

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Κωνσταντίνος Καρπούζης

Διπλ. Ηλεκτρολόγος Μηχανικός Ε.Μ.Π.

ΠΡΟΣΑΡΜΟΖΟΜΕΝΑ ΠΑΡΑΜΕΤΡΙΚΑ ΜΟΝΤΕΛΑ
ΑΝΘΡΩΠΙΝΩΝ ΠΡΟΣΩΠΩΝ ΚΑΙ ΣΩΜΑΤΩΝ
ΚΑΙ ΕΦΑΡΜΟΓΕΣ ΣΤΗΝ
ΕΠΙΚΟΙΝΩΝΙΑ ΑΝΘΡΩΠΟΥ-ΜΗΧΑΝΗΣ

Επιβλέπων Καθηγητής:

Στέφανος Κόλλιας, Καθηγητής ΕΜΠ

Τριμελής Συμβουλευτική Επιτροπή:

Στέφανος Κόλλιας, Καθηγητής ΕΜΠ
Ανδρέας Σταφυλοπάτης, Καθηγητής ΕΜΠ
Παναγιώτης Τσανάκας, Καθηγητής ΕΜΠ

Επταμελής Εξεταστική Επιτροπή

Στέφανος Κόλλιας
Καθηγητής ΕΜΠ

Ανδρέας Σταφυλοπάτης
Καθηγητής ΕΜΠ

Παναγιώτης Τσανάκας
Καθηγητής ΕΜΠ

Εμμανουήλ Σκορδαλάκης
Καθηγητής ΕΜΠ

Βασίλειος Λούμος
Καθηγητής ΕΜΠ

Γεώργιος Παπακωνσταντίνου
Καθηγητής ΕΜΠ

Θεοχάρης Θεοχάρης
Επικ. Καθηγητής ΕΚΠΑ

Περίληψη

Προσαρμοζόμενα Παραμετρικά Μοντέλα Ανθρώπινων Προσώπων και Σωμάτων και Εφαρμογές στην Επικοινωνία Ανθρώπου-Μηχανής

του Κωνσταντίνου Καρπούζη

Η διδακτορική διατριβή έχει σαν κεντρικό αντικείμενο τη σύνθεση και επεξεργασία μοντέλων ανθρώπινων προσώπων και σωμάτων και την ένταξή τους σε διαλογικά περιβάλλοντα με έμφαση στην αποτελεσματική επικοινωνία ανθρώπου – μηχανής. Ο στόχος είναι η ανάπτυξη αλγορίθμων και εργαλείων που βοηθούν στη χρησιμοποίηση και την επεξεργασία τέτοιων αντικειμένων με αφαιρετικό, παραμετρικό τρόπο και την εκμετάλλευσή τους για σκοπούς ανάλυσης από σημασιολογική άποψη. Όλες οι προτεινόμενες ιδέες είναι πλήρως συμβατές με το πρότυπο MPEG-4, γεγονός που τις καθιστά χρησιμοποιήσιμες από μια ευρεία σειρά εφαρμογών.

Η διατριβή είναι οργανωμένη σε οκτώ κεφάλαια: τα δύο πρώτα αφορούν στην αναπαράσταση συνθετικών αντικειμένων και την απλοποίησή τους με αφαίρεση δομικών στοιχείων. Ακολουθεί η περιγραφή της κίνησης σε σκελετικά μοντέλα ανθρώπινου σώματος και η περιγραφή της με σημασιολογικές παραμέτρους. Αυτό κάνει εφικτή την εύκολη αλλαγή του αισθητικού αποτελέσματος με τρόπο που να σέβεται τους περιορισμούς που τίθενται στην κίνηση από τη φυσιολογία του ανθρώπινου σώματος και διευκολύνει τη σχεδίαση κινήσεων με συγκεκριμένο χαρακτήρα. Τα επόμενα κεφάλαια επικεντρώνονται στη μελέτη της φυσιολογίας του ανθρώπινου προσώπου και στην περιγραφή της μορφής και των εκφράσεών του με σημασιολογικές παραμέτρους. Αυτή βρίσκει εφαρμογή στην προσαρμογή γενικής μορφής μοντέλων σε μετρήσεις από πραγματικές εικόνες, στην απόδοσή τους με τεχνικές απεικόνισης υφής και, τέλος, στην απεικόνιση και σύνθεση συναισθημάτων και εκφράσεων.

Πίνακας Περιεχομένων

Ευχαριστίες	vi
Κεφάλαιο 1 Συνθετικά Αντικείμενα στο Πρότυπο MPEG-4.....	7
1.1 Εισαγωγή στο πρότυπο MPEG-4.....	7
1.2 Τα αντικείμενα στο MPEG-4.....	9
1.3 Αναπαράσταση κι επεξεργασία συνθετικών αντικειμένων	15
Κεφάλαιο 2 Απλοποίηση Πολυγωνικών Αντικειμένων.....	26
2.1 Εισαγωγή.....	26
2.2 Ανασκόπηση μεθόδων	28
2.3 Παραδείγματα αλγορίθμων	31
Κεφάλαιο 3 Αλγόριθμος Απλοποίησης με Αντικατάσταση Ακμών	35
3.1 Ορισμοί.....	35
3.2 Αντικατάσταση ακμών.....	37
3.3 Αποτελέσματα.....	42
Κεφάλαιο 4 Αναπαράσταση της Βάδισης.....	46
4.1 Περιγραφή των κωδικοποιημένων δεδομένων.....	46
4.2 Περιγραφή των παραμέτρων BAP.....	48
4.3 Φυσιολογικές μελέτες.....	51
Κεφάλαιο 5 Παραμετρική Αναπαράσταση και Σύνθεση της Βάδισης	55
5.1 Ανάλυση της κίνησης.....	55
5.2 Ανάλυση της βάδισης	61
5.3 Σύνθεση της κίνησης.....	63
5.4 Ο χαρακτήρας της βάδισης.....	66
Κεφάλαιο 6 Μορφολογία και Εκφράσεις Ανθρώπινων Κεφαλών.....	69
6.1 Εισαγωγή.....	69
6.2 Οι μύες του προσώπου.....	70
6.3 Μετρήσεις στο ανθρώπινο πρόσωπο	73
Κεφάλαιο 7 Παραμετρική Αναπαράσταση Ανθρώπινων Κεφαλών	77
7.1 Εισαγωγή.....	77
7.2 Καθορισμός της γεωμετρίας του κεφαλιού	79
7.3 Υπολογισμός των χαρακτηριστικών σημείων	82
7.4 Απεικόνιση υψής και προσαρμογή.....	86
Κεφάλαιο 8 Αναπαράσταση Συναισθημάτων σε Συνθετικά Μοντέλα.....	91
8.1 Εισαγωγή.....	91
8.2 Καθολικά συναισθήματα.....	93
8.3 Περιγραφή με FAPs.....	96

8.4 Σύνθεση ενδιάμεσων εκφράσεων.....	102
Βιβλιογραφία	105

Ευχαριστίες

Η εκπόνηση μιας διδακτορικής διατριβής δεν είναι σχεδόν ποτέ έργο ενός μόνο ατόμου. Θα μπορούσα μάλιστα να πω ότι ένα από τα μεγαλύτερα οφέλη που αποκόμισα στα χρόνια που πέρασα στο Εργαστήριο Ψηφιακής Επεξεργασίας Εικόνας, Βίντεο και Πολυμέσων ήταν η δυνατότητα να συνεργαστώ με καλούς φίλους και επιστήμονες που με βοήθησαν να ολοκληρώσω αυτό το έργο.

Βέβαια, αυτό δε θα ήταν δυνατό χωρίς την καθοριστική συμβολή του Δάσκαλου Στέφανου Κόλλια, με τον οποίο συνεργαστήκαμε άψογα για σχεδόν επτά χρόνια, από τις προπτυχιακές μου σπουδές μέχρι και σήμερα. Η καθοδήγηση και η συμπαράσταση που μου προσέφερε ήταν καταλυτική στην εξέλιξή μου σαν επιστήμονα και άνθρωπο. Στο ίδιο πλαίσιο θα πρέπει να τοποθετήσω και τη συνεργασία μου με τους καλούς φίλους και συναδέλφους Γιάννη Αβρίθη, Βασίλη Αλεξόπουλο και Νικόλα Τσαπατσούλη που με βοήθησαν τόσο στην εκπόνηση, όσο και στη συγγραφή της εργασίας. Σε προσωπικό επίπεδο, η συμβολή της οικογένειάς μου ήταν εξαιρετικά σημαντική, διαθέτοντάς μου περίσσεια από ηθικούς και υλικούς πόρους σε κάθε σημείο της πορείας μου.

Κλείνοντας θα ήθελα να ευχαριστήσω την Ελένη, χωρίς την έμπρακτη και αμέριστη στήριξη της οποίας τίποτα από όλα αυτά δε θα ήταν εφικτό. Το κριτικό, αλλά όχι επικριτικό μάτι της με ωφέλησε απεριόριστα σε θέματα προσανατολισμού και πνευματικής δύναμης για τη συνέχεια.

Ελπίζω να σταθώ αντάξιος όλων αυτών που με βοήθησαν λιγότερο ή περισσότερο κι ο καθένας με τον τρόπο του.

Συνθετικά Αντικείμενα στο Πρότυπο MPEG-4

Σε αντίθεση με τα προηγούμενα πρότυπα MPEG-1 και MPEG-2, τα οποία επικεντρώνονταν στην αποτελεσματική κωδικοποίηση, συμπίεση και μετάδοση κινούμενης εικόνας, το πρότυπο MPEG-4 στοχεύει και στο συνδυασμό πραγματικού και συνθετικού οπτικοακουστικού υλικού, έτσι ώστε να γίνει δυνατή η μετάδοση και απόδοσή του από συστήματα μέσω δικτύων μικρού εύρους ζώνης, αλλά και η διαλογική συμμετοχή του χρήστη – θεατή στη διαμόρφωση του περιεχομένου. Για αυτό το σκοπό, η αναπαράσταση και επεξεργασία συνθετικών αντικειμένων αποτελεί ένα κυρίαρχο τμήμα του προτύπου, με έμφαση στην αφαιρετική αναπαράσταση αντικειμένων όπως το ανθρώπινο πρόσωπο και σώμα, αλλά και χαρακτηριστικών κινήσεων ή ιδιοτήτων τους, όπως η υφή ή η έκφραση.

1.1 ΕΙΣΑΓΩΓΗ ΣΤΟ ΠΡΟΤΥΠΟ MPEG-4

Μια από τις βασικές διαφορές του προτύπου MPEG-4 σε σχέση με τα προηγούμενα αφορά τη θεώρηση της βασικής μονάδας της πληροφορίας που επεξεργάζεται. Στο πρότυπο MPEG-1, τόσο η κινούμενη εικόνα, όσο και το ηχητικό σήμα μπορούν να αναπαραχθούν με την ίδια λογική με την οποία αναπαράγεται μια αναλογική εικόνα VHS σε μια κοινή συσκευή VCR [15]: αυτό πρακτικά σημαίνει ότι δεν υπάρχει η δυνατότητα τυχαίας προσπέλασης στο υλικό, παρά μόνο γρήγορης μετακίνησης προς τα εμπρός ή προς τα πίσω. Στο πρότυπο MPEG-2, η συμπίεση και κωδικοποίηση της κινούμενης εικόνας κάνει δυνατή την αποθήκευση ολόκληρων κινηματογραφικών ταινιών σε ψηφιακούς δίσκους. Παρόλα αυτά, το θέμα της αφαιρετικής αναπαράστασης του υλικού και της εξαγωγής συγκεκριμένων χαρακτηριστικών ή οντοτήτων από αυτό δε θίγεται καθόλου.

Με την πρόοδο των δικτύων μεταφοράς και των οπτικοακουστικών μέσων, τα ζητήματα της ανεξαρτησίας από το φυσικό μέσο, της εύκολης και πιο ορθολογικής μετάδοσης, αλλά και της αυξημένης διαλογικότητας [20] φαίνονται να γίνονται περισσότερο σημαντικά [14]. Έτσι, το πρότυπο θα πρέπει να ανταποκρίνεται αποτελεσματικά σε τομείς όπως:

- η αποτελεσματική αναπαράσταση βίντεο σε ποιότητα που μπορεί να κυμαίνεται από πολύ υψηλή μέχρι σχετικά χαμηλή, δεδομένων μουσικής ή ομιλίας, συνθετικών τρισδιάστατων αντικειμένων γενικής ή εξειδικευμένης μορφής, συνθετικής ομιλίας και κειμένου

- η αντιμετώπιση σφαλμάτων κατά την κωδικοποίηση, ειδικά σε συνθήκες δικτύων χωρίς μεγάλο εύρος ζώνης
- η αναπαράσταση των αντικειμένων με ανεξάρτητο τρόπο, έτσι ώστε να είναι δυνατή η διακριτή επεξεργασία και επαναχρησιμοποίησή τους
- η σύνθεση αντικειμένων ήχου, πραγματικής και συνθετικής εικόνας σε μια συνολική ολοκληρωμένη σκηνή
- η υποστήριξη περιγραφής και διαλογικότητας στα αντικείμενα, καθώς και ενσωμάτωση συνδέσμων περιήγησης στη σκηνή
- η διαχείριση και προστασία των πνευματικών δικαιωμάτων στη σκηνή και τα περιεχόμενά της, έτσι ώστε να έχουν πρόσβαση μόνο οι διαπιστευμένοι χρήστες και να διαπιστώνεται με ασφάλεια η αντιγραφή
- ο καθορισμός μιας ενιαίας μορφοποίησης που να είναι διαφανής ως προς το μέσο, το περιεχόμενο και την υπολογιστική πλατφόρμα.

Η βάση για να καταστεί δυνατή η ικανοποίηση όλων αυτών των αναγκών ήταν η αναπαράσταση της οπτικοακουστικής πληροφορίας με αντικείμενα, τα οποία συνδέονται μεταξύ τους με σχέσεις στο χώρο και το χρόνο. Ένα σημαντικό πλεονέκτημα της αντικειμενοστρεφούς αναπαράστασης είναι ότι διαφορετικοί τύποι αντικειμένων μπορεί να αντιστοιχούν σε διαφορετικές βέλτιστες μεθόδους απόδοσης, χωρίς όμως να εμποδίζεται η αρμονική συνύπαρξή τους σε μια ενιαία σκηνή. Για παράδειγμα, ένας συνθετικός ανθρωπόμορφος χαρακτήρας, του οποίου οι παραμορφώσεις περιγράφονται αφαιρετικά από λεκτικές παραμέτρους μπορεί να συντεθεί με ένα αντικείμενο ψηφιακού βίντεο, το οποίο αναπαριστάται από μια ορθογώνια δομή από pixels και μια μάσκα δύο διαστάσεων για την απόκρυψη ορισμένων περιοχών του. Σε μια τέτοια οργάνωση, η διαλογική επεξεργασία και πλοήγηση μέσα στο περιεχόμενο γίνονται με σαφώς απλούστερο τρόπο και διευκολύνονται εργασίες όπως η επαναχρησιμοποίηση του υλικού, η μεταφορά σε εναλλακτικά σχήματα κωδικοποίησης και η προοδευτική μετάδοση μέσα από δίκτυα χαμηλού εύρους ζώνης, όπως το Internet ή τα ασύρματα δίκτυα.

Επειδή οι ιδέες που ενσωματώνονται στο MPEG-4 μπορούν να ωφεληθούν μια σειρά από διαφορετικές εφαρμογές και περιβάλλοντα, το πρότυπο είναι σχεδιασμένο με τη νοοτροπία της "εργαλειοθήκης", παρά της αυστηρής περιγραφής και προτυποποίησης κάθε ιδέας. Αυτό σημαίνει ότι, ανάλογα με την εφαρμογή και το περιβάλλον στο οποίο θα εφαρμοστεί, μπορούμε να καθορίσουμε μια σειρά από profiles, με κατάλληλες συμβάσεις και ρυθμίσεις για το υλικό, τη μετάδοσή του και την απόδοση – εμφάνιση στον τελικό χρήστη. Το περιεχόμενο αυτών των ρυθμίσεων αφορά περισσότερο τον ίδιο το σχεδιαστή της εφαρμογής, αφού συνήθως είναι αρκετό να εντοπίσει και να χρησιμοποιήσει ένα μικρό υποσύνολο όλων των ιδεών που περιγράφονται στο πρότυπο. Σε αντίθεση με το πρότυπο MPEG-2, το MPEG-4 μπορεί να ωφεληθεί μια σειρά από διαφοροποιημένες χρήσεις, τόσο στον τομέα του ίδιου του υλικού και της κατασκευής του, όσο και αναφορικά με τη μετάδοση και την αναπαράστασή του. Αν και συνήθως ένα πρότυπο θέτει μια σειρά από περιορισμούς με σκοπό να επιβάλλει την τελική συμβατότητα, το MPEG-4 φαίνεται να ακολουθεί μια φιλοσοφία ορισμού του ελάχιστου δυνατού για να επιβάλλει τη μέγιστη χρηστικότητα. Αυτό πρακτικά σημαίνει ότι η υλοποίηση πολλών από τις ιδέες που ενσωματώνονται στο πρότυπο αφήνεται στην ευχέρεια του σχεδιαστή της εφαρμογής ή του διαχειριστή του υλικού. Ένα ενδιαφέρον σημείο αυτής της φιλοσοφίας είναι ότι εναλλακτικοί αλγόριθμοι μπορεί να αναπτυχθούν παράλληλα με το πρότυπο και να ξεχωρίσουν μέσα από διαδικασίες υποστήριξης και ανταγωνισμού.

Στην πρώτη του έκδοση, το πρότυπο MPEG-4 περιλαμβάνει έξι συγκεκριμένα τμήματα:

- **Systems:** περιλαμβάνει την περιγραφή της σκηνής, την πολυπλεξία, το συγχρονισμό, τη διαχείριση και την προστασία της πνευματικής ιδιοκτησίας
- **Visual:** περιγράφει την κωδικοποιημένη αναπαράσταση της πραγματικής και συνθετικής οπτικής πληροφορίας
- **Audio:** περιγράφει την κωδικοποιημένη αναπαράσταση της πραγματικής και συνθετικής ακουστικής πληροφορίας
- **Conformance testing:** καθορίζει τις συνθήκες συμβατότητας για τις συσκευές και το κωδικοποιημένο περιεχόμενο
- **Reference software:** περιλαμβάνει λογισμικό και εργαλεία για τα περισσότερα από τα τμήματα του MPEG-4 και το οποίο μπορεί να χρησιμοποιηθεί για την ανάπτυξη συμβατών εργαλείων
- **Delivery Multimedia Integration Framework (DMIF):** περιγράφει πρωτόκολλα για τη διαχείριση της προοδευτικής μετάδοσης (streaming) της πληροφορίας πολλαπλών μέσων μέσα από τεχνολογίες γενικής μορφής.

Τα τμήματα 1 ως 3, καθώς και το τμήμα 6, αποτελούν τον πυρήνα (core) της τεχνολογίας του MPEG-4, ενώ τα τμήματα 4 και 5 είναι συμπληρωματικά. Ακόμα, τα τρία πρώτα τμήματα είναι ανεξάρτητα από την εμφάνιση της πληροφορίας, αφήνοντας στο τμήμα 6 να διαχειριστεί τις ιδιαιτερότητες του στρώματος αυτού. Γενικά, αν και καθένα από τα τμήματα του προτύπου μπορεί να χρησιμοποιηθεί ξεχωριστά από τα υπόλοιπα, ακόμα και σε συνδυασμό με διαφορετικές τεχνολογίες, το πνεύμα τους είναι τέτοιο που να αποδίδουν καλύτερα όταν υλοποιούνται συνδυαστικά.

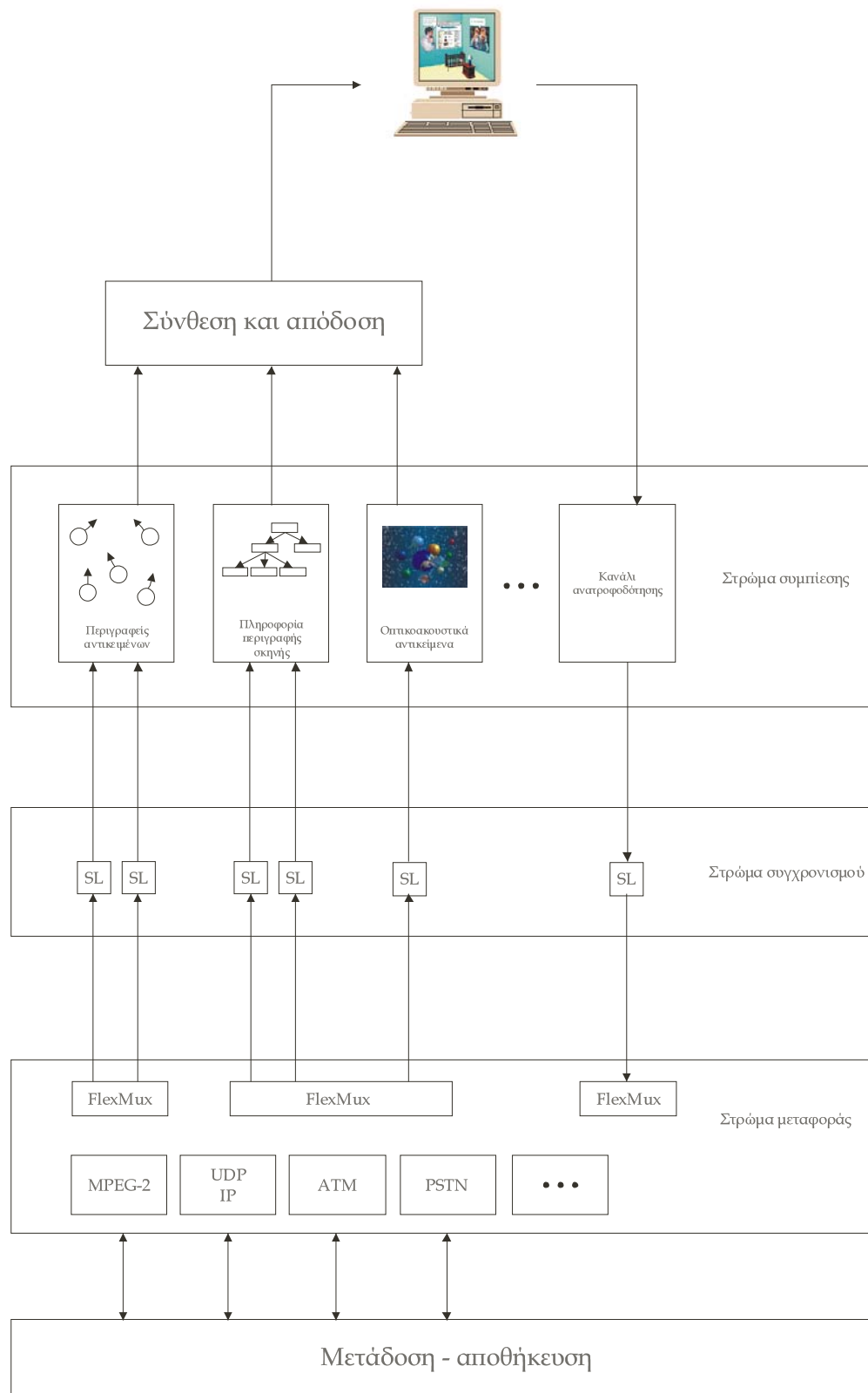
1.2 ΤΑ ΑΝΤΙΚΕΙΜΕΝΑ ΣΤΟ MPEG-4

Σύμφωνα με το πρότυπο MPEG-4, η οπτική και ακουστική πληροφορία σε μια σκηνή μπορεί να περιγραφεί από αντικείμενα, τα οποία μπορούν να συνυπάρχουν σε πολλαπλά στιγμιότυπα. Καθένα από αυτά μπορεί να συμμετάσχει σε ομάδες αντικειμένων, οι οποίες συντίθενται για να ορίσουν σύνολα υψηλότερου επιπέδου και τελικά να καθορίσουν το περιεχόμενο μιας σκηνής. Τα οπτικά αντικείμενα σε μια σκηνή καθορίζονται με σαφή μαθηματικό τρόπο και τοποθετούνται μέσα στο χώρο δύο ή τριών διαστάσεων – αντίστοιχα, τα ηχητικά αντικείμενα μπορούν και αυτά να τοποθετηθούν σε θέσεις με τη χωρική έννοια, μια που το πρότυπο προβλέπει και τη χωρική αναπαράσταση του ηχητικού σήματος. Μετά από την τοποθέτηση των αντικειμένων στο χώρο, οι αλλαγές στη θέση του χρήστη μέσα στη σκηνή αντανakλώνται σε αντίστοιχες μεταβολές των ιδιοτήτων τους, οι οποίες πάντως υπολογίζονται στη συσκευή αναπαραγωγής. Αυτό είναι πολύ χρήσιμο όταν είναι αναγκαία η άμεση ανανέωση του τελικού αποτελέσματος ή όταν δεν υπάρχει κάποιο κανάλι για την ανατροφοδότηση του πομπού με τα νέα δεδομένα, όπως συμβαίνει στην περίπτωση της τηλεοπτικής μετάδοσης.

Σε ότι αφορά την προετοιμασία του υλικού για μετάδοση, το πρότυπο τοποθετεί τα αντικείμενα που απαρτίζουν τη σκηνή στις αντίστοιχες βασικές ροές δεδομένων (elementary streams). Για τους περισσότερους τύπους αντικειμένων, όπως ένα αντικείμενο τύπου βίντεο, έχει νόημα μόνο μια τέτοια βασική ροή – για αντικείμενα κλιμακούμενης κωδικοποίησης πάντως είναι δυνατό να οριστεί μια βασική ροή για το ελάχιστο ποσό πληροφορίας, μαζί με πρόσθετες ροές λεπτομέρειας. Παράλληλα με αυτό, τα δεδομένα ανώτερου επιπέδου, τα οποία περιέχουν και την πληροφορία περιγραφής της σκηνής, περιέχονται στη δική τους ροή δεδομένων. Με αυτόν τον τρόπο διευκολύνεται η επαναχρησιμοποίηση των επιμέρους αντικειμένων, αλλά και η

επιλογή της μετάδοσης ή χρησιμοποίησης ορισμένων μόνο από τα διαθέσιμα αντικείμενα, αν αυτό επιβάλλεται από το περιορισμένο διαθέσιμο εύρος ζώνης.

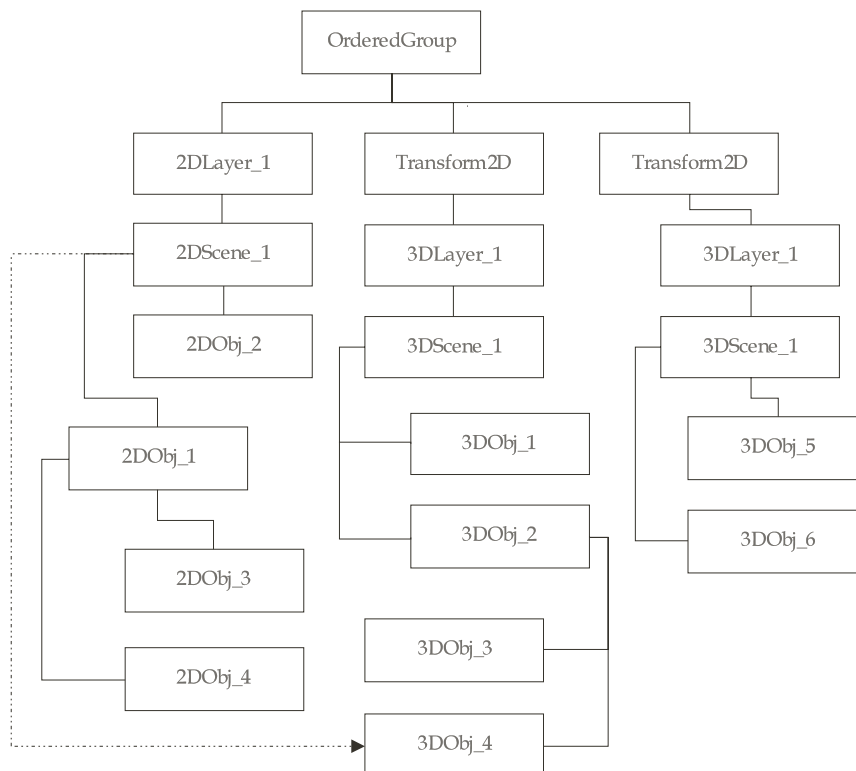
Η απεικόνιση των βασικών ροών δεδομένων σε συγκεκριμένα αντικείμενα γίνεται με βάση τους περιγραφείς αντικειμένων (object descriptors). Με τη σειρά τους, οι περιγραφείς αντικειμένων περιέχουν περιγραφείς βασικών ροών (elementary stream descriptors), δηλαδή οδηγίες προς το σύστημα σχετικά με τους απαραίτητους αποκωδικοποιητές για καθεμιά από τις ροές δεδομένων. Τελικά, η αναπαραγωγή των αντικειμένων και των δεδομένων ελέγχου γίνεται με βάση τις πληροφορίες του στρώματος συγχρονισμού: εδώ οι βασικές ροές διαιρούνται σε πακέτα, τα οποία προετοιμάζονται για τη μεταφορά στο στρώμα μετάδοσης με την πρόσθεση πληροφορίας συγχρονισμού. Συνολικά, αυτή η ροή δεδομένων και πληροφοριών ελέγχου φαίνεται στο Σχήμα 1-1.



Σχήμα 1-1: Αναπαράσταση της επικοινωνίας των στρωμάτων στα οποία περιέχονται τα αντικείμενα

Το πρότυπο MPEG-4 χρησιμοποιεί τη γλώσσα BIFS (Binary Format for Scenes) [18] για την αναπαράσταση και δυναμική ανανέωση του περιεχομένου. Οι σχετικές εντολές και λειτουργίες δεν περιορίζονται στην προσθήκη ή αφαίρεση των αντικειμένων μέσα σε μια σκηνή, αλλά

μπορούν να χρησιμοποιηθούν για τη μεταβολή οπτικών ή ακουστικών ιδιοτήτων για τα αντικείμενα χωρίς να μεταβάλλουν τη δομή ή την περιγραφή του ίδιου του αντικειμένου. Ακόμα, υποστηρίζεται η χρονική μεταβολή και το animation για τα αντικείμενα μιας σκηνής, με σχετικές διακριτές εντολές, αλλά και ο καθορισμός της "συμπεριφοράς", δηλαδή της αντίδρασης του αντικειμένου σε συγκεκριμένα γεγονότα στον αποκωδικοποιητή ή στη χρονική εξέλιξη της σκηνής. Με αυτόν τον τρόπο, είναι ευκολότερη η ανάπτυξη διαλογικών εφαρμογών ή ακόμα και η σύνθεση του αποτελέσματος μιας εφαρμογής σαν υφή ενός αντικειμένου μιας νέας σκηνής. Ένα παράδειγμα συνδυασμού διοδιάστατων και τρισδιάστατων αντικειμένων σε ένα γράφο φαίνεται στο Σχήμα 1-2.



Σχήμα 1-2: Παράδειγμα οργάνωσης μιας σκηνής με τη γλώσσα BIFS

Η γλώσσα BIFS προέρχεται σχηματικά από τη γλώσσα περιγραφής εικονικών κόσμων VRML [19], με τη σημαντική διαφορά ότι εδώ τα τελικά δεδομένα είναι δυαδικής μορφής και όχι μορφής κειμένου. Αυτό έχει σαν συνέπεια τη μείωση του τελικού αρχείου, σε ορισμένες περιπτώσεις κατά 10 ως 15 φορές. Εκτός από αυτό, το πρότυπο MPEG-4 χρησιμοποιεί τα δεδομένα BIFS για να μεταδώσει προοδευτικά την περιγραφή της σκηνής και να ξεκινήσει την απόδοσή τους στον αποκωδικοποιητή, χωρίς να είναι αναγκαία η ολοκλήρωση της μετάδοσής τους. Έτσι, λύνονται ή τουλάχιστον αντιμετωπίζονται αποτελεσματικά τα προβλήματα της προεπισκόπησης του υλικού, αλλά και της αφαιρετικής αναπαράστασης ιεραρχιών υλικού.

Η έννοια των αντικειμένων είναι δυνατό να μεταβάλει τις αναπαραστάσεις ακόμα και κλασικών μορφών ψηφιακών δεδομένων, όπως το βίντεο. Εκτός από την κλασική αναπαράσταση σε πλαίσια ορθογωνικής μορφής είναι δυνατή η κωδικοποίηση αντικειμένων τυχαίας μορφής, τα οποία μπορούν να κωδικοποιηθούν ξεχωριστά από το φόντο της σκηνής και να τοποθετηθούν μπροστά από αυτό κατά την τελική απόδοση. Τέτοια περιγράμματα μπορούν να κωδι-

κοποιηθούν με τη μορφή δυαδικής μάσκας από pixels, τα οποία καθορίζουν την περιοχή που είτε θα εμφανιστεί, είτε θα αποκρυφτεί από το τελικό αποτέλεσμα στον αποκωδικοποιητή. Αν και το αποτέλεσμα μιας τέτοιας ιδέας δεν είναι ιδιαίτερα πρακτικό, μπορεί να είναι χρήσιμο σε περιβάλλοντα χαμηλού εύρους ζώνης, παρά την ιδιαίτερα έντονη εμφάνιση φαινομένων αναδίπλωσης. Καλύτερα αποτελέσματα σε τέτοια περιγράμματα μπορούν να επιτευχθούν με μάσκες που δεν περιγράφονται με δυαδική μορφή, αλλά με περισσότερες διακριτές τιμές διαφάνειας. Αυτό έχει συνήθως σαν αποτέλεσμα την πιο ομαλή σύνθεση του κωδικοποιημένου αντικειμένου με τα υπόλοιπα ή με το φόντο της σκηνής. Ακόμα, αυτό το περίγραμμα μπορεί να καθορίσει τις "σημαντικές" περιοχές ενός αντικειμένου, οι οποίες κωδικοποιούνται με υψηλότερο εύρος ζώνης, αφήνοντας τις υπόλοιπες περιοχές οι οποίες ενσωματώνουν λιγότερη λεπτομέρεια. Πάντως, η εξαγωγή αυτών των περιγραμμάτων δεν προτυποποιείται από το MPEG-4, αφού εδώ δεν περιγράφεται κάποια συγκεκριμένη ιδέα για την κατάτμηση του ψηφιακού βίντεο. Αντίθετα, όπως άλλωστε συμβαίνει και με τα προηγούμενα πρότυπα, το MPEG-4 καθορίζει μόνο τη διαδικασία της αποκωδικοποίησης, αφήνοντας το σχεδιαστή της εφαρμογής να αναπτύξει ή να επιλέξει τους αλγορίθμους υπολογισμού των απαιτούμενων παραμέτρων.

