap	(0,0), (+, +), (-, +), (-, -)

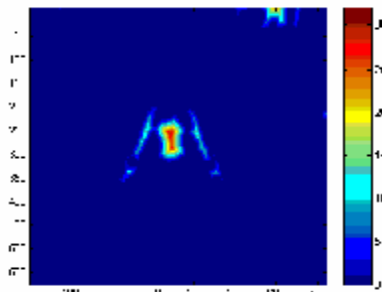
Πίνακας 6.2: Εκτελεσθείσες Χειρονομίες

Τελικά δημιουργήθηκε μία βάση από 140 διαφορετικά βίντεο (7 ηθοποιοί x 20 διαφορετικές εκφράσεις συναισθήματος).

Για τον εντοπισμό της κεφαλής και των χεριών έγινε επιλογή μεταξύ διαφορετικών μεθόδων. Η διαδικασία ανάλυσης κι εντοπισμού της κεφαλής και των χεριών [78]δίνει τελικά για κάθε frame (Εικόνα 6.2) ως αποτέλεσμα μία έγχρωμη μάσκα (Εικόνα 6.3).



Εικόνα 6.2: Στιγμιότυπο ακολουθίας προς εξέταση



Εικόνα 6.3: Έγχρωμη μάσκα

6.1.2. Προσδιορισμός Παραμέτρων Εκφραστικότητας

Για τον προσδιορισμό των παραμέτρων εκφραστικότητας, στη σχετική βιβλιογραφία [73], [74] ορίζονται έξι παράμετροι, η κάθε μία από τις οποίες διαφοροποιείται ανάλογα με το εκφραστικό μέσο στο οποίο εφαρμόζεται. Στην περίπτωση των χειρών, η εκφραστικότητα προσδιορίζεται κατά τις διαφορετικές φάσεις της κάθε χειρονομίας, δηλαδή κατά τη φάση της προετοιμασίας, της κυρίως κίνησης και της διατήρησης της χειρονομίας, καθώς επίσης και από τον τρόπο που συνδυάζονται δύο χειρονομίες [75], [76].

Θεωρούμε έξι παραμέτρους εκφραστικότητας, που, ονομαστικά, είναι οι ίδιες με τις αντίστοιχες του προσώπου (Ενότητα 5.4) και θα αναλυθούν για την περίπτωση του σώματος στην επόμενη ενότητα:

- Γενική Ενεργοποίηση (Overall activation)
- Χωρική έκταση (Spatial extent)
- Χρόνος (Temporal)
- Ρευστότητα (Fluidity)
- Δύναμη/Ενέργεια (Power/Energy)
- Επαναληπτικότητα (Repetitivity)

Γενική Ενεργοποίηση (Overall activation) θεωρείται η «ποσότητα» της κίνησης κατά τη διάρκεια μίας συνομιλίας. Στο σύστημά μας θεωρήσαμε ότι είναι το άθροισμα της νόρμας των διανυσμάτων κίνησης: $OA = \sum_{i=0}^n |\vec{r}(i)| + |\vec{l}(i)|$. Η χωρική έκταση (Spatial extent)

μοντελοποιείται με τη διαστολή ή τη συστολή του χώρου που χρησιμοποιείται για τη χειρονομία μπροστά από τον πράκτορα και υπολογίζεται ως η μέγιστη ευκλίδεια απόσταση της θέσης των δύο χειρών: $SE = \max(|d(\vec{r}(i) - \vec{l}(i))|)$. Επίσης υπολογίζεται και ο μέσος όρος της χωρικής έκτασης για λόγους κανονικοποίησης. Η χρονική παράμετρος της χειρονομίας καθορίζει την ταχύτητα του πιο σημαντικού τμήματος της κίνησης, ενώ συγχρόνως δίνει στοιχεία για το είδος της κίνησης (π.χ. σύντομη ή συνεχή κίνηση). Η ρευστότητα διαφοροποιεί τις ομαλές από τις ξαφνικές κινήσεις, δηλαδή προσπαθεί να απεικονίσει τη συνοχή μεταξύ των μετακινήσεων, οπότε και επιδρά στη συνέχεια της τροχιάς του χεριού, καθώς και στην επιτάχυνση ή επιβράδυνση των άκρων. Για να εξαγάγουμε αυτό το χαρακτηριστικό από τις ακολουθίες εισόδου, υπολογίζουμε το άθροισμα της μεταβολής από τις νόρμες των διανυσμάτων κίνησης. Η Δύναμη ισούται στην ουσία με την πρώτη παράγωγο των διανυσμάτων κίνησης, που υπολογίστηκε στα πρώτα βήματα.

6.2. Σύνθεση Χειρονομιών

6.2.1. Σύνθεση χειρονομιών και εκφράσεων

Το βασικό στοιχείο μίας χειρονομίας είναι το κύριο καρέ που την περιγράφει. Κάθε κύριο καρέ μπορεί να περιλαμβάνει μία ή περισσότερες από τις ακόλουθες παραμέτρους:

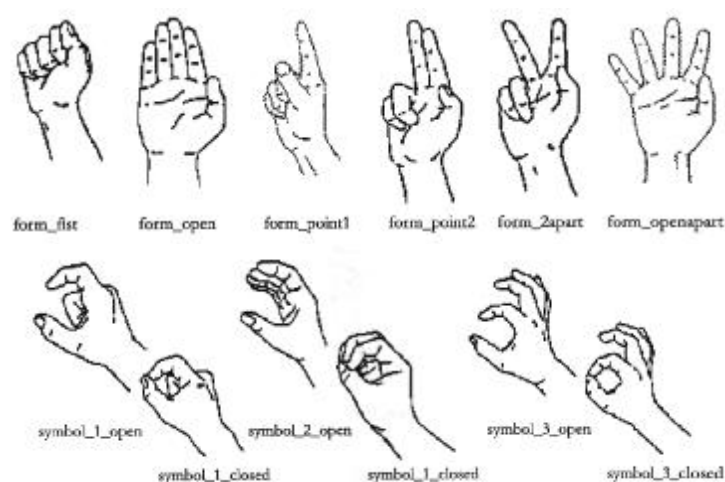
BPAXIONAS: καθορίζει τη διαμόρφωση του κάθε βραχίονα στο χώρο, ουσιαστικά δηλαδή την περιστροφή που εφαρμόζεται στον ώμο και στον αγκώνα. Περιλαμβάνει 3 παραμέτρους για τους 3 άξονες x, y, z, οι οποίες και θα προσδιορίσουν μία περιοχή στο χώρο μπροστά από τον πράκτορα.

ΚΑΡΠΙΟΣ: καθορίζει τον προσανατολισμό της παλάμης με χρήση δύο παραμέτρων: FingerBase (FB) και Palm. Η πρώτη είναι η διεύθυνση του διανύσματος που είναι κατακόρυφο στη βάση των δακτύλων. Η δεύτερη είναι το διάνυσμα που είναι κατακόρυφο στο επίπεδο της παλάμης.

ΧΕΡΙ: επιλέγει για το χέρι μια στατική διαμόρφωση μεταξύ των βασικών μορφών που καθορίζονται από τον McNeill [55].

Μετά τον ορισμό των keyframes, πραγματοποιείται από τον αποκωδικοποιητή MPEG-4 παρεμβολή, ώστε να προκύψει η ολοκληρωμένη χειρονομία.

Η **Εικόνα 6.4** απεικονίζει τα βασικά σχήματα των χεριών:



Εικόνα 6.4: Βασικά Σχήματα Χεριών

6.2.2. Εκφραστικότητα Χειρονομιών

Σε αντιστοιχία με τις παραμέτρους για το πρόσωπο της **Ενότητας 5.4**, έχουν ορισθεί παράμετροι για το σώμα με παρεμφερείς ιδιότητες:

Γενική Ενεργοποίηση: μέτρηση δραστηριότητας (μέτρηση μετακίνησης) με διάφορα εκφραστικά μέσα κατά τη διάρκεια μίας συνομιλίας (π.χ. ταυτόχρονη χρήση έκφρασης προσώπου και χειρονομιών κατά τη διάρκεια συνομιλίας).

Χωρική έκταση: εύρος της κίνησης (π.χ. ο χώρος που χρησιμοποιείται από το σώμα)

Χρόνος: διάρκεια της κίνησης (π.χ. γρήγορη ή διαρκής ενέργεια)

Ρευστότητα: ομαλότητα και συνοχή της συνολικής κίνησης (π.χ., ομαλή, «αέρινη» έναντι ξαφνικής / σπασμωδικής)

Δύναμη/Ενέργεια: δυναμικές ιδιότητες της κίνησης (π.χ. αδύναμη / χαλαρή κίνηση έναντι πομπώδους / έντονης)

Επαναληπτικότητα: τάση προς ρυθμική επανάληψη συγκεκριμένης κίνησης.

Αναλυτικότερα:

Χωρική έκταση

Η βασική ιδέα είναι να επεκταθεί ή να συμπυκνωθεί ο χώρος μπροστά από τον πράκτορα, που χρησιμοποιείται για μία χειρονομία. Αυτός ο χώρος παρουσιάζεται από το διάγραμμα του McNeill [55] σαν ένα σύνολο από τομείς, οι οποίοι χρησιμοποιούνται για τον ορισμό της θέσης των καρπών. Με τη βοήθεια των συντεταγμένων των κέντρων τους, η θέση των τομέων κλιμακώνεται συμμετρικά ή ασύμμετρα με χρήση ενός απλού πίνακα κλιμάκωσης σε ομογενείς συντεταγμένες. Προκειμένου να εντοπίσουμε το σημείο της άρθρωσης για μία χειρονομία, εντοπίζεται ένα σημείο στο δυναμικά προσαρμολζόμενο τεταρτημόριο της κάθε χειρονομίας στο αρχείο ορισμού των χειρονομιών, ώστε στη συνέχεια να υπολογιστεί ο συνδυασμός των γωνιών που απαιτούνται για να επιτευχθεί ο στόχος με αντίστροφη κινηματική (inverse kinematics).

Μία πιο σύνθετη προσέγγιση μπορεί να χρησιμοποιηθεί για το χώρο που «χρειάζονται» εικονικές ή μεταφορικές κινήσεις. Παραδείγματος χάριν, η έννοια της «κυκλικότητας» ή της «σφαιρικότητας» μπορεί να εκφραστεί με τη σχεδίαση μικρών κύκλων στον αέρα από το δείκτη και με σταθερή τη θέση του καρπού ή με την περιστροφή του καρπού και με σταθερό σχήμα χεριού ή ακόμα μετακινώντας ολόκληρο το βραχίονα κυκλικά. Η συγκεκριμένη κατηγορία κινήσεων απαιτεί ολοκληρωμένο έλεγχο σε όλες τις ομάδες βαθμών ελευθερίας του βραχίονα: θέση καρπών, προσανατολισμός παλάμης, θέση δακτύλων.

Τέλος, η γωνία στρώφης αγκώνα (η γωνία μεταξύ ενός διανύσματος που δείχνει προς τα κάτω και του διανύσματος που προκύπτει από την προβολή της θέσης του αγκώνα στη γραμμή που συνδέει τον ώμο με τον καρπό) αλλάζει άμεσα το χώρο που χρησιμοποιεί ο πράκτορας-οι εκτεταμένοι αγκώνες διευρύνουν τη σιλουέτα του πράκτορα.

Χρονικά Χαρακτηριστικά

Όταν εξετάζουμε τα χαρακτηριστικά μίας κίνησης που σχετίζονται με το χρόνο, ερχόμαστε αρχικά αντιμέτωποι με τους περιορισμούς που θέτει η ανάγκη συγχρονισμού του τέλους της κυρίως φάσης της χειρονομίας με την αντίστοιχη φάση της ομιλίας [55], οπότε υπολογίζουμε τους χρόνους για τα frames που προηγούνται ή έπονται σύμφωνα με τους κανόνες μετακίνησης του ανθρώπινου βραχίονα [56] και με το πόσο γρήγορα εκτελούνται οι διάφορες φάσεις της χειρονομίας. Κατά τη φάση προγραμματισμού, η πραγματική μετακίνηση της άρθρωσης του καρπού στο χώρο προσεγγίζεται από γραμμικά τμήματα μέσω χαρακτηριστικών σημείων. Έχοντας στη διάθεσή μας πληροφορίες σχετικά με τα χρονικά χαρακτηριστικά της κάθε φάσης μιας χειρονομίας, μπορούμε να επέμβουμε επιλεκτικά σε κάποια φάση ώστε να διαφοροποιήσουμε τη χειρονομία, π.χ. να την κάνουμε πιο έντονη.

Ρευστότητα

Η παράμετρος αυτή επιδιώκει να συλλάβει τη συνοχή μεταξύ των κινήσεων, οπότε εφαρμόζεται με τροποποίηση της συνέχειας της τροχιάς, καθώς και με επιβράδυνση ή επιτάχυνση των άκρων. Ο κινηματικός έλεγχος επιτυγχάνεται έμμεσα, μέσω της ταχύτητας εκτέλεσης μίας ενέργειας. Για μεταβολή της κίνησης μπορούμε είτε να τροποποιήσουμε τα σημεία ελέγχου των αντίστοιχων καμπυλών Bezier είτε να επιλέξουμε καμπύλες TCB, ώστε να επέμβουμε στην παρεμβολή.

Δύναμη/Ενέργεια

Για να απεικονίσουμε το ποσό ενέργειας και έντασης μίας κίνησης εξετάζουμε τις δυναμικές ιδιότητες των χειρονομιών καθώς και τις φάσεις ακινησίας μεταξύ δύο χειρονομιών. Οι ισχυρές κινήσεις αναμένεται να έχουν μεγαλύτερη επιτάχυνση και επιβράδυνση. Αυτή η συμπεριφορά αποδίδεται μέσω της παραμέτρου έντασης της κινηματικής TCB καμπύλης.

Επαναληπτικότητα

Μερικές φορές, ορισμένες χειρονομίες δεν αποτελούνται αποκλειστικά από μια φάση ενεργοποίησης, μία κύρια φάση και μία τελική, αλλά είναι μία αλληλουχία κυρίων φάσεων, συνήθως πολύ κοντά η μία στην άλλη. Επιπλέον, αυτές οι πολλαπλές κύριες φάσεις παρουσιάζουν συχνά κάποια κοινά χαρακτηριστικά, όπως το ίδιο σχήμα χεριού καθ'όλη τη διάρκεια της χειρονομίας. Στην περίπτωση αυτή ο αποκωδικοποιητής συμπτύσσει το μήκος της κύριας φάσης και το επαναλαμβάνει αρκετές φορές.

Η ανάλυση πραγματικών βίντεο με σκοπό τον προσδιορισμό των συγκεκριμένων παραμέτρων συμβάλλει στη σωστή σύνθεση των χειρονομιών σε έναν αποκωδικοποιητή MPEG-4, αφού όλες οι χειρονομίες βασίζονται στην κίνηση του βραχίονα. Αν επιτύχουμε

την παρακολούθηση της κίνησης των καρπών, μπορούμε στη συνέχεια να προσδιορίσουμε όλες τις υπόλοιπες ιδιότητες, ταχύτητα, επιτάχυνση και αλλαγή κατεύθυνσης μεταξύ άλλων, ώστε να προχωρήσουμε στον προσδιορισμό των έξι παραμέτρων.

6.2.3. Μοντελοποίηση των χειρονομιών με βάση πραγματικές ακολουθίες

Στην παρούσα ενότητα παρουσιάζονται δειγματοληπτικά κάποιες εικόνες από τα δημιουργηθέντα βίντεο, καθώς και χειρονομίες που δημιουργήθηκαν με το λογισμικό POSER και τον MPEG-4 αποκωδικοποιητή Greta.