Σε ότι αφορά τις εφαρμογές χαμηλού εύρους ζώνης, το πρότυπο MPEG-4 μπορεί να χρησιμοποιηθεί αποτελεσματικά ακόμα και για τη μεταφορά βίντεο σε ασύρματες συσκευές, με ρυθμούς μετάδοσης συμβατούς με το πρότυπο GSM, δηλαδή 10 kb/sec. Ένα σημαντικό πρόβλημα σε τέτοιες περιπτώσεις αφορά την πιθανότητα σφαλμάτων που έχουν σχέση με τη μετάδοση. Μια ιδέα που μπορεί να βοηθήσει σε αυτήν την κατεύθυνση είναι η εισαγωγή σημείων συγχρονισμού, τα οποία βοηθούν τον αποκωδικοποιητή να αναπαράγει σωστά τα διάφορα δεδομένα, ακόμα κι αν απολέσει στιγμιαία τη χρονική τους σχέση. Η δεύτερη τεχνική που μπορεί να χρησιμοποιηθεί είναι η κωδικοποίηση αναστρέψιμου κώδικα μεταβλητού μήκους: το αποτέλεσμα αυτής της διεργασίας μπορεί να αποκωδικοποιηθεί αποτελεσματικά, ακόμα κι αν ο αποκωδικοποιητής λάβει τα δεδομένα με αντεστραμμένη σειρά, μέχρι να λάβει το επόμενο σήμα συγχρονισμού.

Γενικά, η κωδικοποίηση σε χαμηλό εύρος ζώνης γίνεται δυνατή με την υποστήριξη κλιμακούμενων αντικειμένων από το πρότυπο MPEG-4. Σε πολλές περιπτώσεις, για παράδειγμα σε μια σελίδα στο web, ο χρήστης καλείται να επιλέξει μόνος του την κωδικοποίηση του υλικού που θέλει να λάβει και κατά συνέπεια το μέγεθος ενός φυσικού αρχείου. Αυτό πρακτικά σημαίνει ότι το ίδιο υλικό πρέπει να κωδικοποιηθεί σε πολλαπλά στιγμιότυπα από τον παροχέα – με την κλιμακωτή κωδικοποίηση, το υλικό μετατρέπεται σε συμβατό αντικείμενο μόνο μια φορά και το σύστημα ανιχνεύει αυτόματα το "σωστό" εύρος ζώνης για τη μετάδοση προς κάθε αποκωδικοποιητή. Πρακτικά, στην κλιμακωτή κωδικοποίηση καθορίζεται ένα βασικό στρώμα πληροφορίας, το οποίο περιέχει μόνο το ελάχιστο επίπεδο ποιότητας – σε αυτό υπερτίθενται διαδοχικά στρώματα βελτίωσης, αν αυτό είναι δυνατό από το διαθέσιμο εύρος ζώνης, δίνοντας έτσι ένα συνολικά καλύτερο αποτέλεσμα. Έτσι, αν μια σκηνή αποτελείται από πολλαπλά αντικείμενα, είναι δυνατό να μεταδοθούν πρώτα ή μόνο κάποια από αυτά ή ακόμα και να χρησιμοποιηθούν διαφορετικά σχήματα κωδικοποίησης ή διόρθωσης σφαλμάτων για καθένα από αυτά. Ακόμα, στοιχεία που δεν αλλάζουν συχνά μπορούν να αναπαρασταθούν από στατικά sprites. Για παράδειγμα, το φόντο μιας στατικής σκηνής όπως ενός εικονικού στούντιο μπορεί να μεταδοθεί μόνο μια φορά, στην αρχή της επικοινωνίας – αν η θέση της κάμερας μεταβληθεί, αντί για την

μετάδοση από την αρχή της νέας θέσης και προσανατολισμού, το σύστημα αρκεί να μεταδώσει μόνο τις νέες συντεταγμένες τεσσάρων σημείων, τα οποία αντιστοιχούν στη θέση του φόντου σαν ένα αντικείμενο τετράπλευρου σχήματος στο βάθος της σκηνής.

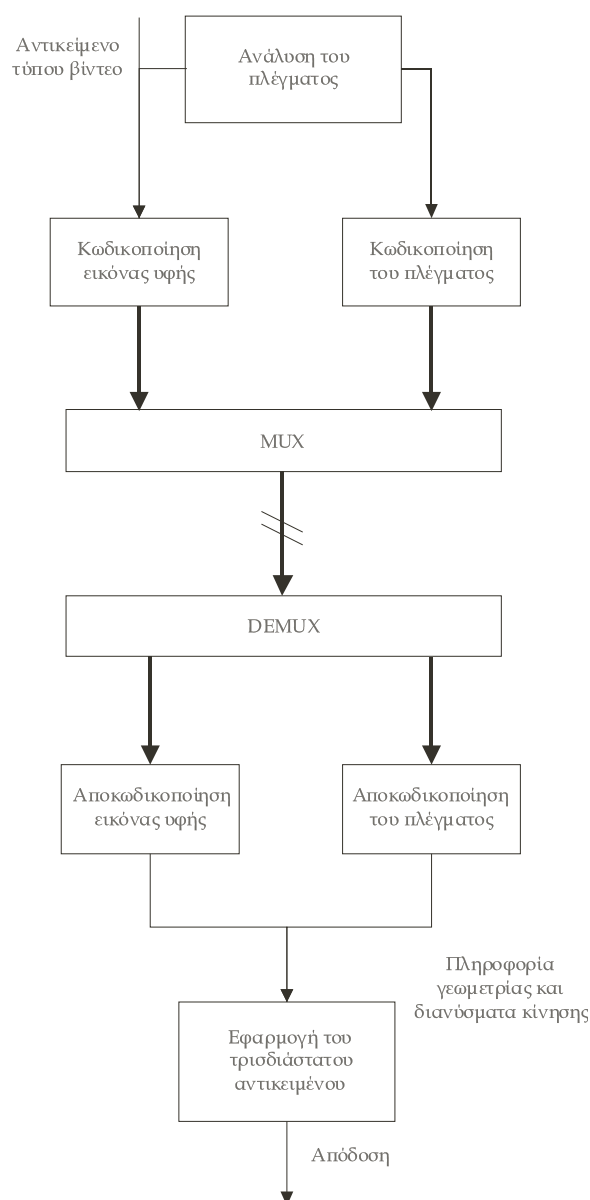
Σε αντίθεση με τα προηγούμενα πρότυπα, ο ήχος στο MPEG-4 έχει την ίδια μεταχείριση με την κινούμενη εικόνα και δεν αντιμετωπίζεται απλά σαν ένα πρόσθετο σήμα που διαπλέκεται με αυτήν. Το εύρος των εργαλείων που προβλέπει το πρότυπο μπορεί να διαχειριστεί την κωδικοποίηση σε εύρος ζώνης από 6 kb/sec μέχρι πάνω από 128 kb/sec, ενώ η κωδικοποίηση τόσο του μονοφωνικού, όσο και του στερεοφωνικού σήματος χρησιμοποιεί πολύ λιγότερα από τα 1411 kb/sec που απαιτούνται από το πρότυπο Red Book για την εγγραφή σε Audio CD. Αν απαιτείται ακόμα καλύτερη ποιότητα, το πρότυπο MPEG-4 περιλαμβάνει τον αλγόριθμο AAC (Advanced Audio Coding) από το πρότυπο MPEG-2, το οποίο έχει ασύγκριτα καλύτερα αποτελέσματα από το MPEG-1 Layer III, δηλαδή τη μορφοποίηση mp3 που επικρατεί στα σύγχρονα συστήματα επικοινωνίας. Σε ότι αφορά την ομιλία, το πρότυπο χρησιμοποιεί ένα σχήμα κωδικοποίησης σε εύρος ζώνης 2 - 4 kb/sec, το οποίο μπορεί να γίνει ακόμα μικρότερο αν οριστεί μεταβλητός ρυθμός μετάδοσης. Εναλλακτικά, μπορεί να χρησιμοποιηθεί το σχήμα CELP (Code Excited Linear Prediction), το οποίο χρησιμοποιεί 4 - 24 kb/sec, αλλά δίνει καλύτερα αποτελέσματα σε υψηλότερους ρυθμούς δειγματοληψίας. Ακόμα, το πρότυπο υποστηρίζει τη δημιουργία ομιλίας από γραπτό κείμενο (TTS - Text to Speech). Το κείμενο αυτό μπορεί να χρησιμοποιηθεί τόσο για λόγους υποτιτλισμού ή περιγραφής, όσο και για να καθορίσει το animation ενός συνθετικού μοντέλου. Η απόδοση του συγκεκριμένου ηχητικού σήματος δεν γίνεται απαραίτητα χωρίς κάποιο χρωματισμό της συνθετικής φωνής, αλλά μπορεί να εκμεταλλευτεί ειδικά σήματα εκφοράς του λόγου, ταχύτητας και συγχρονισμού, ή ακόμα και σημασιολογικών παραμέτρων, όπως το φύλο, η ηλικία ή η τοπική προφορά.

Η αναπαράσταση του ήχου σαν αντικείμενο στο πρότυπο MPEG-4 είναι καθορισμένη με μεγαλύτερη σαφήνεια από ότι συμβαίνει στη σταθερή και την κινούμενη εικόνα και έχει ενιαία μορφή είτε πρόκειται για μονοφωνικά σήματα μικρού εύρους ζώνης, είτε για στερεοφωνικά, λεπτομερή σήματα. Επειδή η περιγραφή της σκηνής γίνεται με βάση έναν ιεραρχικό γράφο, τα ηχητικά αντικείμενα μπορούν να απεικονίσουν διαφορετικά συστήματα εισόδου σε κανάλια εξόδου (για παράδειγμα, συστήματα ηχείων) για διαφορετικά περιβάλλοντα αναπαραγωγής της σκηνής. Ακόμα, η ύπαρξη διακριτών αντικειμένων για τον ήχο, τη μουσική ή την ομιλία δίνει τη δυνατότητα εφαρμογής διαφορετικών εφφέ σε καθένα από αυτά. Αυτό σημαίνει ότι, για παράδειγμα, η ομιλία σε μια κινηματογραφική ταινία μπορεί να ενισχυθεί ανεξάρτητα από το ηχητικό χαλί ή ακόμα και να αντικατασταθεί από ομιλία σε διαφορετική γλώσσα. Μια ενδιαφέρουσα δυνατότητα είναι η συσχέτιση των ηχητικών αντικειμένων με χωρικές θέσεις μέσα στη σκηνή. Αυτό πρακτικά σημαίνει ότι ο αποκωδικοποιητής μπορεί να αποδίδει με διαφορετική ένταση ή διάχυση το ηχητικό αντικείμενο ανάλογα με το πόσο κοντά βρίσκεται η θέση του παρατηρητή σε αυτό, να διαχωρίζει το ηχητικό σήμα και να το κατευθύνει σε διαφορετικά ηχεία, τοποθετημένα κατάλληλα ώστε να δίνουν την αίσθηση της κατεύθυνσης στην προέλευση του ήχου (σύστημα surround) ή ακόμα και να εφαρμόσει εφφέ σαν την ηχώ ή την απόκρυψη του ήχου, ανάλογα με το τρισδιάστατο μοντέλο που μπορεί να περιέχει χωρικά το ηχητικό αντικείμενο. Για παράδειγμα, ένα ηχητικό αντικείμενο που απεικονίζει τον ήχο ενός πολυσύχναστου δρόμου μπορεί να τοποθετηθεί χωρικά "έξω" από ένα τρισδιάστατο αρχιτεκτονικό μοντέλο ενός

σπιτιού: αν τα παράθυρα του μοντέλου είναι ανοιχτά, ο αποκωδικοποιητής μπορεί να αποδίδει το ηχητικό αντικείμενο με μεγαλύτερη ένταση και σαφήνεια, από ότι αν τα παράθυρα ήταν κλειστά. Επειδή αυτή η διεργασία ολοκληρώνεται στον αποκωδικοποιητή, δεν αφιερώνεται σε αυτήν πρόσθετο εύρος ζώνης, ενώ το ηχητικό αντικείμενο μεταδίδεται μόνο μια φορά.

1.3 ΑΝΑΠΑΡΑΣΤΑΣΗ ΚΙ ΕΠΕΞΕΡΓΑΣΙΑ ΣΥΝΘΕΤΙΚΩΝ ΑΝΤΙΚΕΙΜΕΝΩΝ

Το πρότυπο MPEG-4 υποστηρίζει την κωδικοποίηση και το animation για τριγωνικά πολυγωνικά πλέγματα ομοιόμορφης ή ανομοιόμορφης τριγωνοποίησης Delaunay. Το Σχήμα 1-3 απεικονίζει ένα διάγραμμα του σχήματος κωδικοποίησης – αποκωδικοποίησης για τα συνθετικά αντικείμενα, με υποστήριξη για την πιθανή χρήση ενός αντικειμένου τύπου εικόνας (στατικής ή κινούμενης) σαν πληροφορία υψής. Στη συνέχεια θα αναφερθούμε αναλυτικά στα διάφορα τμήματα του σχήματος, όπως την ανάλυση και την κωδικοποίηση του τρισδιάστατου πλέγματος.



Σχήμα 1-3: Απλοποιημένο σχήμα της αρχιτεκτονικής του κωδικοποιητή και αποκωδικοποιητή για τα συνθετικά αντικείμενα

Αναπαράσταση και animation σε πλέγματα δύο διαστάσεων

Ένα τριγωνικό πλέγμα ή επίπεδο αντικειμένου πλέγματος (mesh object plane - MOP) ορίζεται σαν ένας επίπεδος γράφος που διαχωρίζει το επίπεδο του αντικειμένου ή το ελάχιστο παραλληλεπίπεδο του σε τριγωνικά τμήματα. Οι κορυφές καθενός τμήματος ονομάζονται κόμβοι (nodes) – συνολικά, ένα πλεγματοειδές αντικείμενο, το οποίο αποτελείται από μια ακολουθία από επίπεδα αντικειμένων πλέγματος, αναπαριστάται με συμπαγή μορφή από γεωμετρική πληροφορία στα MOPs τύπου I (intra) και τα διανύσματα κίνησης στα MOPs τύπου inter. Η πληροφορία της γεωμετρίας αναφέρεται στη θέση των κόμβων του πλέγματος για καθένα από τα MOPs τύπου intra, ενώ η αίσθηση του animation επιτυγχάνεται με τη διάδοση του πλέγματος σε αυτά τα MOPs χρησιμοποιώντας ένα διάνυσμα κίνησης για κάθε κορυφή και MOP. Τόσο η πληροφορία της γεωμετρίας, όσο και η πληροφορία του animation κωδικοποιούνται με μηχανισμούς πρόβλεψης, δίνοντας σαν αποτέλεσμα μια συμπαγή δυαδική αναπαράσταση.

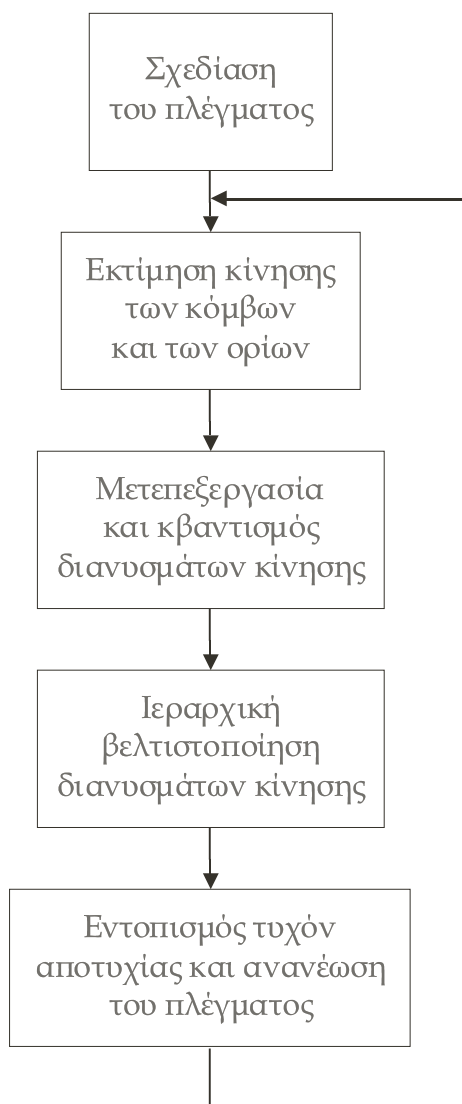
Η μοντελοποίηση της κίνησης στα πολυγωνικά πλέγματα διαφέρει από τη μοντελοποίηση που βασίζεται σε blocks, τεχνική που εφαρμόζεται κατά την κωδικοποίηση πραγματικού βίντεο, αφού εδώ τα τριγωνικά τμήματα δεν επικαλύπτονται ούτε στο πλαίσιο αναφοράς, ούτε στο τρέχον κάθε φορά πλαίσιο. Αντίθετα, τα τριγωνικά τμήματα στο τρέχον πλαίσιο απεικονίζονται σε τμήματα στο πλαίσιο αναφοράς και η πληροφορία υψής για καθένα από αυτά παραμορφώνεται με βάση τα διανύσματα κίνησης για τους κόμβους του. Η απεικόνιση υψής που ορίζεται με αυτόν τον τρόπο αποτελεί ένα κύριο τμήμα του συστήματος υποστήριξης του animation, ενώ η απεικόνιση που αφορά τη γεωμετρία, για τη χρονική στιγμή t ορίζεται από τις σχέσεις:

$$\begin{aligned}x &= a_{1x} + a_{2y} + a_3 \\y &= a_{4x} + a_{5y} + a_6\end{aligned}\tag{1-1}$$

όπου οι παράμετροι κίνησης a_i αντιστοιχούν σε έξι βαθμούς ελευθερίας για την παραμόρφωση καθεμιάς από τις κορυφές ενός τριγώνου πάνω στο επίπεδο. Ακόμα, αν καθορίσουμε κατάλληλους περιορισμούς για την εκτίμηση των παραμέτρων, δηλαδή των διανυσμάτων κίνησης για τους κόμβους) μπορούμε να επιτύχουμε τη συνέχεια της απεικόνισης στα όρια των γειτονικών τριγωνικών πλευρών, κατά την εξέλιξη αυτού του μετασχηματισμού. Από την πλευρά του σήματος, μπορούμε να θεωρήσουμε ότι η μοντελοποίηση των πλεγμάτων αντιστοιχεί ουσιαστικά στη μη σταθερή δειγματοληψία του πεδίου της κίνησης στις θέσεις που αντιστοιχούν στους κόμβους του πλέγματος – έτσι, μπορούμε να εξάγουμε μια συνεχή και τμηματικά αφφινική μορφή του πεδίου κίνησης. Η πρακτική αυτή έχει το πλεονέκτημα της αποτελεσματικής αναπαράστασης πεδίων κίνησης που αντιπροσωπεύουν ήπιες κινήσεις και τα οποία είναι χωρικά συνεχή. Δεν είναι όμως εύκολο να αναπαραστήσουν ασυνέχειες στο πεδίο της κίνησης, όπως μπορεί να συμβεί στην κίνηση ενός αρθρωτού μοντέλου, ή αποκρύψεις της επιφάνειας από συγκεκριμένα τμήματά της.

Η σχεδίαση των δεδομένων του πλέγματος που έχει σχέση με ένα αντικείμενο τύπου βίντεο αφορά τον κωδικοποιητή κι επομένως δεν καθορίζεται με σαφήνεια από το πρότυπο. Εδώ θα περιγράψουμε ένα παράδειγμα τέτοιας διαδικασίας κατά την οποία σχεδιάζονται τα πλέγματα και αποθηκεύονται σε MOPs τύπου intra, ενώ παρακολουθούνται για τον υπολογισμό της πληροφορίας στα MOPs τύπου inter. Συνολικά η διαδικασία περιγράφεται από το Σχήμα 1-4, όπου ο κλάδος της ανατροφοδότησης αυξάνει το μετρητή του τρέχοντος καρέ. Αν η ένδειξη της

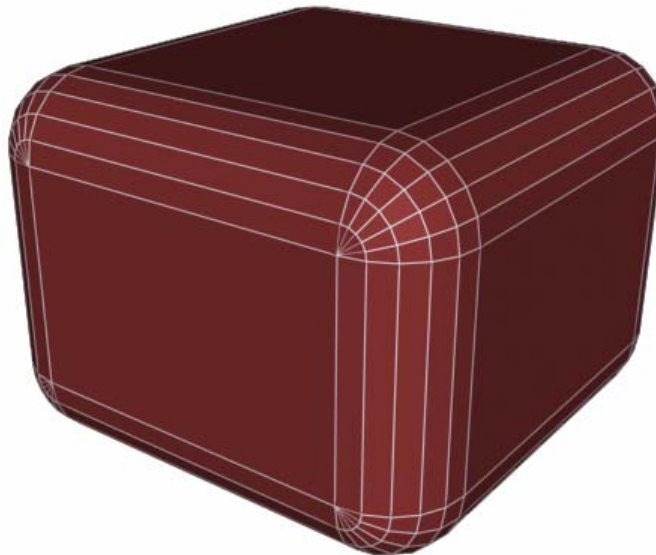
αποτυχίας ξεπεράσει ένα ορισμένο κατώφλι, τότε η διαδικασία σχεδίασης αρχικοποιείται και πάλι και εισάγεται ένα νέο MOP τύπου intra.



Σχήμα 1-4: Σχεδίαση και παρακολούθηση τριγωνικών πλεγμάτων

Η σχεδίαση των πλεγμάτων για τα MOPs τύπου intra μπορεί να ολοκληρωθεί τόσο με ομοιόμορφη κατανομή, όσο και με κατανομή βασισμένη στο περιεχόμενο. Στην πρώτη περίπτωση, το πλέγμα σχεδιάζεται πάνω σε μια τετραπλευρική περιοχή, η οποία συνήθως είναι το ελάχιστο τετράπλευρο του επιπέδου αντικειμένου βίντεο. Για το σκοπό αυτό χρειάζονται μια σειρά από παράμετροι όπως ο αριθμός των κόμβων κατά την οριζόντια και την κατακόρυφη διάσταση, οι διαστάσεις κατά τον οριζόντιο και τον κατακόρυφο άξονα για κάθε κελί ανάμεσα στους κόμβους και η μέθοδος διάσπασης κάθε κελιού για τη διαμόρφωση δύο τριγωνικών πλευρών. Αντίθετα, ένα πλέγμα σχεδιασμένο με βάση το περιεχόμενο μπορεί να σχεδιαστεί έτσι ώστε να καλύπτει ακριβώς το σχετικό VOP, ξεκινώντας από μια προσέγγιση του περιγράμματός του από N_b συνοριακούς κόμβους. Ακολούθως, επιλέγονται N_i εσωτερικά κομβικά σημεία, τα οποία τριγωνοποιούνται κατά Delaunay για να ορίσουν την τοπολογία του πλέγματος. Μια σειρά από ιδέες μπορούν να χρησιμοποιηθούν για την αποτελεσματική προσέγγιση του περιγράμματος

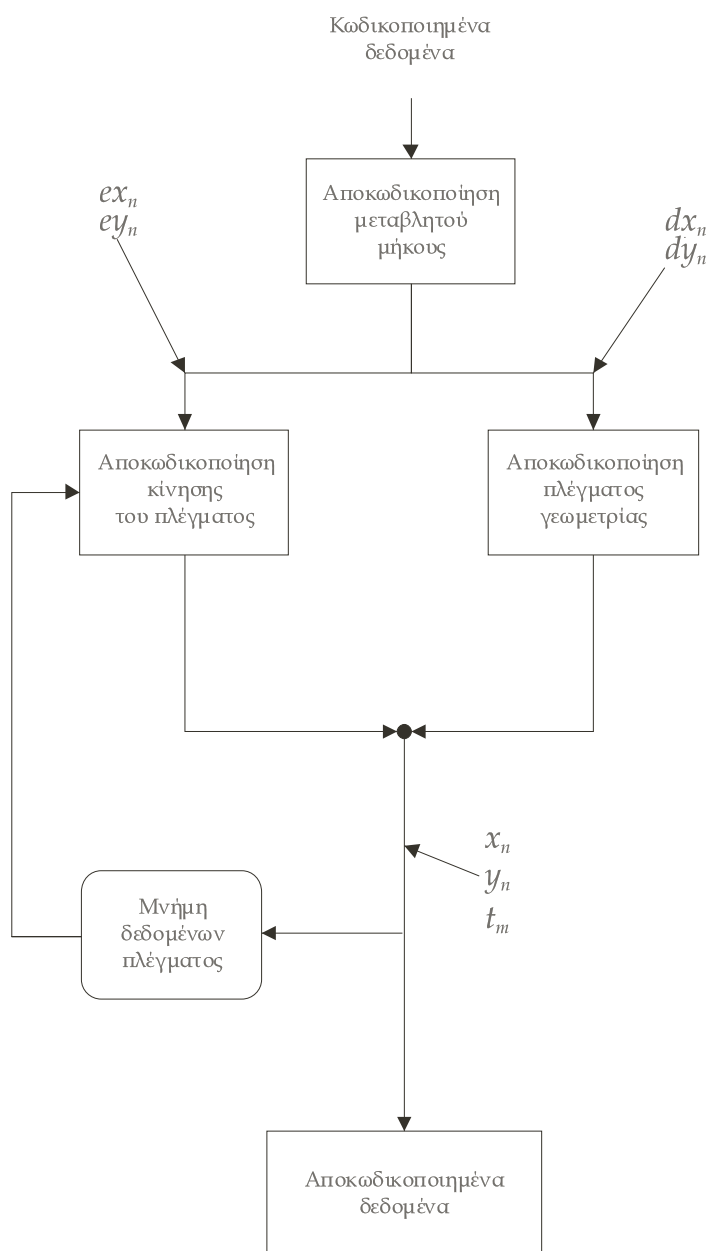
και την επιλογή των εσωτερικών κομβικών σημείων, όπως για παράδειγμα η επιλογή σημείων μεγάλης καμπυλότητας ή αλλαγής της καμπής (Σχήμα 1-5).



Σχήμα 1-5: Παράδειγμα πλέγματος που είναι σχεδιασμένο με βάση το περιεχόμενο

Γενικά, η πληροφορία της κίνησης είναι δυνατό να αντιστοιχεί στην κίνηση ενός πραγματικού αντικειμένου τύπου βίντεο ή να είναι συνθετική, για σκοπούς animation ή μεταβολής μιας εικόνας απεικόνισης υφής. Στην πρώτη περίπτωση, η κίνηση σε ένα αντικείμενο σχετικού τύπου μπορεί να προβλεφθεί με την παρακολούθηση επόμενων πλεγμάτων: εδώ αναζητούμε μέσα σε ολόκληρο το αντικείμενο τις βέλτιστες θέσεις των βασικών κόμβων του πλέγματος που περιγράφεται στο προηγούμενο MOP, παρακολουθώντας με αυτόν τον τρόπο τα χαρακτηριστικά της εικόνας μέχρι το επόμενο MOP τύπου intra.

Η πιο απλή μέθοδος που μπορεί να χρησιμοποιηθεί για το σκοπό αυτό περιλαμβάνει τη δημιουργία blocks γύρω από τη θέση καθενός από τους κόμβους του πλέγματος και κατόπιν την εφαρμογή τεχνικών ταιριάσματος blocks για τον υπολογισμό των διανυσμάτων κίνησης για καθένα από αυτά [16]. Εναλλακτικά, μπορούν να χρησιμοποιηθούν τεχνικές εξαγωνικού ταιριάσματος ή ταιριάσματος κλειστής μορφής για τον υπολογισμό του βέλτιστου διανύσματος κίνησης για καθέναν από τους κόμβους παραμορφώνοντας τα γειτονικά τμήματα με τον περιορισμό της διατήρησης της συνεκτικότητας [17].

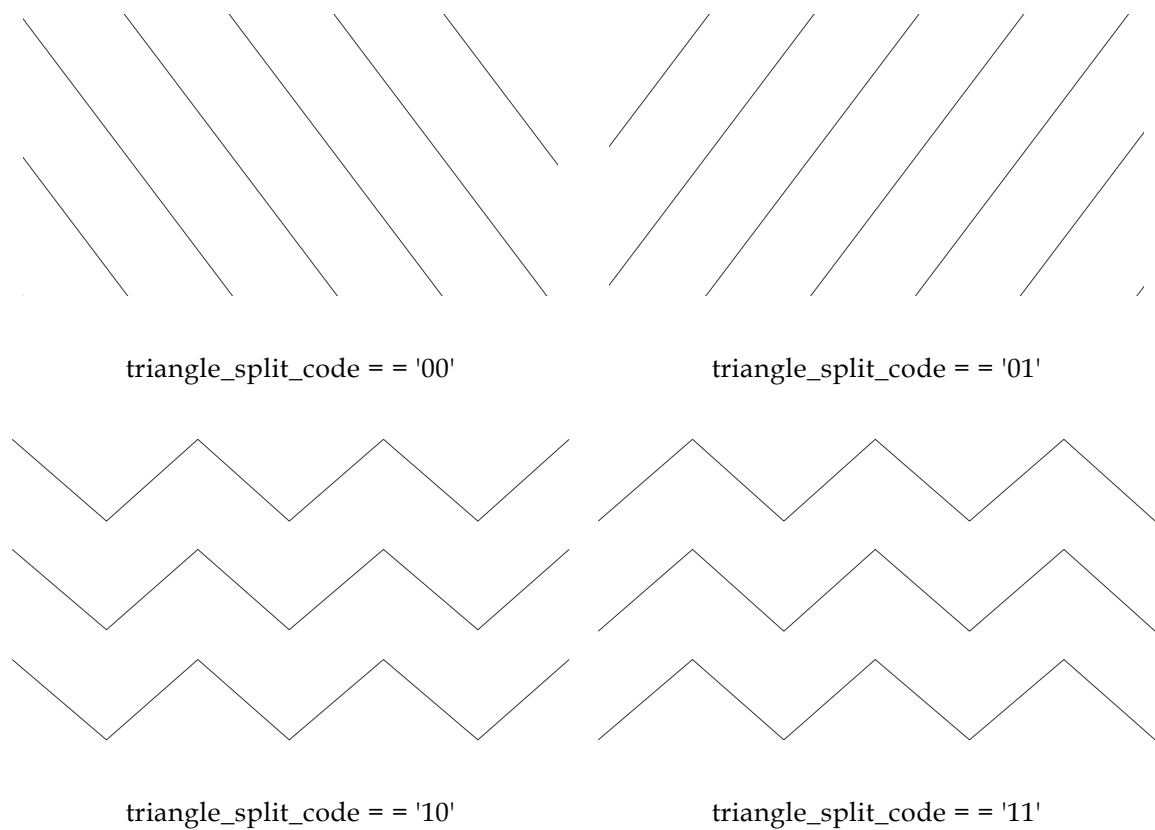
Κωδικοποίηση - αποκωδικοποίηση δισδιάστατων πλεγμάτων

Σχήμα 1-6: Απλοποιημένο διάγραμμα αποκωδικοποιητή πλεγματοειδών αντικειμένων

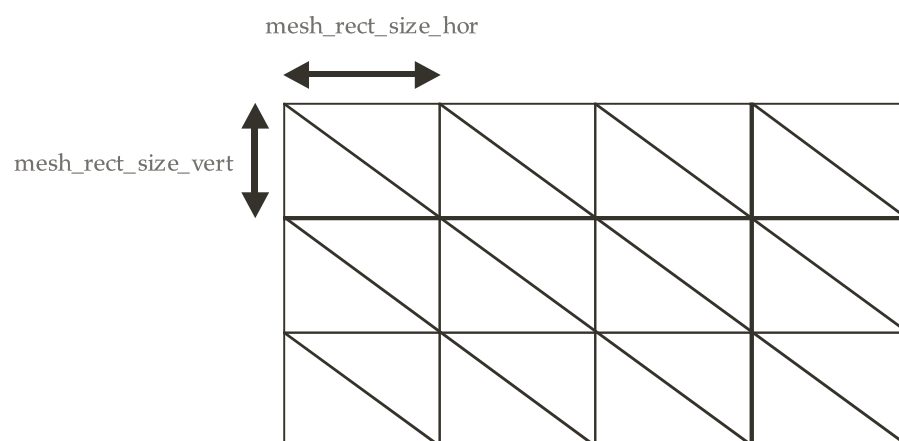
Τα δεδομένα του πλέγματος περιλαμβάνουν τη λίστα των θέσεων (x_n, y_n) των κόμβων και μια λίστα συνεκτικότητας από τριγωνικές πλευρές t_m . Καθεμία από τις πλευρές αυτές ορίζεται από μια τριάδα από δείκτες στη λίστα με τους κόμβους, οι οποίοι αντιστοιχούν και στις κορυφές του τριγώνου. Η σύνταξη της συμπιεσμένης δυαδικής αναπαράστασης των MOPs τύπου intra και inter, καθώς και η σημασιολογική διαδικασία της αποκωδικοποίησης είναι καθορισμένες στα πλαίσια του προτύπου MPEG-4. Κάθε MOP χαρακτηρίζεται από μια ένδειξη που διαχωρίζει εκείνα που περιέχουν πληροφορία γεωμετρίας (MOP τύπου intra), από εκείνα που περιέχουν διανύσματα κίνησης (MOP τύπου inter). Το Σχήμα 1-6 εμφανίζει μια αφαιρετική άποψη της συνολικής διαδικασίας αποκωδικοποίησης ενός αντικειμένου που περιγράφεται από πλέγμα. Πιο συγκεκριμένα, η διαδικασία αποκωδικοποίησης της γεωμετρίας υπολογίζει τις θέσεις των

κορυφών του πλέγματος και με βάση αυτές ανακατασκευάζει το αντικείμενο. Αντίστοιχα, η διαδικασία αποκωδικοποίησης της πληροφορίας κίνησης υπολογίζει τα διανύσματα κίνησης που σχετίζονται με τις κορυφές και τα εφαρμόζει πάνω τους, ανακατασκευάζοντας έτσι το τρέχον αντικείμενο. Το αποτέλεσμα αυτό αποθηκεύεται στη μνήμη των δεδομένων πλέγματος, έτσι ώστε να είναι δυνατό να χρησιμοποιηθεί στην αποκωδικοποίηση της κίνησης στο επόμενο MOP.

Σε ότι αφορά το τμήμα κωδικοποίησης / αποκωδικοποίησης, το κάθε MOP τύπου intra περιέχει σήμανση για το αν το αντικείμενο που περιγράφει έχει σχεδιαστεί με ομοιόμορφη τριγωνοποίηση ή με τριγωνοποίηση Delaunay. Στην πρώτη περίπτωση, το αντικείμενο μπορεί να θεωρηθεί σαν μια οργάνωση από κελιά ορθογώνιου σχήματος, με καθένα από αυτά να υποδιαιρείται σε δύο τρίγωνα από τη διαγώνιό του. Για τον πλήρη καθορισμό της τοπολογίας και των θέσεων των κόμβων χρειάζονται πέντε παράμετροι, με τον πάνω αριστερά κόμβο του πλέγματος να συμπίπτει με την αρχή του τοπικού συστήματος συντεταγμένων. Οι πρώτες δύο παράμετροι, που ονομάζονται `nr_mesh_nodes_hor` και `nr_mesh_nodes_vert`, καθορίζουν τον αριθμό των κόμβων κατά την οριζόντια και κατά την κατακόρυφη διεύθυνση, ενώ οι επόμενες δύο παράμετροι `mesh_rect_size_hor` και `mesh_rect_size_vert` αντιστοιχούν στο μέγεθος των ορθογωνικών κελιών κατά τις δύο κατευθύνσεις. Η τελευταία παράμετρος `triangle_split_code` αντιστοιχεί στη μέθοδο που χρησιμοποιείται για τη διαίρεση του κάθε κελιού σε δύο τριγωνικές πλευρές. Όπως φαίνεται στο Σχήμα 1-7, η παράμετρος αυτή μπορεί να περιγράψει με δύο bits την κατεύθυνση και τη σχέση δύο διαδοχικών διαγώνιων ακμών. Έτσι, η τιμή 00 σημαίνει ότι το κάθε κελί ορθογώνιου σχήματος διαιρείται από μια ακμή που συνδέει την πάνω αριστερά κορυφή με την κάτω δεξιά, η τιμή 01 δημιουργεί ακμές από την πάνω δεξιά κορυφή προς την κάτω αριστερά, ενώ οι τιμές 10 και 11 εναλλάσσουν τις δύο αυτές μορφές σε διαδοχικά κελιά του πλέγματος. Ανάλογα με τη χρήση καθενός αντικειμένου, μπορεί να είναι περισσότερο πρόσφορη κάποια από τις διαθέσιμες αναπαραστάσεις σε σχέση με τις υπόλοιπες – ένας τομέας στον οποίο διαφοροποιείται η απόδοση κάθε τύπου πλέγματος είναι η κατεύθυνση και ο προσανατολισμός ενός μετασχηματισμού που μπορεί να υποστεί. Για παράδειγμα, αν ένα ομοιόμορφο πλέγμα τριγωνοποιείται με την τελευταία παράμετρο να παίρνει την τιμή 01, δεν αποδίδει αποτελεσματικά σε μετασχηματισμούς μεγέθυνσης κατά τον άξονα που ξεκινά από το πάνω αριστερά τμήμα του και κατευθύνεται στο κάτω δεξιά. Για να γίνει δυνατή η σωστότερη αναπαράσταση του μετασχηματισμού θα πρέπει να αυξηθεί αντίστοιχα η πυκνότητα της τριγωνοποίησης, αυξάνοντας όμως παράλληλα την πολυπλοκότητα και τις απαιτήσεις του αντικειμένου σε εύρος ζώνης. Το Σχήμα 1-8 εμφανίζει ένα παράδειγμα ομοιόμορφης οργάνωσης ενός πλέγματος, στο οποίο οι παράμετροι `nr_mesh_nodes_hor` και `nr_mesh_nodes_vert` είναι ίσες με 5 και 4 αντίστοιχα, για να δώσουν ένα σύνολο από είκοσι κορυφές στο πλέγμα, οργανωμένες σε πέντε στήλες και τέσσερις γραμμές. Ακόμα, φαίνονται τα μήκη που αντιστοιχούν στις τιμές των παραμέτρων `mesh_rect_size_hor` και `mesh_rect_size_vert` που καθορίζουν το μέγεθος και τις αναλογίες του κάθε κελιού, ενώ η παράμετρος `triangle_split_code` έχει τιμή ίση με 00.



Σχήμα 1-7: Τύποι ομοιόμορφης τοπολογίας πλέγματος



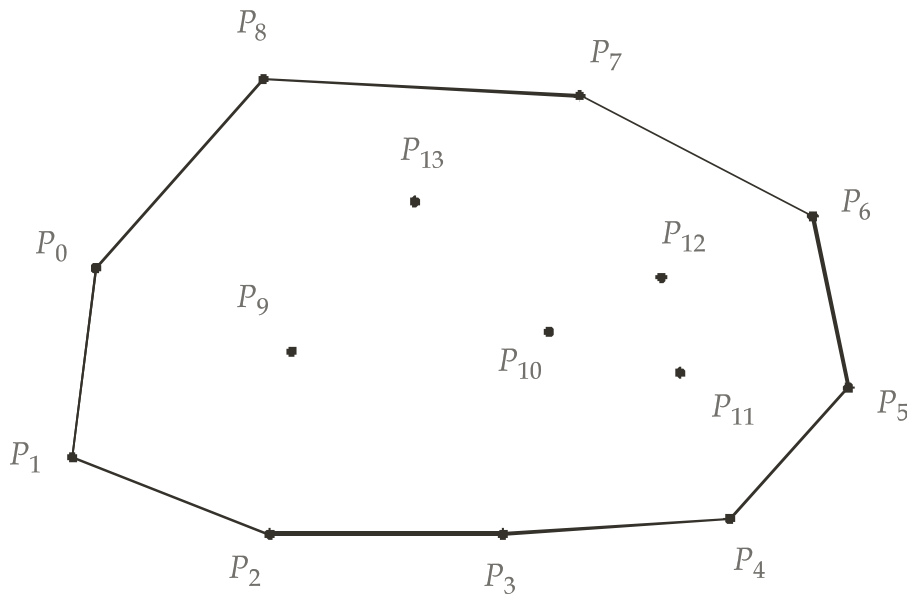
Σχήμα 1-8: Παράδειγμα οργάνωσης ενός πλέγματος ομοιόμορφης τοπολογίας

Η περίπτωση αντικειμένων που έχουν τριγωνοποιηθεί κατά Delaunay είναι μάλλον πιο συνηθισμένη και αρκετά πιο ευέλικτη στην αναπαράσταση και την επεξεργασία. Για να αναπαρασταθεί ένα τέτοιο πλέγμα, χρειάζεται αρχικά η πληροφορία του συνολικού αριθμού των κόμβων του αντικειμένου N , καθώς και ο αριθμός αυτών των κόμβων που βρίσκονται στο όριο του (N_b). Ο μεγαλύτερος όγκος δεδομένων αφορά τις θέσεις $\vec{p}_n = (x_n, y_n)$ όλων των κόμβων του αντικειμένου σε σχέση με την αρχή του τοπικού συστήματος αναφοράς, το οποίο συμπίπτει με το πάνω αριστερά άκρο του ελάχιστου τετραπλεύρου που περικλείει το συγκεκριμένο αντικείμενο.

Στην πράξη, αρκεί να κωδικοποιηθεί αναλυτικά μόνο ο πρώτος από τους κόμβους του πλέγματος $\vec{p}_0 = (x_0, y_0)$, ενώ για την ανακατασκευή των υπόλοιπων κωδικοποιούνται τα μεγέθη Δx_n και Δy_n , έτσι ώστε:

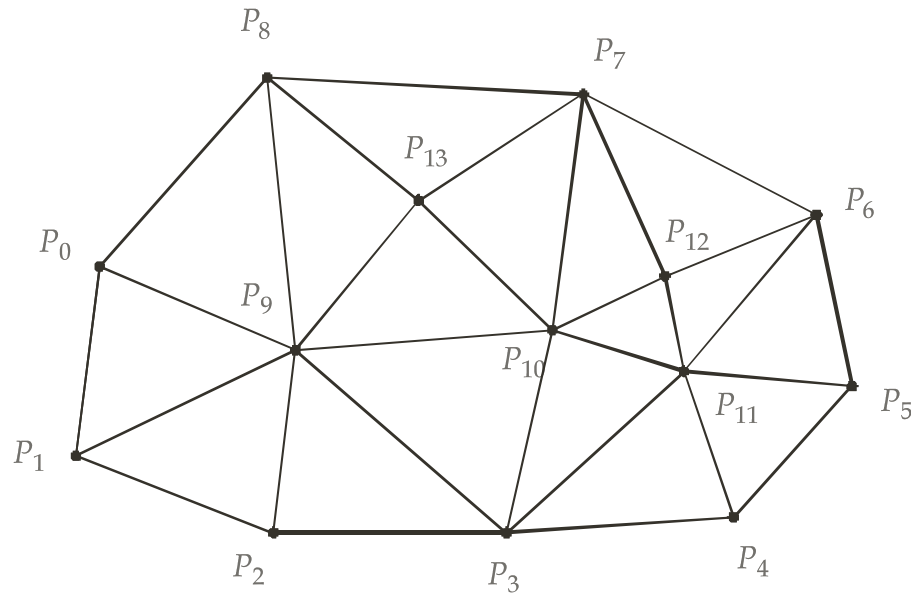
$$x_n = x_{n-1} + \Delta x_n \text{ και } y_n = y_{n-1} + \Delta y_n \quad 1-2$$

Θεωρώντας γνωστό τον αριθμό των κόμβων στο όριο του αντικειμένου και με τον περιορισμό ότι αυτοί κωδικοποιούνται και μεταδίδονται πρώτοι, ο αποκωδικοποιητής μπορεί να διακρίνει και να ανακατασκευάσει το συγκεκριμένο σύνολο χωρίς την ανάγκη για κάποια πρόσθετη πληροφορία σήμανσης, απλά συνδέοντας με ακμές τους διαδοχικούς κόμβους που λαμβάνει. Το Σχήμα 1-9 εμφανίζει ένα ανακατασκευασμένο περίγραμμα για ένα αντικείμενο τύπου πλέγματος, για το οποίο ισχύει $N = 13$ και $N_b = 8$.



Σχήμα 1-9: Ανακατασκευή του ορίου με βάση τους κόμβους που αποκωδικοποιούνται

Τα επόμενα ζεύγη συντεταγμένων που αποκωδικοποιούνται αντιστοιχούν στα εσωτερικά σημεία του πλέγματος, η τελική μορφή του οποίου ανακατασκευάζεται μετά από εφαρμογή τριγωνοποίησης Delaunay σε αυτά, με τον περιορισμό της θέσης των κόμβων που ορίζουν το περίγραμμα του αντικειμένου. Αυτή η διαδικασία θα δημιουργήσει τριγωνικές πλευρές $t_k = \langle \vec{p}_1, \vec{p}_m, \vec{p}_n \rangle$ των οποίων η επιφάνεια είναι ολόκληρη μέσα στην περιοχή που ορίζει το ανακατασκευασμένο περίγραμμα, οργανωμένες έτσι ώστε ο περιγεγραμμένος κύκλος καθεμιάς από αυτές να μην περιέχει κάποιον κόμβο \vec{p}_i , που να είναι "ορατός" από όλες τις κορυφές της, δηλαδή να συνδέεται με ακμή που να ανήκει στο πλέγμα του αντικειμένου. Το Σχήμα 1-10 δείχνει το αποτέλεσμα της τριγωνοποίησης πάνω στους κόμβους του σχήματος 1-9.



Σχήμα 1-10: Το τελικό αποκωδικοποιημένο αντικείμενο μετά τη διαδικασία της τριγωνοποίησης

Κωδικοποίηση - αποκωδικοποίηση της κίνησης

Ένα MOP τύπου *inter* μπορεί να οριστεί από ένα σύνολο δισδιάστατων διανυσμάτων κίνησης $\vec{v}_k = (vx_k, vy_k)$ το οποίο εφαρμόζεται στον κόμβο p_k του πλέγματος στο προηγούμενο MOP - με τη διάδοση αυτών των διανυσμάτων σε όλους τους αντίστοιχους κόμβους του προηγούμενου MOP μπορούμε να ανακατασκευάσουμε τις νέες θέσεις τους, αφού η τοπολογία του πλέγματος δε θα μεταβληθεί, τουλάχιστον μέχρι το επόμενο MOP τύπου *inter*. Η κωδικοποίηση των διανυσμάτων κίνησης γίνεται με πρόβλεψη, η οποία βασίζεται στις θέσεις των δύο προηγούμενων αποκωδικοποιημένων κόμβων, με τη σειρά που ορίζεται διατρέχοντας την τοπολογία κατά βάθος. Με δεδομένη την τοπολογία του πλέγματος, η διαδρομή αυτή αρκεί να καθοριστεί μόνο μια φορά για κάθε MOP τύπου *intra* και αποθηκεύεται στη μνήμη δεδομένων του πλέγματος, όπως περιγράφηκε στο Σχήμα 1-6.

Για να ξεκινήσουμε αυτήν τη διαδρομή στις M πλευρές του πλέγματος, πρέπει αρχικά να εντοπίσουμε τον πάνω αριστερά κόμβο του πλέγματος, ο οποίος βρίσκεται πιο κοντά στην αρχή του συστήματος συντεταγμένων. Έτσι, ο κόμβος n ορίζεται σαν ο κόμβος με το μικρότερο άθροισμα $x_n + y_n$, ενώ αν υπάρχουν περισσότεροι από ένας τέτοιοι κόμβοι, επιλέγεται αυτός με το μικρότερο y_n . Η *αρχική πλευρά* της διαδρομής, η οποία δεικτοδοτείται με την ένδειξη 0, είναι αυτή που περιέχει την ακμή η οποία συνδέει τον κόμβο n με τον επόμενο κόμβο του ορίου του πλέγματος, κατά την ωρολογιακή φορά. Οι υπόλοιπες πλευρές δεικτοδοτούνται με τη σειρά 1, 2, ..., $M-1$, εντοπίζοντας αρχικά τη δεικτοδοτημένη πλευρά με το μικρότερο αριθμητικά δείκτη (*τρέχουσα πλευρά*). Κατόπιν, εντοπίζουμε τη *βασική ακμή* σαν την ακμή της τρέχουσας πλευράς που ανήκει και στη δεικτοδοτημένη πλευρά με τον μικρότερο αριθμητικά δείκτη από όσες γειτονεύουν με την τρέχουσα - στην περίπτωση της αρχικής πλευράς, η βασική ακμή είναι αυτή με βάση την οποία εντοπίστηκε η αρχική πλευρά, δηλαδή εκείνη που συνδέει τον κόμβο n με τον επόμενο κόμβο του πλέγματος κατά την ωρολογιακή φορά. Με βάση τη βασική ακμή, ορίζουμε τη *δεξιά ακμή* της τρέχουσας πλευράς σαν την επόμενη ακμή που ανήκει στην πλευρά με την α-

ντίστροφη ωρολογιακή φορά και την *αριστερή ακμή* σαν την επόμενη με βάση την ωρολογιακή φορά. Αν τώρα η γειτονική πλευρά της τρέχουσας στην οποία ανήκει και η δεξιά ακμή δεν είναι δεικτοδοτημένη, τότε ορίζουμε για αυτήν τον επόμενο αριθμητικά διαθέσιμο δείκτη. Ακολουθώντας, επαναλαμβάνουμε τον ίδιο έλεγχο και για τη γειτονική πλευρά που μοιράζεται την αριστερή ακμή και ορίζουμε τον δείκτη με τον ίδιο τρόπο.

Η ροή δεδομένων που αφορά την κίνηση του πλέγματος αποτελείται από τα διανύσματα σφαλμάτων πρόβλεψης $\vec{e}_k = (ex_k, ey_k)$, στα οποία έχει εφαρμοστεί κωδικοποίηση μεταβλητού μήκους. Η πρώτη πληροφορία της ροής αφορά το διάνυσμα κίνησης για την πάνω αριστερά κορυφή n_0 , για την οποία δε χρησιμοποιείται η τεχνική της πρόβλεψης, κι έτσι ισχύει:

$$\vec{v}_k = \vec{e}_k \quad 1-3$$

Ακολουθώντας, ανακτάται το διάνυσμα σφάλματος πρόβλεψης e_{n1} για τον επόμενο με την ωρολογιακή φορά κόμβο του ορίου του πλέγματος και εφαρμόζεται η σχέση:

$$\vec{v}_{n1} = \vec{v}_{n0} + \vec{e}_{n1} \quad 1-4$$

Έτσι, ανακατασκευάζουμε τα δεδομένα για τη βασική ακμή της αρχικής πλευράς και τελικά κάθε πλευράς του πλέγματος και χρησιμοποιούμε τα διανύσματα κίνησης που υπολογίζονται έτσι για να ορίσουμε την πρόβλεψη για την τρίτη κορυφή κάθε πλευράς, αν αυτή δεν ανήκει στη βασική ακμή μιας από τις γειτονικές. Το διάνυσμα πρόβλεψης ορίζεται από τη σχέση:

$$\vec{w}_n = 0.5 (\lfloor vx_m + vx_m + 0.5 \rfloor, \lfloor vy_m + vy_m + 0.5 \rfloor) \quad 1-5$$

οπότε και το τελικό διάνυσμα κίνησης υπολογίζεται από τη σχέση:

$$\vec{v}_k = \vec{w}_k + \vec{e}_k \quad 1-6$$

Απλοποίηση Πολυγωνικών Αντικειμένων

Τα πολυγωνικά αντικείμενα χρησιμοποιούνται σε ένα σημαντικό εύρος εφαρμογών, που εντάσσονται τόσο στους τομείς της διασκέδασης και της επικοινωνίας, όσο και σε τεχνικές ή κατασκευαστικές χρήσεις. Ειδικά με την έλευση του προτύπου MPEG-4 και τις αυξημένες δυνατότητες που αυτό προσφέρει όσον αφορά την υβριδική κωδικοποίηση συνθετικών και πραγματικών αντικειμένων, γίνεται πλέον ιδιαίτερα σημαντική η βέλτιστη και χρηστική αναπαράσταση φυσικών ή τεχνητών αντικειμένων, έτσι ώστε η ανάπτυξη να επικεντρωθεί στο περιεχόμενο κι όχι στην παρουσίαση. Στο επόμενο κεφάλαιο παρουσιάζεται το πρόβλημα και οι εφαρμογές στις οποίες εμφανίζεται και αναπτύσσεται μια ανασκόπηση των πιο συνηθισμένων μεθόδων απλοποίησης πολυγωνικών αντικειμένων. Οι μέθοδοι αυτοί κατηγοριοποιούνται με βάση τη μεθοδολογία της δράσης τους και τα χαρακτηριστικά τους, ενώ στη συνέχεια παρουσιάζονται κάποια συγκεκριμένα παραδείγματα ανάπτυξής τους.

2.1 ΕΙΣΑΓΩΓΗ

Οι σύγχρονες εφαρμογές απόδοσης (rendering) πολυγωνικών αντικειμένων σε πραγματικό χρόνο γίνονται ολοένα και πιο απαιτητικές σε ό,τι αφορά την ποιότητα των μοντέλων που χρησιμοποιούνται. Αυτό αφορά τόσο το χρόνο απόδοσης, όσο και την ταχύτητα με την οποία τα απαιτούμενα δεδομένα μπορούν να μεταδοθούν μέσα από δίκτυα χαμηλού εύρους ζώνης, όπως οι τηλεφωνικές συνδέσεις Internet ή το πρωτόκολλο WAP (Wireless Application Protocol) για την ασύρματη σύνδεση συσκευών.

Ο χρόνος απόδοσης ενός πολυγωνικού αντικειμένου είναι τουλάχιστον ανάλογος προς τον αριθμό των πλευρών που το αποτελούν, στην περίπτωση τουλάχιστον που δεν εφαρμοστεί κάποια μέθοδος βελτιστοποίησης ή προοδευτικής απόδοσης. Επειδή ακριβώς το πιο χρονοβόρο τμήμα των αλγορίθμων απόδοσης είναι ο υπολογισμός των μοναδιαίων διανυσμάτων για τις κορυφές και τις πλευρές της σκηνής, ο σκοπός της βελτιστοποίησης είναι η μείωση του αριθμού των πολυγώνων που θα πρέπει να αποδοθούν. Ακόμα όμως κι αν γίνει αυτό, η παραμικρή αλλαγή σε κάποια από τις παραμέτρους της σκηνής, για παράδειγμα στο φωτισμό ή τον προσανατολισμό του αντικειμένου, θα έχει σαν αποτέλεσμα την επανάληψη της διαδικασίας από την αρχή. Εξ άλλου, η διαδικασία της απόδοσης δε λαμβάνει σε κανένα σημείο υπόψη της τη θέση

και τη σημασία ενός πολυγωνικού αντικειμένου μέσα σε μια σκηνή. Για παράδειγμα, αν πρέπει να αποδώσουμε μια σκηνή που περιέχει ανθρώπινα πρόσωπα σε διαφορετικές αποστάσεις από την κάμερα και ακόμα διάφορα αντικείμενα στο φόντο της, θα πρέπει να δώσουμε περισσότερη προσοχή σε αντικείμενα που βρίσκονται κοντά στον οπτικό άξονα της κάμερας και πιο κοντά σε αυτήν – ο λόγος για αυτή τη διάκριση είναι ότι συνήθως αυτά τα αντικείμενα τραβούν το ενδιαφέρον του θεατή, ενώ τα υπόλοιπα σπάνια παρατηρούνται σε κανονικές συνθήκες.

Για να εξασφαλιστεί λοιπόν σταθερός ρυθμός απόδοσης σε εφαρμογές πραγματικού χρόνου, είναι πιθανό να χρειαστεί να απλοποιηθούν κάποια από τα αντικείμενα του φόντου, έτσι ώστε να μειωθεί ουσιαστικά και ο χρόνος που απαιτείται για την απόδοση της συνολικής σκηνής. Αυτό θα έχει σαν αποτέλεσμα τη μείωση της ποιότητας στην εικόνα αυτών των αντικειμένων – σε αυτό το σημείο, κάνουμε την παραδοχή ότι αυτή η μείωση της ποιότητας δε γίνεται αντιληπτή από το θεατή, ειδικά αν τα συγκεκριμένα αντικείμενα κινούνται. Σε κάθε περίπτωση, η απλοποίηση αυτή είναι απαραίτητη για συγκεκριμένες εφαρμογές, όπως η τηλεδιάσκεψη σε περιβάλλοντα χαμηλού εύρους ζώνης ή σε συσκευές που χρησιμοποιούν αλγορίθμους απόδοσης σε ενσωματωμένα (embedded) κυκλώματα. Μια εφαρμογή στην οποία δεν ενδιαφέρει τόσο η ποιότητα, όσο η ταχύτητα απόδοσης ενός πολυγωνικού μοντέλου είναι η παρουσίαση μιας "πρόχειρης" μορφής του αντικειμένου, μέχρι να μεταδοθεί ή να αποδοθεί ολόκληρο. Αυτό έχει ιδιαίτερη σημασία σε εφαρμογές που λειτουργούν στο Internet και μπορεί να εκτελούνται κυρίως στο υπολογιστικό σύστημα του χρήστη.

Μια πρώτη προσέγγιση της λειτουργικότητας της απλοποίησης ενός πολυγωνικού αντικειμένου έγινε ήδη από το πρότυπο VRML. Εδώ, για καθένα από τα μοντέλα που χρησιμοποιούνται μπορούν να καθοριστούν διαφορετικές μορφές του σε διαφορετικά επίπεδα λεπτομέρειας (levels of detail – LOD) και οι οποίες αντικαθιστούν δυναμικά το αρχικό μοντέλο όταν η θέση παρατήρησης απομακρυνθεί από αυτό. Αν και αυτά τα μοντέλα θα πρέπει να μεταδοθούν έτοιμα κι αλλιώς στην εφαρμογή απόδοσης, η χρήση τους μπορεί να "ελαφρύνει" κατά πολύ το υπολογιστικό βάρος μιας σκηνής κι έτσι να προσφέρει μια έντονα διαλογική μορφή περιήγησης. Η μετάβαση από τη μια μορφή LOD στην επόμενη γίνεται, σύμφωνα με το πρότυπο, όταν η απόσταση από τη θέση παρατήρησης ξεπεράσει ένα καθορισμένο από το σχεδιαστή κατώφλι – αυτό σημαίνει ότι κατά την περιήγηση του χρήστη μέσα στη σκηνή, τα αντικείμενα αυτά μπορεί να αλλάξουν μορφή απότομα, μια που οι διακριτές μορφές LOD είναι περιορισμένες σε αριθμό. Έτσι, φαίνεται περισσότερο χρηστική μια ιδέα προοδευτικής μείωσης της λεπτομέρειας ενός πολυγωνικού μοντέλου, τόσο με βάση την απόσταση από την κάμερα, όσο και με βάση την πραγματική ανάγκη παρουσίας της λεπτομέρειας, κάτι που είναι περισσότερο εύκολο να μετρηθεί αντικειμενικά.

Αυτό το κεφάλαιο περιγράφει μια σειρά από μεθόδους απλοποίησης πολυγωνικών μοντέλων και τις διακρίνει σε δύο μεγάλες ομάδες: στην πρώτη από αυτές, κάθε βήμα του αλγορίθμου απλοποίησης υπολογίζει το επόμενο δομικό τμήμα του μοντέλου που πρέπει να αντικατασταθεί με βάση κάποιο "γεωμετρικό ιστορικό". Αυτό μας δίνει τη δυνατότητα να υπολογίζουμε σε κάθε στιγμή τη γεωμετρική απόσταση του απλοποιημένου μοντέλου από το αρχικό ή από το προηγούμενο στάδιο και να διακόψουμε την εξέλιξη της διαδικασίας όταν η διαφορά (σφάλμα) φτάσει σε κάποιο προκαθορισμένο κατώφλι. Αν και αυτή η δυνατότητα διακοπής είναι αρκετά χρήσιμη, μπορεί να οδηγήσει σε καταστάσεις όπου κάποια από τα δομικά στοιχεία του μοντέ-

λου χαρακτηρίζονται από ιδιότητες που υπερβαίνουν το κατώφλι που έχουμε ορίσει, ενώ τα υπόλοιπα όχι. Ακόμα, οι απαιτήσεις τέτοιων αλγορίθμων σε μνήμη είναι αρκετά μεγάλες, μια που τα δεδομένα που πρέπει να φυλάσσονται, παράλληλα με την τρέχουσα μορφή είναι αρκετά και χρησιμοποιούνται μόνο για τους υπολογισμούς σφάλματος.

Η δεύτερη ομάδα αλγορίθμων που παρουσιάζεται και που περιλαμβάνει και τον προτεινόμενο αλγόριθμο αναφέρεται συχνά στη βιβλιογραφία σαν απλοποίηση χωρίς μνήμη: ο λόγος που γίνεται αυτό είναι ότι κατά τη διάρκεια ενός βήματος απλοποίησης δε γίνεται κάποια σύγκριση με το αρχικό μοντέλο ή το προηγούμενο βήμα, παρά μόνο μετά το τέλος του βήματος. Με αυτόν τον τρόπο αποκομίζουμε τεράστια οφέλη σε αποθηκευτικό χώρο και χρόνο ολοκλήρωσης της διαδικασίας, ενώ διατηρώντας το αντίστοιχο ιστορικό απλοποιήσεων μπορούμε εύκολα να μεταφερθούμε δυναμικά και χωρίς απώλειες σε οποιοδήποτε προηγούμενο βήμα της διαδικασίας.

2.2 ΑΝΑΣΚΟΠΗΣΗ ΜΕΘΟΔΩΝ

Αν και η απλοποίηση – συμπίεση πολυγωνικών αντικειμένων είναι ένας από τους τομείς που προσελκύουν μεγάλο ερευνητικό ενδιαφέρον, είναι ταυτόχρονα και ένα αρκετά παλιό πρόβλημα για τη συνθετική εικόνα. Ήδη από το 1976, άρχισε να εκφράζεται η ιδέα για την ιεραρχική αναπαράσταση ενός μοντέλου σε πολλαπλούς βαθμούς λεπτομέρειας, έτσι ώστε να μπορεί να χρησιμοποιείται κάθε φορά η κατάλληλη [3]. Η έκρηξη στο ενδιαφέρον για τέτοιες ιδέες που παρατηρήθηκε από το 1992 και μετά οφείλεται σε μεγάλο βαθμό στην ανάπτυξη εξειδικευμένου υλικού που επιταχύνει τη διαχείριση και απόδοση πολύπλοκων σκηνών με πολυγωνικά μοντέλα. Άλλωστε, η απλότητα του μαθηματικού μοντέλου που περιγράφει μια πολυγωνική τοπολογία και τις διάφορες πράξεις που μπορούν να εφαρμοστούν πάνω της είναι αυτή που, ακόμα και σήμερα, τα τοποθετεί στην κορυφή της προτίμησης, τόσο για επιστημονικές, όσο και για πιο πρακτικές εφαρμογές. Στην πράξη, ακόμα και αναπαραστάσεις που υπερέχουν των πολυγωνικών σε αρκετούς και σημαντικούς τομείς, όπως η αναπαράσταση με *splines* ή με *voxels*, μπορούν εύκολα να μετατραπούν σε πολυγωνικές έτσι ώστε να υποβοηθηθεί η απόδοση και η επεξεργασία τους.

Από την άλλη, η αδυναμία της πολυγωνικής αναπαράστασης να περιγράψει αποτελεσματικά μια μη επίπεδη επιφάνεια έχει οδηγήσει σε μια σειρά από εναλλακτικές προσεγγίσεις, έτσι ώστε να βελτιωθεί η ποιότητα και η πληρότητα της παρεχόμενης πληροφορίας, χωρίς να θυσιάζουν οι ευκολίες της βασικής ιδέας. Στα πλαίσια αυτά, εκτός από την απλοποίηση των πολυγωνικών αντικειμένων που βρίσκονται σε απόσταση από την κάμερα ή που αποδίδονται από συσκευές χωρίς εξειδικευμένο υλικό, συναντάμε αλγορίθμους που ενσωματώνουν στην τελική εικόνα πρόσθετες οπτικές πληροφορίες, εκτός από την εμφάνιση του μοντέλου (για παράδειγμα, *texture mapping* ή σκίαση *Phong*) και αλγορίθμους που αποκόπτουν τμήματα του μοντέλου που δε φαίνονται σε κάποια συγκεκριμένη όψη γιατί κρύβονται από άλλα τμήματα της γεωμετρίας του.