Στις **Εικόνες 6.5, 6.6** παρουσιάζονται κάποια frames από δύο από τα βίντεο που αναφέρθηκαν στην **Ενότητα 6.1**, που απεικονίζουν τη χειρονομία «Θεέ μου!» (*“oh!my god”*). Οι τιμές των έξι παραμέτρων, όπως αυτές προέκυψαν από το Σύστημα Ανάλυσης που παρουσιάστηκε στην **Ενότητα 6.1**, απεικονίζονται στο διάγραμμα της **Εικόνας 6.7**. Για κάθε μία από τις έξι περιπτώσεις, η πρώτη ράβδος αντιστοιχεί στο άτομο της **Εικόνας 6.5**, ενώ η δεύτερη στο άτομο της **Εικόνας 6.6**. Εδώ πρέπει να σημειωθεί ότι οι τιμές που παρουσιάζονται στο διάγραμμα είναι κανονικοποιημένες.

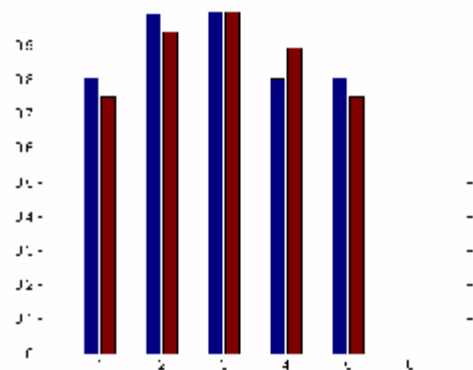
Οι τιμές των είκοσι χειρονομιών για τις παραμέτρους : α)ολική ενεργοποίηση, β) χωρική έκταση, γ) ρευστότητα, δ) δύναμη/ενέργεια απεικονίζονται στις **Εικόνες 6.8 (α-δ)**, ενώ η **Εικόνα 6.9** απεικονίζει τις μέσες τιμές των έξι παραμέτρων εκφραστικότητας για τρία άτομα και οι **Εικόνες 6.10 (α-β)** απεικονίζουν αντίστοιχα τις μέσες τιμές της *Ολικής Ενεργοποίησης* και της *Δύναμης* για θετικές και αρνητικές τιμές της ενεργοποίησης (activation). Όπως ήταν αναμενόμενο, οι τιμές των χειρονομιών που τοποθετούνται στο πρώτο και το δεύτερο τεταρτημόριο του Τροχού Συναισθημάτων της Whissel (θετικές τιμές activation) είναι υψηλότερες.



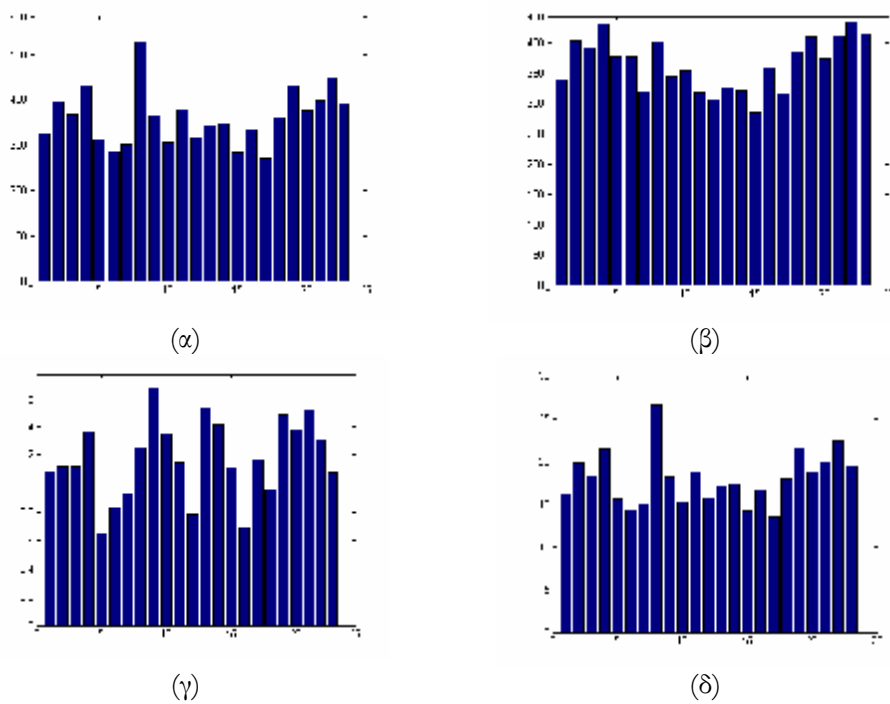
Εικόνα 6.5: Frames του υποκειμένου 4 από τη χειρονομία «Θεέ μου!»



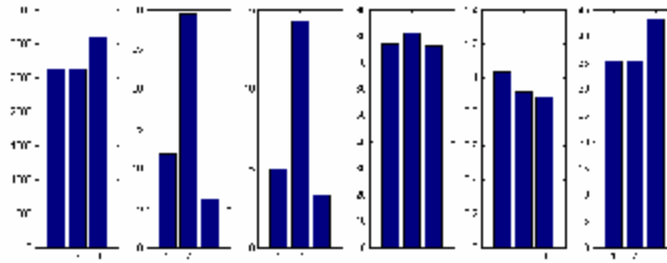
Εικόνα 6.6: Frames του υποκειμένου 21 από τη χειρονομία «Θεέ μου!»



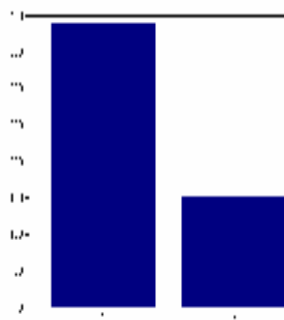
Εικόνα 6.7: Πειραματικά αποτελέσματα για τις έξι παραμέτρους της Ενότητας 6.1.2



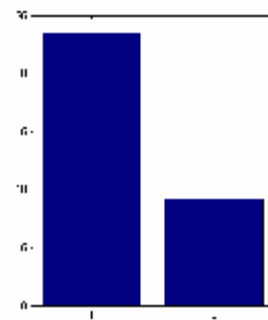
Εικόνα 6.8: Πειραματικά αποτελέσματα για τις είκοσι χειρονομίες για τις παραμέτρους : α)ολική ενεργοποίηση, β) χωρική έκταση, γ) ρευστότητα, δ) δύναμη/ενέργεια



Εικόνα 6.9: Μέση τιμή των έξι παραμέτρων για τρία άτομα



(α) μέση τιμή Ολικής Εννεργοποίησης



(β) μέση τιμή Δύναμης

Εικόνα 6.10

Μερικά από τα καρέ του βίντεο της δημιουργηθείσας έκφρασης απεικονίζονται στην **Εικόνα 6.11**.

Το εργαλείο που χρησιμοποιήθηκε για τη σύνθεση είναι ο MPEG-4 αποκωδικοποιητής GretaPlayer (Ενότητα 7.1).



Εικόνα 6.11: Frames της συντεθειμένης έκφρασης.

Κεφάλαιο 7

Αναπαράσταση Σώματος και Προσώπου σε Εκφραστικούς Πράκτορες - Εφαρμογές

7.1. Greta

Η Greta αποτελεί έναν Πράκτορα με σώμα και δυνατότητα ομιλίας (ECA - Embodied Conversational Agent), απόλυτα συμβατό με το πρότυπο MPEG-4. Η δημιουργία της έγινε από το *GRETA Group: Embodied Conversational Agent Research at the IUT Montreuil* [90]. Η συμβολή της παρούσας διατριβής στο συγκεκριμένο αποκωδικοποιητή έγκειται στον εμπλουτισμό της βιβλιοθήκης που χρησιμοποιεί, τόσο για την απεικόνιση των εκφράσεων του προσώπου, όσο και για τη χρήση χειρονομιών. Τα προφίλ που περιγράφηκαν στο **Κεφάλαιο 5** μετατράπηκαν, σε συνεργασία με την ομάδα της Greta, σε μορφή κατάλληλη (γλώσσα APML), ώστε η Greta να γίνει πιο εκφραστική και πιο αληθοφανής.

Η πρώτη φάση δημιουργίας της Greta περιελάμβανε ένα τρισδιάστατο μοντέλο προσώπου [91], ικανό να προσομοιώνει γρήγορα και αληθοφανώς τις κινήσεις του ανθρωπίνου προσώπου, ακολουθώντας το πρότυπο MPEG-4. Το μοντέλο αυτό χρησιμοποιεί μία ψευδομυϊκή προσέγγιση της συμπεριφοράς του δέρματος του προσώπου και περιέχει και ειδικά χαρακτηριστικά, όπως ρυτίδες και ουλές. Για τις κινήσεις των σημείων του προσώπου χρησιμοποιεί τα FAPs, όπως αυτά περιγράφηκαν στο **Κεφάλαιο 3**. Σε κάθε ακολουθία εκφράσεων που δημιουργείται, αρχικά το πρόσωπο έχει ουδέτερη έκφραση και όλες οι παράμετροι εκφράζονται ως μετακινήσεις από την αρχική αυτή θέση (Εικόνα 7.1).

Η επιφάνεια του προσώπου δημιουργήθηκε από πολύγωνα με τη βοήθεια της OpenGL, ενώ δόθηκε ιδιαίτερη προσοχή στη δημιουργία των περιοχών του προσώπου που είναι οι πιο εκφραστικές, άρα και πιο σημαντικές κατά την απεικόνιση των εκφράσεων, όπως η περιοχή του στόματος. Όπως και όλοι οι αποκωδικοποιητές MPEG-4, έχει

αποθηκευμένες τις έξι βασικές εκφράσεις, σε περίπτωση που είναι ενεργοποιημένο το *FAP2* (Κεφάλαιο 3).

Σε επόμενο στάδιο, και προκειμένου να δημιουργηθεί ένας ολοκληρωμένος ECA, η Greta «εμπλουτίστηκε» με σώμα, ώστε να κοινωνεί τα συναισθήματά της όχι μόνο μέσω εκφράσεων προσώπου – και λόγου-, αλλά και μέσω χειρονομιών και στάσης σώματος. Δημιουργήθηκε ένα σώμα [92] συμβατό με το πρότυπο MPEG-4 και μία γλώσσα περιγραφής χειρονομιών. Χρησιμοποιήθηκε ως βάση ένα πλήρως αρθρωμένο μοντέλο σώματος, όπως αυτό ορίστηκε στις προδιαγραφές του H-Anim 1.1 [89], ενώ για την υφή του δέρματος χρησιμοποιήθηκε αρχικά το μοντέλο του λογισμικού Alias | Wavefront's Maya™ 3.0 suite, αλλά μπορεί να χρησιμοποιηθεί και οποιοδήποτε άλλο μοντέλο που ταιριάζει στις αναλογίες του σκελετού (Εικόνα 7.2).



Εικόνα 7.1: Ουδέτερη έκφραση του προσώπου της Greta



Εικόνα 7.2: Μοντέλο σώματος H-Anim και το αντίστοιχο μοντέλο προσομοίωσης υφής δέρματος.

7.2. Συνδυασμός Προσώπου και Σώματος, Ανάλυσης και Σύνθεσης

Καθ' όλη τη διάρκεια της παρούσας έρευνας έγινε φανερή η αλληλεξάρτηση ανάλυσης και σύνθεσης εκφράσεων. Σε κάποιες περιπτώσεις η σύνθεση χρησιμοποιούσε την ανάλυση, ώστε να δημιουργήσει τα «προφίλ» των εκφράσεων του προσώπου ή τις καινούριες χειρονομίες, ενώ σε άλλες περιπτώσεις η ανάλυση χρησιμοποιούσε τα αποτελέσματα της σύνθεσης, ώστε να εντοπιστούν σωστά τα χαρακτηριστικά του προσώπου, τα χέρια και φυσικά στη συνέχεια οι εκφράσεις και οι χειρονομίες αντίστοιχα. Πλέον όμως, εκτός από τη «χείρα βοηθείας» που προσφέρει η μία στην άλλη, στις σύγχρονες εφαρμογές η ανάλυση και η σύνθεση οφείλουν να συνυπάρχουν, ώστε να οδηγήσουν σε ένα εντυπωσιακότερο αποτέλεσμα. Στο σενάριο που χρησιμοποιήσαμε στην έρευνά μας [93], ένας πράκτορας έχει την ικανότητα να «καταλαβαίνει», να «ερμηνεύει» και τέλος να αντιγράφει εκφράσεις προσώπου και χειρονομίες ενός πραγματικού ανθρώπου. Χρησιμοποιούμε τις τεχνικές

ανάλυσης που παρουσιάστηκαν, ώστε ο πράκτοράς μας να μπορέσει να «κατανοήσει» τις εκφράσεις και τις χειρονομίες του ατόμου που έχει μπροστά του. Με τον όρο «κατανόηση» δεν εννοούμε την πλήρη ανάλυση και σύνθεση των εκφράσεων, αλλά την ανάλυση και σύνθεση της διάθεσης του χρήστη.

Αρχικά το σύστημα προσπαθεί να ερμηνεύσει την έκφραση που βλέπει ο πράκτορας – για παράδειγμα να μετατρέψει τα FAPs που έχει προσδιορίσει, στην αντίστοιχη έκφραση του θυμού- την οποία μπορεί να προσαρμόσει – ώστε να απεικονίσει μία πιο θυμωμένη έκφραση- ή να αποκτήσει μία έκφραση που θα είναι αποκλειστικά για τον πράκτορα – ίσως ένα άλλο σύνολο FAPs που θα απεικονίζουν, όμως, πάλι την έκφραση του θυμού. Στην περίπτωση που η έκφραση που καλείται να απεικονίσει ο πράκτορας δεν είναι μία από τις βασικές εκφράσεις, θα χρησιμοποιήσει στην σύνθεση της νέας έκφρασης τις τιμές των FAPs, όπως αυτές προέκυψαν από τη διαδικασία της ανάλυσης. Όσον αφορά στη χειρονομία, στην παρούσα φάση η χειρονομία δίνεται στον πράκτορα, δηλαδή προσδιορίζεται το σχήμα των χεριών, και, μέσω της ανάλυσης, δίνεται πληροφορία για τις παραμέτρους εκφραστικότητας, ώστε ο πράκτορας να «εμπλουτίσει» τη δεδομένη χειρονομία κάνοντας την να φαίνεται πιο θυμωμένη, πιο χαρούμενη κ.ο.κ.

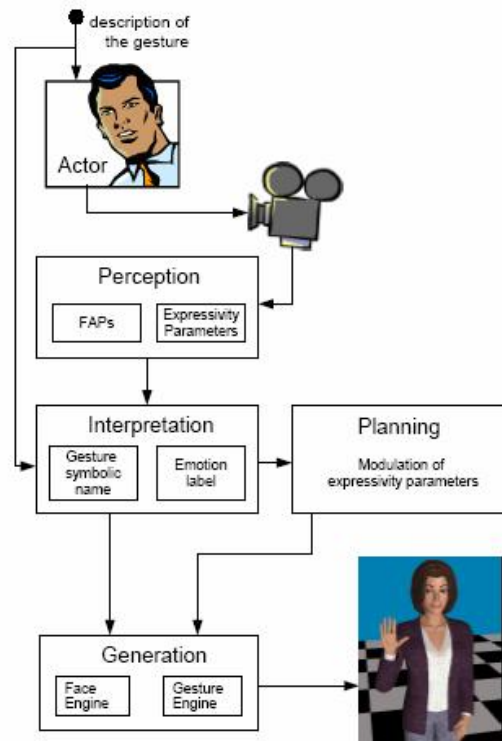
Στον **Πίνακα 7.1** φαίνεται η επίδραση που έχουν οι διάφορες παράμετροι εκφραστικότητας (Ενότητες 5.4.1 και 6.2.2) στο κεφάλι, στο πρόσωπο και τα χέρια.