Γενικά, μπορούμε να κατηγοριοποιήσουμε τους αλγορίθμους απλοποίησης με βάση μια σειρά από κριτήρια, ανάλογα με την επέμβαση στο αρχικό μοντέλο ή το εύρος των εφαρμογών. Πρακτικά αυτό σημαίνει ότι η απλοποίηση πολυγωνικών μοντέλων δεν είναι ένα μονοσήμαντα

λυμένο πρόβλημα, αλλά και ότι δεν υπάρχει ένας αλγόριθμος ή μια οικογένεια αλγορίθμων που να είναι καλύτερη με βάση όλα τα κριτήρια ή όλες τις συνθήκες προβλημάτων. Στη συνέχεια θα παρουσιάσουμε μια σειρά από ιδέες για την κατηγοριοποίηση των διαθέσιμων αλγορίθμων και τις εφαρμογές στις οποίες εμφανίζουν τα περισσότερα πλεονεκτήματα

Μηχανισμός απλοποίησης

Η βασικότερη μέθοδος κατηγοριοποίησης αφορά το μηχανισμό με τον οποίο αφαιρούνται τα πολύγωνα από το μοντέλο. Η μεγάλη πλειοψηφία των αλγορίθμων μπορεί να ενταχθεί σε μια από τις παρακάτω κατηγορίες:

- **Αλλαγή δειγματοληψίας:** εδώ η απλοποίηση γίνεται με τη μεταβολή της πυκνότητας των δομικών στοιχείων του μοντέλου σε σχέση με το πρωτότυπο. Τα δομικά αυτά στοιχεία μπορεί να είναι απλά σημεία – κορυφές πάνω σε ένα μοντέλο που περιγράφεται από κυρτή επιφάνεια ή voxels μέσα σε ένα τρισδιάστατο πλέγμα το οποίο περιβάλλει το μοντέλο. Ανάλογα με το στόχο που θέτουμε στον αλγόριθμο όσον αφορά τον τελικό αριθμό δομικών στοιχείων ή την απόσταση από το πρωτότυπο, μπορούμε να επιλέξουμε το κατάλληλο κάθε φορά βήμα δειγματοληψίας και να δημιουργήσουμε από την αρχή μια καινούρια προσέγγιση της αρχικής επιφάνειας.
- **Προσαρμοζόμενη υποδιαίρεση (adaptive subdivision):** εδώ η διαδικασία ξεκινά από ένα πολύ απλό βασικό μοντέλο, ακόμα και κάποιο από τα πιο απλά Πλατωνικά στερεά, το οποίο περιγράφεται συνήθως από τριγωνικές ή τετραγωνικές πλευρές. Καθεμιά από αυτές τις πλευρές υποδιαιρείται με επαναληπτικό τρόπο, μέχρι κάποιο στάδιο της διαδικασίας να μας δώσει ένα μοντέλο αρκετά κοντά στο αρχικό ή το ζητούμενο. Ένα σημαντικό πρόβλημα που μπορεί να συναντήσουμε στους αλγορίθμους αυτής της οικογένειας αφορά στην επιλογή του βασικού μοντέλου, το οποίο συνήθως δημιουργείται από κάποιο σχήμα μετάβασης από την αρχική τοπολογία. Ακόμα, οι τεχνικές ομοιόμορφης κατάτμησης των δομικών στοιχείων είναι πολύ πιθανό ότι δε θα μπορέσουν να προσεγγίσουν κάποιες τοπικές λεπτομέρειες του μοντέλου ή απότομες αλλαγές στη μορφή της επιφάνειάς του, χωρίς να αυξήσουν δραματικά τον αριθμό των πολυγώνων του. Από την άλλη, ένα αρκετά σημαντικό πρακτικό πλεονέκτημα των αλγορίθμων υποδιαίρεσης είναι ότι, με κάποιες απλές μετατροπές, μπορούν να διατηρήσουν το ιστορικό των απλοποιήσεων και των υποδιαίρεσεων κι έτσι να μεταβαίνουν δυναμικά από τη μια μορφή του μοντέλου σε κάποια πιο λεπτομερή και αντίστροφα – για να γίνει αυτό βέβαια, θα πρέπει κάθε φορά να δημιουργηθούν και όλα τα ενδιάμεσα βήματα.
- **Αποκοπή:** σε αυτήν την περίπτωση, ο αλγόριθμος αφαιρεί με επαναληπτικό τρόπο δομικά στοιχεία του πολυγωνικού μοντέλου (κορυφές ή πλευρές) και τριγωνοποιεί εκ νέου την οπή που προκύπτει. Πρακτικά αυτή η διαδικασία απλοποιεί το μοντέλο χωρίς να επιτρέπει την αφαίρεση κάποιας κορυφής ή πλευράς που θα έχει σαν αποτέλεσμα την αλλαγή της τοπολογίας σε τοπικό επίπεδο κι έτσι το συνολικό κέρδος από την απλοποίηση είναι συνήθως πολύ μικρό.
- **Συγχώνευση κορυφών:** οι αλγόριθμοι αυτής της κατηγορίας απλοποιούν το μοντέλο συγχωνεύοντας δύο ή περισσότερες κορυφές του αρχικού μοντέλου σε μια καινούρια, η οποία μπορεί να λάβει μέρος και η ίδια σε μια παρόμοια διαδικασία. Η συγχώνευση αυτή θα έχει σαν αποτέλεσμα τον εκφυλισμό κάποιων από τις πλευρές του μοντέλου, οι οποίες στη συνέχεια αφαιρούνται, μειώνοντας έτσι την πολυπλοκότητα της επιφάνειας. Στις περισσότερες περιπτώσεις, οι αλγόριθμοι αυτοί μπορούν να απλοποιήσουν αρκετά πολύπλοκες επιφάνειες, ακόμα και αναδιπλούμενες – τότε, μόνο δύο κορυφές μπορούν να συγχωνευτούν σε κάθε βήμα, σε μια διαδικασία που ονομάζεται απλοποίηση ακμής (edge collapse).

Εξάρτηση από την όψη

Μια πιο σύγχρονη ιδέα για την απλοποίηση ενός πολυγωνικού μοντέλου εξαρτά την πορεία του αλγορίθμου από τη θέση του παρατηρητή ή της κάμερας. Πιο συγκεκριμένα, οι αλγόριθμοι που είναι ανεξάρτητοι από την όψη υπολογίζουν διαφορετικές εκδόσεις LOD για τα μοντέλα μιας σκηνής – τα μοντέλα αυτά χρησιμοποιούνται δυναμικά σε μια σκηνή, συνήθως ανάλογα με την απόστασή τους από τη θέση της κάμερας. Όσο περισσότερο πλησιάζει το μοντέλο σε αυτή, το σύστημα επιλέγει μια πιο λεπτομερή αναπαράσταση για την απόδοση, αντικαθιστώντας την προηγούμενη. Αντίθετα, οι αλγόριθμοι που εξαρτώνται από την όψη (view-dependent) αντικαθιστούν την πολλαπλή έκφραση του μοντέλου από μια συγκεκριμένη, η οποία μεταβάλλεται δυναμικά. Κατά την απόδοση, ο αλγόριθμος υπολογίζει την τοπολογία ανάλογα με τη θέση και τον προσανατολισμό της κάμερας – αυτό σημαίνει ότι κάποια τμήματα του μοντέλου που είναι πιο κοντά στην κάμερα είναι περισσότερο λεπτομερή από άλλα. Έτσι, η συνολική τοπολογία του μοντέλου μπορεί να καλύπτει διαφορετικά στιγμιότυπα της απλοποίησης, ειδικά αν ο κύριος άξονας του μοντέλου είναι παράλληλος ή σχεδόν παράλληλος με τον οπτικό, οπότε και το αντικείμενο εκτείνεται σε μεγάλο τμήμα κάθετα στο επίπεδο παρατήρησης.

Υπολογισμός του σφάλματος

Μια σειρά από αλγορίθμους απλοποίησης υπολογίζουν τις διάφορες μορφές LOD με βάση κάποιο τελικό στόχο ή κάποιο κατώφλι σε ότι αφορά τα δομικά στοιχεία. Συνήθως πάντως το σφάλμα σε σχέση με το αρχικό μοντέλο ή το προηγούμενο βήμα παίζει σημαντικό ρόλο στην επιλογή του επόμενου βήματος και του δομικού στοιχείου που θα απλοποιηθεί. Οι αλγόριθμοι που βασίζονται στην πιστότητα της διαδικασίας χρησιμοποιούν σαν είσοδο κάποια έκφραση της σχέσης ή της απόκλισης με το αρχικό μοντέλο και προσπαθούν να μειώσουν το συνολικό αριθμό πολυγώνων χωρίς να παραβούν αυτήν την έκφραση. Μια πολύ συνηθισμένη τέτοια έκφραση αφορά τη μέγιστη απόσταση της απλοποιημένης επιφάνειας, σε τοπικό ή συνολικό επίπεδο, από την αρχική. Οι αλγόριθμοι προσαρμοζόμενης υποδιαίρεσης προσφέρονται για τέτοιου είδους έλεγχο, αφού το είδος και το αποτέλεσμα της απλοποίησης μπορεί εύκολα να καθοριστεί σε τοπικό επίπεδο και με αναλυτική, μετρήσιμη μορφή.

Από την άλλη, ο στόχος της απλοποίησης μπορεί να εκφραστεί με βάση το συνολικό αριθμό πολυγώνων του μοντέλου, ο οποίος είναι συνήθως γνωστός σαν polygon budget. Η λογική εδώ είναι ότι η απλοποίηση θα σταματήσει όταν ο συνολικός αριθμός των πολυγώνων φτάσει στο ζητούμενο, διατηρώντας όσο είναι δυνατό την ποιότητα του αρχικού μοντέλου. Η διαδικασία αποκοπής είναι εδώ περισσότερο ταιριαστή, αφού μπορεί να εφαρμοστεί επιλεκτικά στα δομικά στοιχεία του μοντέλου, και μάλιστα όσες φορές αυτό είναι επιθυμητό. Γενικά πάντως είναι επιθυμητός ένας συνδυασμός και των δύο ιδεών, μια που η πιστότητα του τελικού μοντέλου ανακλάται πάνω στην εικόνα που προκύπτει από την απόδοσή του, ενώ ο αριθμός πολυγώνων επηρεάζει την ταχύτητα της απόδοσης και της επεξεργασίας και είναι ιδιαίτερα σημαντικός σε εφαρμογές που λειτουργούν σε πραγματικό χρόνο.

2.3 ΠΑΡΑΔΕΙΓΜΑΤΑ ΑΛΓΟΡΙΘΜΩΝ

Ένας από τους πιο διαδομένους αλγόριθμους αποκοπής παρουσιάστηκε το 1992 σαν συμπλήρωμα ενός αλγορίθμου Marching Cubes, για την εξαγωγή ισοδυναμικών επιφανειών από χωρικά δεδομένα [5]. Επειδή τα αντικείμενα που προκύπτουν σαν αποτέλεσμα των αλγορίθμων Marching Cubes περιέχουν λεπτομέρεια η οποία είναι σε μεγάλο βαθμό άχρηστη, με βάση τη δομή του μοντέλου, είναι απαραίτητο να απορριφθούν οι πρόσθετες πλευρές που βρίσκονται σε επίπεδες επιφάνειες και δεν προσφέρουν κάτι στην περιγραφή τους. Στη συγκεκριμένη υλοποίηση, αυτό επιτυγχάνεται εξετάζοντας κάθε φορά μια κορυφή του μοντέλου για το αν η αφαίρεσή της θα μεταβάλλει την τοπολογία της γειτονιάς στην οποία ανήκει – αν η απάντηση είναι αρνητική, τότε η κορυφή και όλα τα σχετιζόμενα με αυτήν δομικά στοιχεία απορρίπτονται από το μοντέλο. Η οπή που προκύπτει σαν αποτέλεσμα της απόρριψης κλείνει με τη χρήση αλγορίθμων loop-splitting. Η διαδικασία επαναλαμβάνεται μέχρι το σημείο όπου δεν είναι δυνατή πλέον η απόρριψη άλλων πλευρών.

Ένα ενδιαφέρον πλεονέκτημα που έχει η συγκεκριμένη μέθοδος αφορά στην επαναχρησιμοποίηση στοιχείων που αφορούν τις κορυφές του μοντέλου, όπως οι συντεταγμένες για την απεικόνιση υψής (mapping coordinates) ή ακόμα και τα μοναδιαία διανύσματά τους. Επειδή οι κορυφές που τελικά απομένουν στο μοντέλο είναι ένα υποσύνολο των αρχικών, δεν είναι απαραίτητο να υπολογίσουμε εκ νέου αυτά τα στοιχεία, με δεδομένο ότι δεν έχει αλλάξει η θέση τους μέσα στο μοντέλο, ούτε η τοπολογία της γειτονιάς τους.

Μια ενδιαφέρουσα πρόταση για το συνδυασμό των ιδεών της δειγματοληψίας και της αποκοπής [6], η οποία εφαρμόζεται πιο αποτελεσματικά σε ομαλές επιφάνειες, χωρίς απότομες ακμές ή αλλαγές στην καμπυλότητα, χρησιμοποιεί την ιδέα του polygon budget, αλλά όχι κάποιο προφανές ποιοτικό κριτήριο. Εδώ, ο ζητούμενος αριθμός κορυφών κατανέμεται με τυχαίο τρόπο πάνω στην επιφάνεια του αρχικού μοντέλου και κατόπιν μετακινείται με βάση δυνάμεις απώθησης ανάμεσα στις πιο κοντινές από αυτές. Έτσι, τελικά οι κορυφές θα συγκλίνουν σε θέσεις τέτοιες, ώστε η κατανομή τους να είναι ομοιόμορφη πάνω στην τελική επιφάνεια. Κατόπιν, ο αλγόριθμος κατασκευάζει μια *ενδιάμεση* επιφάνεια, ανάμεσα στην αρχική και την απλοποιημένη, η οποία περιέχει τόσο τα δομικά στοιχεία της πρώτης, όσο και τις καινούριες κορυφές. Μετά από ένα βήμα βελτιστοποίησης των τριγωνικών πλευρών που προκύπτουν, οι κορυφές της αρχικής επιφάνειας απορρίπτονται, αφήνοντας μόνο αυτές που τοποθετήθηκαν σε αυτήν από τον αλγόριθμο απλοποίησης. Ένα πλεονέκτημα αυτής της ιδέας είναι η εύκολη και ομαλή μετάβαση από τη μια μορφή LOD στην αμέσως επόμενη, κάτι που διαφοροποιεί την τελική εμφάνιση σε σχέση με πρότυπα περιορισμένων μοντέλων LOD, όπως τη VRML.

Σε περιπτώσεις όπου η ταχύτητα της υλοποίησης είναι περισσότερο σημαντική από την αισθητική ποιότητα του τελικού αποτελέσματος είναι αρκετά συνηθισμένοι οι αλγόριθμοι απλοποίησης με συγχώνευση κορυφών. Μια προσέγγιση που συνδυάζει την ταχύτητα και απλότητα, με τη διατήρηση μιας μετρικής σφάλματος σε σχέση με το αρχικό μοντέλο [7] ταξινομεί τις κορυφές με σειρά σημασίας και επικεντρώνεται στην πιο σημαντική από αυτές. Εκεί τοποθετείται ένα πλέγμα του οποίου το μέγεθος καθορίζεται από το χρήστη, και το οποίο χρησιμεύει στο να περικλείσει όλες τις κορυφές που θα απλοποιηθούν με συγχώνευση. Κατόπιν, η "σημαντική" κορυφή τοποθετείται στο γεωμετρικό κέντρο του πλέγματος και η διαδικασία επαναλαμβάνεται.

Με αυτήν τη λογική, διατηρείται η ανεξαρτησία από τον προσανατολισμό του μοντέλου, ενώ η τελική ποιότητα και ο αριθμός των πολυγώνων μπορούν να ελεγχθούν εύκολα με βάση το μέγεθος και την πυκνότητα του πλέγματος.

Επειδή οι ακμές που αντιστοιχούν στο περίγραμμα (silhouette) ενός μοντέλου επηρεάζουν σημαντικά την εμφάνισή του κατά την απόδοση, μια που ορίζουν τα όριά του σε σχέση με το φόντο, θεωρούνται από τον αλγόριθμο σαν πολύ σημαντικές. Οι πρώτες υλοποιήσεις της ιδέας χρησιμοποιούν σαν μέτρο για τη σημασία μιας κορυφής την ποσότητα $1/\theta$, όπου θ η μέγιστη γωνία που σχηματίζουν ανά δύο οι ακμές που συναντώνται σε μια κορυφή – στη συγκεκριμένη περίπτωση πάντως, θεωρείται ότι η ποσότητα $\cos(\theta/2)$ είναι μάλλον ένα πιο ασφαλές κριτήριο για το αν η συγκεκριμένη κορυφή βρίσκεται στο περίγραμμα. Η λογική εδώ είναι ότι όσο πιο απότομη είναι η κλίση της επιφάνειας σε ένα συγκεκριμένο σημείο, τόσο πιο πιθανό είναι αυτή η "ακίδα" να προεξέχει από το κυρίως σώμα και να αποδίδεται στο περίγραμμα του αντικειμένου.

Γενικά, οι αλγόριθμοι που διατηρούν την αρχική τοπολογία του μοντέλου έχουν σχετικά περιορισμένη επιτυχία σε ότι αφορά την ποσοτική απλοποίηση της γεωμετρίας ενός αντικειμένου. Από την άλλη, η απουσία μιας μετρικής για την ποιότητα της διαδικασίας και την απόσταση από την αρχική επιφάνεια μπορεί να φέρει αποτελέσματα που να είναι μεν σωστά με τη μαθηματική ή την υπολογιστική έννοια, αλλά όχι χρηστικά από την αισθητική άποψη. Οι προσεγγίσεις που χρησιμοποιούν voxels προσδοκούν στο να απλοποιήσουν την τοπολογία του μοντέλου με ελεγχόμενο και προοδευτικό τρόπο. Αρχικά δημιουργείται ένα τρισδιάστατο ομοίωμα του μοντέλου, με την υπέρθεση σε αυτό ενός τρισδιάστατου πλέγματος, στο οποίο η πυκνότητα των πολυγώνων χρησιμεύει στον καθορισμό της τιμής για καθένα από τα voxels. Κατόπιν, εφαρμόζονται διαδικασίες αναδειγματοληψίας και βαθυπερατού φιλτραρίσματος, δημιουργώντας έτσι μια αναπαράσταση χαμηλότερης διακριτότητας. Το φίλτρο εξασφαλίζει ότι οι υψηλές συχνότητες, οι οποίες αντιστοιχούν στις απότομες και τοπικές αλλαγές στην καμπυλότητα και τη μορφή της επιφάνειας θα εξαλειφθούν. Τα τελικά voxels χρησιμεύουν σαν είσοδος σε έναν αλγόριθμο Marching Cubes για να δώσουν την απλοποιημένη πολυγωνική γεωμετρία – με τη σειρά της, αυτή απλοποιείται περαιτέρω με κάποιον από τους συνηθισμένους αλγορίθμους διατήρησης της τοπολογίας, για να εξαλειφθεί η υπερβολική λεπτομέρεια από τον αλγόριθμο Marching Cubes.

Μια σημαντική ιδέα [8] για την υλοποίηση μιας σειράς αλγορίθμων χρησιμοποιεί αντίγραφα της επιφάνειας μετακινημένα το πολύ κατά απόσταση ϵ από αυτήν. Τα αντίγραφα αυτά σχηματίζουν ένα περίγραμμα απλοποίησης (simplification envelope) στο οποίο το εξωτερικό σκέλος υπολογίζεται από τη μεταφορά κάθε κορυφής του αρχικού μοντέλου κατά ϵ , κατά μήκος του σχετικού μοναδιαίου διανύσματος, ενώ το εσωτερικό από την μεταφορά κατά $-\epsilon$. Αν η καμπυλότητα της επιφάνειας είναι τέτοια που τα δύο σκέλη διασταυρώνονται, τότε τοπικά το ϵ ελαττώνεται μέχρι να πάψει να συμβαίνει αυτό.

Από αυτό το σημείο και μετά, το περίγραμμα απλοποίησης καθοδηγεί τη διαδικασία αποκοπής, διατηρώντας σε κάθε βήμα την ενδιάμεση επιφάνεια μέσα σε αυτό. Έτσι, εξασφαλίζεται ότι η συνολική τοπολογία διατηρείται και ότι η απόσταση από την πρωτότυπη επιφάνεια είναι το πολύ ίση με ϵ , κάτι που δίνει γενικά πολύ καλά αποτελέσματα. Τα προβλήματα που σχετίζονται με αυτήν την ιδέα έχουν περισσότερο να κάνουν με τη δυσκολία υλοποίησης ενός σταθε-

ρού συστήματος κατασκευής του περιγράμματος απλοποίησης από την αρχική επιφάνεια και με το περιορισμό στο εύρος της απλοποίησης που τίθεται από τον καθορισμό του ϵ . Πάντως, η χρήση αλγορίθμων αποκοπής είναι περισσότερο ένα συγκεκριμένο παράδειγμα που παρουσιάστηκε, παρά περιορισμός, μια που δεν υπάρχει λόγος να μη χρησιμοποιηθεί κάποιο σχήμα απλοποίησης με συγχώνευση κορυφών.

Μια σημαντική ιδέα που ξεκίνησε από την ανάγκη απλοποίησης μοντέλων που δημιουργήθηκαν από τρισδιάστατους σαρωτές και τα οποία περιέχουν σημαντικό αριθμό άχρηστων δομικών στοιχείων, ξεκινά από την ανακατασκευή της επιφάνειας από ένα σύνολο από σημεία [2]. Εδώ η απλοποίηση είναι απλά ένα βήμα βελτιστοποίησης της συνολικής διαδικασίας και χρησιμοποιεί τον αριθμό των κορυφών και την απόστασή τους από το πρωτότυπο σαν μια συνάρτηση ενέργειας που πρέπει να ελαχιστοποιηθεί. Ξεκινώντας από ένα τυχαίο τοποθετημένο σύνολο δειγμάτων και κορυφών, δοκιμάζεται το αποτέλεσμα ενός μετασχηματισμού σε κάποια από τις ακμές του αρχικού μοντέλου – ο μετασχηματισμός μπορεί να αφορά την απόρριψη της ακμής, τη διάσπασή της σε δύο ή τη μεταβολή του προσανατολισμού της. Με βάση το αποτέλεσμα, επιχειρείται να ελαχιστοποιηθεί η συνάρτηση ενέργειας για το επόμενο βήμα. Αν αυτό δεν είναι δυνατό, τότε ο μετασχηματισμός αναιρείται.

Γενικά, ο συγκεκριμένος αλγόριθμος είναι πολύ προσεκτικός στις αλλαγές που επιφέρει, διατηρώντας σε μεγάλο βαθμό την τοπολογία του μοντέλου. Αν και αυτό έχει σαν αποτέλεσμα να μην είναι δυνατή η απλοποίηση σε μεγάλο βαθμό, οι απότομες αλλαγές στην επιφάνεια εντοπίζονται αρκετά εύκολα και οι πλευρές που προκύπτουν είναι πάντοτε καλά ορισμένες. Πάντως, η υλοποίηση είναι εξαιρετικά αργή, κάτι που περιορίζει τη χρήση του σε off-line εφαρμογές.

Η παραπάνω τεχνική χρησιμοποιείται για την εισαγωγή της ιδέας του progressive mesh [9], δηλαδή ενός μοντέλου στο οποίο διατηρείται το ιστορικό και μπορεί να χρησιμοποιηθεί για τη μετάβαση σε μια διαφορετική μορφή LOD. Πιο συγκεκριμένα, η διαδικασία ξεκινά από ένα πολύ απλό βασικό μοντέλο, το οποίο δημιουργείται από μια σειρά από απορρίψεις ακμών, ακολουθούμενη από διασπάσεις κορυφών, διαδικασία δυϊκή της απόρριψης μιας ακμής. Η λογική εδώ είναι ότι το βασικό μοντέλο δημιουργείται από τη μείωση της λεπτομέρειας, με την απόρριψη ακμών, ενώ οι πιο λεπτομερείς αναπαραστάσεις ενσωματώνουν πρόσθετη λεπτομέρεια με τη βοήθεια της διάσπασης κορυφών. Επειδή οι δύο μετασχηματισμοί είναι δυϊκοί, η διαδικασία είναι πλήρως αντιστρέψιμη κι έτσι η εφαρμογή όλων των υπολογισμένων διασπάσεων θα μας οδηγήσει από το βασικό μοντέλο στο αρχικό – αν η διαδικασία σταματήσει κάπου ενδιάμεσα, τότε θα μας δώσει μια απλοποιημένη μορφή του πρωτοτύπου. Έτσι, είναι πολύ εύκολος ο καθορισμός κάποιου στόχου, με την έννοια του polygon budget, αλλά και ο καθορισμός μιας αντιπροσωπευτικής συνάρτησης ενέργειας για τον καθορισμό της επόμενης ακμής που θα απλοποιηθεί. Αυτός μπορεί να γίνει με βάση όχι μόνο γεωμετρικά κριτήρια, αλλά και ποιοτικά ή σχετικά με την απόδοση του μοντέλου, όπως το υλικό ή η απεικόνιση υφής.

Μια προσέγγιση που έχει σημαντικό μαθηματικό υπόβαθρο, μια που μεταφέρει την ιδέα των wavelets στα τρισδιάστατα μοντέλα κατασκευάζει πολυδιακριτικές (multiresolution) αναπαραστάσεις για ένα μοντέλο, χωρίς απαραίτητα να φτάνει ποτέ σε μια μορφή που να ταυτίζεται με το πρωτότυπο. Το σημαντικό πλεονέκτημα αυτής της μεθόδου [11] είναι ότι είναι πολύ εύκολο να μεταφερθούμε με ομαλό τρόπο από τη μια αναπαράσταση LOD στην επόμενη, αφαι-

ρώντας ή προσθέτοντας συντελεστές wavelet. Επειδή ο υπολογισμός της απόστασης από το αρχικό μοντέλο είναι εσωτερικό χαρακτηριστικό του μαθηματικού μοντέλου είναι πολύ εύκολο να σταματήσουμε τη διαδικασία μετάδοσης κι ενσωμάτωσης των συντελεστών αυτών όταν φτάσουμε ή ξεπεράσουμε κάποιο τέτοιο όριο.

Η διαδικασία ξεκινά από ένα βασικό μοντέλο, το οποίο παρουσιάζει συγκεκριμένες ιδιότητες συνεκτικότητας ανάμεσα στα δομικά του στοιχεία. Πιο συγκεκριμένα, επειδή τα διάφορα στάδια της πολυδιακριτικής ανάλυσης κατασκευάζονται με την κατάτμηση μιας πλευράς σε τέσσερις, θα πρέπει η τοπολογία του βασικού μοντέλου να υποστηρίζει αυτήν τη διαδικασία – αντίστροφα, η κατασκευή του βασικού μοντέλου γίνεται από το αρχικό, χωρίς όμως να λαμβάνεται υπόψη απαραίτητα η πληροφορία της σύνδεσης των κορυφών. Έτσι, κατασκευάζεται ένα διάγραμμα Voronoi σε τρεις διαστάσεις από το οποίο εντοπίζονται οι κορυφές που θα ενωθούν στη συνέχεια με ακμές με τριγωνοποίηση Delaunay, σαν χαρακτηριστικές κορυφές γειτονικών τμημάτων (sites). Από εκεί και πέρα, το νέο μοντέλο χρησιμοποιείται για τον υπολογισμό του βασικού, με διαδοχικό διαχωρισμό σε τμήματα πληροφορίας και λεπτομέρειας. Το πιο σημαντικό πρόβλημα της ιδέας των wavelets είναι ότι δεν είναι αποτελεσματικά σε σημεία απότομων αλλαγών στην καμπυλότητα της επιφάνειας, αν και τέτοιοι αλγόριθμοι δε χρησιμοποιούνται τόσο για την απλοποιημένη αναπαράσταση ενός μοντέλου, όσο για την προοδευτική μετάδοσή του από ένα μέσο χαμηλού εύρους ζώνης ή την επεξεργασία του σε πιο χοντρικό επίπεδο.

Αλγόριθμος Απλοποίησης με Αντικατάσταση Ακμών

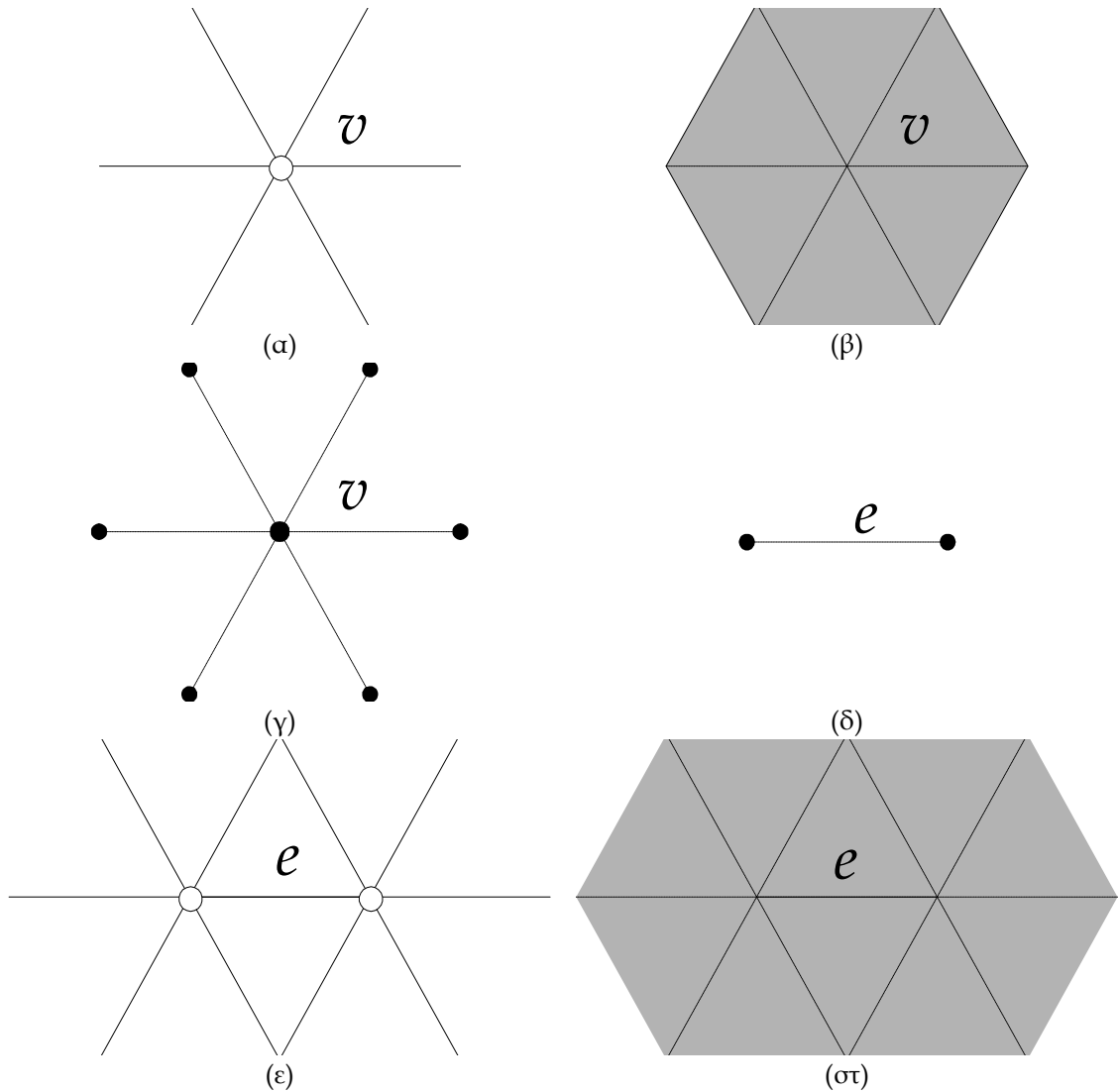
Ανάλογα με την εφαρμογή στην οποία εντάσσεται η διαδικασία απλοποίησης ενός πολυγωνικού αντικειμένου, θα πρέπει να ανταποκριθεί σε μια σειρά από συγκεκριμένες απαιτήσεις. Έτσι, σε περιπτώσεις επεξεργασίας μοντέλων με μεγάλους αριθμούς πολυγώνων, όπως για παράδειγμα αυτά που προέρχονται από τρισδιάστατους σαρωτές ή μεθόδους ανακατασκευής, θα πρέπει να γίνεται μαζική απλοποίηση της τοπολογίας, χωρίς η διατήρηση της μορφής ή του περιγράμματος να αποτελεί βασική προτεραιότητα. Αντίθετα, σε περιπτώσεις απόδοσης μοντέλων με χαμηλό αριθμό πολυγώνων σε υβριδικά περιβάλλοντα συνθετικών και πραγματικών αντικειμένων είναι προτιμότερη η πιο συντηρητική απλοποίηση, με παράλληλη διατήρηση όμως του περιγράμματος, έτσι ώστε να μείνει κατά το δυνατό αναλλοίωτη η μορφή του στον παρατηρητή. Η μέθοδος που αναπτύσσεται εδώ μπορεί να χρησιμοποιηθεί σε γενικής μορφής πολυγωνικά αντικείμενα με κλειστό περίγραμμα, δίνοντας αποτελέσματα σε σχεδόν πραγματικό χρόνο και με κύριο βάρος στη διατήρηση του περιγράμματος ανεξάρτητα από τη συγκεκριμένη όψη. Μετά τον προτεινόμενο αλγόριθμο παρατίθεται μια σειρά αποτελεσμάτων, τόσο σε απλά και σύνθετα πολυγωνικά αντικείμενα γενικού ενδιαφέροντος, όσο και σε μοντέλα ανθρώπινων κεφαλιών, τα οποία αποτελούν το πεδίο ενδιαφέροντος της διατριβής.

3.1 ΟΡΙΣΜΟΙ

Πριν προχωρήσουμε στην παρουσίαση του προτεινόμενου αλγορίθμου, ας ορίσουμε μια σειρά από έννοιες και σύμβολα που θα χρησιμοποιήσουμε στη συνέχεια. Η βασική δομική οντότητα ενός πολυγωνικού μοντέλου είναι η *κορυφή* (vertex), που εδώ θα παριστάνεται ως v , ενώ η γεωμετρική της αναπαράσταση αντιστοιχεί στο τρισδιάστατο διάνυσμα \mathbf{x} . Στο επόμενο επίπεδο, μια μη προσανατολισμένη *ακμή* \bar{e} ορίζεται σαν το σύνολο $\{v_0^e, v_1^e\} = \{[e]_0, [e]_1\}$, όπου το $[s]$ δηλώνει τις $n-1$ πλευρές ενός n -simplex s . Στην περίπτωση που χρησιμοποιούμε προσανατολισμένες πλευρές, τότε αυτές θα παριστάνονται ως διατεταγμένα ζεύγη $\vec{e} = ([e]_0, [e]_1)$. Στα επόμενα επίπεδα, τα simplices θεωρούνται διατεταγμένα, εκτός κι αν αναφέρεται διαφορετικά, οπότε θα χρησιμοποιούνται τα σύμβολα \bar{s} και \vec{s} . Έτσι, μια *πλευρά* ορίζεται σαν ένα σύνολο από προσα-

νατολισμένες πλευρές, δηλαδή $t = \{\vec{e}_0^t, \vec{e}_1^t, \vec{e}_2^t\} = \{(v_0^t, v_1^t), (v_1^t, v_2^t), (v_2^t, v_0^t)\}$. Για λόγους συντομίας, ο τελευταίος όρος μπορεί να αντικαθίσταται από την έκφραση (v_0^t, v_1^t, v_2^t) .

Ο τελεστής $\lceil s \rceil$ αντιστοιχεί στα $n+1$ simplices στα οποία ένα n -simplex είναι υποσύνολο: για παράδειγμα, το σύμβολο $\lceil v \rceil$ υποδηλώνει τις ακμές που συναντώνται στην κορυφή v . Μπορούμε να γενικεύσουμε τους παραπάνω ορισμούς σε γενικά σύνολα, με την παραδοχή ότι $\lfloor S \rfloor = \bigcup_{s \in S} \lfloor s \rfloor$. Με αυτήν τη λογική, ο τελεστής $\lfloor S \rfloor$ μειώνει τις διαστάσεις του S κατά μια, ενώ αντίστροφα ο τελεστής $\lceil S \rceil$ προσθέτει μια ακόμα διάσταση. Έτσι, μπορούμε να αναπαραστήσουμε το όριο ενός συνόλου S σαν ένα σύνολο από προσανατολισμένες πλευρές $\partial S = \{\vec{e} \in S : |\lceil e \rceil| = 1\}$. Όλες αυτές οι συμβάσεις φαίνονται στο Σχήμα 3-1.