	Κεφάλι	Πρόσωπο	Χέρια (Χειρονομίες)
Χωρική Έκταση	μεγαλύτερη/μικρότερη κίνηση	αυξημένη/μειωμένη μυϊκή συστολή	μεγαλύτερη/μικρότερη κίνηση
Χρονικά Χαρακτηριστικά	μεγαλύτερη / μικρότερη ταχύτητα κίνησης	μικρότερο / μεγαλύτερο αρχικό και τελικό στάδιο	μικρότερο/μεγαλύτερο στάδιο προετοιμασίας και κυρίως στάδιο
Ρευστότητα	αύξηση/μείωση συνοχής της κίνησης του κεφαλιού	αύξηση/μείωση συνοχής της σύσπασης των μυών	αύξηση/μείωση συνοχής μεταξύ διαδοχικών χειρονομιών
Δύναμη / Ενέργεια	μεγαλύτερη / μικρότερη ανύψωση κεφαλής	μεγαλύτερη / μικρότερη ένταση μυϊκής συστολής	μεγαλύτερη / μικρότερη ένταση χειρονομίας
Επαναληπτικότητα	περισσότερες / λιγότερες επαναλήψεις νευμάτων	-	περισσότερες / λιγότερες επαναλήψεις της ίδιας κίνησης

Πίνακας 7.1: Επίδραση παραμέτρων εκφραστικότητας σε κεφάλι, εκφράσεις προσώπου και χειρονομίες

Στην **Εικόνα 7.3** δίνεται ένα διάγραμμα του προτεινόμενου συστήματος. Το τμήμα *Αντίληψη (Perception)* αναλύει το βίντεο κι εξάγει τις παραμέτρους εκφραστικότητας για τις χειρονομίες (Κεφάλαιο 6) και τις μετακινήσεις των χαρακτηριστικών του προσώπου, από τις οποίες θα προκύψουν οι τιμές των αντίστοιχων FAPs. Τα αποτελέσματα προωθούνται στο τμήμα *Κατανόησης (Interpretation)*, το οποίο έχει τη δυνατότητα να προσδιορίσει αν τα FAPs του προηγούμενου σταδίου δίνουν κάποια από τις γνωστές εκφράσεις, ενώ στη συνέχεια το

τμήμα *Προετοιμασία (Planning)* καθορίζει με ποιο τρόπο θα απεικονιστούν η έκφραση του προσώπου και η χειρονομία στον πράκτορα, ορίζοντας και μικρές μετατροπές που απαιτείται να γίνουν στις προσδιορισμένες παραμέτρους εκφραστικότητας, με σκοπό τον αριότερο συνδυασμό προσώπου και σώματος. Στη συνέχεια, το επιλεγμένο σύνολο των FAPs και των BAPs θα απεικονιστεί στην Greta μέσω της *Σύνθεσης (Generation)*.



Εικόνα 7.3: Διάγραμμα προτεινόμενου συστήματος

7.3. Η Μέθοδος Πρόκλησης Συναισθημάτων SAL

Η μέθοδος SAL (Sensitive Artificial Listener) αναπτύχθηκε στο πλαίσιο των προγραμμάτων ERMIS [86] και HUMAINE [87] και αποτελεί την επέκταση του συστήματος ELIZA [88] που δημιουργήθηκε ως αποτέλεσμα έρευνας στο πεδίο της τεχνητής νοημοσύνης από το Weizenbaum, τη δεκαετία του '60. Το σύστημα ELIZA προσομοιώνει τη μέθοδο θεραπείας του Carl Rogers κατά την οποία οι ασθενείς αναφέρουν τα προβλήματά τους σε έναν ακροατή, ο οποίος απαντά κατάλληλα με στόχο να προκαλέσει τη συνέχιση της συζήτησης χωρίς όμως να εκφέρει κρίση ή σχόλια. Στις πρώτες υλοποιήσεις της ιδέας του Weizenbaum οι χρήστες αλληλεπιδρούσαν με το σύστημα χρησιμοποιώντας μηνύματα κειμένου μέσω του πληκτρολογίου και ο υπολογιστής απαντούσε με μηνύματα κειμένου στην οθόνη. Αργότερα το σύστημα εμπλουτίστηκε με σύνθεση φωνής και συνθετικά μοντέλα προσώπων. Στην πλευρά της εισόδου, το πληκτρολόγιο παραμένει η

βασική συσκευή εισαγωγής πληροφορίας, αφού η αυτόματη αναγνώριση φυσικού λόγου δεν έχει ωριμάσει ακόμα σε επίπεδο που να είναι ικανή να αναγνωρίζει συναισθηματικά φορτισμένη ομιλία.

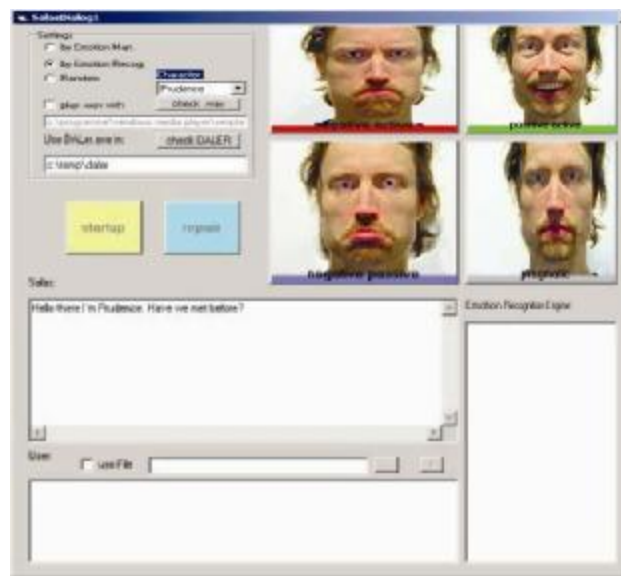
Ο τρόπος καταγραφής αποτελεί εγγενώς προβληματική περιοχή αυτής της μεθόδου. Λαμβάνοντας υπόψη ότι πρέπει να καταγράφεται και η φωνή αλλά και η εικόνα, είναι φανερό η ανάγκη συμβιβασμού μεταξύ των απαιτήσεων ποιότητας εγγραφής που ορίζει η επεξεργασία σήματος και της μη-παρεμβολής που απαιτεί η ψυχολογία. Εάν κάποιος είναι πολύ προσεκτικός με την ποιότητα εγγραφής, οι χρήστες θα αισθανθούν περιορισμένοι και είναι μάλλον απίθανο να εμφανίσουν την επιθυμούμενη καθημερινή ήπια συναισθηματική έκφραση. Από την άλλη, οι αλγόριθμοι ανάλυσης εικόνας και ομιλίας δεν αναμένεται να λειτουργήσουν ικανοποιητικά σε αυθόρμητες κινήσεις του κεφαλιού και των χεριών, μειωμένο φωτισμό και ήπια μουσική υπόκρουση.

Η υλοποίηση του SAL αποτελείται κυρίως από μια εφαρμογή λογισμικού η οποία σχεδιάστηκε ώστε να επιτρέπει στους χρήστες να μεταπηδούν σε διαφορετικές συναισθηματικές καταστάσεις. Περιέχει τέσσερις «προσωπικότητες» (*Εικόνα 7.4*), οι οποίες ακούν το χρήστη και απαντούν σε αυτό που λέει ανάλογα με το «χαρακτήρα» που έχει δοθεί στην κάθε μία. Αυτό που προσδιορίζει τις προσωπικότητες αυτές είναι ότι κάθε μία προσπαθεί να οδηγήσει το χρήστη προς μία κατάσταση-στόχο:

- Η Poppy είναι χαρούμενη και προσπαθεί να οδηγήσει το χρήστη προς μία θετική συναισθηματική κατάσταση.
- Ο Spike είναι επιθετικός και προσπαθεί να κάνει το χρήστη να υιοθετήσει μία εριστική ή και οργισμένη διάθεση.
- Ο Obadiah είναι καταθλιπτικός και προσπαθεί να οδηγήσει το χρήστη προς μία παθητική - απαισιόδοξη στάση.
- Η Prudence είναι πραγματίστρια και προσπαθεί να «προσγειώσει» το χρήστη.

Ο πυρήνας του συστήματος αποτελείται από μια σειρά απαντήσεων, οι οποίες χρησιμοποιούνται ανάλογα με την «προσωπικότητα» με την οποία έχει επιλέξει να συνομιλήσει ο χρήστης.

Τα δεδομένα που παρήχθησαν με τη βοήθεια του SAL χρησιμοποιήθηκαν για την εξαγωγή χαρακτηριστικών του προσώπου και την εκτίμηση συναισθήματος, ενώ με περαιτέρω μελέτη τα αποτελέσματα του πειράματος χρησιμοποιούνται και για σύνθεση, αφού, με τη βοήθεια του SAL, παράγονται εκφράσεις που δεν ανήκουν στις έξι βασικές, οπότε λαμβάνουμε δεδομένα για απεικόνιση ενδιάμεσων, πιο σύνθετων συναισθηματικών καταστάσεων.



Εικόνα 7.4: Διαπροσωπεία SAL

Κεφάλαιο 8

Σύνθεση Εκφράσεων και Οντολογίες

Ο πραγματικός κόσμος ενός ανθρώπου μπορεί να μεταφερθεί στην εικονική πραγματικότητα μέσω ενός πράκτορα (agent), οπότε ένας τέτοιος πράκτορας επιβάλλεται να έχει τη δυνατότητα έκφρασης συναισθημάτων μέσω του προσώπου ή του σώματος και ιδανικά η χρήση του MPEG-4 θα συνέβαλλε στον καλύτερο έλεγχο των εκφράσεων αυτών.

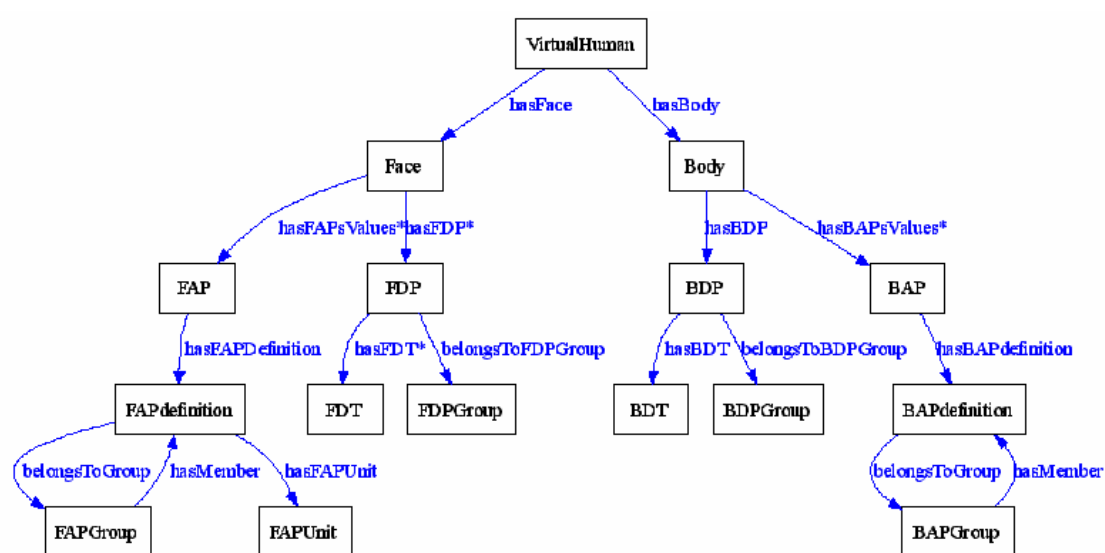
Μία μέθοδος αποδοτική στο να προσφέρει ευκολότερη χρήση των όσων περιγράφηκαν στα προηγούμενα κεφάλαια, από ερευνητικούς τομείς που δεν είναι ειδικοί στο αντικείμενο αλλά και από τους υπολογιστές, είναι η χρήση οντολογιών.

Η εφαρμογή αυτή είναι ακόμα σε πειραματικό στάδιο και πραγματοποιείται σε συνεργασία με το **Εργαστήριο Εικονικής Πραγματικότητας** του **Ελβετικού Ινστιτούτου Τεχνολογίας** (Virtual Reality Lab (VRlab), Swiss Federal Institute of Technology (EPFL)), το οποίο ειδικεύεται στους εικονικούς ανθρώπους και γενικότερα στην εικονική πραγματικότητα. Έχει ήδη δημιουργήσει μία Οντολογία που την ονομάζει Virtual Human Ontology [79], η οποία μοντελοποιεί τη γνώση κατασκευής εικονικού ανθρώπου, λαμβάνοντας υπόψη τη σχετική πολυπλοκότητα. Ως συνέχεια και συμπλήρωση της προαναφερθείσας Οντολογίας δημιουργήσαμε την οντολογία των εκφράσεων του ανθρωπίνου προσώπου και του σώματος. Δημιουργήθηκαν δύο ξεχωριστές οντολογίες, μία για το πρόσωπο και μία για το σώμα, με έμφαση στην πρώτη, δεδομένης της πολυπλοκότητας ορισμού του σώματος στα πλαίσια του προτύπου MPEG-4. Για τη δημιουργία τους χρησιμοποιήθηκε η Ontology Web Language (OWL) [80]. Η OWL είναι μία επίσημη γλώσσα, η οποία καθορίζει με συγκεκριμένο και αυστηρό τρόπο τις σημασιολογικές σχέσεις μεταξύ των οντοτήτων, και προσφέρει καλύτερη διαχείριση του δικτυακού περιεχομένου σε σχέση με τις άλλες γλώσσες (XML, RDF, RDF Schema-RDFS), αφού παρέχει επιπλέον λεξιλόγιο. Η γλώσσα αυτή, συγκρινόμενη με παλιότερες γλώσσες, διατυπώνει περισσότερες σχέσεις και λογικές πράξεις μεταξύ των οντοτήτων. Για το σχεδιασμό της οντολογίας χρησιμοποιήθηκε ο Protégé ontology editor [81], ο οποίος είναι open source και βασίζεται στη γνώση.

8.1. MPEG-4: FBA Node

Αρχικά, δημιουργήθηκε το κομμάτι της οντολογίας που παρουσιάζει τη δομή του σχεδιασμού κίνησης προσώπου και σώματος, σύμφωνα με τον ορισμό του προτύπου MPEG-4, όπως αυτός παρουσιάστηκε στο **Κεφάλαιο 3**. Τα FAPs που μπορούν να χρησιμοποιηθούν μαζί παρουσιάζονται στην οντολογία σε ομάδες, προκειμένου να διευκολυνθεί η διαδικασία σχεδιασμού κίνησης [82].

Αντίστοιχα με το πρόσωπο, δημιουργήθηκε και η ανάλογη δομή για το σώμα (*Ενότητα 8.3*). Η περιγραφόμενη δομή στο πλαίσιο της οντολογίας απεικονίζεται στην **Εικόνα 8.1**.



Εικόνα 8.1: Απεικόνιση του αντικείμενου Σχεδιασμού Κίνησης Προσώπου και Σώματος του MPEG-4 στην οντολογία

8.2. Πρόσωπο

Προκειμένου να γίνει κατανοητή η χρησιμότητα της Οντολογίας για το πρόσωπο, θεωρήσαμε δύο διαφορετικά σενάρια.

Το πρώτο αφορά στην ανάκτηση συνθετικών ή πραγματικών βίντεο, τα οποία είναι χαρακτηρισμένα με υψηλού επιπέδου περιγραφείς, π.χ. με συγκεκριμένη κίνηση των χαρακτηριστικών του προσώπου, η οποία εκφράζει κάποιο συναίσθημα. Απώτερος σκοπός αυτής της διαδικασίας είναι η δημιουργία επαναχρησιμοποιούμενου, κλιμακούμενου και προσαρμόσιμου περιεχομένου στα εικονικά περιβάλλοντα.

Το δεύτερο σενάριο αφορά στις περιπτώσεις κατά τις οποίες η οντολογία μπορεί να παρέχει ένα είδος ειδίκευσης του τομέα της γνώσης. Αφού η παρούσα οντολογία μπορεί να περιλαμβάνει όλα τα συστατικά μέρη της δομής του προσώπου, της κίνησής του και της

σχέσης αυτής με τα συναισθήματα, ο χρήστης μπορεί να επιδιώξει να δημιουργήσει κάποια συγκεκριμένη έκφραση, αναζητώντας τις απαραίτητες παραμέτρους κτλ.

Δεδομένου ότι η σύνθεση των εκφράσεων περιλαμβάνει πολλούς τομείς γνώσης, όπως ψυχολογία, έλεγχος εμφύχωσης (animation control), πρότυπα κτλ, η χρήση οντολογιών αποτελεί μία καλή λύση για τον προσδιορισμό μίας κοινής, μοιραζόμενης γνώσης [83].

8.2.1. Δημιουργία Οντολογίας

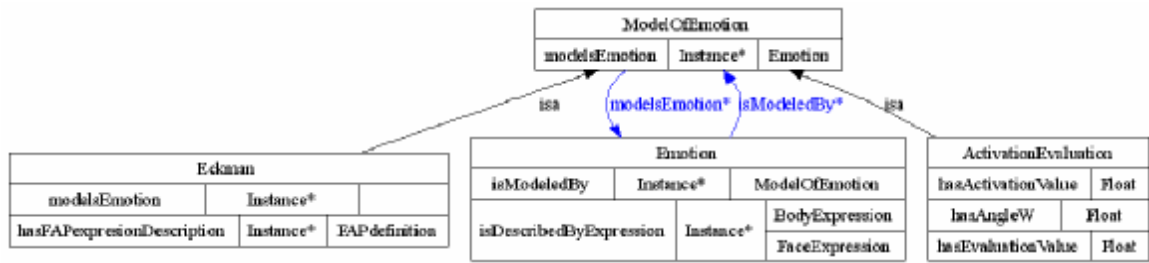
Αρχικά έπρεπε να ορίσουμε τη γνώση του αντικειμένου του σχεδιασμού κίνησης του προσώπου σύμφωνα με το πρότυπο MPEG-4 μέσα στην οντολογία, ώστε να περιγράψουμε τη δομή σχεδιασμού κίνησης του εικονικού ανθρώπου (Virtual Human-VH). Αυτή η δομή μπορεί να χρησιμοποιηθεί στην εγκαθίδρυση σχέσης μεταξύ των συστατικών μερών της κίνησης και των παραμέτρων της έκφρασης μέσω των Προφίλ των Εκφράσεων (Κεφάλαιο 5). Οι εκφράσεις άλλωστε έχουν μία στενή σχέση με τα συναισθήματα, οπότε πρέπει και αυτές να περιγραφούν στο πλαίσιο της οντολογίας μέσω ψυχολογικών μοντέλων.

Αυτή η οντολογία έχει σκοπό να δίνει απαντήσεις σε ερωτήματα που μπορούν να παρέχουν γνώση σχετικά με την εκφραστική κίνηση ανθρώπινων χαρακτήρων. Η οντολογία που δημιουργήσαμε δηλαδή πρέπει και έχει τη δυνατότητα να απαντήσει σε ερωτήματα του είδους: *Πώς κινούνται τα χαρακτηριστικά του προσώπου ώστε να εκφραστεί λύπη; Με ποιους διαφορετικούς τρόπους μπορεί να εκφραστεί ο θυμός; Ποιες είναι οι τιμές των FAPs για τα συναισθήματα ανήσυχος, τρομοκρατημένος; Με δεδομένο ένα σύνολο από FAPs, ποιο συναίσθημα μπορούμε να δημιουργήσουμε;*

8.2.2. Αναπαράσταση Συναισθήματος

Για την καλύτερη και συνεχή απεικόνιση του συναισθήματος, η οντολογία περιέχει και πληροφορίες σχετικά με τις μεθόδους απεικόνισής του, π.χ. σχετικά με τον τροχό συναισθημάτων της Whissel (Κεφάλαιο 2) και περιέχει τις τιμές ενεργοποίησης (activation) και αποτίμησης (evaluation) για διάφορα συναισθήματα.

Η αναπαράσταση συναισθήματος, έτσι όπως αυτή έχει εισαχθεί στην οντολογία, απεικονίζεται στην **Εικόνα 8.2**. Ένα συναίσθημα μπορεί να μοντελοποιηθεί από ένα ή περισσότερα Μοντέλα Συναισθήματος (Models of Emotions). Εδώ χρησιμοποιείται τα μοντέλα των Ekman και Whissel. Ένα Συναίσθημα (Emotion) μπορεί να περιγραφεί από μία Έκφραση Προσώπου (Face Expression).



Εικόνα 8.2: Περιγραφή των ψυχολογικών μοντέλων του συναισθήματος στην οντολογία

8.2.3. Προφίλ Εκφράσεων Προσώπου (Βασικών και Ενδιάμεσων)

Στην οντολογία βασική γνώση αποτέλεσαν τα προφίλ των εκφράσεων, βασικών και μη, όπως αυτά ορίστηκαν στο **Κεφάλαιο 5**. Για παράδειγμα, τα δεδομένα που παρουσιάζονται στον **Πίνακα 5.5** παίρνουν τη μορφή του **Πίνακα 8.1** (όπου εμφανίζεται μόνο ένα τμήμα των ενεργοποιούμενων FAPs) προκειμένου να εισαχθούν στην οντολογία.

Expression	F_{3min}	F_{3max}	F_{5min}	F_{5max}	F_{19min}	F_{19max}	...
Fear ($P_F^{(4)}$)	400	560	-200	-160	-630	-570	...
Depressed ($P_D^{(0)}$)	160	230	-100	-65	-110	-310	...
Sad ($P_S^{(0)}$)	0	0	0	0	-265	-41	...

Πίνακας 8.1: Ενδιάμεσο Συναισθήμα Απογοητευμένος και τα βασικά συναισθήματα από τα οποία προέκυψε, με χρήση των μετρήσεων ενεργοποίησης και αποτίμησης

8.2.4. Οντολογία Εκφράσεων Προσώπου

Στην **Εικόνα 8.3** παρουσιάζεται το διάγραμμα της οντολογίας το σχετικό με τις βασικές και τις ενδιάμεσες εκφράσεις προσώπου και τη μεταξύ τους σχέση. Η Έκφραση Προσώπου (FaceExpression) μπορεί να είναι βασική ή ενδιάμεση ανάλογα με τον τύπο της έκφρασης. Όλα τα Προφίλ Εκφράσεων Προσώπου (FaceExpressionProfile) περιλαμβάνουν ένα διάστημα τιμών για κάθε FAP, σύμφωνα με τους κανόνες που διατυπώθηκαν στο **Κεφάλαιο 5**. Το Προφίλ Ενδιάμεσης Έκφρασης (IntermediateProfile) περιλαμβάνει δύο βασικές εκφράσεις και ένα Διάστημα Τιμών FAPs (FAPRangeVariation). Για τις ενδιάμεσες εκφράσεις ορίσαμε περιορισμούς στο πλαίσιο της οντολογίας, βασιζόμενοι στους κανόνες που χρησιμοποιήσαμε για τη δημιουργία τους. Για την καλύτερη περιγραφή αυτής της γνώσης, παρουσιάζονται οι ακόλουθοι ισχυρισμοί:

1. Η ΈκφρασηΠροσώπου μπορεί να οριστεί από ένα προφίλ βασικής ή ενδιάμεσης έκφρασης.

2. Κάθε ενδιάμεση έκφραση πρέπει να περικλείει δύο βασικές εκφράσεις

IntermediateProfile encloses ArchetypalProfiles = 2

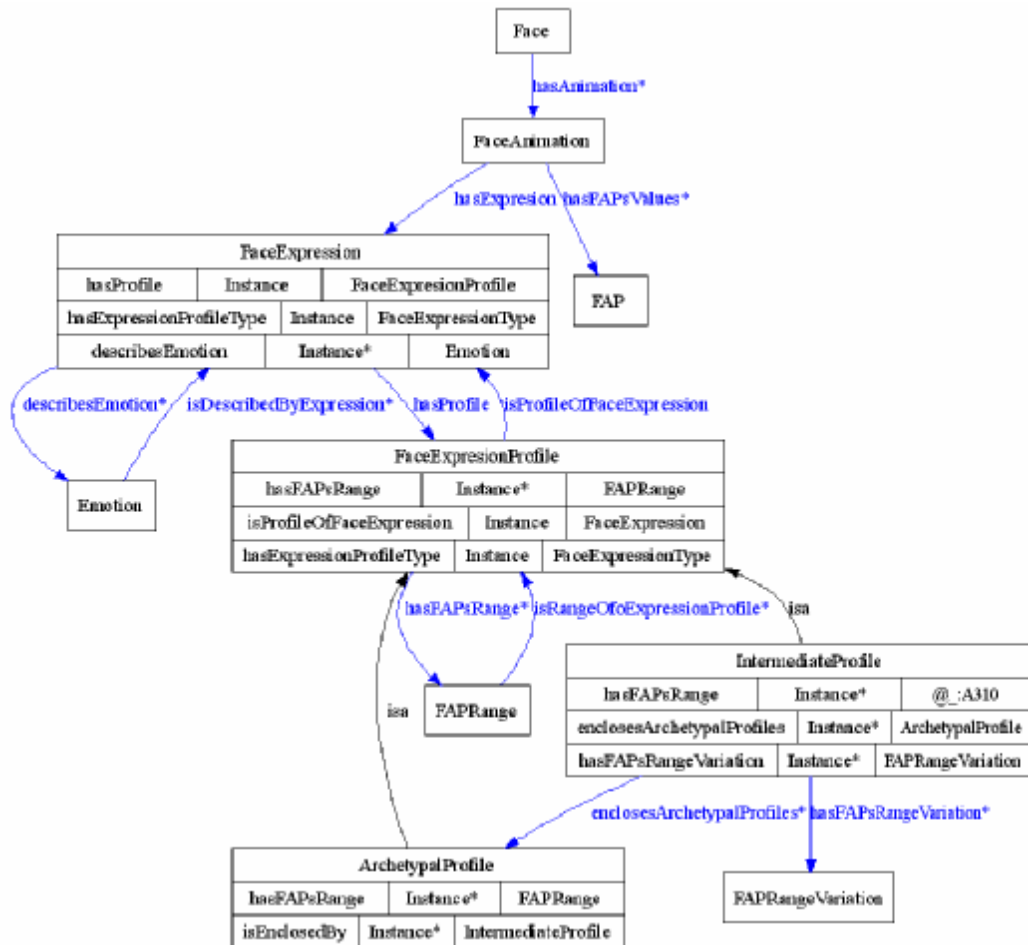
3. Κάθε προφίλ περιλαμβάνει το αντίστοιχο διάστημα τιμών για κάθε FAP (FAPRange). Η τάξη FAPRange (FAPRange class) περιλαμβάνει το min και max των τιμών των FAPs.

4. Το διάστημα τιμών των FAPs των ενδιάμεσων εκφράσεων είναι η ένωση των διαστημάτων των βασικών εκφράσεων που περικλείονται στη συγκεκριμένη ενδιάμεση έκφραση. Αυτό μπορεί να εκφραστεί ως περιορισμός για την ιδιότητα έχειΔιάστημαΤιμώνFAPs (hasFAPRange) ως εξής:

$\forall \text{ hasFAPsRange } (FAPRange \cap ((\exists \text{ isRangeOfExpressionProfile } ArchetypalProfile) \cap (\text{isEnclosedBy } IntermediateProfile)))$

5. Η ΜεταβολήΔιαστήματοςΤιμώνFAPs (FAPRangeVariation) για τις ενδιάμεσες εκφράσεις υπολογίζεται με βάση τρεις εναλλακτικές:

α) FAPs περιλαμβανόμενα και στις δύο βασικές εκφράσεις, αλλά με αντίθετο πρόσημο, β) FAPs περιλαμβανόμενα και στις δύο βασικές εκφράσεις, με ίδιο πρόσημο, γ) αμοιβαίως αποκλειόμενα FAPs. Ο υπολογισμός των τιμών αυτών γίνεται μέσω του γεννήτορα εκφράσεων κι εξηγείται στην **Ενότητα 8.2.5**.



Εικόνα 8.3: Περιγραφή των προφίλ των βασικών και των ενδιάμεσων εκφράσεων στην οντολογία

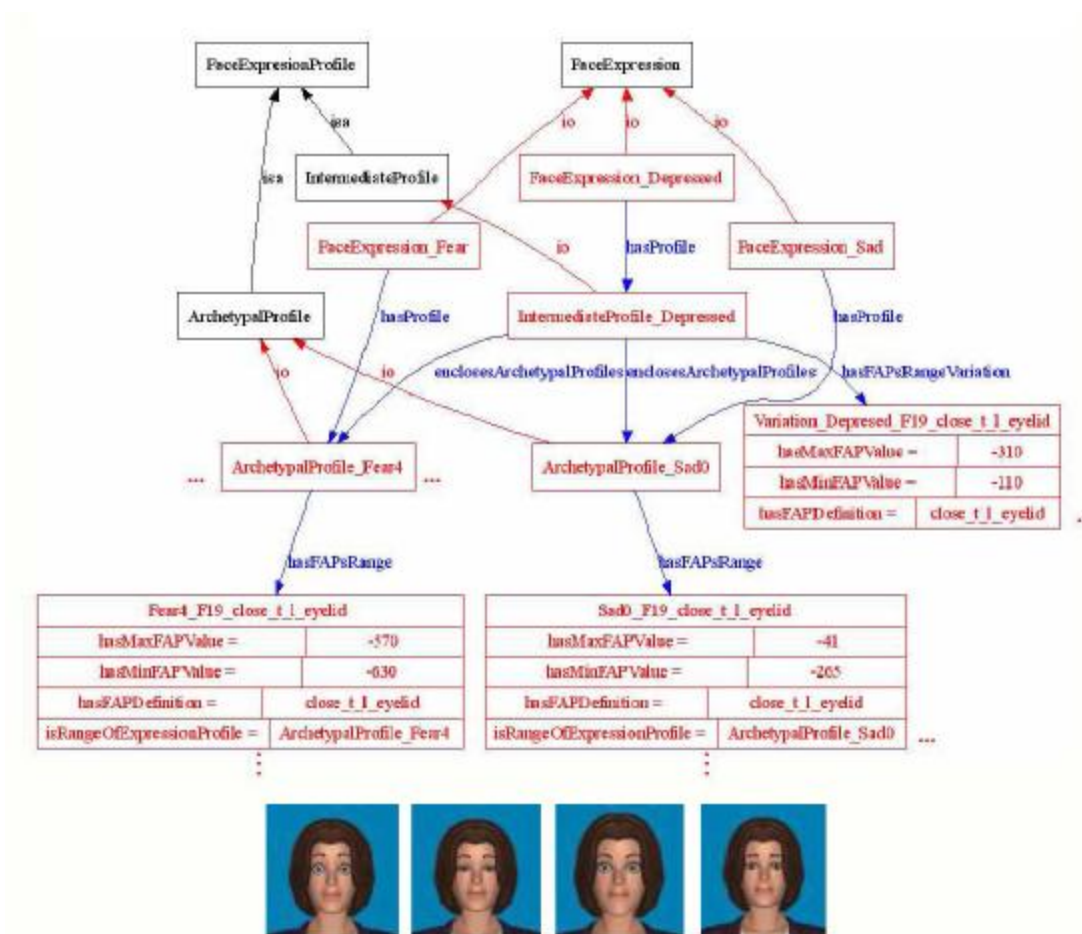
8.2.5. Παραγωγή Κανόνων στην Οντολογία των Εκφράσεων

Ως παράδειγμα περιγράφεται στην **Εικόνα 8.4** η δημιουργία, μέσω της οντολογίας, της έκφρασης *απογοητευμένος* μέσω των προφίλ του *φοβισμένου* και του *λυπημένου*.