Σχήμα 3-1: Χρήση των simplices $\lfloor s \rfloor$ και $\lceil s \rceil$. (α) $\lceil v \rceil$: οι ακμές που καταλήγουν στην κορυφή v , (β) $\lfloor v \rfloor$: οι πλευρές που περιέχουν την κορυφή v , (γ) $\lfloor v \rfloor$: οι κορυφές που είναι γειτονικές με την v , (δ) $\lfloor e \rfloor$: οι κορυφές που ορίζουν την ακμή e , (ε) $\lceil e \rceil$: οι ακμές που είναι γειτονικές με την e , (στ) $\lceil e \rceil$: οι πλευρές που έχουν κοινά σημεία με την e

Γενικά, χρησιμοποιούμε τα σύμβολα L , A , και V για να αναφερθούμε στις έννοιες του μήκους, του εμβαδού και του όγκου αντίστοιχα, ενώ όλα τα διανύσματα ορίζονται σε μορφή στήλης. Ακόμα, χρησιμοποιούμε εναλλακτικά την αναπαράσταση ενός σημείου σαν διάνυσμα, υποθέτοντας ότι το σημείο εκφράζεται σαν απόσταση από την αρχή των αξόνων ενός συστήματος αναφοράς. Μια άλλη σύμβαση αφορά τη χρήση ομογενών συντεταγμένων για τη συμπαγή αναπαράσταση αφινικών μετασχηματισμών και σχέσεων μεταξύ διανυσμάτων. Με αυτόν τον τρόπο, μπορούμε να αναπαραστήσουμε πράξεις όπως η πρόσθεση διανυσμάτων, το εσωτερικό και το εξωτερικό γινόμενο με κλειστό και ομοιόμορφο τρόπο. Για παράδειγμα,

$$(\mathbf{A}\mathbf{b})\bar{\mathbf{x}} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{pmatrix} = \mathbf{A}\mathbf{x} + \mathbf{b}$$

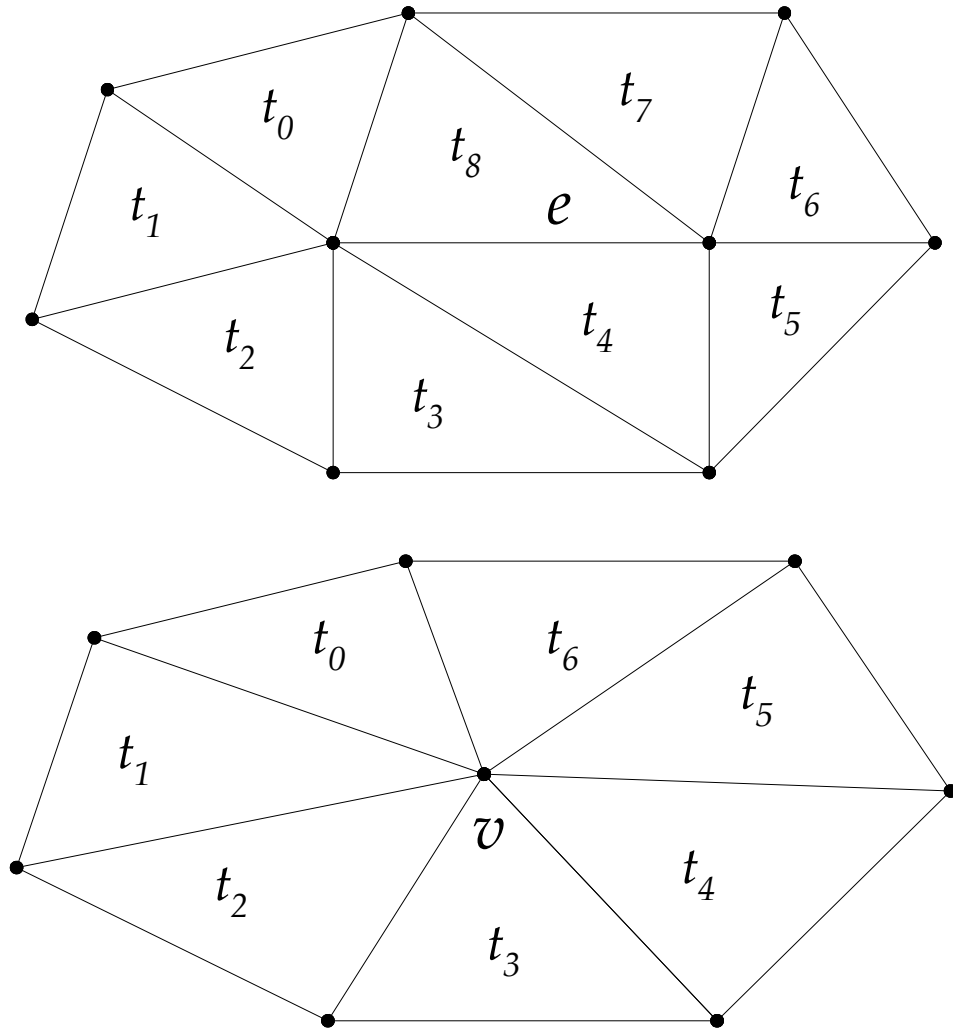
όπου η παράσταση $(\mathbf{A}\mathbf{b})$ υποδηλώνει την προσθήκη μιας ακόμα στήλης \mathbf{b} στον πίνακα \mathbf{A} διαστάσεων 3×3 .

3.2 ΑΝΤΙΚΑΤΑΣΤΑΣΗ ΑΚΜΩΝ

Η γενική ιδέα του αλγορίθμου συνίσταται στην αντικατάσταση μιας ακμής ή μιας ομάδας ακμών του πολυγωνικού μοντέλου και στον υπολογισμό μιας καινούριας συνάρτησης κόστους για τα δομικά στοιχεία που επηρεάζονται από την απλοποίηση αυτή. Η επόμενη ακμή που θα απλοποιηθεί θα είναι κάθε φορά αυτή για την οποία η συνάρτηση κόστους θα είναι η μικρότερη. Πιο συγκεκριμένα, κάθε εφαρμογή του αλγορίθμου απλοποίησης εφαρμόζεται πάνω σε μια ακμή $e = \{v_0, v_1\}$ και την αντικαθιστά με μια νέα κορυφή v . Σαν αποτέλεσμα, οι πλευρές $[e]$ εκφυλίζονται σε ακμές και αφαιρούνται από το μοντέλο, ενώ οι υπόλοιπες γειτονικές ακμές και πλευρές αναδιοργανώνονται ώστε όλες οι αναφορές των κορυφών v_0 και v_1 να αντικατασταθούν από τη νέα κορυφή v (Σχήμα 3-2).

Γενικά, το πρώτο βήμα σε αυτήν τη διαδικασία είναι ο υπολογισμός της συνάρτησης κόστους για καθεμιά από τις ακμές του μοντέλου μας. Σε καθένα από τα διαδοχικά βήματα, η πρώτη από τις διαθέσιμες ακμές ελέγχεται για την εγκυρότητα του αποτελέσματος που θα δώσει η απλοποίησή της και ενδέχεται να μη συμμετάσχει στη διαδικασία στο συγκεκριμένο βήμα, αν δεν ικανοποιεί κάποια από τις δεδομένες απαιτήσεις. Αυτός ο έλεγχος είναι αναγκαίος μια που η "τυφλή" εφαρμογή της διαδικασίας μπορεί να οδηγήσει σε ενδιάμεσα μοντέλα με εκφυλισμένες πλευρές, οπές ή μοντέλα που να είναι γενικά προβληματικά στην απόδοση με τους συνηθισμένους αλγορίθμους. Στην περίπτωση μας αναφερόμαστε μόνο σε κλειστά πολυγωνικά μοντέλα, οπότε δεν έχει ιδιαίτερο νόημα να ελέγξουμε αν μια πλευρά ορίζει μια οπή στην επιφάνεια του μοντέλου. Πάντως, ακόμα κι αν μια ακμή θεωρηθεί άκυρη για τη διαδικασία σε κάποια φάση της, ο υπολογισμός της συνάρτησης κόστους θα γίνει κανονικά στο επόμενο βήμα του αλγορίθμου και είναι πιθανό η συγκεκριμένη ακμή να απλοποιηθεί αργότερα. Αυτό μπορεί να συμβεί μια που οι λόγοι που την έκαναν άκυρη στο συγκεκριμένο σημείο της διαδικασίας απλοποίησης μπορεί να πάψουν να υφίστανται λόγω της αναδιοργάνωσης των δομικών στοιχείων στη γειτονιά της ακμής. Γενικά, η ακμή e θα είναι υποψήφια για απλοποίηση αν ικανοποιούνται οι συνθήκες [2] :

- οι κορυφές v_i του συνόλου $\llbracket e \rrbracket$ που είναι γειτονικές σχηματίζουν μεταξύ τους πλευρά
- η ακμή (v_i, v_j) ανήκει στο όριο της επιφάνειας αν και μόνο αν οι κορυφές v_i και v_j ανήκουν στο όριο και
- η περιοχή $\llbracket e \rrbracket$ περιέχει πάνω από τέσσερις κορυφές αν καμιά από τις κορυφές $\lceil e \rceil$ δεν ανήκει στο όριο της επιφάνειας ή πάνω από τρεις κορυφές αν κάποια από τις κορυφές $\lceil e \rceil$ ανήκει στο όριο.



Σχήμα 3-2: Διαδικασία απόρριψης της ακμής e και δημιουργία της κορυφής v . Οι πλευρές t_7 και t_8 εκφυλίζονται και αφαιρούνται από το μοντέλο

Η συνάρτηση κόστους εξαρτάται από τα τρίγωνα που περιέχουν την υπό εξέταση ακμή, αλλά και τα simplices κατώτερης τάξης, δηλαδή τις κορυφές που την ορίζουν. Ακολουθώντας τη σύμβαση που περιγράψαμε πιο πάνω, οι πλευρές αυτές καλύπτονται από το σύμβολο $\llbracket e \rrbracket$. Επειδή δεν έχουμε κάποια πρότερη γνώση της τοπολογίας του μοντέλου μας σε κάθε συγκεκριμένο βήμα της διαδικασίας απλοποίησης, θεωρούμε ότι είναι πιθανό να επηρεαστούν όλες οι πλευρές του μοντέλου μας. Αυτό σημαίνει ότι για όλες οι ακμές $\{e_j\}$ για τις οποίες ισχύει η σχέση $\llbracket e_j \rrbracket \cap \llbracket e \rrbracket \neq \emptyset$ θα πρέπει να υπολογιστεί εκ νέου η συνάρτηση κόστους, πριν προχωρήσουμε στο επόμενο βήμα.

Η συνολική διαδικασία μπορεί να χωριστεί σε δύο τμήματα, τη επιλογή της μετρικής για τον υπολογισμό της συνάρτησης κόστους για την απλοποίηση κάθε ακμής e και την επιλογή της θέσης \mathbf{x} στην οποία θα πρέπει να τοποθετηθεί η καινούρια κορυφή v που αντικαθιστά την "παλιά" ακμή. Η θέση αυτή μπορεί να καθοριστεί ανάλογα με τις απαιτήσεις που τίθενται στον αλγόριθμο σε ό,τι αφορά τη διατήρηση κάποιων γεωμετρικών χαρακτηριστικών ή ακόμα και τη σαφή διατήρηση κάποιων από τις κορυφές σε συγκεκριμένα σημεία. Η τελευταία περίπτωση είναι περισσότερο συνηθισμένη όταν κάποιες από τις κορυφές έχουν τοποθετηθεί σε θέσεις οι οποίες αντιστοιχούν σε πραγματικά δεδομένα, για παράδειγμα από την ανάλυση μιας εικόνας ή μιας ακολουθίας εικόνων, και είναι σημαντική η διατήρησή τους στη συγκεκριμένη θέση. Βλέποντας τη διαδικασία από τη σκοπιά της διατήρησης της μορφής του μοντέλου, όσο αυτό είναι δυνατό, μπορούμε να ορίσουμε μια συνάρτηση $f_c(\mathbf{x})$ που αντιστοιχεί στο κόστος της αντικατάστασης της ακμής e με την κορυφή \mathbf{x} και να ζητήσουμε την ελαχιστοποίησή της, λαμβάνοντας υπόψη παράγοντες όπως η μεταβολή στον όγκο και την επιφάνεια της περιοχής που επηρεάζεται.

Η βασική μας προτεραιότητα για τον υπολογισμό του \mathbf{x} αφορά στο συνδυασμό τριών ή περισσότερων περιορισμών που εκφράζονται από τις γραμμικές εξισώσεις $\hat{\mathbf{a}}_i^T \mathbf{x} = \hat{b}_i$. Αυτό πρακτικά σημαίνει ότι το σημείο \mathbf{x} αντιστοιχεί στην τομή μη παράλληλων επιπέδων στο χώρο \mathbb{R}^3 . Εκτός από αυτό, μπορούμε να ζητήσουμε την ικανοποίηση και κάποιων πρόσθετων συνθηκών, για να καλύψουμε την πιθανότητα δύο ή περισσότεροι περιορισμοί να είναι γραμμικά εξαρτημένοι. Στην περίπτωση που δύο ή περισσότερα από αυτά τα επίπεδα είναι σχεδόν παράλληλα, μικρές μεταβολές στις τιμές των παραμέτρων που τα ορίζουν θα έχουν σαν αποτέλεσμα μεγάλες μεταβολές στο τελικό αποτέλεσμα. Επειδή τέτοιες καταστάσεις είναι συχνές, για παράδειγμα σαν αποτέλεσμα περιορισμένης αριθμητικής ακρίβειας ή στρογγυλοποιήσεων, προσθέτουμε έναν ακόμα περιορισμό (\mathbf{a}_n, b_n) , ο οποίος θα ισχύσει μόνο στην περίπτωση που το μοναδιαίο διάνυσμα του επιπέδου \mathbf{a}_n δε βρίσκεται μέσα σε μια περιοχή γωνίας α σε σχέση με το γραμμικό συνδυασμό των μοναδιαίων διανυσμάτων των επιπέδων που αντιστοιχούν στους προηγούμενους συνδυασμούς. Αυτό σημαίνει ότι με δεδομένους $n-1$ συνδυασμούς, δεχόμαστε το (\mathbf{a}_n, b_n) αν:

- $n = 1 : \hat{\mathbf{a}}_1 \neq 0$ 3-1

- $n = 2 : (\hat{\mathbf{a}}_1^T \hat{\mathbf{a}}_2)^2 < \hat{\mathbf{a}}_1^T \hat{\mathbf{a}}_1 \hat{\mathbf{a}}_2^T \hat{\mathbf{a}}_2 \cos^2(\alpha)$ 3-2

- $n = 3 : [\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \hat{\mathbf{a}}_3]^2 > (\hat{\mathbf{a}}_1 \times \hat{\mathbf{a}}_2)^T (\hat{\mathbf{a}}_1 \times \hat{\mathbf{a}}_2) \hat{\mathbf{a}}_3^T \hat{\mathbf{a}}_3 \sin^2(\alpha)$ 3-3

όπου η παράσταση $[\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \hat{\mathbf{a}}_3]$ αντιστοιχεί στο τριπλό διανυσματικό γινόμενο. Αν ο περιορισμός που εξετάζουμε ικανοποιεί αυτές τις σχέσεις, τότε είναι συμβατός με το σύνολό τους και προστίθεται σε αυτό. Όταν οι περιορισμοί που ικανοποιούνται φτάσουν τους τρεις, τότε πλέον το \mathbf{x} υπολογίζεται από τη σχέση $\mathbf{x} = \mathbf{A}^{-1} \mathbf{b}$.

Μια σειρά από περιορισμούς προέρχεται από την προσπάθεια ελαχιστοποίησης μιας τετραγωνικής συνάρτησης που εξαρτάται από ένα σύνολο γραμμικών μεταξύ τους περιορισμών. Στην περίπτωσή μας, η συνάρτηση που πρέπει να ελαχιστοποιηθεί μπορεί να γραφεί με τη μορφή:

$$f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{b}^T \mathbf{x} + \frac{1}{2}$$

$$\begin{aligned}
&= \frac{1}{2} \begin{pmatrix} \mathbf{x}^T & 1 \end{pmatrix} \begin{pmatrix} \mathbf{A} & -\mathbf{b} \\ -\mathbf{b}^T & c \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix} \\
&= \frac{1}{2} \bar{\mathbf{x}}^T \bar{\mathbf{A}} \bar{\mathbf{x}}
\end{aligned} \tag{3-4}$$

Σε αυτήν τη μορφή, το \mathbf{A} εκφράζει τη συμμετρική καθορισμένη Hessian της f και το $\bar{\mathbf{A}}$ ένα θετικό συμμετρικό πίνακα διαστάσεων 4×4 . Ο στόχος μας είναι να ελαχιστοποιήσουμε τη συνάρτηση f σε σχέση με το σύνολο των περιορισμών $\hat{\mathbf{A}}\mathbf{x} = \hat{\mathbf{b}}$. Στο γραμμικό αυτό πρόβλημα, ως υποθέσουμε ότι με δεδομένους n περιορισμούς, \mathbf{Q} είναι ένας πίνακας διαστάσεων $(3-n) \times 3$ και του οποίου οι γραμμές είναι ορθογώνιες μεταξύ τους και ως προς τα διανύσματα $\hat{\mathbf{a}}_i$. Τότε, για τους υπόλοιπους $3 - n$ περιορισμούς ισχύει η σχέση:

$$\mathbf{Q}(\mathbf{Ax}-\mathbf{b})=0 \tag{3-5}$$

όπου $\mathbf{Ax}-\mathbf{b} = \nabla f$. Πρακτικά, αυτό σημαίνει ότι το ελάχιστο της συνάρτησης f , υπό τους δεδομένους περιορισμούς βρίσκεται στο σημείο όπου εξαφανίζεται η προβολή της βαθμίδας της στο χώρο που καλύπτει ο πίνακας \mathbf{Q} . Οι υπόλοιποι $3-n$ γραμμικοί μεταξύ τους περιορισμοί που προκύπτουν από τη σχέση 3-5 προστίθενται με την προϋπόθεση ότι ικανοποιούν τους κανόνες συμβατότητας.

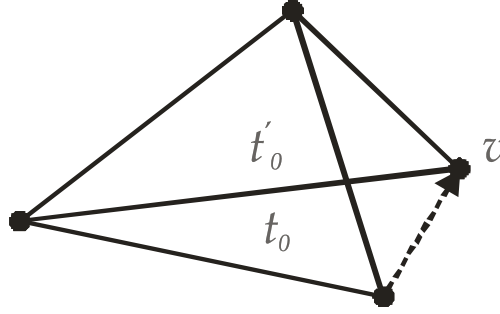
Ένας πολύ σημαντικός περιορισμός που αφορά την τοποθέτηση της νέας κορυφής v αφορά τη διατήρηση του συνολικού όγκου του μοντέλου. Μετά από κάθε αντικατάσταση μιας ακμής, η γειτονική της περιοχή αναδιοργανώνεται, με αποτέλεσμα την αλλαγή της τοπολογίας και του όγκου. Αν η καινούρια θέση της κορυφής v είναι, για παράδειγμα, στο μέσο της "παλιάς" ακμής e , τότε μπορούμε να δείξουμε ότι το μοντέλο θα παρουσιάσει μια τοπική συστολή για ακμές οι οποίες αντιστοιχούν σε οξείες στερεές γωνίες, ενώ αντίθετα θα διασταλεί αν η e αντιστοιχεί σε αμβλεία στερεή γωνία. Επειδή γενικά τα πολυγωνικά μοντέλα δεν είναι συμπαγή, παρά μόνο χρησιμεύουν για να ορίσουν ένα υποσύνολο του χώρου, είναι επιθυμητό να διατηρήσουμε τόσο την εμφάνιση του μοντέλου σε τρεις διαστάσεις, όσο και την προβολή του στο επίπεδο.

Αν και η γενική ιδέα του αλγορίθμου απλοποίησης μπορεί να εφαρμοστεί τόσο σε εσωτερικές ακμές, όσο και σε ακμές οι οποίες ανήκουν σε όρια μιας οπής στην επιφάνεια, εδώ θα αρκεστούμε στην εφαρμογή σε κλειστά πολυγωνικά μοντέλα. Αυτό σημαίνει ότι τα δομικά στοιχεία που ανήκουν στην περιοχή $\llbracket e \rrbracket$ είναι εσωτερικά, και μάλιστα κάθε ακμή ανήκει σε δύο και μόνο πλευρές.

Ας παρατηρήσουμε την τοπική αλλαγή στην τοπολογία του μοντέλου κατά την απλοποίηση της ακμής e : αν η πλευρά $t = (v_i^e, v_1^t, v_2^t)$ αντικατασταθεί από την πλευρά $t' = (v, v_1^t, v_2^t)$, τότε η γραμμική κίνηση της κορυφής v_i^e προς τη "νέα" κορυφή v ορίζει έναν όγκο, σε συνδυασμό με την πλευρά t (Σχήμα 3-3). Γεωμετρικά, αυτός ο όγκος μπορεί να περιγραφεί από το τετράεδρο (v, v_i^e, v_1^t, v_2^t) - αν το τετράεδρο αυτό βρίσκεται πάνω από το επίπεδο της πλευράς t , τότε ο όγκος του θεωρείται θετικός και το μοντέλο διαστέλλεται τοπικά. Για να μπορέσουμε λοιπόν να διατηρήσουμε αυτόν τον όγκο σταθερό, θέτουμε:

$$\sum_i V(v, v_0^{t_i}, v_1^{t_i}, v_2^{t_i}) = \sum_i \frac{1}{6} \begin{vmatrix} v_x & v_{0x}^{t_i} & v_{1x}^{t_i} & v_{2x}^{t_i} \\ v_y & v_{0y}^{t_i} & v_{1y}^{t_i} & v_{2y}^{t_i} \\ v_z & v_{0z}^{t_i} & v_{1z}^{t_i} & v_{2z}^{t_i} \\ 1 & 1 & 1 & 1 \end{vmatrix} = 0 \quad 3-6$$

και λύνουμε ως προς v . Η εξίσωση αυτή περιορίζει τον υπολογισμό του \mathbf{x} σε ένα επίπεδο. Αυτό σημαίνει ότι μπορούμε πλέον να προσθέσουμε τη σχέση αυτή στο σύνολο των περιορισμών, με την προϋπόθεση ότι δε δίνει σαν αποτέλεσμα κάποια εκφυλισμένη πλευρά.



Σχήμα 3-3: Το ιδεατό τετράεδρο που σχηματίζεται από την κίνηση της κορυφής v

Επειδή η σχέση 3-6 συνοψολογίζει όλες τις πλευρές του μοντέλου οι οποίες εμπλέκονται στην τοπική αναδιοργάνωσή του και εξασφαλίζει τη διατήρηση του όγκου σε τοπικό επίπεδο, μπορούμε να πούμε ότι διατηρείται ακόμα και ο συνολικός όγκος του μοντέλου, μια που οι υπόλοιπες πλευρές του δεν επηρεάζονται σε αυτό το βήμα. Εκτός από αυτό, είναι επιθυμητό να ελαχιστοποιήσουμε τον όγκο του καθενός από τα στοιχειώδη τετράεδρα και σε ό,τι αφορά την απόλυτη τιμή του, έτσι ώστε να διατηρηθεί, όσο είναι δυνατό η εμφάνιση του μοντέλου σε τοπικό επίπεδο. Για το λόγο αυτό, θα πρέπει να ελαχιστοποιήσουμε τη σχέση

$$f_{V_0}(\mathbf{x}) = \frac{1}{2} \bar{\mathbf{x}}^T \bar{\mathbf{A}}_{V_0} \bar{\mathbf{x}} \quad 3-7$$

Η σχέση αυτή, όπως και η σχέση 3-5 είναι παρόμοιες με αυτές που αναφέρονται στην εργασία [1], με τη διαφορά ότι εκεί οι τετραγωνικές εξισώσεις διατηρούνται κατά την απλοποίηση, ενώ στους αλγορίθμους απλοποίησης χωρίς μνήμη οι συντελεστές αυτοί υπολογίζονται σε κάθε καινούριο βήμα. Γενικά, αν οι κορυφές $\llbracket e \rrbracket$ είναι συνεπίπεδες, τότε η σχέση 3-7 θα δώσει άπειρες λύσεις, μια που ο όγκος του τετραέδρου είναι μηδενικός. Ακόμα, αν οι κορυφές αυτές είναι σχεδόν συνεπίπεδες, τότε η εξίσωση μπορεί να δώσει αποτελέσματα που να μην είναι συμβατά με τους προϋπάρχοντες περιορισμούς – ακόμα και σε αυτήν την περίπτωση πάντως, το σφάλμα στον καθορισμό του νέου όγκου είναι απειροελάχιστο μπροστά στο όφελος από την απλοποίηση, ειδικά αν αυτή συνδυαστεί και με τη βελτιστοποίηση της νέας τοπολογίας.

Ένα τελευταίο βήμα που έχει περισσότερο σχέση με τη βελτιστοποίηση αυτή, αφορά το σχήμα και τις αναλογίες που θα έχουν οι πλευρές που θα προκύψουν από την αναδιάταξη της γειτονιάς της ακμής που απλοποιείται. Σε περιπτώσεις όπου οι κορυφές $\llbracket e \rrbracket$ είναι συνεπίπεδες ή οι κορυφές $\llbracket \partial \llbracket e \rrbracket \rrbracket$ είναι συγγραμμικές, τότε οι περιορισμοί που έχουμε θέσει δεν είναι όλοι συμβατοί ή δε μας δίνουν μοναδική λύση. Μπορούμε να εκμεταλλευτούμε μια τέτοια περίπτωση για να βελτιστοποιήσουμε τη μορφή των νέων πλευρών: γενικά, οι αλγόριθμοι απόδοσης έχουν καλύτερα και πιο προβλέψιμα αποτελέσματα όταν εφαρμόζονται πάνω σε πλευρές με

σχεδόν ίσες ακμές, παρά σε πλευρές που είναι εξαιρετικά μακρόστενες. Έτσι, ο αντικειμενικός σκοπός είναι να ελαχιστοποιήσουμε την έκφραση $f_s(\mathbf{x}) = \sum_i L(v, v_i)^2$, δηλαδή το άθροισμα των μηκών των ακμών $[v]$ που συμπίπτουν στην κορυφή v , υψωμένων στο τετράγωνο. Έτσι τελικά, ο λόγος της επιφάνειας προς την περίμετρο για καθεμιά από τις πλευρές μεγιστοποιείται. Επιλύοντας την παραπάνω σχέση, μπορούμε να δείξουμε ότι το μέγεθος f_s αυξάνει με την αύξηση της απόστασης της κορυφής v από το κέντρο του τετραέδρου v_i .

Αν και όλοι οι αλγόριθμοι απλοποίησης που παρουσιάζονται χρησιμοποιούν κάποια λογική αντικατάστασης ακμών, η διαφορά της οικογένειας των αλγορίθμων απλοποίησης χωρίς μνήμη έγκειται στο ότι η προτεραιότητα της κάθε ακμής (δηλαδή η συνάρτηση κόστους) για την απλοποίησή της στο επόμενο βήμα υπολογίζεται σε σχέση με το αμέσως προηγούμενο, και όχι σε σχέση με το αρχικό πολυγωνικό μοντέλο. Μπορούμε να γράψουμε τη συνάρτηση κόστους για κάθε ακμή σαν:

$$f_c(\mathbf{x}) = \lambda f_{v_0}(\mathbf{x}) + (1 - \lambda) L(e)^2 f_{B_0}(\mathbf{x})$$

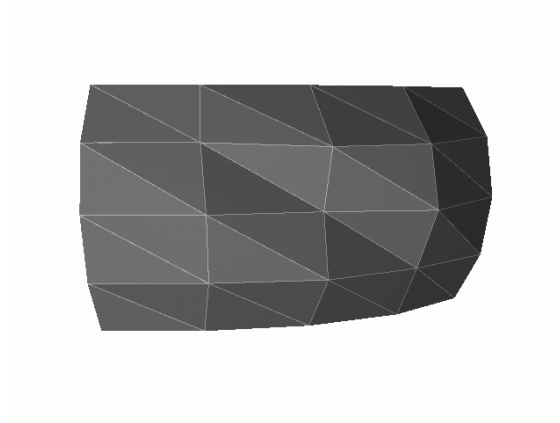
όπου ο όρος $f_{B_0}(\mathbf{x})$ μπορεί να χρησιμοποιηθεί για τον υπολογισμό των αλλαγών στο όριο μιας επιφάνειας. Επειδή στην περίπτωση μας θεωρούμε όλα τα αντικείμενα κλειστά, ο όρος αυτό μηδενίζεται κι έτσι η παραπάνω σχέση γίνεται

$$f_c(\mathbf{x}) = f_{v_0}(\mathbf{x}) \quad 3-8$$

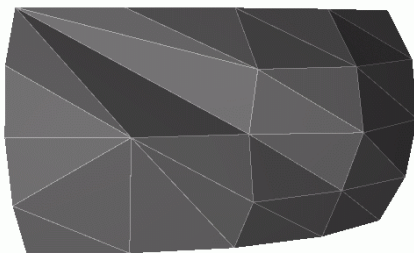
3.3 ΑΠΟΤΕΛΕΣΜΑΤΑ

Ο αλγόριθμος που περιγράφηκε παραπάνω εφαρμόζεται στη συνέχεια σε μια σειρά από πολυγωνικά αντικείμενα. Το πρώτο από αυτά τα αντικείμενα είναι ένα πολυγωνικό αντικείμενο τύπου patch, το οποίο έχει υποστεί μικρή παραμόρφωση έτσι ώστε η επιφάνειά του να παρουσιάσει καμπυλότητα. Η αρχική διάταξη των δομικών στοιχείων του patch αποτελείται από πέντε σειρές και στήλες κορυφών, ενωμένες από γραμμές και στήλες ακμών που διαμορφώνουν όμοιες τριγωνικές πλευρές. Αυτή η ελάχιστη μορφή δεν προσφέρεται ιδιαίτερα για απλοποίηση, αφού δεν υπάρχουν δομικά στοιχεία που να αποτελούν "πλεονασμό" στην περιγραφή της επιφάνειας. Αυτό σημαίνει ότι η αφαίρεση δομικών στοιχείων έχει σαν αποτέλεσμα την εισαγωγή αριθμητικού σφάλματος στην επιφάνεια – παρόλα αυτά, η γενική μορφή και ειδικά το προφίλ της επιφάνειας διατηρείται, ειδικά αν λάβουμε υπόψη μας ότι οι συνηθισμένοι αλγόριθμοι απόδοσης εξομαλύνουν τις ατέλειες στην επιφάνεια, χωρίς να χρειαστεί να διατεθούν πρόσθετα πολύγωνα για το σκοπό αυτό.

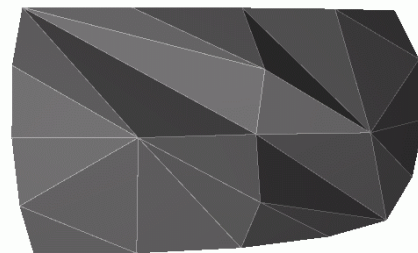
Η αρχική επιφάνεια, η οποία φαίνεται στο Σχήμα 3-4α, αποτελείται από κορυφές και πλευρές, οι οποίες προκύπτουν από την κανονική διάταξη των δομικών στοιχείων της. Στα επόμενα στιγμιότυπα, διατηρείται η μορφή και η διάταξη στα όρια της επιφάνειας, λόγω του περιορισμού στη διαχείριση κλειστών επιφανειών. Αυτός ο περιορισμός είναι πολύ σημαντικός στη διατήρηση του προφίλ του αντικειμένου που απλοποιείται, μια που έτσι διατηρείται η γενική του μορφή κατά την απόδοση.