Όπως λοιπόν φαίνεται στην **Εικόνα 8.4**, το συναίσθημα του φόβου εκφράζεται μέσω της ΈκφρασηςΠροσώπουΦόβος (FaceExpressionFear). Αυτή η έκφραση θεωρείται βασική, οπότε έχει και ΒασικόΠροφίλ (ArchetypalProfile): Βασική_Έκφραση_Φόβος0 (ArchetypalProfile_Fear0), Βασική_Έκφραση_Φόβος4 (ArchetypalProfile_Fear4). Κάθε προφίλ περιέχει τις αντίστοιχες τιμές FAPs. Με τον ίδιο τρόπο λαμβάνουμε το προφίλ της βασικής έκφρασης Λύπη (ArchetypalProfile_Sad0). Προκειμένου να ορίσουμε την Απογοήτευση ως ενδιάμεση έκφραση μεταξύ Φόβου και Λύπης, δημιουργήθηκε η ΈκφρασηΠροσώπουΑπογοήτευση (DepressionFaceExpression) ως ενδιάμεση έκφραση, με προφίλ που περιλαμβάνει τις Βασικές Εκφράσεις Φόβος και Λύπη. Το ΔιάστημαΤιμών (FAPRange) για το προφίλ που δημιουργείται, ορίζεται ως η ένωση του διαστήματος τιμών κάθε FAP στα προφίλ του Φόβου και της Λύπης. Στο συγκεκριμένο παράδειγμα

απεικονίζεται το FAP19 close_t_l_eyelid, το οποίο περιέχεται και στις δύο βασικές εκφράσεις με το ίδιο πρόσημο, δηλαδή χρησιμοποιεί το δεύτερο κανόνα της **Ενότητας 5.3.1**.

Κατά τον ίδιο τρόπο εισάγονται στην οντολογία όλα τα προφίλ των εκφράσεων, βασικών ή ενδιάμεσων.



Εικόνα 8.4: Παράδειγμα δημιουργίας προφίλ Έκφρασης Προσώπου και απεικόνισης τελικού αποτελέσματος στον MPEG-4 αποκωδικοποιητή GretaPlayer

8.2.6. Σενάρια Χρήσης Οντολογίας

Η μέθοδος με την οποία λαμβάνουμε πληροφορίες από τις Οντολογίες είναι η πραγματοποίηση ερωτημάτων. Για το λόγο αυτό χρησιμοποιήθηκε το plug-in new Racer Query Language interface [84] for OWL ontologies in Protg. Αυτή η γλώσσα ερωτημάτων είναι πολύ κοντινή στη φυσική γλώσσα. Η γλώσσα nRQL παρέχεται από τη RacerPro [85].

Στη συνέχεια παρουσιάζονται στην RQL, τα ερωτήματα που δημιουργήθηκαν. Η πρώτη ερώτηση είναι μέρος του σεναρίου, κατά το οποίο μπορούμε να λάβουμε

σχεδιασμούς κίνησης από την οντολογία, δηλαδή στο πρώτο σενάριο θέλουμε να βρούμε τρόπους έκφρασης ενός συγκεκριμένου συναισθήματος.

Πώς μπορούμε να εκφράσουμε το συναίσθημα της απογοήτευσης;

(What is the facial animation for expressing depressed emotion?)

nRQL:

(retrieve (?a)

(and (?a | FaceAnimation |)

(?a ?b | hasExpression |)

(?b | FaceExpression |)

(?b ?c | describesEmotion |)

(c? | Depressed |)))

Result: (((?A | FaceAnimation Depressed |)))

Βρήκαμε ότι υπάρχει η ΈκφρασηΠροσώπουΑπογοήτευση (FaceAnimationDepressed) ορισμένη στην οντολογία.

Στην περίπτωση του δεύτερου σεναρίου, θέλουμε, για παράδειγμα, να γνωρίζουμε την απαραίτητη πληροφορία για τον ορισμό της συγκεκριμένης έκφρασης προσώπου, δηλαδή: *Αυτή η έκφραση είναι βασική ή ενδιάμεση; Αν είναι ενδιάμεση, από ποιες βασικές εκφράσεις προκύπτει; Ποια είναι τα αντίστοιχα διαστήματα τιμών των FAPs; κτλ.*

Ποιο είναι το διάστημα τιμών των FAPs για την έκφραση του απογοητευμένου?

(What is the range of FAP values for the face expression depressed?)

nRQL:

(retrieve (?H)

(and

(?A | FaceAnimation |)

(or (and (?A ?B | hasExpression |)

(?B | FaceExpression |)

(?B (some | describesEmotion | | CONCEPT -FOR- Depressed |)))

(?B ?F | hasProfile |)

(?F | IntermediateProfile |)

(?F ?G | enclosesArchetypalProfiles |)

(?G | ArchetypalProfile |)

(?G ?H | hasFAPsRange |)

(?H | FAPRange |))))

Result: (((?H | Sad0F 19 |)) ((?H | Fear4F 3 |))) ...

Το τελευταίο ερώτημα έχει ως αποτέλεσμα την ένωση του συνόλου των FAPs των βασικών εκφράσεων, οι οποίες δίνουν την ενδιαμέση έκφραση της *απογοήτευσης*.

8.3. Σώμα

8.3.1. Δημιουργία Οντολογίας

Το στάδιο που πρέπει να προηγηθεί στη διαδικασία δημιουργία της Οντολογίας είναι, όπως αναφέρθηκε και στην **Ενότητα 8.1**, ο ορισμός, εντός της οντολογίας, του τμήματος του MPEG-4 που πραγματεύεται το σχεδιασμό κίνησης για το σώμα (FBA). Με χρήση της δομής του ανθρώπινου σώματος, όπως αυτή ορίζεται από το FBA, μπορούμε να ορίσουμε μία σχέση μεταξύ των στοιχείων που χρησιμοποιούνται κατά το σχεδιασμό κίνησης και των παραμέτρων εκφραστικότητας, όπως αυτές ορίζονται στο **Κεφάλαιο 6**. Στον ορισμό αυτό συμμετέχουν και τα ψυχολογικά μοντέλα περιγραφής συναισθήματος που επίσης περιέχονται στην Οντολογία (*Εικόνα 8.2*). Το μοντέλο που επιλέγουμε στην περίπτωση μας είναι ο Τροχός Συναισθημάτων της Whissel, ενώ παράλληλα εισάγουμε ως έννοιες στην οντολογία τις προαναφερθείσες παραμέτρους εκφραστικότητας.

8.3.2. Παράμετροι Σώματος

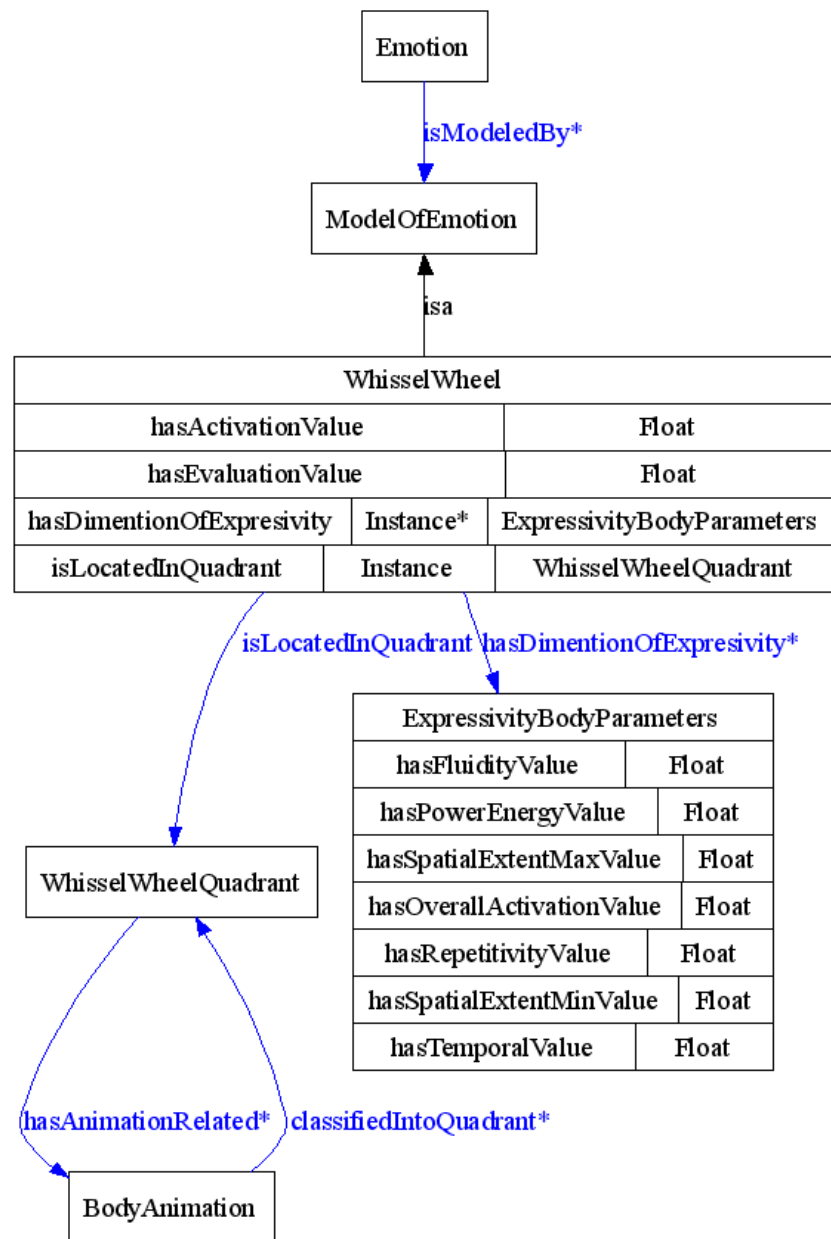
Οι εκφράσεις του σώματος, κυρίως οι χειρονομίες, εξαρτώνται από το ευρύτερο πλαίσιο έκφρασης του συναισθήματος. Η κυριότερη διαφορά τους από τις εκφράσεις προσώπου είναι ότι είναι αδύνατο να προσδιορίσουμε το συναίσθημα, δίχως να έχουμε στη διάθεσή μας το πρόσωπο ή έστω το γενικότερο πλαίσιο μέσα στο οποίο πραγματοποιείται η χειρονομία. Η σύνθεση κίνησης για το σώμα θεωρείται εξαιρετικά δύσκολη λόγω της πολυπλοκότητας της δομής του και των πολλών βαθμών ελευθερίας που διαθέτει. Όπως αναφέρθηκε και στα προηγούμενα κεφάλαια, καλύτερη απόδοση επιτυγχάνεται με χρήση πληροφορίας εξαγόμενης από πραγματικά βίντεο.

Εντός οντολογίας, οι χειρονομίες μπορούν να ταξινομηθούν με σύνδεσή τους με συναισθήματα, για παράδειγμα το χειροκρότημα μπορεί άμεσα να συνδεθεί με τη χαρά, ενώ το σταύρωμα των χεριών μπορεί να συνδεθεί με το θυμό ή την έλλειψη άνεσης κτλ. Αυτή η διαδικασία της ταξινόμησης προέκυψε από τα βίντεο που παρουσιάζονται στο **Κεφάλαιο 6**. Με χρήση της ταξινόμησης του **Πίνακα 6.2**, μπορούμε να αναζητήσουμε χειρονομίες που αντιστοιχούν σε συγκεκριμένο πλαίσιο ή συναισθηματική κατάσταση. Στη συνέχεια, και πέραν αυτής της ταξινόμησης, χρησιμοποιήσαμε τις παραμέτρους εκφραστικότητας (Γενική Ενεργοποίηση, Χωρική Έκταση, Χρονικά Χαρακτηριστικά, Ρευστότητα, Δύναμη / Ενέργεια και Επαναληπτικότητα - *Κεφάλαιο 6*).

8.3.3. Οντολογία Σώματος

Οι χειρονομίες, όπως αυτές έχουν ταξινομηθεί, και οι αντίστοιχες παράμετροι εκφραστικότητας περιέχονται στην οντολογία της **Εικόνας 8.5**. Κάθε κίνηση του σώματος έχει ταξινομηθεί σε ένα ή περισσότερα τεταρτημόρια του Τροχού Συναισθημάτων της Whissel, σύμφωνα με τον **Πίνακα 6.2**, ενώ προφανώς υπάρχει αντιστοιχία μεταξύ συναισθημάτων και τεταρτημορίων. Κάθε συναισθηματικό σύνδεσμο με συγκεκριμένες τιμές των παραμέτρων εκφραστικότητας, οι οποίες σχετίζονται άμεσα και με τις αντίστοιχες χειρονομίες.

Στην οντολογία περιλήφθηκαν οι χειρονομίες που περιγράφονται στο **Κεφάλαιο 6** (explain, oh!my God!, Wave κτλ) συνοδευόμενες από το αντίστοιχο τεταρτημόριο του Τροχού Συναισθημάτων. Επίσης στην ίδια οντολογία περιλαμβάνονται πληροφορίες σχετικά με τη διακύμανση των τιμών των παραμέτρων εκφραστικότητας και πώς αυτή συνδέεται με τα τεταρτημόρια του Τροχού.



Εικόνα 8.5: Διάγραμμα της οντολογίας για το σώμα

8.3.4. Σενάρια Χρήσης

Όπως και στην περίπτωση της Οντολογίας για το Πρόσωπο, για την υποβολή ερωτημάτων χρησιμοποιήθηκε το plug-in new Racer Query Language interface [84] for OWL ontologies in Protg.

Στη συνέχεια παρουσιάζονται στην RQL, τα ερωτήματα που καλείται να απαντήσει η οντολογία.

Ποιες είναι οι κινήσεις που εκφράζουν χαρά; (ή Κινήσεις που ανήκουν στο ίδιο τεταρτημόριο με το συναίσθημα της χαράς (1^ο τεταρτημόριο))

What are the animations that express joy? (or Animations that belongs to the same quadrant as emotion joy (1st quadrant))

(retrieve (?c) (and (?a | WhisselWheel|)

(?a ?b | isLocatedInQuadrant|)

(?b | WhisselWheelQuadrant|)

(?b ?c| hasAnimationRelated|)

(?c | BodyAnimation|)

(?a (some | modelsEmotion| | CONCEPT – FOR – Joy|))))

Result:

claping

explain

oh_my_God

right_hand_wave

Ποιες είναι οι εκφραστικές παράμετροι που θα «ενισχύσουν» αυτή την εμφύχωση με το συναίσθημα της χαράς;

What are the parameters of expression to enhance this animation with joy?