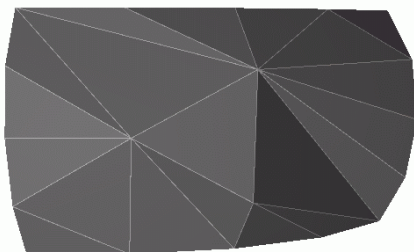
(α)



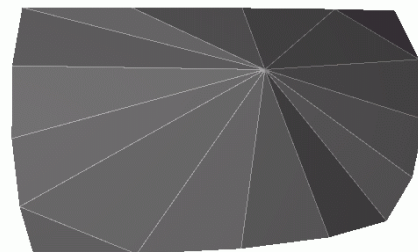
(β)



(γ)



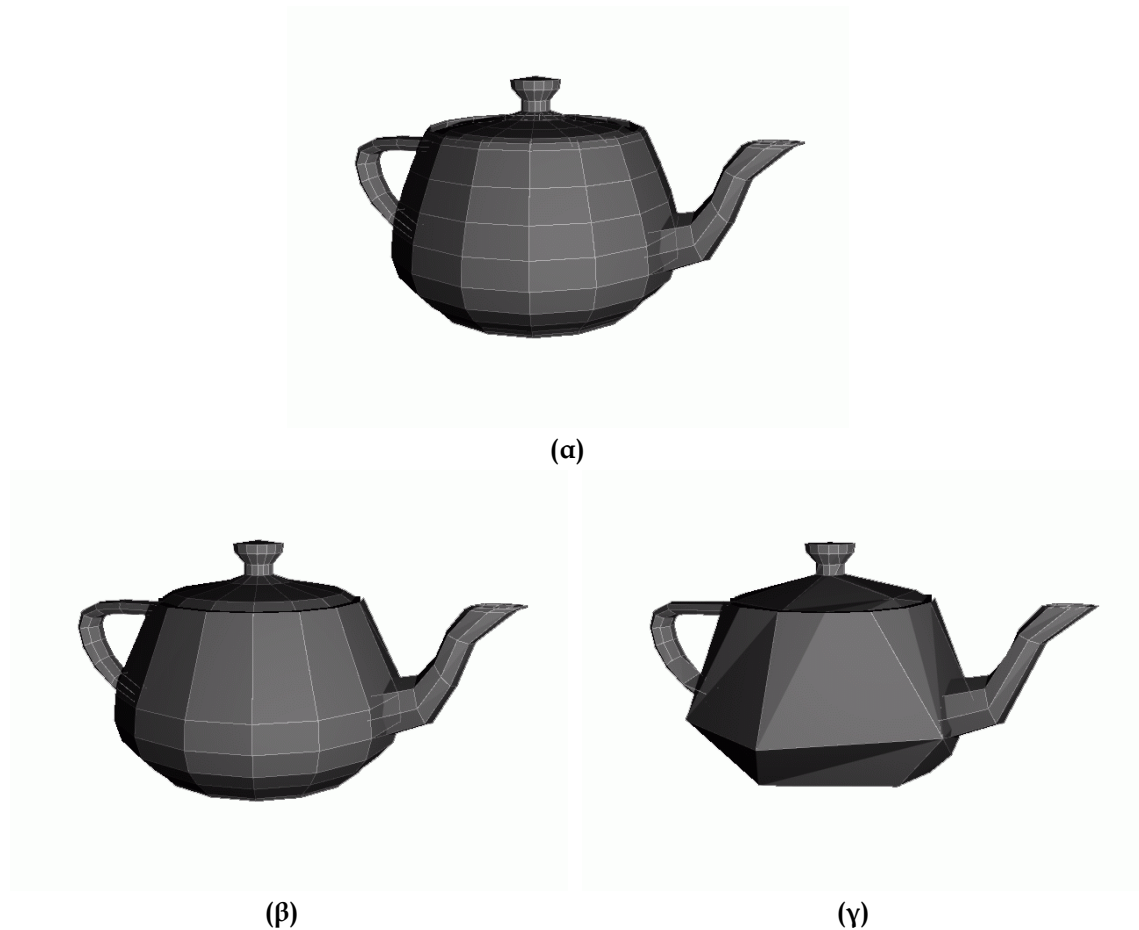
(δ)



(ε)

Σχήμα 3-4: Απλοποίηση ενός πολυγωνικού patch. (α): Το πρωτότυπο μοντέλο, (β) – (ε): διαδοχικά στάδια απλοποίησης

Το επόμενο αντικείμενο που απλοποιείται είναι η κλασική τσαγιέρα του Πανεπιστημίου της Utah, μετά τη μετατροπή της σε πολυγωνικό αντικείμενο σε δύο επίπεδα τριγωνοποίησης των splines που την ορίζουν. Το μοντέλο της τσαγιέρας χρησιμοποιείται κυρίως για την επίδειξη αλγορίθμων σκίασης, λόγω της ιδιόμορφης γεωμετρίας που το χαρακτηρίζει, αλλά και λόγω της ποικιλίας στις μορφές σύνδεσης των επιμέρους τμημάτων. Στο μοντέλο που χρησιμοποιείται εδώ, τα επιμέρους τμήματα της τσαγιέρας, όπως το χερούλι ή το καπάκι, έχουν απλοποιηθεί σαν ξεχωριστά αντικείμενα.



Σχήμα 3-5: Απλοποίηση ενός μοντέλου τσαγιέρας. (α): Το πρωτότυπο μοντέλο, (β) – (γ): διαδοχικά στάδια απλοποίησης

Το τελευταίο αντικείμενο που απλοποιήθηκε ήταν ένα πολυγωνικό μοντέλο ανθρώπινου κεφαλιού, συμβατό με το πρότυπο MPEG-4 και τις απαιτήσεις του για τα σχετικά συνθετικά αντικείμενα [12]. Η διαδικασία απλοποίησης μπορεί να βοηθήσει σημαντικά σε εφαρμογές υβριδικής κωδικοποίησης συνθετικών και πραγματικών αντικειμένων [13], αφού έτσι μειώνεται κατά πολύ το εύρος ζώνης που απαιτείται για τη μετάδοσή τους, αλλά και οι υπολογιστικές απαιτήσεις για την απόδοσή τους σε πραγματικό χρόνο. Το αρχικό μοντέλο, μαζί με την εικόνα απεικόνισης υψής, φαίνεται σε εμπρόσθια όψη στο Σχήμα 3-6α, και αποτελείται από 2141 πλευρές, οι οποίες απλοποιούνται προοδευτικά σε 1929 (Σχήμα 3-6β), 1307 πλευρές (Σχήμα 3-6δ), και τέλος 680 πλευρές (Σχήμα 3-6γ). Ακόμα και στην τελική μορφή, το μοντέλο δεν παρουσιάζει ιδιάιτερη παραμόρφωση, ειδικά στα σημεία που ενσωματώνουν τη λεπτομέρεια της έκφρασης, δηλαδή τις περιοχές γύρω από τα μάτια και το στόμα [40].



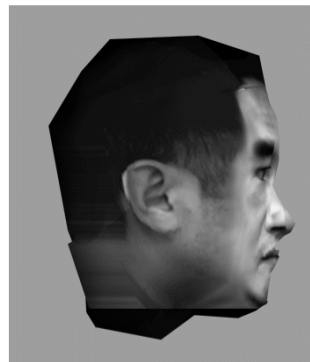
(α)



(β)



(γ)



(δ)

Σχήμα 3-6: Απλοποίηση πολυγωνικού μοντέλου κεφαλιού. (α): Το πρωτότυπο μοντέλο, (β) – (δ): διαδοχικά στάδια απλοποίησης

Αναπαράσταση της Βάδισης

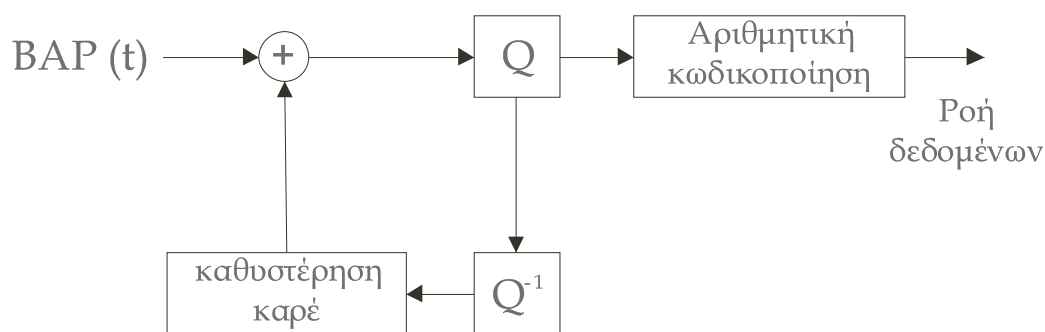
Μια από τις πιο σημαντικές καινοτομίες του προτύπου MPEG-4 αφορά την υβριδική κωδικοποίηση πραγματικών και συνθετικών αντικειμένων. Ειδικά σε ό,τι αφορά τα συνθετικά μοντέλα ανθρώπινων προσώπων και σωμάτων, το πρότυπο υποστηρίζει την παραμετρική αναπαράσταση και επεξεργασία τους με τρόπο πιο πρόσφορο για τη σημασιολογική τους θέση σε μια σκηνή. Έτσι, δίνεται η δυνατότητα στο σχεδιαστή του υλικού να επικεντρωθεί στην ποιότητα του περιεχομένου, χωρίς να χρειαστεί να ασχοληθεί ιδιαίτερα με εργασίες χαμηλού επιπέδου στα πολυγωνικά αντικείμενα της σκηνής του. Ακόμα, η χρήση τέτοιων παραμέτρων μειώνει κατά πολύ το απαιτούμενο εύρος ζώνης για τη μετάδοση της συνθετικής εικόνας, χωρίς αυτό να έχει απαραίτητα αντίκτυπο στην ποιότητα του αποδιδόμενου αποτελέσματος. Αντίστροφα, επειδή οι παραμετρικές αναπαραστάσεις είναι πιο κοντά στην οργάνωση σκελετικών αντικειμένων, από ότι η επεξεργασία σε επίπεδο πολυγώνων, διευκολύνεται η αναπαραγωγή στον αποκωδικοποιητή και μάλιστα με εύκολη προσαρμογή στο περιβάλλον και τις απαιτήσεις της εφαρμογής.

4.1 ΠΕΡΙΓΡΑΦΗ ΤΩΝ ΚΩΔΙΚΟΠΟΙΗΜΕΝΩΝ ΔΕΔΟΜΕΝΩΝ

Τα αντικείμενα τύπου Body, δηλαδή τα παραμετρικά μοντέλα που μπορούν να χρησιμοποιηθούν σε μια συνθετική σκηνή για να αναπαραστήσουν το ρόλο ενός ανθρώπινου ή ανθρωπόμορφου σώματος, μπορούν να θεωρηθούν σαν ένα οργανωμένο σύνολο από κόμβους στο γράφο της σκηνής, οι οποίοι εμψυχώνονται από παραμέτρους στη σχετική ροή δεδομένων. Η γεωμετρική αναπαράσταση και η θέση του σώματος ελέγχονται από δύο ξεχωριστές ροές δεδομένων, η πρώτη από τις οποίες περιέχει στιγμιότυπα από παραμέτρους ορισμού σώματος (Body Definition Parameters – BDP) και η δεύτερη τα δεδομένα που χρησιμοποιούνται για την εμψύχωση, δηλαδή τις παραμέτρους εμψύχωσης σώματος (Body Animation Parameters – BAP) – σε ανώτερο επίπεδο, η πρώτη ροή αντιστοιχεί στη γλώσσα BIFS [18], ενώ η δεύτερη αναφέρεται ως Face and Body Animation (FBA). Ο ρόλος των παραμέτρων αυτών είναι τέτοιος ώστε αν οι παράμετροι BAP ερμηνευτούν σωστά από τον αποκωδικοποιητή, το αισθητικό και λειτουργικό αποτέλεσμα θα είναι αρκετά υψηλής ποιότητας σε ό,τι αφορά τη στάση και την εμψύχωση διαφορετικών σκελετικών μοντέλων, ενώ οι παράμετροι BDP μπορούν να χρησιμοποιηθούν για να προσαρμόσουν ένα γενικής μορφής μοντέλο σε συγκεκριμένες διαστάσεις, αναλογίες και υφή.

Στην αρχική του μορφή, το αντικείμενο τύπου Body περιέχει ένα γενικό μοντέλο ανθρώπινου σώματος, στη "φυσιολογική" του στάση. Αυτό σημαίνει ότι τα πόδια είναι προσανατολισμένα προς τα εμπρός, οι βραχίονες είναι παράλληλα με τον κορμό του σώματος και οι παλάμες είναι προσανατολισμένες προς τον κορμό. Οι προκαθορισμένες τιμές των BAPs, οι οποίες καθορίζουν αυτήν την προκαθορισμένη στάση, μπορούν να μεταβληθούν άμεσα με βάση τη σχετική ροή δεδομένων, αν και αυτό δεν είναι απαραίτητο για να ξεκινήσει η απόδοση του μοντέλου από τον αποκωδικοποιητή. Αν και το ανθρώπινο σώμα επιβάλλει μια σειρά από περιορισμούς στην κίνηση και τον προσανατολισμό των τμημάτων του σκελετικού μοντέλου, το πρότυπο MPEG-4 δεν τους ακολουθεί απαραίτητα, κάνοντας με αυτόν τον τρόπο δυνατή την απόδοση υπερβολικών κινήσεων και στάσεων που μπορεί να είναι χρήσιμες σε εφαρμογές ψυχαγωγίας.

Σε επίπεδο ροής δεδομένων, ένα αντικείμενο τύπου body σχηματίζεται από μια χρονική ακολουθία από επίπεδα αντικειμένων body (Body Object Planes – BOP), με καθένα από αυτά τα επίπεδα να αποτελείται από μια ροή δεδομένων BAP. Καθένα από τα αντικείμενα που περιγράφονται με αυτόν τον τρόπο, αποτελεί έναν κόμβο στο γράφο της σκηνης, με την πιθανή προσθήκη ενδιάμεσων χρονικών και χωρικών σχέσεων με άλλους τέτοιους κόμβους. Γενικά, τα δεδομένα BAP αντιστοιχούν στις γωνίες των συνδέσμων ανάμεσα σε διαδοχικά μέλη του σκελετικού μοντέλου, όπως οι φάλαγγες στα δάκτυλα των χεριών, τον καρπό και τον πήχη. Οι θετικές τιμές σε αυτές τις παραμέτρους αντιστοιχούν σε περιστροφές με φορά αντίστροφη από αυτή των δεικτών του ρολογιού, ενώ στην προκαθορισμένη στάση οι παράμετροι θεωρούνται ότι έχουν μηδενική τιμή. Οι περιστροφές που περιγράφονται στη ροή δεδομένων των BAPs ολοκληρώνονται γύρω από άξονες που κινούνται μαζί με το αντικείμενο τύπου Body και οι οποίοι περιστρέφονται παράλληλα με το μετασχηματισμό του μέλους στο οποίο είναι συνδεδεμένοι – αυτό σημαίνει ότι οι άξονες των τοπικών συστημάτων συντεταγμένων κάθε μέλους δεν είναι γενικά παράλληλοι με τους άξονες του συνολικού συστήματος αναφοράς του αντικειμένου ή με αυτούς του παγκόσμιου συστήματος αναφοράς.



Σχήμα 4-1: Κωδικοποίηση των BAPs με πρόβλεψη

Σε ότι αφορά την κωδικοποίηση των τιμών των παραμέτρων BAP, το πρότυπο MPEG-4 χρησιμοποιεί κι εδώ ένα σχήμα αριθμητικής κωδικοποίησης με πρόβλεψη. Αυτό σημαίνει ότι είναι απαραίτητη η τιμή της παραμέτρου στο προηγούμενο BOP και ο συνδυασμός της με το σφάλμα της πρόβλεψης, το οποίο κωδικοποιείται και τοποθετείται στη ροή δεδομένων, όπως φαίνεται στο Σχήμα 4.1. Ουσιαστικά, το σχήμα αυτό εμποδίζει, μέσω των προβλέψεων, το σφάλμα κωδικοποίησης να διαδοθεί και να μεγαλώσει.

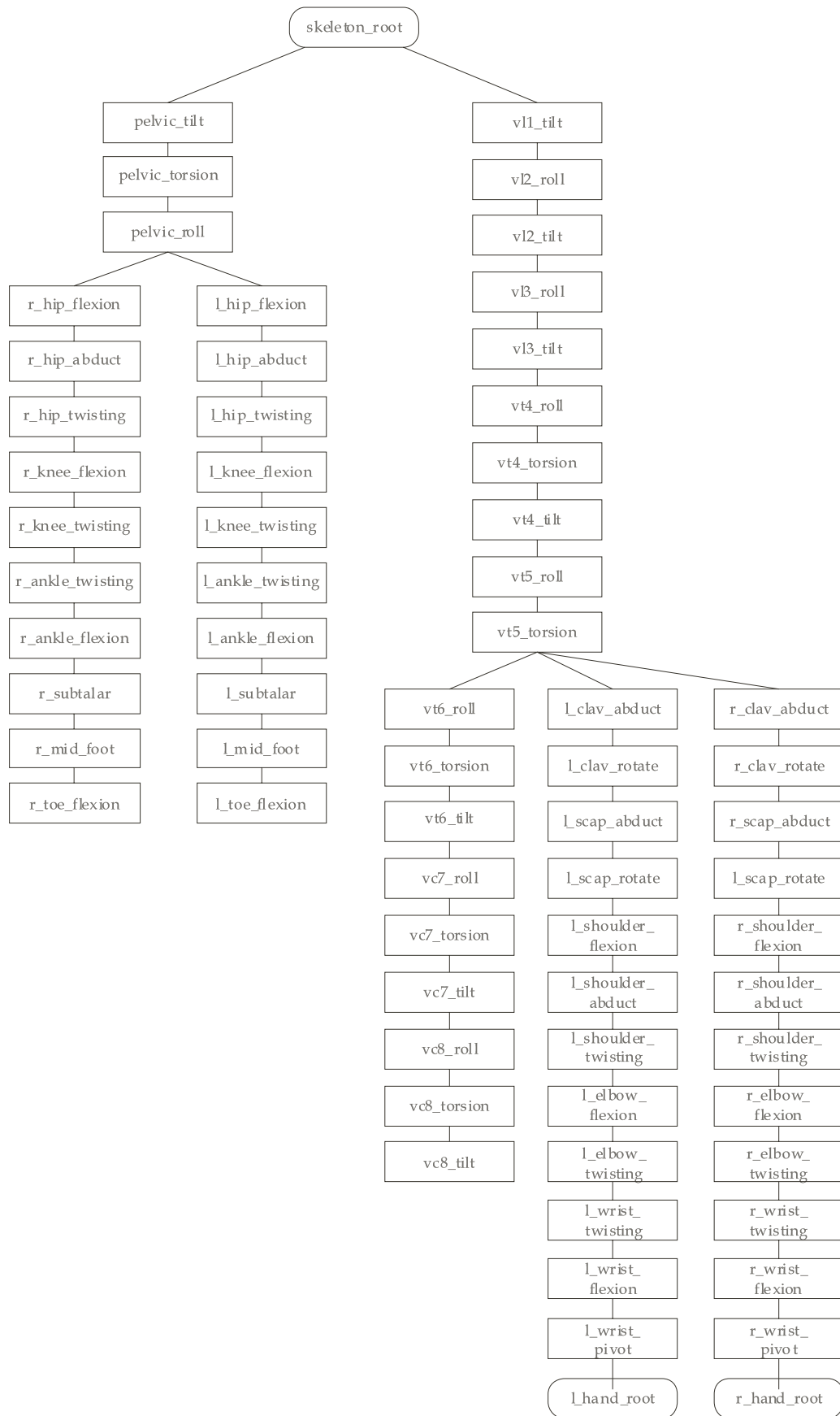
4.2 ΠΕΡΙΓΡΑΦΗ ΤΩΝ ΠΑΡΑΜΕΤΡΩΝ ΒΑΡ

Επειδή οι παράμετροι ΒΑΡ περιγράφουν την κίνηση και τη συμπεριφορά του μοντέλου τύπου Body με σημασιολογική έννοια, περιέχουν μια σειρά από βαθμούς ελευθερίας (Degrees of Freedom – DOF), οι οποίοι είναι αρκετοί για να περιγράψουν τη στάση και την εμπύχωση του σκελετικού μοντέλου. Έξι από αυτούς τους βαθμούς ελευθερίας αφορούν τη θέση και τον προσανατολισμό του συνολικού μοντέλου, ενώ οι υπόλοιποι 66 καθορίζουν την κατάσταση των εσωτερικών συνδέσμων, η τοπολογία των οποίων φαίνεται στο Σχήμα 4-2. Σε αυτήν την αναπαράσταση, οι σύνδεσμοι που ανήκουν στα χέρια θεωρούνται προαιρετικοί και παραλείπονται για λόγους απλότητας.

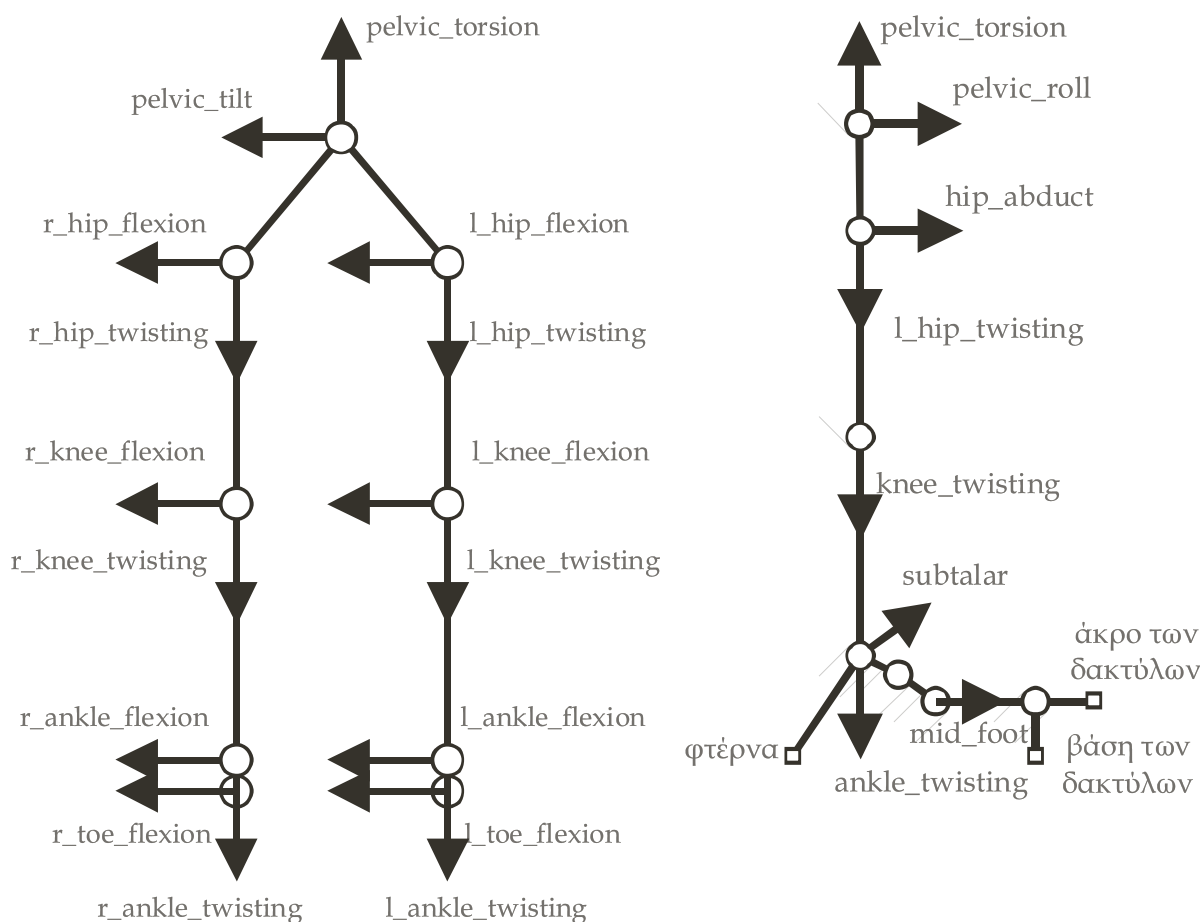
Ξεκινώντας από το βασικό κόμβο `skeleton_root`, το μοντέλο περιλαμβάνει τρεις βαθμούς ελευθερίας για τον προσανατολισμό της λεκάνης και εννιά βαθμούς σε καθένα από τα πόδια, από το σύνδεσμο του γοφού στο σύνδεσμο των δακτύλων. Γενικά, ο βασικός κόμβος της σκελετικής οργάνωσης ταυτίζεται με τη λεκάνη για να διευκολυνθεί η συνολική μετακίνηση του αντικειμένου, μια που ο συγκεκριμένος κόμβος μένει σταθερός σε ότι αφορά τη μεταφορά [24]. Η ονοματολογία των κινήσεων (περιστροφών) της λεκάνης, χρησιμοποιείται και για τους αντίστοιχους βαθμούς των κόμβων της σπονδυλικής στήλης. Έτσι, ο βαθμός `torsion` περιγράφει την περιστροφή γύρω από τον κατακόρυφο άξονα, ο οποίος περνάει από το βασικό κόμβο, ο βαθμός `roll` την περιστροφή παράλληλα με τον άξονα που είναι παράλληλος με τη διεύθυνση της κίνησης του αντικειμένου, ενώ ο βαθμός `tilt` την περιστροφή παράλληλα με τον άξονα που είναι κάθετος στην κίνηση.

Προχωρώντας προς το κάτω μέρος της ιεραρχίας, ο βαθμός `flexion` στο γοφό, το γόνατο και τον αστράγαλο αντιστοιχεί στην έκταση της άρθρωσης προς τα εμπρός, ενώ ο βαθμός `twisting` στην περιστροφή που είναι παράλληλη προς τον άξονα του μέλους που ξεκινά από το σύνδεσμο. Το τμήμα της ιεραρχίας που αφορά την πατούσα είναι σχετικά πιο πολύπλοκο, σε ότι αφορά τους συνδέσμους, μπορεί να περιγραφεί όμως συνολικά με τρεις βαθμούς ελευθερίας για τη θέση και τον προσανατολισμό. Έτσι, ο βαθμός `toe_flexion` περιγράφει την περιστροφή των συνδέσμων στα δάκτυλα, ο βαθμός `subtalar` την περιστροφή γύρω από τον αστράγαλο και, τέλος, ο βαθμός `mid_foot` το τόξο που μπορεί να διαγράψει η πατούσα. Το Σχήμα 4-3 εμφανίζει τους βαθμούς ελευθερίας στους κόμβους των ποδιών, καθώς τους άξονες περιστροφής που αντιπροσωπεύουν.

Η περιγραφή και οι κινητικές δυνατότητες του πάνω τμήματος του κορμού είναι σχετικά πιο πολύπλοκες, λόγω των ιδιαιτεροτήτων στους ώμους και τις ξεχωριστές αλυσίδες των χεριών. Στην προκαθορισμένη κατάσταση των κόμβων, οι βαθμοί ελευθερίας είναι όμοιοι με αυτούς που ορίζονται για τα πόδια, αν ο βραχίονας είναι έτσι περιστρεμμένος, ώστε η παλάμη να βλέπει προς τα πίσω.



Σχήμα 4-2: Τοπολογία αντικειμένων τύπου Body με απλοποιημένη σπονδυλική στήλη



Σχήμα 4-3: Βαθμοί ελευθερίας στους κόμβους των ποδιών του σκελετικού μοντέλου

Περιγραφή των παραμέτρων BDP

Οι παράμετροι BDP μπορούν να χρησιμοποιηθούν από τον κωδικοποιητή για να προσαρμόσουν ένα γενικό μοντέλο τύπου Body σε συγκεκριμένες μετρήσεις ή για να ολοκληρώσουν τη μεταφορά ενός συγκεκριμένου μοντέλου μαζί με την πληροφορία εμφύχωσης. Σε κανονικές συνθήκες, η μετάδοση των παραμέτρων BDP γίνεται μόνο μια φορά και ακολουθείται από τη ροή δεδομένων που περιέχει τις παραμέτρους BAP – αν η ροή δεδομένων δεν περιέχει BDPs, τότε ο αποκωδικοποιητής εφαρμόζει τις παραμέτρους BAP σε ένα μοντέλο γενικής μορφής.

Γενικά, η ροή δεδομένων BDP περιέχει πληροφορίες που καθορίζουν τη μορφή της γεωμετρίας του μοντέλου και τις συντεταγμένες απεικόνισης υψής, αν έχει χρησιμοποιηθεί αυτή η τεχνική. Η μορφή και οι αναλογίες της γεωμετρίας καθορίζονται σύμφωνα με το πρότυπο VRML 2.0 και το σχετικό κόμβο H-Anim Proto Segment [21]. Ακόμα, μεταδίδονται οι θέσεις που αντιστοιχούν στους συνδέσμους, σύμφωνα με το μηχανισμό Proto Joint, ενώ αν έχει εφαρμοστεί απεικόνιση υψής, οι σχετικές εικόνες μεταδίδονται σαν τμήμα του κόμβου PROTO SEGMENT. Αν το συνολικό μοντέλο δεν είναι διαθέσιμο, τότε η ροή δεδομένων μπορεί να χρησιμοποιηθεί για να προσαρμόσει ένα γενικό μοντέλο, χρησιμοποιώντας ένα υποσύνολο του ορισμού του σκελετικού μοντέλου σύμφωνα με τη σύνταξη H-Anim. Αυτό ουσιαστικά γίνεται με τη μεγέθυν-

ση των μελών στο γενικό μοντέλο, σύμφωνα με τις θέσεις των συνδέσμων που μεταδίδονται και οι οποίες αντιστοιχούν στο πεδίο center του κόμβου Joint. Έτσι, συνολικά το πρότυπο ορίζει τρία διακριτά προφίλ για την εμπύχωση του αντικειμένου τύπου Body: το προφίλ Simple, στο οποίο αγνοείται από τον αποκωδικοποιητή η πληροφορία των BDPs, το προφίλ Calibration στο οποίο το γενικό μοντέλο στον αποκωδικοποιητή προσαρμόζεται σύμφωνα με τα δεδομένα της ροής και το προφίλ Predictable, κατά το οποίο ο αποκωδικοποιητής χρησιμοποιεί συνολικά το μοντέλο που μεταδίδεται, αντικαθιστώντας αυτό που είναι αποθηκευμένο τοπικά.

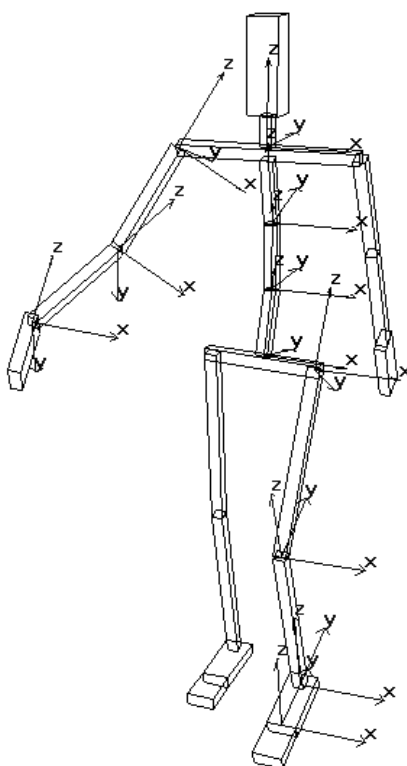
Εκτός από τις πραγματικές τιμές των παραμέτρων BDP που μεταδίδονται με τη ροή δεδομένων, το πρότυπο χρησιμοποιεί μια σειρά από υποθέσεις που αφορούν την αρχική θέση του μοντέλου, αλλά και το συνολικό σύστημα συντεταγμένων του. Έτσι, η αρχή του συστήματος ορίζεται στο έδαφος, όπου $y = 0$, και ανάμεσα στα πόδια του μοντέλου, με τον άξονα x να δείχνει προς το αριστερό μέρος του, τον άξονα y προς τα πάνω και τον άξονα z προς τα εμπρός. Συνολικά, για την κωδικοποίηση των παραμέτρων BAP, το πρότυπο MPEG-4 χρησιμοποιεί το σχήμα προσαρμοζόμενης αριθμητικής κωδικοποίησης που περιγράφεται στο [22].

Σε ότι αφορά τα γενικά χαρακτηριστικά και την ιεραρχία του μοντέλου, η προσέγγιση ενός ανθρώπινου σώματος από μια σκελετική δομή παρουσιάζει μια σειρά από πλεονεκτήματα, τα οποία διευρύνουν τη γενικότητα και εξασφαλίζουν την συστηματικότητα της μελέτης. Γενικά, οι άνθρωποι διαφέρουν σημαντικά ως προς την εξωτερική τους εμφάνιση - εκτός από τις διαφορές που υπάρχουν στα μορφολογικά χαρακτηριστικά των δύο φύλων, παρατηρούνται επίσης διαφορές στην κατανομή του βάρους και τη μυϊκή διάπλαση. Παρόλα αυτά, και με την επιφύλαξη παθολογικών καταστάσεων, υπάρχουν σημαντικά λιγότερες διαφορές σε ότι αφορά τη σκελετική δομή, αφού οι αναλογίες των διάφορων μελών παραμένουν γενικά οι ίδιες [25] και οι διαφορές στο ύψος είναι σημαντικά πιο εύκολο να μοντελοποιηθούν σε μια σκελετική ιεραρχική δομή. Επίσης, παρά το γεγονός ότι το αίτιο των κινήσεων σε ένα ανθρώπινο σώμα είναι η δράση των μυών, το δυναμικό αποτέλεσμα εφαρμόζεται στα οστά, τα οποία με την σειρά τους μετακινούν τα υπόλοιπα μέρη του σώματος. Φαίνεται λοιπόν να δικαιολογείται απόλυτα η χρήση γωνιών ανάμεσα στους συνδέσμους για να περιγραφεί η στάση του ανθρώπινου σώματος, αλλά και η περιγραφή των κινήσεων που παρατηρούνται με βαθμούς ελευθερίας που αντιστοιχούν σε περιστροφές. Κάτι τέτοιο διευκολύνει σημαντικά και την επιβολή περιορισμών στις κινήσεις των συνδέσμων, οι οποίοι αν και δεν προβλέπονται από το πρότυπο MPEG-4, είναι αναγκαίοι για τη φυσικότητα της κίνησης, ειδικά σε διαλογικές εφαρμογές. Ακόμα, η χρήση ενός σκελετικού μοντέλου, χωρίς την απαραίτητη συνύπαρξη ενός γεωμετρικού μοντέλου, κάνει δυνατή τη σχεδίαση της εμπύχωσης που θα πρέπει να αποδοθεί ακόμα και σε οικιακά υπολογιστικά συστήματα, αφού δεν είναι απαραίτητη η υπολογιστική δύναμη του αποκωδικοποιητή για την απόδοση ενός πολύπλοκου πολυγωνικού αντικειμένου. Έτσι, τα δεδομένα της κίνησης μπορούν εύκολα να επαναχρησιμοποιηθούν, κάτι που αποτελεί βασικό χαρακτηριστικό του προτύπου MPEG-4, με την προϋπόθεση ότι η σκελετική δομή παραμένει σταθερή [23].