(retrieve (?b) (and (?a | WhisselWheel|)

(?a ?b | hasDimentionOfExpresivity|)

(?b | ExpressivityBodyParameters|)

(?a (some | modelsEmotion| | CONCEPT – FOR – Joy|))))

Result:

ExpressivityBodyParameters_Joy2

ExpressivityBodyParameters_Joy1

Πρέπει να σημειωθεί ότι κάθε ένα από τα δύο τελευταία στοιχεία περιλαμβάνει και τις έξι παραμέτρους εκφραστικότητας.

Κεφάλαιο 9

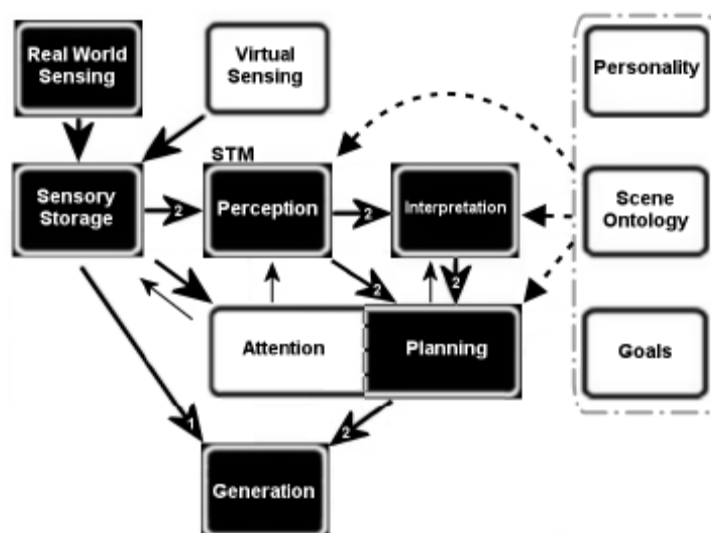
Συμπεράσματα-Μελλοντική Έρευνα και Περιορισμοί

Οι άνθρωποι κοινωνούν τα συναισθήματά τους όχι μόνο λεκτικά, αλλά και με την ασυνείδητη ή συνειδητή επιλογή των λέξεων, την εκφορά του λόγου, τις εκφράσεις του προσώπου τους, το βλέμμα τους, τη στάση του σώματος, τις χειρονομίες τους και φυσικά τις ενέργειές τους. Όλα αυτά μπορούν να αγνοηθούν όταν σκοπός είναι η γνωστοποίηση γεγονότων, αλλά γίνονται ιδιαίτερα σημαντικά όταν μιλάμε για συναίσθημα. Πληροφορίες για το συναίσθημα μπορούμε να λάβουμε από το λόγο, το πρόσωπο και το σώμα. Πολλές προσπάθειες ανάλυσης έχουν γίνει για κάθε ένα από τα παραπάνω χαρακτηριστικά χωριστά. Ειδικότερα το πρόσωπο και το σώμα βρίσκονται σε άμεση σχέση μεταξύ τους, με το δεύτερο να παίζει συμπληρωματικό αλλά ουσιαστικό ρόλο στην έκφραση του συναισθήματος.

Η παρούσα διατριβή ασχολήθηκε αρχικά με το πρόσωπο, παραθέτοντας στοιχεία της ανάλυσης, αλλά επικεντρώνοντας στη σύνθεση. Το σώμα αντιμετωπίστηκε με τον ίδιο τρόπο. Παρουσιάστηκαν στοιχεία της ανάλυσης που χρησιμοποιήθηκαν κατά μεγάλο ποσοστό σε εφαρμογές σύνθεσης. Για τη σύνθεση χρησιμοποιήθηκε ο Πράκτορας με σώμα και δυνατότητα ομιλίας (ECA) Greta, ο οποίος είναι απόλυτα συμβατός με το πρότυπο MPEG-4 και παρέχει ικανά στοιχεία ελέγχου για τις εφαρμογές της παρούσας διατριβής, η οποία δεν περιορίστηκε στις βασικές, συνηθισμένες εκφράσεις, αλλά ασχολήθηκε και με εκφράσεις προσώπου, στις οποίες δεν μπορεί να αποδοθεί εύκολα χαρακτηρισμός συναισθήματος, παρ'όλαυτά κάνουν έντονη την παρουσία τους στην καθημερινότητά μας. Οι εκφράσεις αυτές, που δεν ανήκουν σε κάποια κατηγορία, μοντελοποιήθηκαν με συνδυασμό γνωστών εκφράσεων και η πλειοψηφία τους πέρασε επιτυχώς τον έλεγχο ειδικών (ψυχολόγων), αλλά και κριτών χωρίς κάποια ειδίκευση. Άλλωστε τα αποτελέσματα αυτής της διατριβής στοχεύουν σε εφαρμογές απευθυνόμενες στο ευρύ κοινό. Ένας οποιοσδήποτε χρήστης που βλέπει έναν εικονικό ξεναγό στην οθόνη του ή έναν εικονικό συμπαίκτη ή ακόμα και έναν εικονικό συνομιλητή, επιθυμεί να βλέπει κάποιον που θα πλησιάζει όσο το δυνατόν περισσότερο στους ανθρώπους που συναντά καθημερινά και για να γίνει αυτό, ο

εικονικός ξεναγός-συμπαίκτης-συνομιλητής οφείλει να έχει τη δυνατότητα να εκφράζει συναισθήματα μέσα από υιοθέτηση αληθοφανών εκφράσεων. Η σύνθεση των ενδιάμεσων εκφράσεων και η δημιουργία κανόνων για τη μοντελοποίησή τους θα αποτελέσει τη βάση σημαντικών εφαρμογών στο χώρο της Επικοινωνίας Ανθρώπου-Μηχανής.

Ένα ζητούμενο της μελλοντικής έρευνας είναι ο σωστός συγχρονισμός εκφράσεων προσώπου, χειρονομιών και λόγου σε συνδυασμό με τη μελέτη της χρονικής εξέλιξης αυτών. Η ικανότητα ενός πράκτορα να κατανοεί, να «χρησιμοποιεί» ή να προσαρμόζεται στο συναισθήμα του χρήστη κρίνεται απαραίτητη σε συστήματα πλήθους περιοχών – διδασκαλία, πληροφόρηση, εμπόριο, τεχνική ή άλλου είδους υποστήριξη, διασκέδαση. Τα μελλοντικά συστήματα δεν πρέπει να περιορίζονται στην αναγνώριση κι έκφραση συγκεκριμένων συναισθημάτων, πρέπει και να μπορούν να τα χρησιμοποιούν ή να βρίσκουν τρόπους να τα αποφεύγουν ή να προσαρμόζουν τη διάθεσή τους στη διάθεση του χρήστη. Σε αυτή την κατεύθυνση και ως συνέχεια της δουλειάς που παρουσιάστηκε στο **Κεφάλαιο 7**, ενδεικτικά αναφέρεται το σύστημα της **Εικόνας 9.1**, η οποία αποτελεί την ολοκληρωμένη έκδοση του συστήματος της **Εικόνας 7.3**. Τα μεγάλα βέλη υποδεικνύουν την κατεύθυνση της πληροφορίας, τα μικρότερα βέλη τα σήματα ελέγχου, ενώ τα βέλη με τις διακεκομμένες γραμμές υποδεικνύουν ένωση με βάσεις δεδομένων εκτός συστήματος.



Εικόνα 9.1: Ολοκληρωμένο σύστημα αντιγραφής συναισθηματικής κατάστασης χρήστη από ECA

Με σκοπό την ευρεία χρήση των αποτελεσμάτων της παρούσας έρευνας, δημιουργήθηκε και η οντολογία που παρουσιάζεται στο **Κεφάλαιο 8** και η οποία έχει τη δυνατότητα να εμπλουτίζεται με καινούρια προφίλ εκφράσεων και φιλοδοξεί να αποτελέσει ένα χρήσιμο εργαλείο για όποιον θα ήθελε να αναζητήσει τρόπους απεικόνισης, αλλά και αναγνώρισης διαφορετικών εκφράσεων, χωρίς να περιορίζεται στις έξι βασικές εκφράσεις.

Βέβαια, στο σημείο αυτό οφείλουμε να αναφερθούμε στους περιορισμούς που παρουσιάζει ένα τέτοιο αντικείμενο έρευνας εξαιτίας της φύσης του, της ενασχόλησής του δηλαδή με τη συναισθηματική κατάσταση ανθρώπων.

Καταρχήν υπάρχουν οι περιορισμοί τεχνικής φύσεως. Για τη διαδικασία της ανάλυσης απαιτούνται καλές συνθήκες φωτισμού, αλλά και αποδοτικός συνδυασμός των διαφορετικών μορφών έκφρασης συναισθήματος. Η τοποθέτηση μικροφώνου για τη λήψη του ήχου δεν πρέπει να εμποδίζει την αναγνώριση της έκφρασης του προσώπου, ενώ η όλη διαδικασία πρέπει να γίνεται με τέτοιο τρόπο, ώστε να επηρεάζει όσο το δυνατόν λιγότερο τη συναισθηματική κατάσταση του υποκειμένου. Η διαδικασία της σύνθεσης παρουσιάζει προβλήματα κυρίως στη διαδικασία του συγχρονισμού προσώπου-σώματος-λόγου.

Πέραν όμως των περιορισμών τεχνικής φύσεως, υπάρχουν και ηθικοί περιορισμοί κυρίως λόγω της φύσης των συγκεκριμένων συστημάτων. Αν και τα ηθικά προβλήματα που ανακύπτουν δεν είναι τόσο σοβαρά, όσο υποδεικνύουν στο ευρύ κοινό τα έργα επιστημονικής φαντασίας στη λογοτεχνία ή στον κινηματογράφο (Frankenstein, 1984, Blade Runner, Η Οδύσσεια του Διαστήματος κ.ο.κ.), προβληματισμός επί του θέματος κρίνεται απαραίτητος.

Όταν ελέγχεις τη συναισθηματική κατάσταση κάποιου ατόμου, πρέπει να είσαι σίγουρος ότι γνωρίζει ότι παρακολουθείται και ότι έχει τη δυνατότητα, ανά πάσα στιγμή να εγκαταλείψει την εφαρμογή. Άλλωστε οι συνήθεις εφαρμογές αφορούν σε εντοπισμό ανίας ή εκνευρισμού λόγω της εφαρμογής, όμως προσοχή πρέπει να δοθεί σε περιπτώσεις παρακολούθησης συναισθηματικής κατάστασης υπαλλήλων σε μία επιχείρηση, σε περιπτώσεις έκτακτης ανάγκης σε θέματα υγείας καθώς και σε περιπτώσεις παρακολούθησης ατόμων με ειδικές ανάγκες, όπου οι γραμμές διαχωρισμού μεταξύ βοηθητικού χαρακτήρα του συστήματος παρακολούθησης της συναισθηματικής κατάστασης / απόκρισης ECA κι εκμετάλλευσης ιδιαιτέρων συνθηκών είναι πολύ λεπτές. Οφείλουμε λοιπόν να είμαστε ιδιαίτερα προσεκτικοί και να μην επαναπαυόμαστε στην αυτονόητη ευθυνότητα του ερευνητή, αλλά να φροντίζουμε τα συστήματά μας να είναι σύμφωνα με διεθνείς κανονισμούς και να έχουμε πλήρη επίγνωση σε ποια κατηγορία χρηστών απευθυνόμαστε κάθε φορά.

Παράρτημα – Αποδόσεις όρων

FACS	Facial Action Coding System	Σύστημα Κωδικοποίησης Κινήσεων Προσώπου
AUs	Action Units	Μονάδες δράσης
FBA	Facial and Body Animation	Σχεδιασμός Κίνησης Προσώπου και Σώματος
FDP	Facial Definition Parameters	Παράμετροι Προσδιορισμού Προσώπου
FAP	Facial Animation Parameters	Παράμετροι Κίνησης Προσώπου
FAPU	Facial Animation Parameter Units	Μονάδες Κίνησης Χαρακτηριστικών Προσώπου
FP	Facial Points	Σημεία Προσώπου
BDP	Body Definition Parameters	Παράμετροι Προσδιορισμού Σώματος
BAP	Body Animation Parameters	Παράμετροι Κίνησης Σώματος
BBA	Bone-Based Animation	Κίνηση Βασιζόμενη στα Οστά
ECA	Embodied Conversational Agent	Πράκτορας με Σώμα και δυνατότητα ομιλίας
SAL	Sensitive Artificial Listener	

Βιβλιογραφία

- [1]. P. Ekman and W. Friesen. Unmasking the Face: A guide to recognizing emotions from facial clues. Prentice-Hall, Inc., 1975.
- [2]. P. Ekman. Emotion in the human face. Cambridge University Press, 1982
- [3]. P. Ekman and W. Friesen. Felt, false, miserable smiles. Journal of Nonverbal Behaviour, 6(4):238–251, 1982
- [4]. W. H. Dittrich, Facial motion and the recognition of emotions. Psychologische Beitrage 33 366-377, 1991
- [5]. A.W. Young and V. Bruce. Perceptual categories and the computation of “grandmother”. In V. Bruce, editor, Face Recognition, volume 1 of 3, pages 5–49. Lawrence Erlbaum Associates, a special issue of the European journal of cognitive psychology edition, 1991
- [6]. D.W. Massaro and M.M. Cohen. Perception of synthesized audible and visible speech. Psychological Science, 1(1):55–63, January 1990.
- [7]. C. Benoit, T. Lallouache, T. Mohamedi, A. Tseva, and C. Abry. Nineteen (+two) french *visemes* for visual speech synthesis. In Proceedings of the ESCA Workshop on Speech Synthesis, Autrans, 1990.
- [8]. Q. Summerfield. Lipreading and audio-visual speech perception. Philosophical Transactions of the Royal Society of London, 335(1273):71–78, 1992.
- [9]. P. Ekman and W. Friesen. Unmasking the Face: A guide to recognizing emotions from facial clues. Prentice-Hall, Inc., 1975
- [10]. H. G. Wallbott and Klaus R. Scherer. Cues and channels in emotion recognition. Journal of Personality and Social Psychology, 51(4):690–699, 1986
- [11]. H. G. Wallbott. Bodily expression of emotion. European Journal of Social Psychology, 28:879–896, 1998
- [12]. S. Kshirsagar, T. Molet, N. Magnanthalman, Principal components of expressive speech animation. In Proc. ComputerGraphics International, IEEE Computer Society, pp. 38–44, 2001
- [13]. S. Marsella, W.L. Johnson, and K. LaBore. Interactive pedagogical drama. In Proceedings of the 4th International Conference on Autonomous Agents, Barcelona, Spain, June 2000.

- [14]. S. Marsella and J. Gratch. Modeling coping behaviour in virtual humans: Don't worry, be happy. In proceedings of the 2nd International Conference on Autonomous Agents and Multiagent Systems, Melbourne, Australia, 2003.
- [15]. T. Lebourque, S. Gibet, High level specification and control of communication gestures: The gessyca system. In Proceedings of the Computer Animation 1999, IEEE Computer Society, p. 24, 1999.
- [16]. J. Cassell, H. Vilhjálmsón, and T. Bickmore. BEAT: The behaviour expression animation toolkit. In Eugene Fiume, editor, SIGGRAPH 2001, Computer Graphics Proceedings, pages 477–486. ACM Press / ACM SIGGRAPH, 2001.
- [17]. C. Pelachaud, V. Carofiglio, B. De Carolis, and F. de Rosis. Embodied contextual agent in information delivering application. In First International Joint Conference on Autonomous Agents & Multi-Agent Systems (AAMAS), Bologna, Italy, July 2002.
- [18]. I. Poggi. Mind markers. In C. Mueller and R. Posner, editors, The Semantics and Pragmatics of Everyday Gestures. Berlin Verlag Arno Spitz, Berlin, 2001.
- [19]. G. Ball and J. Breese. Emotion and personality in a conversational agent. In: S. Prevost, J. Cassell, J. Sullivan, E. and Churchill, (eds.), *Embodied Conversational Characters*. MIT Press, Cambridge, MA, 2000.
- [20]. J. Bates, The role of emotion in believable agents. *Communications of the ACM*, 37, 7, 1994.
- [21]. J. Cassell, J. Bickmore, M. Billinghurst, L. Campbell, K. Chang, H. Vilhjálmsón, and H. Yan, Embodiment in conversational interfaces: Rea, *CHI'99*, Pittsburgh, PA, pp. 520-527, 1999.
- [22]. I. Poggi and C. Pelachaud, Emotional meaning and expression in animated faces. In: A. Paiva (Ed): *Affective Interactions*. Springer, New York, 2000.
- [23]. I. S. Pandzic and R. Forchheimer eds., *MPEG-4 Facial Animation - The standard, implementations and applications*. John Wiley & Sons, Chichester, 2002.
- [24]. B. Krenn, M. Grice, P. Piwek, M. Schröder, M. Klesen, S. Baumann, H. Pirker, K. van Deemter, E. Gstrein: Generation of Multi-modal Dialogue for Net Environment, In Proceedings of KONVENS-02, 30 September - 2 October 2002, pp.91-98, Saarbrücken, Germany.
- [25]. D. Chi, M. Costa, L. Zhao, N. Badler, The EMOTE model for effort and shape. In Proceedings of the 27th annual conference on Computer graphics and

- interactive techniques, ACM Press/Addison-Wesley Publishing Co., pp. 173–182, 2000.
- [26]. M. Byun, N. I. Badler, FacEMOTE: qualitative parametric modifiers for facial animations. In Proceedings of the 2002 ACM SIGGRAPH/ Eurographics symposium on Computer animation, ACM Press, pp. 65– 71, 2002.
 - [27]. F. A. Barrientos, Controlling Expressive Avatar Gesture. PhD thesis, University of California, Berkeley, EECS Department, Computer Science Division, 2002.
 - [28]. ISO/IEC JTC1/SC29/WG11 N3205, "Multi-users technology (Requirements and Applications)", December 1999, Maui
 - [29]. M. Tekalp, "Face and 2-D Mesh Animation in MPEG-4", Tutorial Issue On The MPEG-4 Standard, *Image Communication Journal*, Elsevier, 1999.
 - [30]. P. Ekman and W. Friesen, *The Facial Action Coding System*, Consulting Psychologists Press, San Francisco, CA, 1978 (<http://www.paulekman.com>)
 - [31]. EC TMR Project PHYSTA Report, “Development of Feature Representation from Facial Signals and Speech,” January 1999.
 - [32]. P. Ekman, “Facial expression and Emotion,” *Am. Psychologist*, vol. 48 pp.384-392, 1993
 - [33]. P. Ekman and W.V. Friesen, “Pictures of Facial Affect”, Palo Alto, CA: Consulting Psychologists Press, 1978.
 - [34]. C. Izard, L. Dougherty, and E.A. Hembree, “A System for Identifying Affect Expressions by Holistic Judgements”, *Technical Report, Univ. of Delaware*, 1983.
 - [35]. C. M. Whissel, *The dictionary of affect in language*, R. Plutchnik and H. Kellerman (Eds) *Emotion: Theory, research and experience: vol 4, The measurement of emotions*. Academic Press, New York, 1989.
 - [36]. R. Plutchik, *Emotion: A psychoevolutionary synthesis*, Harper and Row, New York, 1980.
 - [37]. R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz and J. Taylor, “Emotion Recognition in Human-Computer Interaction”, *IEEE Signal Processing Magazine*, January 2001
 - [38]. N. Tsapatsoulis, K. Karpouzis, G. Stamou, F. Piat and S. Kollias, “A Fuzzy System for Emotion Classification based on the MPEG-4 Facial Definition Parameter Set,” in *Proceedings of the EUSIPCO-2000*, Tampere, Finland, September 2000.

- [39]. F. Parke and K. Waters, *Computer Facial Animation*, A K Peters, 1996
- [40]. G. Faigin, "The Artist's Complete Guide to Facial Expressions", Watson-Guptill, New York, 1990
- [41]. <http://www.media.mit.edu/>
- [42]. EC TMR Project PHYSTA Report, "Review of Existing Techniques for Human Emotion Understanding and Applications in Human-Computer Interaction", October 1998.
- [43]. P. Chellapa, C. Wilson and S. Sirohey, "Human and Machine Recognition of Faces: A Survey," *Proceedings of IEEE*, vol.83, no. 5, pp. 705-740, 1995.
- [44]. H.A. Rowley, S. Baluja, and T. Kanade, "Neural Network-Based Face Detection", *IEEE Trans. on PAMI*, vol.20, no. 1, pp.23-28. Jan. 1998.
- [45]. K. Karpouzis, N. Tsapatsoulis and S. Kollias, "Moving to Continuous Facial Expression Space using the MPEG-4 Facial Definition Parameter (FDP) Set," *in Proc. of the Electronic Imaging 2000 Conference of SPIE*, San Jose, CA, USA, January 2000.
- [46]. Kin-Man Lam and Hong Yan, "An Analytic-to-Holistic Approach for Face Recognition Based on a Single Frontal View," *IEEE Trans. on PAMI*, vol. 20, no. 7, July 1998. <http://www.computer.org/tpami/tp1998/i7toc.htm>
- [47]. J. Ostermann and E. Haratsch, "An animation definition interface: Rapid design of MPEG-4 compliant animated faces and bodies", Int'l. Workshop on Synthetic – Natural Hybrid Coding and Three Dimensional Imaging, Rhodes, Greece, 5-9 September 1997, pp.216 – 219.
- [48]. F. Lavagetto and R. Pockaj, "The facial animation engine: Toward a high-level interface for the design of MPEG-4 compliant animated faces", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, pp.277-289, Mar. 1999.
- [49]. ISTFACE - MPEG-4 Facial Animation System – Version 3.3.1 Gabriel Abrantes, (Developed in the context of the European Project ACTS MoMuSys © 97-98 Instituto Superior Tecnico).
- [50]. George J.Klir and Bo Yuan, "Fuzzy Sets and Fuzzy Logic, Theory and Applications", Prentice Hall, New Jersey, 1995.
- [51]. A. Raouzaïou, N. Tsapatsoulis and S. Kollias, "Parameterized Facial Expression Synthesis for Videoconferencing Applications", *Proc. of Workshop on Very Low Bitrate Video Coding (VLBV01)*, Athens, Greece, October 2001 (<http://www.image.ntua.gr/vlbv01>).

- [52]. M. Preda, "Virtual Character Animation within MPEG-4", Workshop and Exhibition on MPEG-4, San Jose, June 2002.
- [53]. ISO/IEC IS 14496-2 Visual, 1999
- [54]. S. Ioannou, "Αποτελεσματικές τεχνικές εξαγωγής χαρακτηριστικών και ανάλυσης εκφράσεων στην επικοινωνία ανθρώπου μηχανής", Ph.D. Thesis, School of Electrical and Computer Engineering, National Technical University of Athens
- [55]. D. McNeill, Hand and Mind - What gestures reveal about thought. The University Chicago Press, Chicago, IL, 1992.
- [56]. S. Gibet, J.-F. Kamp, F. Poirier, Gesture analysis: Invariant laws in movement. In Gesture-Based Communication in Human- Computer Interaction - GW 2003, Camurri A., Volpe G., (Eds.), no. 2915
- [57]. S. N. Steketee, N. I. Badler, Parametric keyframe interpolation incorporating kinetic adjustment and phrasing control. In Proceedings of the 12th annual conference on Computer graphics and interactive techniques (1985), ACM Press, pp. 255–262. 7
- [58]. B. Hartmann, M. Mancini, C. Pelachaud., Formational parameters and adaptive prototype instantiation for mpeg-4 compliant gesture synthesis. In Proceedings of the Computer Animation 2002 (2002), IEEE Computer Society, p. 111. 3, 6, 7
- [59]. Isabella Poggi, Catherine Pelachaud, Emanuela Magno Caldognetto: Gestural Mind Markers in ECAs. Gesture Workshop 2003: 338-349
- [60]. E. Bevacqua, M. Mancini, C. Pelachaud. Speaking with Emotions AISB 2004 Convention: Motion, Emotion and Cognition, University of Leeds, UK, 2004
- [61]. B. Hartmann, M. Mancini, C. Pelachaud. HUMAINE WP3 Workshop Geneva: Expressivity Control - Enabling Emotional Displays and Individualized Actions in Embodied Conversational Agents, Poster
- [62]. Vincent Maya, Myriam Lamolle, Catherine Pelachaud: Influences and Embodied Conversational Agents. AAMAS 2004: 1306-1307
- [63]. J. W. Young, Head and Face Anthropometry of Adult U.S. Civilians, FAA Civil Aeromedical Institute, 1963-1993 (final report 1993).
- [64]. R. W. Picard, *Affective Computing*. MIT Press, Cambridge, MA, (1997)
- [65]. T. Wehrle, S. Kaiser, S. Schmidt and K. R. Scherer, Studying the dynamics of emotional expression using synthesized facial muscle movements. *Journal of Personality and Social Psychology*, 78 (1) (2000) 105-119

- [66]. D. Sander, D. Grandjean and K. R. Scherer, A systems approach to appraisal mechanisms in emotion. *Neural Networks*, 18, 317-352 (2005)
- [67]. I. J. Roseman and C. A. Smith, Appraisal Theory: Overview, Assumptions, Varieties, Controversies. In Scherer, K.R., Schorr, A., & Johnstone, T., (Eds) *Appraisal Processes in Emotion: Theory Methods, Research*. Oxford, New York: Oxford University Press, 3-19 (2001)
- [68]. A. Ortony, G. L. Clore and A. Collins, *The Cognitive Structure of Emotions*, Cambridge, UK: Cambridge University Press (1988)
- [69]. A. J. Pegna, A. Khateb, F. Lazeyras and M. L. Seghier, Discriminating emotional faces without primary visual cortices involves the right amygdala. *Nature Neuroscience*, 8(1), 24-25 (2004)
- [70]. K. R. Scherer, Appraisal Considered as a Process of Multilevel Sequential Checking. In Scherer, K.R., Schorr, A., & Johnstone, T., (Eds) *Appraisal Processes in Emotion: Theory Methods, Research*. Oxford, New York: Oxford University Press, 92-129 (2001)
- [71]. K. R. Scherer, On the Nature and Function of Emotion: A Component Process Approach. In Scherer, K.R., & Ekman, P., (Eds) *Approaches to Emotion*. Hillsdale, New Jersey, London: Lawrence Erlbaum Associates, Publishers. 293-318 (1984)
- [72]. T. Wehrle, S. Kaiser, S. Schmidt and K. R. Scherer, Studying the dynamics of emotional expression using synthesized facial muscle movements. *Journal of Personality and Social Psychology*, 78 (1), 105-119, 2000.
- [73]. B. Hartmann, M. Mancini, C. Pelachaud, Implementing Expressive Gesture Synthesis for Embodied Conversational Agents. *Gesture Workshop* (2005) Vannes
- [74]. Harald G. Wallbott. Bodily expression of emotion. *European Journal of Social Psychology*, 28:879-896, 1998.
- [75]. Jinni A. Harrigan. Listener's body movements and speaking turns. *Communication Research*, 12(2):233-250, April 1985.
- [76]. P. Gallaher, Individual differences in nonverbal behavior: Dimensions of style. *Journal of Personality and Social Psychology* 63 (1992)
- [77]. <http://xface.itc.it/> : **Xface** is a set of open source tools for creation of MPEG-4 based 3D Talking Heads.

- [78]. G. Caridakis, A. Raouzaïou, K. Karpouzis, S. Kollias, "Synthesizing Gesture Expressivity Based on Real Sequences", Workshop on multimodal corpora: from multimodal behaviour theories to usable models, LREC 2006 Conference, Genoa, Italy, 24-26 May.
- [79]. Mario Gutiérrez, Daniel Thalmann, Frédéric Vexo, Laurent Moccozet, Nadia Magnenat-Thalmann, Michela Mortara, and Michela Spagnuolo. An ontology of virtual humans: incorporating semantics into human shapes. In *Proceedings of Workshop towards Semantic Virtual Environments (SVE05), March 2005*, pages 57–67, 2005.
- [80]. W3C. Owl web ontology language. <http://www.w3.org/tr/owl-features/>.
- [81]. Protégé. (c) 2005 Stanford Medical Informatics. <http://protege.stanford.edu/index.html>.
- [82]. ISO/IEC JTC1/SC29/WG11 N4668. Overview of the mpeg-4 standard, March 2002.
- [83]. T. R. Gruber. The role of a common ontology in achieving sharable, reusable knowledge bases. In *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning*, pages 601–602, 1991.
- [84]. nRQL. new racer query language (nrql) interface to owl ontologies in protégé. http://www.cs.concordia.ca/~k_bhoopa/nrql.html.
- [85]. RacerPro. Renamed abox and concept expression reasoner. <http://www.racersystems.com>.
- [86]. ERMIS, Emotionally Rich Man-Machine Intelligent System. IST-2000-29319 (<http://www.image.ece.ntua.gr/ermis>)
- [87]. HUMAINE NoE, Human-Machine Interaction Network onEmotion, 2004-2007 (<http://emotion-research.net>)
- [88]. J. Weizenbaum, ELIZA – A Computer Program For the Study of Natural Language Communication Between Man and Machine, Communications of the ACM, Volume 9, Number 1, 1966, pp.36-45.
- [89]. Web3D Consortium Humanoid Animation Working Group, H-Anim 1.1 specification, <http://h-anim.org/spec1.1/>, 1999.
- [90]. <http://www.iut.univ-paris8.fr/greta/>
- [91]. S. Pasquariello, C. Pelachaud, Greta: A Simple Facial Animation Engine, 6th Online World Conference on Soft Computing in Industrial Applications, Session on Soft Computing for Intelligent 3D Agents, Septembre

- [92]. B. Hartmann, M. Mancini, C. Pelachaud, Formational Parameters and Adaptive Prototype Instantiation for MPEG-4 Compliant Gesture Synthesis. Proceedings of Computer Animation 2002, IEEE Computer Society Press, 2002.
- [93]. Elisabetta Bevacqua, Amaryllis Raouzaïou, Christopher Peters, George Caridakis, Kostas Karpouzis, Catherine Pelachaud, Maurizio Mancini: Multimodal Sensing, Interpretation and Copying of Movements by a Virtual Agent. Perception and Interactive Technologies, Kloster Irsee, Germany, June 19 - 21, 2006
- [94]. Leslie G. Farkas, “Anthropometry of the Head and Face”, Lippincott Williams & Wilkins, September 1994.