4.3 ΦΥΣΙΟΛΟΓΙΚΕΣ ΜΕΛΕΤΕΣ

Τα διάφορα μέλη που αποτελούν το συνθετικό μοντέλο είναι ταξινομημένα σε μία ιεραρχική δομή, στην οποία καθένα από αυτά έχει ένα μοναδικό τμήμα ως "πρόγονο" του, ενώ μπορεί να έχει έναν ή περισσότερα τμήματα ως "απογόνους". Επίσης, όπως φαίνεται στο Σχήμα 4-4, για

κάθε μέλος, κι επομένως και για κάθε άρθρωση, έχει οριστεί ένα τοπικό σύστημα αξόνων, η αρχή του οποίου βρίσκεται στο σημείο που αυτό ενώνεται με τον προηγούμενο του, δηλαδή στο σημείο της άρθρωσης, με την κατεύθυνση του z άξονα να συμπίπτει με αυτή του άξονα του μέλους. Σαν αποτέλεσμα της ταξινόμησης των μελών σε αυτήν την ιεραρχική δομή, οι μετασχηματισμοί κάθε μέλους μεταφέρεται σε όλους τους απογόνους του, ενώ αντίθετα ο πρόγονός του μένει ανεπηρέαστος. Για παράδειγμα, αν η ροή δεδομένων περιέχει το ΒΑΡ για την περιστροφή του βραχίονα, τότε τα μέλη που βρίσκονται πιο χαμηλά από αυτόν στην ιεραρχία, δηλαδή ο πήχης και η παλάμη θα ακολουθήσουν αυτή την περιστροφή. Αν η ρεαλιστική απεικόνιση είναι απαραίτητη, η εφαρμογή θα πρέπει να έχει την πρόσθετη δυνατότητα να περιορίσει τους άξονες του τοπικού συστήματος αξόνων κατά τους οποίους επιτρέπεται να μετακινηθεί ή να περιστραφεί ο κάθε σύνδεσμος.



Σχήμα 4-4: Απλοποιημένη αναπαράσταση της ιεραρχίας του σκελετικού μοντέλου και των τοπικών συστημάτων συντεταγμένων για τα μέλη

Έχοντας περιγράψει την έννοια της ιεραρχίας, μπορούμε να περιγράψουμε τον τρόπο με τον οποίο μοντελοποιούνται οι αρθρώσεις σε ένα συνθετικό μοντέλο. Με τον όρο "άρθρωση" υποδεικνύονται τα σημεία του ανθρώπινου σώματος στα οποία συνδέονται δύο μέλη (τμήματα) του ανθρώπινου σκελετού και στα οποία υπάρχει η δυνατότητα κίνησης. Για παράδειγμα, στην άρθρωση του αγκώνα συνδέεται ο βραχίονας με τον πήχη, μέλη που μοντελοποιούνται με τους αντίστοιχους κόμβους. Αυτό πρακτικά σημαίνει ότι όταν μετακινείται ο βραχίονας θα πρέπει να ακολουθήσει την μετακίνηση αυτή και ο πήχης, ενώ ο πήχης έχει την δυνατότητα να περιστρέφεται γύρω από την άρθρωση του αγκώνα, χωρίς ταυτόχρονα να μετακινείται ο βραχίονας. Αντίθετα, δεν επιτρέπεται η ανεξάρτητη μετατόπιση του, αφού κάτι τέτοιο θα είχε σαν αποτέλε-

σμα την αποκόλληση του βραχίονα από την άρθρωση. Σε ότι αφορά τους περιορισμούς που είναι αναγκαίοι για να αποφύγουμε την ανεξάρτητη μετακίνηση του πήχη, δεν επιτρέπουμε στο αντίστοιχο μέλος να μετατοπίζεται στο τοπικό σύστημα αξόνων, παρά μόνο να περιστραφεί γύρω από την άρθρωση του αγκώνα. Το ίδιο συμβαίνει σε όλες τις αρθρώσεις του σκελετικού μοντέλου, με τρεις πολύ βαθμούς ελευθερίας να ορίζονται για κάθε άρθρωση και οι οποίοι αφορούν την περιστροφή και όχι την μετατόπιση των μελών που καταλήγουν στο σύνδεσμο.

Στην πραγματικότητα, ο πήχης δεν είναι ελεύθερος να περιστρέφεται σε οποιαδήποτε κατεύθυνση, αλλά μόνο σε αυτήν της έκτασης. Έχει λοιπόν νόημα να περιορίσουμε την δυνατότητα περιστροφής, να μειώσουμε δηλαδή τους βαθμούς ελευθερίας, απαγορεύοντας στον αντίστοιχο σύνδεσμο να περιστρέφεται γύρω από τους υπόλοιπους άξονες του τοπικού συστήματος αξόνων. Επιπλέον η οποιαδήποτε περιστροφή, μέσα στα πλαίσια των δεδομένων βαθμών ελευθερίας σε κάθε άρθρωση, περιορίζεται και από φυσικούς παράγοντες. Για παράδειγμα στην άρθρωση του αγκώνα δεν είναι δυνατή η υπερέκταση, δηλαδή η κάμψη της άρθρωσης προς τα πίσω, ενώ από την άλλη η κάμψη του πήχη περιορίζεται από το γεγονός ότι το φυσικό μέλος δεν μπορεί στην πράξη να διαπεράσει το βραχίονα. Ανάλογοι περιορισμοί ισχύουν για τις περισσότερες από τις υπόλοιπες αρθρώσεις, σε φυσιολογικές τουλάχιστον συνθήκες. Για να γίνει δυνατή η μοντελοποίηση των φυσικών αυτών περιορισμών, θέτουμε συγκεκριμένα όρια για την κάθε περιστροφή, περιορίζοντας έτσι την περιστροφή του κάθε συνδέσμου, μέσα στα πλαίσια των επιτρεπόμενων βαθμών ελευθερίας. Ο πίνακας 4-1 εμφανίζει τις συνηθισμένες τιμές των ορίων για φυσιολογικές κινήσεις [26].

ΜΕΛΟΣ	ΒΑΘΜΟΙ ΕΛΕΥΘΕΡΙΑΣ	ΤΟΠΙΚΟ ΣΥΣΤΗΜΑ ΑΞΟΝΩΝ		
		Χ	Υ	Ζ
Λεκάνη	6	Ελεύθερη	Ελεύθερη	Ελεύθερη
Μηρός	3	-120° - 20°	-31° - 53°	-30° - 31°
Κνήμη	1	0° - 113°	0°	0°
Πέλμα	3	-35° - 38°	-43° - 35°	-24° - 23°
Δάκτυλα	1	-45° - 10°	0°	0°
Σ. Στήλη 1	3	-17° - 29°	-18° - 18°	-6° - 6°
Σ. Στήλη 2	3	-12° - 24°	-25° - 25°	-6° - 6°
Σ. Στήλη 3	3	-5° - 10°	-20° - 20°	-31° - 31°
Μπράτσο	3	-188° - 61°	-48° - 134°	-60° - 80°
Βραχίονας	3	-142° - 0°	0°	-120° - 20°
Παλάμη	3	-27° - 47°	-99° - 90°	-113° - 77°
Λαιμός	3	-61° - 60°	-41° - 41°	-79° - 79°

Πίνακας 4-1: Χαρακτηριστικές τιμές των ορίων για την περιστροφή των κύριων συνδέσμων ενός ανθρώπινου μοντέλου

Στον πίνακα 4-1 δεν αναφέρονται τα μέλη που αντιστοιχούν στο κεφάλι και τους ώμους του μοντέλου. Αυτό συμβαίνει γιατί στους συνδέσμους που αφορούν τα αντίστοιχα τμήματα του ανθρώπινου σκελετού, δεν παρουσιάζονται ανεξάρτητες περιστροφές: όλες οι κινήσεις του κεφαλιού οφείλονται στις κινήσεις του λαιμού, ενώ οι κινήσεις που παρουσιάζονται στους ώμους είναι στην πραγματικότητα το αποτέλεσμα κινήσεων στην σπονδυλική στήλη, και κυρίως στο τμήμα "Σ. Στήλη 3" που αποτελεί και τον πρόγονο του συγκεκριμένου τμήματος.

Με βάση τα παραπάνω, μπορούμε να ορίσουμε περισσότερο πολύπλοκα μοντέλα για την πιο λεπτομερή περιγραφή των διαφόρων κινήσεων στο ανθρώπινο σώμα. Για παράδειγμα, στο μοντέλο που περιγράφεται έχει συμπεριληφθεί μόνο ένα τμήμα για τα δάκτυλα των ποδιών, το οποίο όμως είναι αρκετό για να μελετηθούν κινήσεις όπως το περπάτημα και το τρέξιμο των ποδιών.

Παραμετρική Αναπαράσταση και Σύνθεση της Βάδισης

Η αναπαράσταση της κίνησης ενός σκελετικού μοντέλου με τα απαιτούμενα δεδομένα για τις τιμές των περιστροφών στους συνδέσμους είναι συνήθως αρκετά πολύπλοκη και περιέχει περίσσεια δεδομένων στις περιπτώσεις περιοδικών κινήσεων, όπως η φυσιολογική βάδιση. Έτσι, είναι αναγκαίο να ορίσουμε ένα αφαιρετικό σχήμα περιγραφής της κίνησης, στο οποίο θα πρέπει να είναι δυνατή η προσαρμογή συγκεκριμένων παραμέτρων, έτσι ώστε να αλλάξουν τα χαρακτηριστικά της και το τελικό αισθητικό αποτέλεσμα. Στην προτεινόμενη ανάλυση, η βάδιση περιγράφεται με ένα μικρό σύνολο χαρακτηριστικών στάσεων του σκελετικού μοντέλου, ενώ τα ενδιάμεσα δεδομένα παράγονται με τη γραμμική παρεμβολή στις τιμές των γωνιών στις αρθρώσεις στις χαρακτηριστικές στάσεις. Έτσι, τα δεδομένα που απαιτούνται μειώνονται κατά πολύ, ενώ παράλληλα δίνεται στο σχεδιαστή ή τον κωδικοποιητή η δυνατότητα μεταβολής μιας σειράς από παραμέτρους στις χαρακτηριστικές στάσεις, αλλάζοντας την αίσθηση της κίνησης.

5.1 ΑΝΑΛΥΣΗ ΤΗΣ ΚΙΝΗΣΗΣ

Παρά το γεγονός ότι είναι αρκετά εύκολο για το ανθρώπινο μάτι να αναγνωρίσει μια κίνηση και να την ταυτοποιήσει ως βάδιση, είναι σχετικά πιο δύσκολο να δοθεί ένας συστηματικός ορισμός για την κίνηση αυτή. Αυτό συμβαίνει γιατί τα ποιοτικά χαρακτηριστικά της κίνησης διαφέρουν σημαντικά από άνθρωπο σε άνθρωπο, αλλά κι εξαρτώνται από την ψυχολογική και σωματική κατάσταση ακόμη και για ένα συγκεκριμένο άτομο: η κίνηση αποτυπώνεται διαφορετικά για άτομα που βιάζονται, είναι αγχωμένα ή ήρεμα, χαρούμενα ή λυπημένα. Χρειάζεται λοιπόν να αναζητήσουμε έναν ορισμό για το ποιες κινήσεις ονομάζουμε "βάδιση", όπως επίσης να αναζητήσουμε και μία κοινή βάση, να ορίσουμε με άλλα λόγια μια κίνηση που θα ονομάζουμε "φυσιολογικό" περπάτημα και με βάση αυτή στην συνέχεια να επεκτείνουμε την μελέτη και σε άλλα είδη βάδισης.

Κινησιολογικές μελέτες ορίζουν τη βάδιση σαν τη "μεταφορά του κέντρου βάρους του σώματος δια μέσου του χώρου κατά μήκος ενός δρόμου που απαιτεί την ελάχιστη δαπάνη ενέργειας, ως αποτέλεσμα των εναλλάξ στροφικών κινήσεων των κάτω άκρων". Ο ορισμός αυτός είναι μάλλον γενικός και μπορεί να περιλαμβάνει όλες τις κινήσεις των κάτω άκρων που έχουν ως αποτέλεσμα την μετακίνηση και μεταφορά του κέντρου βάρους του σώματος, το οποίο στην

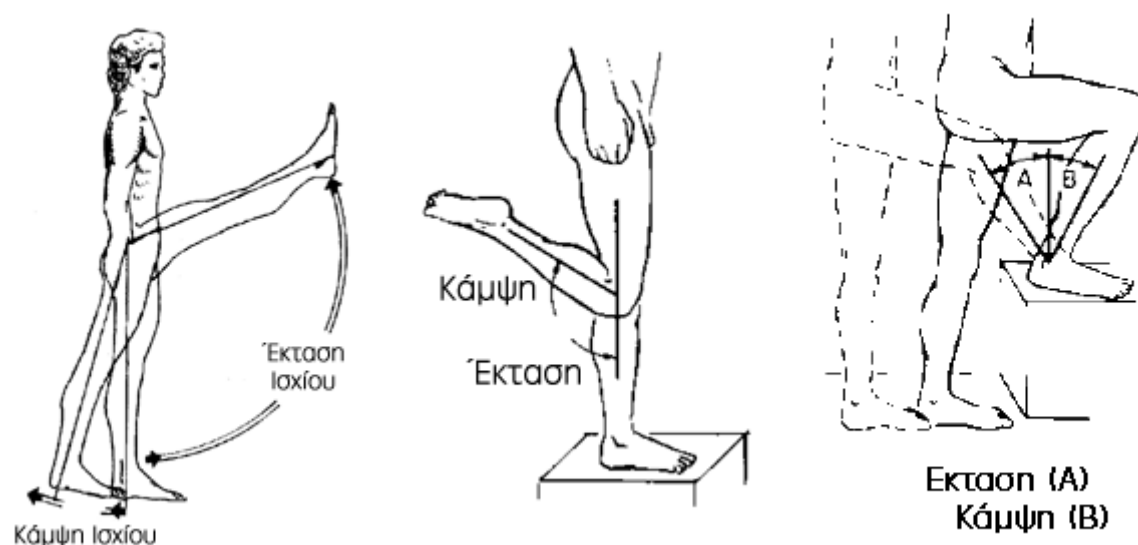
πραγματικότητα βρίσκεται λίγο πιο πάνω από την μέση και στην περίπτωση μας θεωρούμε ότι αντιπροσωπεύεται από τον μέλος "λεκάνη". Η απαίτηση της ελάχιστης ενέργειας ικανοποιείται με την αποφυγή περιττών ή υπερβολικών κινήσεων, αλλά είναι μάλλον αναγκαίο να αναζητηθεί ένας πιο πρακτικός ορισμός. Σε αυτόν θα πρέπει τουλάχιστον να συμπεριληφθεί και η απαίτηση να βρίσκεται τουλάχιστον ένα από τα δύο πόδια σε επαφή με το έδαφος σε όλη την διάρκεια της κίνησης, έτσι ώστε να γίνει δυνατή η διάκριση του τρεξίματος. Έτσι και με βάση την καθημερινή εμπειρία θα μπορούσαμε να πούμε ότι: Βάδιση είναι η μεταφορά του κέντρου βάρους του σώματος διαμέσου του χώρου, ως αποτέλεσμα των στροφικών κινήσεων των κάτω άκρων, κατά την διάρκεια της οποίας, σε κάθε χρονική στιγμή, τουλάχιστον ένα από τα κάτω άκρα βρίσκεται στο έδαφος.

Από μελέτες στον χώρο της κινησιολογίας προκύπτει ότι κατά την διάρκεια της βάδισης η κίνηση στο ανθρώπινο σώμα εμφανίζεται κυρίως στα πόδια και στην λεκάνη. Το υπόλοιπο σώμα είτε ακολουθεί την κίνηση της λεκάνης, όπως για παράδειγμα ο κορμός, είτε κινείται με τέτοιο τρόπο ώστε να διατηρηθεί η ισορροπία, όπως συμβαίνει με την αυθόρμητη κίνηση των χεριών. Αυτό ουσιαστικά σημαίνει ότι για να αναλύσουμε και να ορίσουμε την "φυσιολογική" βάδιση, αλλά και οποιοδήποτε άλλο είδος βάδισης, θα πρέπει να επικεντρώσουμε την μελέτη μας στον τρόπο με τον οποίο κινούνται τα πόδια και η λεκάνη, και ειδικότερα στον τρόπο με τον οποίο μεταβάλλονται οι γωνίες στους συνδέσμους ανάμεσα στα διάφορα μέλη των τμημάτων αυτών του ανθρώπινου σώματος.

Παρατηρώντας πραγματικά κινησιολογικά δεδομένα, φτάνουμε στο συμπέρασμα ότι η βάδιση είναι μια περιοδική, επαναληπτική κίνηση. Με άλλα λόγια, και τα δύο πόδια κατά την διάρκεια της βάδισης επαναλαμβάνουν τις ίδιες κινήσεις ανά τακτά χρονικά διαστήματα, με την περίοδο αυτής της κίνησης να ονομάζεται "κύκλος βάδισης". Οι κινησιολόγοι χωρίζουν συνήθως τον κάθε κύκλο σε δύο βασικές περιόδους, σε δύο φάσεις, για κάθε πόδι:

- φάση της στήριξης: αρχίζει την στιγμή που η πτέρνα του ποδιού ακουμπά στο έδαφος και τελειώνει την στιγμή που τα δάκτυλα του ίδιου σκέλους αφήνουν το έδαφος.
- φάση της αιώρησης: αρχίζει την στιγμή που τα δάκτυλα του ποδιού αφήνουν το έδαφος και τελειώνει την στιγμή που η πτέρνα του ίδιου σκέλους ακουμπά στο έδαφος.

Η περίοδος της στήριξης καλύπτει συνήθως το 60% του κύκλου βάδισης, ενώ η περίοδος της αιώρησης καλύπτει το υπόλοιπο 40%. Στα πλαίσια των παραπάνω μελετών [27] έχει αποδειχθεί ότι κατά την διάρκεια της βάδισης, στις αρθρώσεις των ποδιών εμφανίζονται κυρίως τρεις κινήσεις: κάμψη και έκταση των αρθρώσεων του ισχίου, του γόνατος και του αστραγάλου. Στο Σχήμα 5-1 φαίνεται η αναπαράσταση των κινήσεων αυτών, ενώ στον πίνακα 5-1 φαίνονται οι περιστροφές των αντίστοιχων συνδέσμων στο μοντέλο στις οποίες αντιστοιχούν οι παραπάνω κινήσεις.

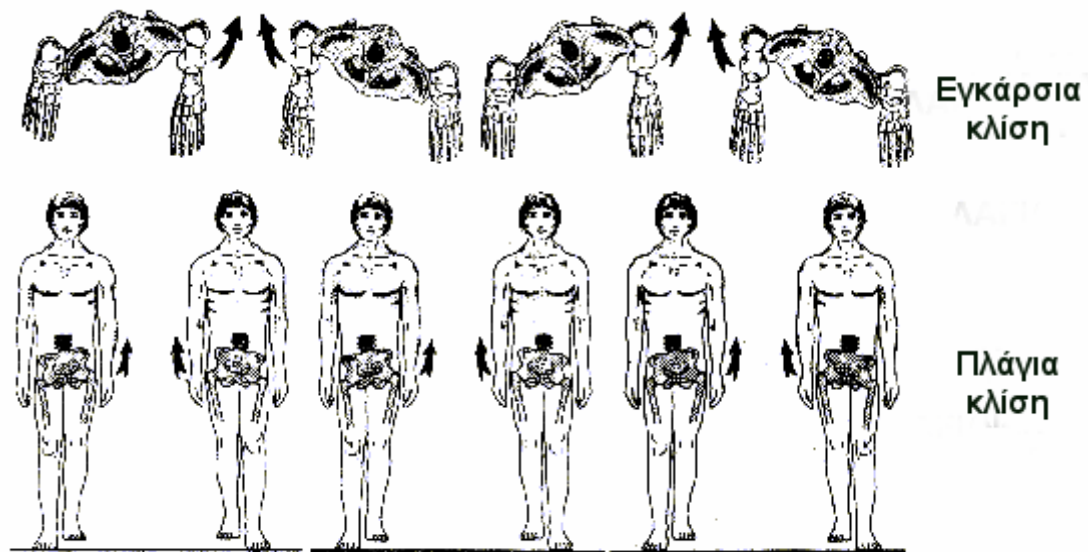


Σχήμα 5-1: Αναπαράσταση των κινήσεων στους συνδέσμους των ποδιών κατά τη φυσιολογική βάδιση

ΚΙΝΗΣΕΙΣ ΣΤΗ ΒΑΔΙΣΗ		ΚΙΝΗΣΕΙΣ ΣΤΟ ΜΟΝΤΕΛΟ	
ΑΡΘΡΩΣΗ	ΚΙΝΗΣΗ	ΣΥΝΔΕΣΜΟΣ	ΚΙΝΗΣΗ
ΙΣΧΙΟ	Κάμψη	ΜΗΡΟΣ	Θετική περιστροφή κατά Χ
	Έκταση		Αρνητική περιστροφή κατά Χ
ΓΟΝΑΤΟ	Κάμψη	ΚΝΗΜΗ	Θετική περιστροφή κατά Χ
	Έκταση		Αρνητική περιστροφή κατά Χ
ΑΣΤΡΑΓΑΛΟΣ	Κάμψη	ΦΤΕΡΝΑ	Αρνητική περιστροφή κατά Χ
	Έκταση		Θετική περιστροφή κατά Χ

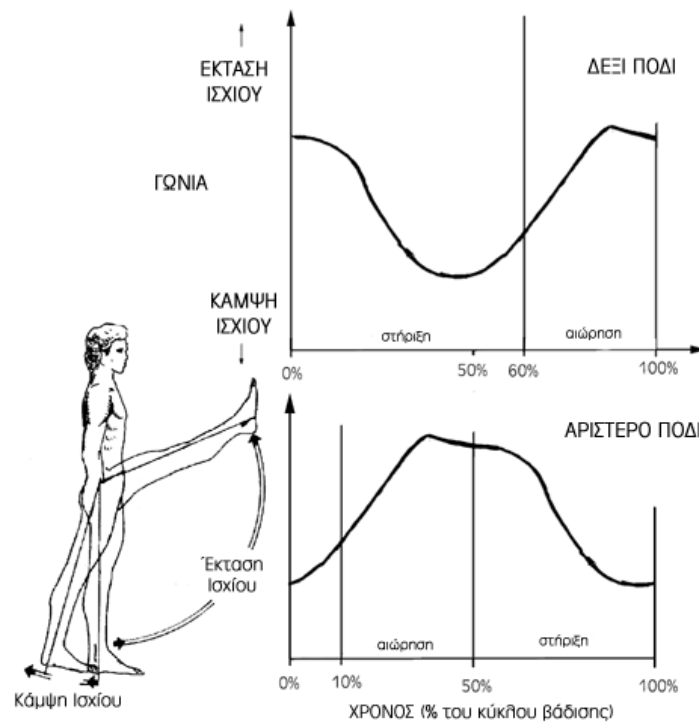
Πίνακας 5-1: Σχέση πραγματικών κινήσεων στους συνδέσμους των ποδιών και μετασχηματισμών στο μοντέλο

Σε ότι αφορά την περιοχή της λεκάνης, κατά την διάρκεια της βάδισης εμφανίζονται κατά βάση δύο είδη κινήσεων, εκτός της μεταφορικής, και οι οποίες αντιστοιχούν σε εγκάρσια και πλάγια κλίση. Οι κινήσεις αυτές περιγράφονται στο Σχήμα 5-2, και αντιστοιχούν σε περιστροφή των αντίστοιχων συνδέσμων στο μοντέλο κατά τον τοπικό Z και Y άξονα αντίστοιχα. Εκτός από αυτό, κατά την περιστροφή της λεκάνης παράλληλα με τους άξονες Z και Y, οι σύνδεσμοι που αφορούν τους μηρούς περιστρέφονται κατά τους αντίστοιχους άξονες στην αντίθετη κατεύθυνση, ώστε να διατηρήσουν τον προσανατολισμό τους προς τα εμπρός, προς την κατεύθυνση δηλαδή της κίνησης.

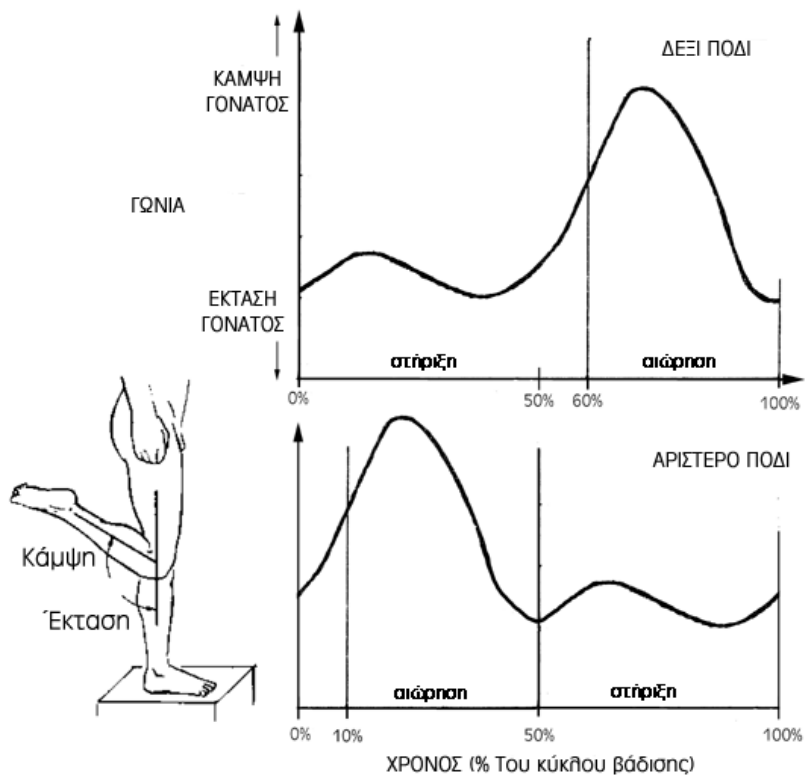


Σχήμα 5-2: Κλίση της λεκάνης κατά τη διάρκεια της βάδισης

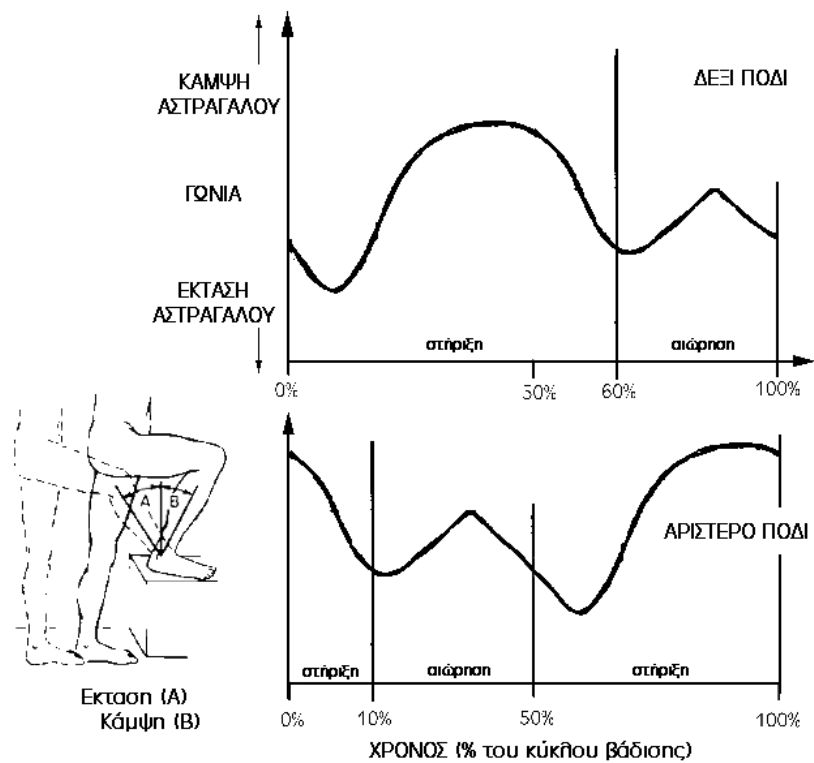
Συνολικά, αποδεικνύεται ότι οι παραπάνω επιμέρους περιστροφές μπορούν να τυποποιηθούν κατά την διάρκεια του κύκλου της κανονικής βάδισης. Στα σχήματα 5-3, 5-4 και 5-5 φαίνονται οι χαρακτηριστικές καμπύλες που περιγράφουν τις περιστροφές που εμφανίζονται στους συνδέσμους του ισχίου, του γονάτου και του αστραγάλου για τα δύο πόδια και για έναν κύκλο βάδισης που αρχίζει και τελειώνει όταν το δεξί πόδι ακουμπά στο έδαφος (αρχή της φάσης στήριξης για το πόδι αυτό). Επίσης φαίνεται για κάθε πόδι η περίοδος κατά την οποία βρίσκεται αυτό στη φάση στήριξης ή αιώρησης, ενώ στο Σχήμα 5-6, φαίνονται οι καμπύλες που περιγράφουν τις περιστροφές που εμφανίζονται στην λεκάνη, κατά την ίδια περίοδο.



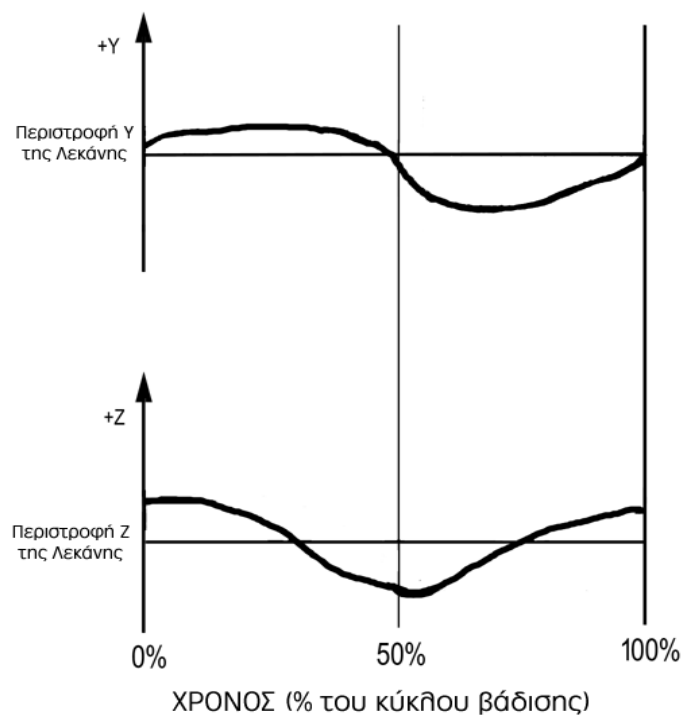
Σχήμα 5-3: Χαρακτηριστική καμπύλη για την περιστροφή του συνδέσμου στο ισχίο κατά τον κύκλο της κανονικής βάρδισης



Σχήμα 5-4: Χαρακτηριστική καμπύλη για την περιστροφή του συνδέσμου στο γόνατο κατά τον κύκλο της κανονικής βάρδισης



Σχήμα 5-5: Χαρακτηριστική καμπύλη για την περιστροφή του συνδέσμου στον αστράγαλο κατά τον κύκλο της κανονικής βάδισης



Σχήμα 5-6: Χαρακτηριστική καμπύλη για την περιστροφή του συνδέσμου στη λεκάνη κατά τον κύκλο της κανονικής βάδισης

Στα σχήματα αυτά, δεν εμφανίζονται οι πραγματικές τιμές που λαμβάνει η κάθε γωνία σε κάθε χρονική στιγμή. Αυτό συμβαίνει διότι οι πραγματικές τιμές εξαρτώνται σε μεγάλο βαθμό

από τις αναλογίες του συγκεκριμένου ανθρώπου ή του μοντέλου που περπατάει. Στην πραγματικότητα, αρκεί να περιγράψουμε την ποιοτική συμπεριφορά των γωνιών στις αρθρώσεις κατά την διάρκεια ενός μέσου, κανονικού βαδίσματος, και κυρίως για τις θέσεις στις οποίες εμφανίζονται τα, ολικά ή τοπικά, ακρότατα στις καμπύλες αυτές, και όχι τόσο για τις πραγματικές τιμές τους.

5.2 ΑΝΑΛΥΣΗ ΤΗΣ ΒΑΔΙΣΗΣ

Οι παραπάνω καμπύλες περιγράφουν τον τρόπο με τον οποίο μεταβάλλονται οι γωνίες στις αρθρώσεις του ποδιού και στην λεκάνη κατά την διάρκεια ενός συνηθισμένου περπατήματος. Θα μπορούσε λοιπόν κανείς να αναπαράγει μία τέτοια κίνηση σε ένα συμβατό συνθετικό μοντέλο, απλά μεταβάλλοντας τις γωνίες ανάμεσα στους συνδέσμους του κατά τρόπο ώστε η καμπύλη της μεταβολής να ακολουθεί την προδιαγεγραμμένη. Στην πράξη όμως αυτό μπορεί να παράγει μεγάλο όγκο δεδομένων, τα οποία δε μπορούν εύκολα να μεταβληθούν με σημασιολογικά κριτήρια. Για παράδειγμα, για να μπορέσει ο σχεδιαστής της κίνησης να αλλάξει το φυσικό μήκος που διανύει το μοντέλο σε κάθε βήμα, θα πρέπει να ξανασχεδιάσει ολόκληρο τον κύκλο, εκτός από την επεξεργασία και την προσαρμογή του μοντέλου. Είναι λοιπόν αναγκαίο να παράγουμε μια αφαιρετική αναπαράσταση του κύκλου, έτσι ώστε να εκφράσουμε τα μορφολογικά χαρακτηριστικά της με παραμετρικό τρόπο.