Κατάλογος Δημοσιεύσεων της συγγραφέως

Διεθνή Περιοδικά με διαδικασία κρίσης

- [J1] A. Raouzaïou, G. Caridakis, L. Malatesta, K. Karpouzis and S. Kollias, "Emotion Theory and Multimodal Synthesis of Affective ECAs", International Journal of Humanoid Robotics, submitted for publication
- [J2] M. Wallace, S. Ioannou, A. Raouzaïou, K. Karpouzis, S. Kollias, "Dealing with Feature Uncertainty in Facial Expression Recognition Using Possibilistic Fuzzy Rule Evaluation", International Journal of Intelligent Systems Technologies and Applications, vol. 1, nos ³/₄, pp. 409-429, 2006
- [J3] S. Ioannou, A. Raouzaïou, V. Tzouvaras, T. Mailis, K. Karpouzis, S. Kollias, "Emotion recognition through facial expression analysis based on a neurofuzzy network", Special Issue on Emotion: Understanding & Recognition, Neural Networks, Elsevier, Volume 18, Issue 4, May 2005, Pages 423-435
- [J4] A. Raouzaïou, N. Tsapatsoulis, K. Karpouzis and S. Kollias, "Parameterized facial expression synthesis based on MPEG-4 ", Eurasip Journal on Applied Signal Processing, Vol. 2002, No 10, pp. 1021-1038, 2002
- [J5] K. Karpouzis, N. Tsapatsoulis, A. Raouzaïou, G. Mosxovitis and S. Kollias, "Enhancing Nonverbal Human Computer Interaction with Expression Recognition", ACM SIGCAPH Bulletin, Issue 67, June 2000, pp. 1-9, ACM Press, NY, USA

Κεφάλαια σε βιβλία

- [B6] A. Raouzaïou, E. Spyrou, K. Karpouzis and S. Kollias, "An Intermediate Expressions Generator System in the MPEG-4 Framework", L. Atzori, D. Giusto, R. Leonardi, F. Pereira. (Eds.): Visual Content Processing and Representation 2005, LNCS, Vol. 3893, pp. 129 – 136, Springer-Verlag, 2006.
- [B7] K. Karpouzis, A. Drosopoulos, S. Ioannou, A. Raouzaïou, N. Tsapatsoulis, S. Kollias, "Facial and Body Feature Extraction for Emotionally-Rich HCI", Mehdi Khosrow-Pour (ed.), Encyclopedia of Information Science and Technology, Idea Group Publishers, pp. 1180 - 1185, 2005

- [B8] K. Karpouzis, A. Raouzaïou, A. Drosopoulos, S. Ioannou, T. Balomenos, N. Tsapatsoulis and S. Kollias, "Facial expression and gesture analysis for emotionally-rich man-machine interaction", N. Sarris, M. Strintzis, (eds.), 3D Modeling and Animation: Synthesis and Analysis Techniques, pp. 175-200, Idea Group Publ., 2004
- [B9] A. Raouzaïou, K. Karpouzis and S. Kollias, "Online Gaming and Emotion Representation", N. Garcia, J. Martinez, L. Salgado (eds.), Visual Content Processing and Representation, Lecture Notes in Computer Science, Springer-Verlag, Vol. 2849, pp. 298 - 305, 2004
- [B10] K. Karpouzis, A. Raouzaïou, St. Kollias, "'Moving' avatars: Emotion Synthesis in Virtual Worlds", C. Stephanidis, J. Jacko (eds.), Human – Computer Interaction: Theory and Practice, vol. 2, pp. 503 – 507, Lawrence Erlbaum Associates, Inc., 2003.
- [B11] N. Tsapatsoulis, A. Raouzaïou, S. Kollias, R. Cowie and E. Douglas-Cowie, "Emotion Recognition and Synthesis based on MPEG-4 FAPs", in MPEG-4 Facial Animation, Igor Pandzic, R. Forchheimer (eds), John Wiley & Sons, UK, 2002.

Συνέδρια με διαδικασία κρίσης

- [C12] G. Caridakis, A. Raouzaïou, K. Karpouzis, S. Kollias, "Synthesizing Gesture Expressivity Based on Real Sequences", Workshop on multimodal corpora: from multimodal behaviour theories to usable models, LREC 2006 Conference, Genoa, Italy, 24-26 May.
- [C13] L. Malatesta, A. Raouzaïou, K. Karpouzis and S. Kollias, "MPEG-4 Facial Expression Synthesis based on Appraisal Theory", rd IFIP conference in Artificial Intelligence Applications and Innovations, AIAI 2006
- [C14] E. Bevacqua, A. Raouzaïou, C. Peters, G. Caridakis, K. Karpouzis, C. Pelachaud, M. Mancini, "Multimodal Sensing, Interpretation and Copying of Movements by a Virtual Agent", Perception and Interactive Technologies, Kloster Irsee, Germany, June 19 - 21, 2006
- [C15] A. Garcia-Rojas, F. Vexo, D. Thalmann, A. Raouzaïou, K.Karpouzis, S. Kollias, "Emotional Body Expression Parameters In Virtual Human Ontology", In Proceedings of 1st Int. Workshop on Shapes and Semantics, Matsushima, Japan, June 2006, pp. 63-70
- [C16] A. Garcia-Rojas, F. Vexo, D. Thalmann, A. Raouzaïou, K.Karpouzis, S. Kollias, L. Moccozet, N. Magnenat-Thalmann., "Emotional Face Expression Profiles Supported by Virtual Human Ontology", Computer Animation and Social Agents (CASA) 2006, Geneva, July 5th-7th, 2006

- [C17] S. Ioannou, M. Wallace, K. Karpouzis, A. Raouzaïou and S. Kollias, "Combination of Multiple Extraction Algorithms in the Detection of Facial Features", Proceedings of the IEEE International Conference on Image Processing (ICIP), Genova, Italy, September 2005
- [C18] S. Ioannou, M. Wallace, K. Karpouzis, A. Raouzaïou and S. Kollias, "Confidence-Based Fusion of Multiple Feature Cues for Facial Expression Recognition", FUZZ-IEEE 2005, May 22-25, Reno, Nevada, USA
- [C19] A. Raouzaïou, E. Spyrou, K. Karpouzis and S. Kollias, "Emotion Synthesis: an Intermediate Expressions' Generator System in the MPEG-4 Framework", International Workshop VLBV05, 15-16 September 2005, Sardinia, Italy
- [C20] M. Mancini, B. Hartmann, C. Pelachaud, A. Raouzaïou, K. Karpouzis, "Expressive Avatars in MPEG-4", IEEE International Conference on Multimedia & Expo (ICME) 2005, July 6-8, Amsterdam, The Netherlands
- [C21] A. Raouzaïou, K. Karpouzis, S. Kollias, "Emotion Synthesis in the MPEG-4 Framework", IEEE International workshop on multimedia signal processing (MMSP), Siena, Italy
- [C22] A. Raouzaïou, S. Ioannou, G. Akrivas, K. Karpouzis and S. Kollias, "Adaptation of Expression Analysis Based on Evaluation Principles", Fourth European Symposium on Intelligent Technologies and their implementation on Smart Adaptive Systems (Eunite 2004), Aachen, Germany, June 2004.
- [C23] S. Ioannou, A. Raouzaïou, K. Karpouzis, M. Pertselakis, N. Tsapatsoulis, S. Kollias, "Adaptive Rule-Based Facial Expression Recognition", G. Vouros, T. Panayiotopoulos (Eds.), Lecture Notes in Artificial Intelligence, Vol. 3025, Springer-Verlag, pp. 466 - 475, 2004.
- [C24] S. Ioannou, A. Raouzaïou, K. Karpouzis and S. Kollias, "Adaptation of Facial Feature Extraction and Rule Generation in Emotion-Analysis Systems", International Joint Conference on Neural Networks (IJCNN 2004), Budapest, Hungary, July 2004.
- [C25] T. Balomenos, A. Raouzaïou, S. Ioannou, A. Drosopoulos, K. Karpouzis, and S. Kollias, "Emotion Analysis In Man-Machine Interaction Systems", Workshop on Multimodal Interaction and Related Machine Learning Algorithms (MLMI04), Geneva
- [C26] M. Wallace, A. Raouzaïou, N. Tsapatsoulis, S. Kollias, "Facial Expression Classification Based on MPEG-4 FAPs: The Use of Evidence and Prior Knowledge for Uncertainty Removal", Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Budapest, Hungary, July 2004, vol.1, pp 51-54

- [C27] T. Balomenos, A. Raouzaïou, S. Ioannou, A. Drosopoulos, K. Karpouzis, S. Kollias, "Emotion Analysis in Man-Machine Interaction Systems", Samy Bengio, Hervé Bourlard (Eds.), Machine Learning for Multimodal Interaction, Lecture Notes in Computer Science, Vol. 3361, 2004, pp. 318 - 328, Springer-Verlag .
- [C28] A. Raouzaïou, K. Karpouzis and S. Kollias, "Emotion Synthesis in Virtual Environments", 6th International Conference on Enterprise Information Systems, April 2004, Porto, Portugal.
- [C29] A. Raouzaïou, K. Karpouzis and S. Kollias, "Online Gaming and Emotion Representation", International Workshop Very Low Bitrate Video Coding 2003 (VLBV03), 18-19 September, Madrid, Spain
- [C30] A. Raouzaïou, K. Karpouzis and S. Kollias, "Emotion Representation for Online Gaming", IEEE International Conference on Multimedia & Expo (ICME), July 6-9, Baltimore, MD, USA
- [C31] K. Karpouzis, A. Raouzaïou, N. Tsapatsoulis and S. Kollias, "Emotion Representation for Virtual Environments", ConTEL 2003, 7th International Conference on Telecommunications, Zagreb, Croatia, June 11-13, 2003.
- [C32] K. Karpouzis, A. Raouzaïou, P. Tzouveli and S. Kollias, "MPEG-4: One multimedia standard to unite all", IEEE International Conference on Multimedia & Expo (ICME), July 6-9, Baltimore, MD, USA
- [C33] A. Raouzaïou, S. Ioannou, K. Karpouzis, N. Tsapatsoulis, S. Kollias, R. Cowie, "An Intelligent Scheme for Facial Expression Recognition", O. Kaynak (ed.), Artificial Neural Networks and Neural Information Processing, Lecture notes in Computer Science 2714, Springer, pp. 1109 – 1116, 2003.
- [C34] T. Balomenos, A. Raouzaïou, K. Karpouzis, S. Kollias and R. Cowie, "An Introduction to Emotionally Rich Man-Machine Intelligent Systems", Eunite 2003, Third European Symposium on Intelligent Technologies, Hybrid Systems and their implementation on Smart Adaptive Systems, Oulu, Finland
- [C35] A. Raouzaïou, S. Ioannou, K. Karpouzis, S. Kollias, R. Cowie , "Facial Expression Analysis", Eunite 2003, Third European Symposium on Intelligent Technologies, Hybrid Systems and their implementation on Smart Adaptive Systems, Oulu, Finland
- [C36] K. Karpouzis, A. Raouzaïou, S. Kollias, "'Moving' avatars: Emotion Synthesis in Virtual Worlds", Human-Computer Interaction International 2003, 22 - 27 June, Crete, Greece, vol. 2, pp. 503 - 507

- [C37] Th. Balomenos, A. Drosopoulos, A. Raouzaïou, K. Karpouzis, St. Kollias, "Use of the finite element method in 3D structure and motion estimation of non-rigid objects", SPIE Electronic Imaging 2003, 20–24 January 2003, Santa Clara, California, USA
- [C38] K. Karpouzis, A. Raouzaïou and S. Kollias, "Emotionally-rich Virtual Environment", PCI2003, 9th Panhellenic Conference in Informatics, Thessaloniki, Greece, November 2003
- [C39] A. Raouzaïou, S. Ioannou, K. Karpouzis, N. Tsapatsoulis, S. Kollias, R. Cowie , "An Intelligent Scheme for Facial Expression Recognition", Proceedings of the International Conference on Artificial Neural Networks (ICANN), Istanbul, Turkey, June 2003
- [C40] A. Raouzaïou, N. Tsapatsoulis, V. Tzouvaras, G. Stamou and S. Kollias , "A Hybrid Intelligence System for Facial Expression Recognition ", Eunit 2002, European Symposium on Intelligent Technologies, Hybrid Systems and their implementation on Smart Adaptive Systems, Algarve, Portugal
- [C41] A. Raouzaïou, N. Tsapatsoulis and S. Kollias, "Parameterized Facial Expression Synthesis for Videoconferencing Applications", Proc.of Intl Workshop on Very Low Bitrate Video Coding (VLBV01), Athens, Greece, October 2001
- [C42] A. Raouzaïou, K. Karpouzis and S. Kollias, "Facial Expression Analysis in MPEG-4 Sequences", Proceedings of ICAV 3D, Mykonos, Greece, August 2001

Βιογραφικό Σημείωμα

Η υποψήφια διδάκτορας Αμαρυλλίς-Ολγα Ραουζαίου γεννήθηκε στην Αθήνα στις 23 Ιουλίου 1977. Το 1995 αποφοίτησε από την Ελληνογαλλική Σχολή Ουρσουλινών και την ίδια χρονιά έγινε δεκτή στη Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του ΕΜΠ. Ακολούθησε την κατεύθυνση Πληροφορικής και αποφοίτησε το 2000 με βαθμό 8,10. Το Δεκέμβριο του ίδιου χρόνου έγινε δεκτή από τη Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του ΕΜΠ για εκπόνηση διδακτορικής διατριβής στον τομέα της Ψηφιακής Επεξεργασίας Εικόνας υπό την επίβλεψη του καθηγητή κ. Σ. Κόλλια.

Ξένες Γλώσσες

Αγγλικά: Certificate of Proficiency in English

Γαλλικά: Diplôme d'études françaises (2^e degré) (Sorbonne II) - Option: Littérature
Diplôme Approfondi de Langue Française (D.A.L.F.) - Option: Mathématiques et Sciences de la Matière

Κατά τη διάρκεια εκπόνησης της διατριβής της, η κ. Ραουζαίου εργάστηκε ως ερευνήτρια σε διάφορα ερευνητικά και αναπτυξιακά έργα και υποστήριξε εργαστηριακά μαθήματα του τομέα Τεχνολογίας Πληροφορικής και Υπολογιστών υπό την επίβλεψη του καθηγητή Σ. Κόλλια.

Ειδικότερα

Επικουρικό Διδακτικό Έργο:

1. Γραφική με Υπολογιστές [2006, 2005, 2004, 2003, 2002, 2001]
2. Ψηφιακή Επεξεργασία Εικόνας [2000]
3. Τεχνολογία Εικόνας σε Συστήματα Πολυμέσων [2001]
4. Ψηφιακή Επεξεργασία Εικόνας και Ανάλυση Video [2002]

Συμμετοχή σε Ερευνητικά - Αναπτυξιακά Προγράμματα

- **HUMAINE:** Human-Machine Interaction Network on Emotion, Network of Excellence, IST-2002-2.3.1.6 Multimodal Interfaces Συμμετοχή στους τομείς που σχετίζονται με τη σύνθεση εκφράσεων προσώπου και τους εικονικούς χαρακτήρες, 2004-2008.
- **ERMIS:** Emotionally Rich Man – Machine Interaction Systems, IST 2000-29319 (2001-2004) Συμμετοχή στους τομείς που σχετίζονται με τη δημιουργία προφίλ για την ανάλυση συναισθήματος από εικόνες και τη σύνθεση εκφράσεων προσώπου, 2001-2004.
- **ΠΥΘΑΓΟΡΑΣ** ΥΠΕΠΘ, Πρόγραμμα ΕΠΕΑΕΚ, Ευφύια και Συναισθημα σε διαδραστικά περιβάλλοντα. Συμμετοχή στον τομέα του συναισθήματος, 2004-2006.