Παρατηρώντας τις χαρακτηριστικές καμπύλες των περιστροφών στους συνδέσμους της λεκάνης και των ποδιών, φαίνεται ότι μπορούμε να διαιρέσουμε έναν κύκλο βάδισης σε δύο τμήματα, με το δεύτερο από αυτά να είναι είτε ένα αντίγραφο του πρώτου, σε ότι αφορά τις κινήσεις στις αρθρώσεις του ισχίου, του γόνατος και του αστραγάλου, είτε το αποτέλεσμα καθρεπτισμού για τις κινήσεις στη λεκάνη. Στο δεύτερο μισό του κύκλου, σε ότι αφορά τις αρθρώσεις του ισχίου, του γόνατος και του αστραγάλου, το αριστερό πόδι κάνει τις ίδιες κινήσεις που έκανε το δεξί πόδι στο πρώτο μισό, και αντίστροφα. Πιο συγκεκριμένα, οι καμπύλες που φαίνονται στα σχήματα 5-3, 5-4 και 5-5 δείχνουν ότι οι κινήσεις του αριστερού ποδιού για το δεύτερο μισό του κύκλου (πέρα από το 50% στο χρονικό άξονα) είναι ίδιες με αυτές που αντιστοιχούν στις κινήσεις του δεξιού ποδιού για το πρώτο μισό. Το ίδιο συμβαίνει και για την αντίστοιχη καμπύλη για το δεύτερο μισό του κύκλου στο δεξί πόδι, που είναι ίδια με αυτή του αριστερού ποδιού στο πρώτο μισό. Αντίστοιχα, σε ότι αφορά τις καμπύλες που περιγράφουν τις κινήσεις της λεκάνης (Σχήμα 5-6), το δεύτερο μισό δεν είναι παρά το συμμετρικό του πρώτου μισού ως προς τον άξονα του χρόνου (καθρεπτισμός), και για τις δύο καμπύλες, για την περιστροφή κατά X και Z. Έτσι, συνολικά μπορούμε να συμπεράνουμε ότι η κανονική βάδιση μπορεί να οριστεί και να περιγραφεί, παρατηρώντας και αναλύοντας την κίνηση και των δύο ποδιών, μόνο για το μισό του κύκλου βάδισης.

Προχωρώντας σε μια περισσότερο αφαιρετική αναπαράσταση, παρατηρούμε ότι για το πρώτο μισό του κύκλου βάδισης όλα τα τοπικά ή ολικά ακρότατα και για τις τέσσερις καμπύλες και για τα δύο πόδια μπορούν, με την επιφύλαξη μια μικρής χρονικής μετατόπισης, να εντοπιστούν σε τέσσερις χαρακτηριστικές χρονικές στιγμές, οι οποίες ορίζονται με ποσοστά του κύκλου βάδισης. Έτσι, με βάση τις μέγιστες και ελάχιστες τιμές των διαφόρων γωνιών που εμφανίζονται σε κάθε ένα από αυτά τα χαρακτηριστικά χρονικά σημεία, ορίζουμε τέσσερις χαρακτηρι-

στικές στάσεις για το ανθρώπινο σώμα ή για ένα συνθετικό μοντέλο κατά την διάρκεια ενός κύκλου βάδισης [29].

Ονομάζουμε τις τέσσερις αυτές χαρακτηριστικές στάσεις [28] ως εξής: στάση έκτασης (extended position), η οποία αντιστοιχεί στη χρονική θέση 0%, δηλαδή στην αρχή του κύκλου βάδισης, τη στάση ανάκλισης (recoil position) στο 12% του κύκλου βάδισης, τη στάση μετάβασης (passing position) στο 25% του κύκλου και, τέλος, τη φάση αναμονής (anticipation phase), της οποίας η αρχή τοποθετείται χρονικά στο 37% του κύκλου βάδισης. Οι τέσσερις αυτές στάσεις, καθώς και τα τοπικά ή ολικά ακρότατα που εμφανίζονται σε κάθε μία, φαίνονται στον πίνακα 5-2.

				ΠΟΔΙ ΣΤΗ ΦΑΣΗ ΣΤΗΡΙΞΗΣ				ΠΟΔΙ ΣΤΗ ΦΑΣΗ ΑΙΩΡΗΣΗΣ			
	ΠΕΡΙΣΤΡΟΦΗ ΔΕΚΑΝΗΣ			ΠΕΡΙΣΤΡΟΦΗ ΚΑΤΑ Χ							
	Χ	Υ	Ζ	ΜΗ-ΡΟΣ	ΚΝΗΜΗ	ΦΤΕΡΝΑ	ΔΑΚΤΥΛΑ	ΜΗΡΟΣ	ΚΝΗΜΗ	ΦΤΕΡΝΑ	ΔΑΚΤΥ-ΛΑ
ΣΤΑΣΗ ΕΚΤΑΣΗΣ	0	0	MAX		MIN		0	MAX			MIN
ΣΤΑΣΗ ΑΝΑΚΛΗΣΗΣ	0				Τοπικό MAX	Τοπικό MAX	0			MAX	0
ΣΤΑΣΗ ΜΕΤΑΒΑΣΗΣ	0	MAX	0		Τοπικό MIN		0		MAX		0
ΦΑΣΗ ΑΝΑΜΟΝΗΣ	0					MIN		MIN		Τοπικό MIN	0

Πίνακας 5-2: Ποιοτικός ορισμός χαρακτηριστικών φάσεων σε έναν κύκλο βάδισης

Σχετικά με τον παραπάνω πίνακα, θα πρέπει να παρατηρήσουμε ότι σε κάθε φάση το ένα πόδι είναι στο έδαφος, δηλαδή στην φάση της στήριξης, ενώ το άλλο πόδι βρίσκεται στην φάση της αιώρησης. Μόνη εξαίρεση αποτελεί η στάση έκτασης, στην οποία και τα δύο πόδια είναι στο έδαφος. Στην περίπτωση αυτή ως "πόδι στην φάση της αιώρησης" στον πίνακα αναφέρεται καταχρηστικά αυτό που είναι πίσω, αυτό δηλαδή που στην επόμενη φάση (στάση ανάκλισης) θα περάσει στην φάση της αιώρησης.

Στον πίνακα αυτό περιγράφονται τα τοπικά και ολικά ακρότατα, που εμφανίζονται σε κάθε φάση, όσον αφορά την περιστροφή των συνδέσμων του μοντέλου παράλληλα με τον άξονα Χ, μετασχηματισμός που αποτελεί το βασικό είδος κίνησης που εμφανίζεται σε έναν κύκλο βάδισης. Στις περιστροφές αυτές, η μέγιστη αρνητική περιστροφή παρουσιάζεται ως ελάχιστο στον παραπάνω πίνακα: για παράδειγμα, κατά τη φάση αναμονής παρουσιάζεται ολικό μέγιστο στην κάμψη στην άρθρωση του αστραγάλου του ποδιού που βρίσκεται στο έδαφος, αλλά επειδή η κάμψη στον αστράγαλο αντιστοιχεί σε αρνητική περιστροφή στο σύνδεσμο της φτέρνας στο μοντέλο μας, το ακρότατο αυτό εμφανίζεται ως ελάχιστο στον πίνακα. Στην πράξη, τα δάκτυλα περιστρέφονται μονάχα στην περίπτωση που στο πόδι που βρίσκεται στη στάση της στήριξης, έχει σηκωθεί η φτέρνα, ώστε να διατηρήσουν την οριζόντια θέση τους. Αυτό συμβαίνει μόνο

στις χαρακτηριστικές στάσεις της φάσης αναμονής και της στάση έκτασης, ενώ η μέγιστη κάμψη (μέγιστη αρνητική περιστροφή κατά X) εμφανίζεται στη δεύτερη στάση.

Ανάλογες με τις παραπάνω χαρακτηριστικές στάσεις μπορούν να εντοπιστούν [28] και για άλλα είδη κινήσεων, όπως για παράδειγμα το τρέξιμο. Εξάλλου επειδή το σύνολο των στάσεων είναι χαρακτηριστικό για κάθε κίνηση, η περιγραφή του μπορεί να χρησιμοποιηθεί για τη διάκριση μεταξύ των διαφόρων ειδών κίνησης. Για παράδειγμα, παρακολουθώντας έναν άνθρωπο να τρέχει, μπορούμε να διαπιστώσουμε ότι στην κίνηση αυτή εμφανίζεται μία χαρακτηριστική φάση, κατά την οποία και τα δύο πόδια βρίσκονται στον αέρα. Μάλιστα αυτό συμβαίνει περίπου στο ίδιο ποσοστό του κύκλου με αυτό στο οποίο εμφανίζεται η φάση αναμονής σε έναν κύκλο βάρδιας. Επειδή όμως κατά την βάρδια σε όλες τις στάσεις υπάρχει τουλάχιστον ένα πόδι στο έδαφος, είναι δυνατόν να διαχωρίσουμε τα δύο αυτά είδη κινήσεων, συγκρίνοντας τις αντίστοιχες φάσεις.

Πιο συγκεκριμένα, σε ένα φυσιολογικό κύκλο βάρδιας, η αρχή της στάσης έκτασης σηματοδοτεί την αρχή του κύκλου βάρδιας και εντοπίζεται στη στιγμή κατά την οποία το μπροστινό πόδι ακουμπά στο έδαφος. Αντίστοιχα, ο εντοπισμός της στάσης ανάκλησης μπορεί να διευκολυνθεί με τον εντοπισμό του τοπικού ελαχίστου στην γωνία που σχηματίζεται στο γόνατο του ποδιού που βρίσκεται στην φάση της στήριξης, ενώ ο εντοπισμός της στάσης μετάβασης στηρίζεται στον εντοπισμό της μέγιστης γωνίας στο γόνατο του ποδιού που βρίσκεται στην φάση της στήριξης. Επίσης στη στάση αυτή παρουσιάζεται το ολικό ελάχιστο για τον κύκλο βάρδιας για την γωνία στο γόνατο του ποδιού που βρίσκεται στην φάση της αιώρησης. Επομένως στην περίπτωση που το ορατό πόδι θα ήταν αυτό που βρίσκεται στη φάση της αιώρησης, ο εντοπισμός της στάσης αυτής μπορεί να στηριχθεί στον εντοπισμό του ελαχίστου στην γωνία του γόνατος. Τέλος, ο εντοπισμός της φάσης αναμονής μπορεί να στηριχθεί στον εντοπισμό της μέγιστης έκτασης στο ισχίο του ποδιού που βρίσκεται στην φάση της αιώρησης. Γενικά, αυτή η φάση είναι η δυσκολότερη στον εντοπισμό, αφού αυτός στηρίζεται στην παρακολούθηση της γωνίας σε μία δυσδιάκριτη πρακτικά άρθρωση, όπως αυτή του ισχίου. Παρόλα αυτά, ο εντοπισμός θα μπορούσε να στηριχθεί στο γεγονός ότι η μέγιστη έκταση στο ισχίο έχει ως αποτέλεσμα το σήκωμα του μηρού του ποδιού που βρίσκεται σε αιώρηση στο μέγιστο ύψος για έναν κύκλο βάρδιας.

5.3 ΣΥΝΘΕΣΗ ΤΗΣ ΚΙΝΗΣΗΣ

Βασισμένοι στα αποτελέσματα της ανάλυσης του κύκλου της φυσιολογικής βάρδιας, μπορούμε να εκφράσουμε μια αφαιρετική αναπαράσταση της κίνησης, έτσι ώστε να γίνει πιο εύκολη και πρακτική η αναπαραγωγή της σε ένα συνθετικό μοντέλο. Πιο συγκεκριμένα, οι τέσσερις χαρακτηριστικές στάσεις, που ορίζονται από τις ακρότατες τιμές, χρησιμοποιούνται ως χαρακτηριστικά καρέ (key frames) σε μια συνθετική ακολουθία, ενώ τα ενδιάμεσα καρέ μπορούν να προκύψουν με γραμμική παρεμβολή ανάμεσα στις τιμές στα χαρακτηριστικά καρέ.

Στην πράξη, ο καθορισμός των γωνιών στους συνδέσμους του συνθετικού μοντέλου ολοκληρώνεται με τη μέθοδο της ευθείας κινηματικής, με την εξαίρεση του τμήματος της ιεραρχίας που αφορά τα χέρια, για το οποίο χρησιμοποιείται ανάστροφη κινηματική. Η επιλογή της μεθόδου της ευθείας κινηματικής για τη διαδικασία αυτή στηρίζεται στην ανάγκη του απόλυτου ελέγχου των τιμών των περιστροφών στους συνδέσμους, αφού οι τέσσερις χαρακτηριστικές στά-

σεις ορίζονται από την θέση των ακροτάτων σε ότι αφορά τις περιστροφές στα πόδια και στην λεκάνη. Από την άλλη, η κίνηση των χεριών δεν μπορεί να μοντελοποιηθεί με ανάλογο τρόπο κατά την διάρκεια της βάδισης, αφού ορίζεται με βάση το στόχο στην άκρη της σχετικής αλυσίδας, κι έτσι δεν είναι απαραίτητο να διατηρηθεί άμεσος έλεγχος στις γωνίες των περιστροφών στις συγκεκριμένες αρθρώσεις. Έτσι, σε κάθε στάση καθορίζουμε την επιθυμητή θέση για τον σύνδεσμο της παλάμης σε κάθε χέρι, με τις γωνίες των περιστροφών να υπολογίζονται αυτόματα από την εξίσωση της ανάστροφης κινηματικής για τους υπόλοιπους συνδέσμους μέχρι και τους ώμους.

Για τις ανάγκες της διαδικασίας σύνθεσης, στο ιεραρχικό μοντέλο προστίθενται δύο επιπλέον σύνδεσμοι, ένας για κάθε πόδι, στην αντίστοιχη φυσική θέση του πίσω μέρους της φτέρνας. Οι δύο σύνδεσμοι χρησιμοποιούνται και από το πρότυπο MPEG-4 (Σχήμα 4-3) και χρησιμοποιούνται για να καθορίζουν την θέση του πίσω μέρους της φτέρνας του κάθε ποδιού του μοντέλου, καθώς και στην επιβολή περιορισμών που έχουν σχέση με το έδαφος πάνω στο οποίο εξελίσσεται η κίνηση.

Ο κύκλος βάδισης που θα κατασκευάσουμε έχει συνολικό μήκος 32 καρέ. Αυτό σημαίνει ότι η στάση έκτασης θα πρέπει να εμφανίζεται στο καρέ 0, η στάση ανάκλησης στο καρέ 4, η στάση μετάβασης στο καρέ 8 και η έναρξη της φάσης αναμονής στο καρέ 12. Με βάση αυτές τις χρονικές θέσεις μπορούμε να συνθέσουμε το πρώτο μισό του κύκλου βάδισης, στο οποίο περιέχεται όλη η απαραίτητη πληροφορία για την κατασκευή ολόκληρου του κύκλου. Τα αντίστοιχα καρέ για το δεύτερο μισό, κατασκευάζονται αντιγράφοντας τις κινήσεις του ενός ποδιού στο άλλο και με καθρεπτισμό για την κίνηση της λεκάνης.

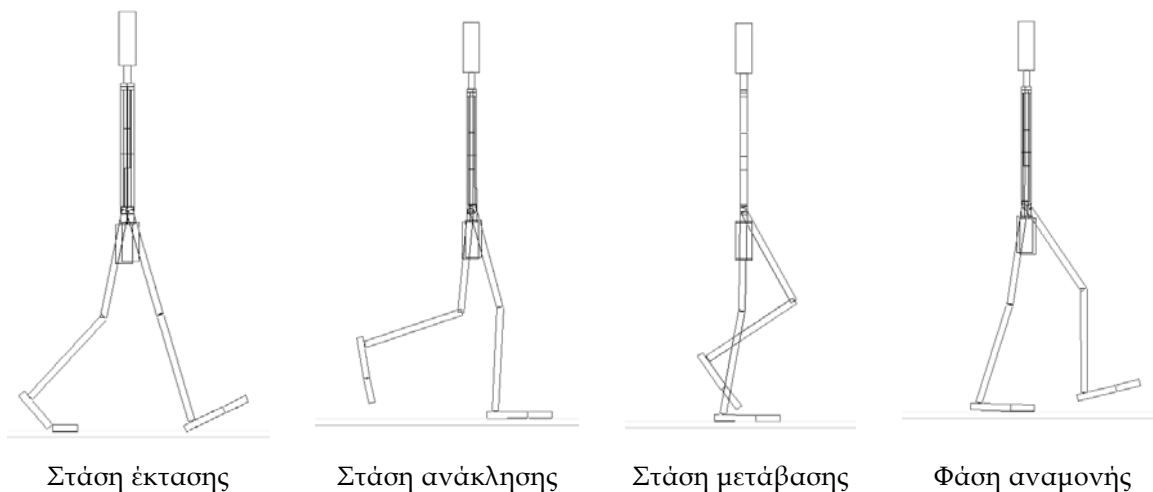
Για τον καθορισμό της στάσης του μοντέλου σε κάθε ένα από τα τέσσερα καρέ κλειδιά ακολουθούμε την παρακάτω διαδικασία:

- Αρχικά περιστρέφουμε κατάλληλα, κατά τον τοπικό Y και Z άξονα, τον σύνδεσμο της λεκάνης, ενώ ταυτόχρονα οι σύνδεσμοι των μηρών περιστρέφονται κατά την αντίθετη κατεύθυνση ώστε να διατηρούν τον προσανατολισμό τους προς την κατεύθυνση της κίνησης.
- Στην συνέχεια, δημιουργούνται οι χαρακτηριστικές στάσεις του μοντέλου στα τέσσερα καρέ κλειδιά, καθορίζοντας τις κατάλληλες περιστροφές για τους συνδέσμους του μηρού, της κνήμης και της φτέρνας σε κάθε πόδι. Στη φάση αναμονής, για το πόδι που βρίσκεται σε στήριξη και στη στάση έκτασης για το πόδι που βρίσκεται σε αιώρηση, περιστρέφουμε κατάλληλα και τον σύνδεσμο στα δάκτυλα, έτσι ώστε να αυτός να διατηρεί την οριζόντια θέση του επάνω στο επίπεδο του εδάφους.
- Στο τρίτο βήμα της διαδικασίας, ρυθμίζεται η κατακόρυφη θέση του συνδέσμου της λεκάνης σε κάθε χαρακτηριστική στάση, έτσι ώστε τα πόδια του μοντέλου να μην βρίσκονται κάτω από το προκαθορισμένο επίπεδο του εδάφους. Αυτή η διαδικασία διευκολύνεται από την ιεραρχική δομή του μοντέλου: αφού υπολογιστεί η κατακόρυφη θέση του συνδέσμου στη φτέρνα ή του συνδέσμου στα δάκτυλα για τη φάση αναμονής και τη στάση έκτασης, αν αυτή αντιστοιχεί σε θέση κάτω από το επίπεδο του εδάφους, ο σύνδεσμος της λεκάνης μετασχηματίζεται κατά την αντίθετη κατεύθυνση. Λόγω της ιεραρχικής δομής του μοντέλου, όλα οι υπόλοιποι σύνδεσμοι θα ακολουθήσουν την κίνηση της λεκά-

νης και έτσι ολόκληρο το μοντέλο θα βρεθεί επάνω από το επίπεδο που έχει οριστεί σαν έδαφος.

- Στο σημείο αυτό, έχουμε δημιουργήσει τις τέσσερις χαρακτηριστικές στάσεις για το μοντέλο και έχουμε εξασφαλίσει ότι τα πόδια δεν διαπερνούν το έδαφος. Στο τελευταίο αυτό βήμα ρυθμίζεται η οριζόντια θέση και κίνηση του συνδέσμου "λεκάνη", έτσι ώστε τα πόδια του μοντέλου να μη γλιστράνε κατά την κίνηση. Για να γίνει αυτό δυνατό, θα πρέπει κατά την μετάβαση από την μία χαρακτηριστική στάση στην άλλη οι κατάλληλοι σύνδεσμοι του ποδιού που βρίσκεται κάθε φορά στην φάση της στήριξης να διατηρούν σταθερή την θέση τους. Για να γίνει αυτό δυνατό, αρκεί να υπολογίσουμε τη μετατόπιση που υφίσταται ο καθένας, ως αποτέλεσμα των μετασχηματισμών που έχουν γίνει μέχρι τώρα στο μοντέλο, και στην συνέχεια να μετατοπίσουμε το σύνδεσμο της λεκάνης κατά την αντίθετη κατεύθυνση. Οι υπόλοιποι σύνδεσμοι θα ακολουθήσουν την μετατόπιση χάρη στην ιεραρχική δομή του μοντέλου.

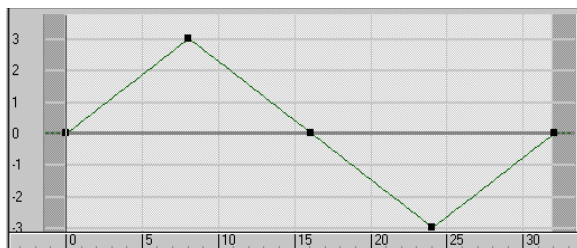
Συνοψίζοντας την παραπάνω διαδικασία, γίνεται αντιληπτό ότι τα δύο πρώτα στάδια αφορούν τη σχεδίαση της στάσης του μοντέλου στις χαρακτηριστικές στάσεις, ενώ τα επόμενα στοχεύουν στη σωστή απόδοση της κίνησης σε ότι αφορά την φυσική προσομοίωση και τους σχετικούς νόμους. Το Σχήμα 5-7 απεικονίζει τη σύνθεση των στάσεων που αφορούν τα χαρακτηριστικά καρέ της φυσιολογικής βάδισης, για έναν κύκλο που ξεκινάει με την έναρξη της φάσης στήριξης για το δεξιό πόδι.



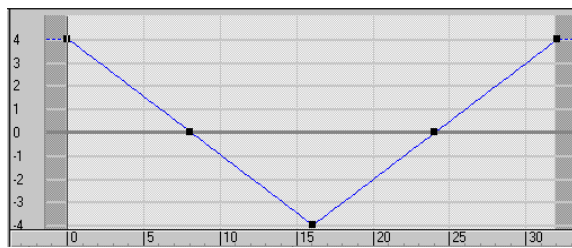
Σχήμα 5-7: Σύνθεση των χαρακτηριστικών στάσεων στη φυσιολογική βάδιση

Για να αποδοθεί η αίσθηση της συνεχούς κίνησης, θα πρέπει να συνθέσουμε τα ενδιάμεσα καρέ για το συνθετικό αντικείμενο, με βάση τα τέσσερα χαρακτηριστικά καρέ που σχεδιάσαμε, δηλαδή να υπολογίσουμε τις τιμές για τις γωνίες στους συνδέσμους ανάμεσα στα μέλη του σκελετικού μοντέλου. Με βάση τα πειράματα που διεξήχθησαν, φαίνεται ότι αρκεί η απλούστερη περίπτωση της γραμμικής παρεμβολής για να αποδοθεί μια ρεαλιστική ακολουθία συνθετικών καρέ. Έτσι, τα συνολικά δεδομένα που περιγράφουν την κίνηση μειώνονται από το σύνολο των γωνιών ανάμεσα στους συνδέσμους σε κάθε καρέ, έστω και μετά την αριθμητική κωδικοποίηση, σε τέσσερα σύνολα σε συμμετρικές χρονικές στιγμές της κίνησης. Οι καμπύλες που περιγράφουν

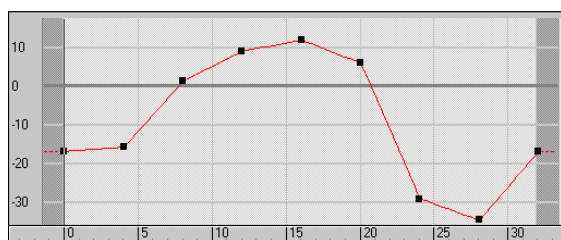
τις περιστροφές που συμβαίνουν στις αρθρώσεις του μοντέλου κατά την διάρκεια του κύκλου βάδισης που κατασκευάσαμε φαίνονται στα σχήματα .



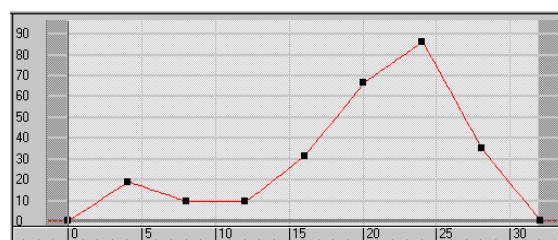
Περιστροφή της λεκάνης στον άξονα y



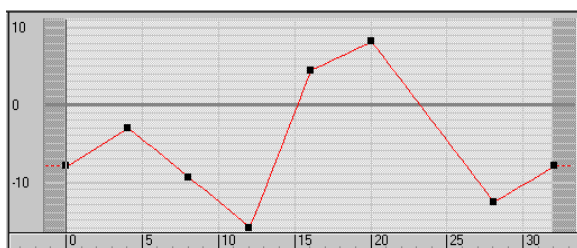
Περιστροφή της λεκάνης στον άξονα z



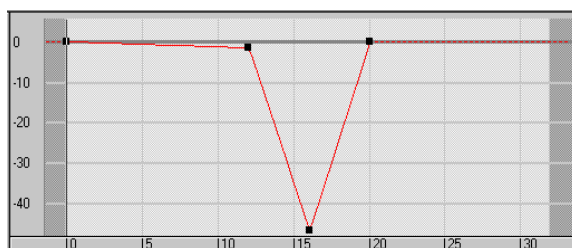
Περιστροφή των μηρών στον άξονα x



Περιστροφή του γονάτου στον άξονα x



Περιστροφή του αστράγαλου στον άξονα x



Περιστροφή των δακτύλων στον άξονα x

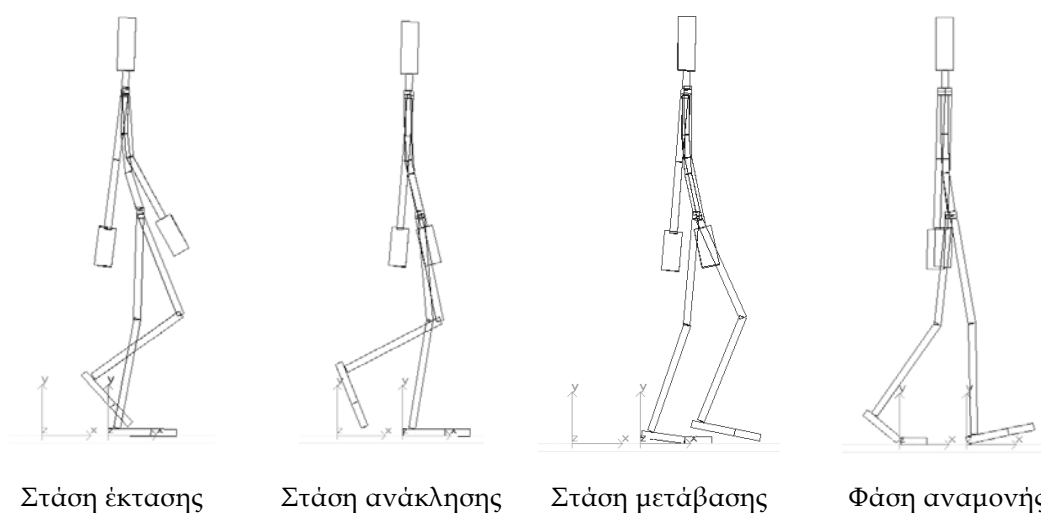
Σχήμα 5-8: Περιστροφή στους συνδέσμους των ποδιών κατά τη σύνθεση της φυσιολογικής βάδισης

5.4 Ο ΧΑΡΑΚΤΗΡΑΣ ΤΗΣ ΒΑΔΙΣΗΣ

Η περιοδική κίνηση που περιγράφηκε από τις παραπάνω χαρακτηριστικές στάσεις εντάσσεται στην έννοια της φυσιολογικής βάδισης, όπως τουλάχιστον αυτή γίνεται αντιληπτή από ανθρώπους που ανήκουν σε διαφορετικό πολιτιστικό υπόβαθρο. Η παραμετρική αναπαράσταση του κύκλου βάδισης σημαίνει ότι μεταβάλλοντας τις περιστροφές στις αρθρώσεις του μοντέλου στις τέσσερις χαρακτηριστικές φάσεις, είναι δυνατό να συνθέσουμε ένα σύνολο από κύκλους βάδισης που να διαφέρουν από το φυσιολογικό ή συνηθισμένο. Σε κάθε περίπτωση, οι πραγματικές τιμές για τις περιστροφές στις αρθρώσεις που περιγράφουν την στάση του μοντέλου σε κάθε μία από τις χαρακτηριστικές φάσεις εξαρτώνται από τις αναλογίες του μοντέλου. Αλλάζοντας τις τιμές των ακροτάτων, διατηρώντας όμως ταυτόχρονα το σχήμα των καμπυλών που περιγράφουν την συμπεριφορά των γωνιών στις αρθρώσεις του ποδιού, όπως αυτές φαίνονται στο Σχήμα 5-8, μπορούμε να αποδώσουμε ένα συγκεκριμένο "χαρακτήρα" στον κύκλο βάδισης.

Για την δημιουργία των χαρακτηριστικών καρέ που θα περιγράψουν το συνθετικό κύκλο βάδισης θα πρέπει να ακολουθείται η διαδικασία που ορίζεται από τα τέσσερα βήματα της προηγούμενης παραγράφου, έτσι ώστε να επιβεβαιώνεται η σωστή φυσική εξέλιξη του κύκλου βάδισης. Σε κάθε περίπτωση αφού κάνουμε τις απαραίτητες αλλαγές στις γωνίες των αρθρώσεων του μοντέλου, τα δύο τελευταία βήματα επαναλαμβάνονται, έτσι ώστε να διορθώσουμε τα προβλήματα που μπορεί να έχουν ως αποτέλεσμα να γλιστράνε τα πόδια του μοντέλου, ή να βρίσκονται κάτω από το επίπεδο του εδάφους.

Συνολικά, με τη μεταβολή μιας παραμέτρου που αφορά γωνίες στους συνδέσμους του ποδιού που στην συγκεκριμένη στάση βρίσκεται σε στήριξη, όλες οι άλλες γωνίες στο πόδι αυτό θα πρέπει να προσαρμοστούν, έτσι ώστε η φτέρνα ή τα δάχτυλα να παραμείνουν οριζόντια. Ο παραπάνω περιορισμός σημαίνει ότι το άθροισμα των γωνιών στις αρθρώσεις του ισχίου και του γονάτου θα πρέπει να είναι ίσο με την γωνία στον αστράγαλο. Αντίστοιχα, στη φάση αναμονής το άθροισμα των γωνιών στο ισχίο, στο γόνατο και στον αστράγαλο, θα πρέπει να είναι ίσο με την γωνία στα δάχτυλα, ενώ το ίδιο ισχύει και για το πίσω πόδι στη φάση έκτασης. Στη φάση αυτή οι σύνδεσμοι των ποδιών μπορούν να βρεθούν και κάτω από το επίπεδο του εδάφους, αλλά το πρόβλημα αυτό διορθώνεται με την επανάληψη των δύο τελευταίων βημάτων της διαδικασίας προσαρμογής, αφού γίνουν οι επιθυμητές αλλαγές. Αντίστοιχα, όταν θα πρέπει να μεταβληθεί η γωνία σε μία άρθρωση του ποδιού που βρίσκεται σε αιώρηση, αυτή περιορίζεται μόνο από όρια περιστροφής που έχουν αρχικά τεθεί κατά την κατασκευή του μοντέλου. Πάντως, οι σύνδεσμοι του ποδιού αυτού δεν θα πρέπει να βρίσκονται κάτω από το οριζόντιο επίπεδο του ποδιού που βρίσκεται στην ίδια στάση στη φάση της στήριξης, γιατί τότε θα προκύψουν λάθη στο τέλος της διαδικασίας προσαρμογής σε σχέση με το δάπεδο και την ολίσθηση.



Σχήμα 5-9: Απεικόνιση των χαρακτηριστικών καρέ στη σύνθεση της βαριεστημένης κίνησης

Συνήθως οι αλλαγές που αφορούν το χαρακτηριστικό καρέ που ορίζει τη φάση έκτασης περιορίζονται στον ορισμό του μήκους του διασκελισμού, γιατί σε αυτή την στάση και τα δύο πόδια βρίσκονται στο έδαφος και επομένως για οποιαδήποτε άλλη αλλαγή θα πρέπει να ρυθμιστούν οι γωνίες και στα δύο πόδια ώστε τα πέλματα να βρίσκονται στο ίδιο οριζόντιο επίπεδο. Αυτό συμβαίνει γιατί η μεταβολή μιας μόνο παραμέτρου θα είχε ως αποτέλεσμα, ακόμα και με-

τά την επανάληψη των δύο τελευταίων βημάτων της διαδικασίας προσαρμογής, το ένα από τα δύο πόδια να βρίσκεται κάτω από το επίπεδο του εδάφους. Εκτός από αυτό, η αλλαγή στο μήκος του διασκελισμού έχει ως αποτέλεσμα και την αλλαγή της ταχύτητας της κίνησης, αφού οι θέσεις στο χρόνο των χαρακτηριστικών στάσεων και των καρέ κλειδιών που τις περιγράφουν είναι σταθερές. Το Σχήμα 5-9 εμφανίζει τα χαρακτηριστικά καρέ για τη βαριεστημένη βάδιση, μετά τη διαδικασία προσαρμογής στο δάπεδο.

Μορφολογία και Εκφράσεις Ανθρώπινων Κεφαλών

Η αποτελεσματική σημασιολογική αναπαράσταση ανθρώπινων κεφαλών είναι μια από τις ταχύτερα εξελισσόμενες ιδέες στις εφαρμογές επικοινωνίας ανθρώπου – μηχανής. Αυτό οφείλεται στη δυνατότητα των σύγχρονων οικιακών υπολογιστών να αποδίδουν σε πραγματικό χρόνο τα σχετικά μοντέλα με ικανοποιητικό βαθμό ρεαλισμού και να συνεργάζονται αποτελεσματικά ακόμα και μέσω δικτύων χαμηλού εύρους ζώνης, αλλά κυρίως στην αυξημένη ζήτηση για εφαρμογές που παρουσιάζουν ένα ανθρώπινο πρόσωπο σαν τμήμα της διασύνδεσης με το χρήστη. Πραγματικά, συγκριτικές δοκιμές [35] δείχνουν ότι δικτυακά sites που προσφέρουν υπηρεσίες ηλεκτρονικού εμπορίου και χρησιμοποιούν τέτοια μοντέλα για την επικοινωνία με το χρήστη – πελάτη προσελκύουν πολύ μεγαλύτερο ενδιαφέρον και συνεχόμενες επισκέψεις, από ότι τα συνηθισμένα sites που χρησιμοποιούν συστήματα επιλογής βασισμένα στο κείμενο. Εκτός από αυτό, η αποτελεσματική συνθετική αναπαράσταση ανθρώπινων προσώπων μπορεί να χρησιμοποιηθεί και σε μια σειρά από άλλες εφαρμογές, είτε γενικού ενδιαφέροντος όπως η διασκέδαση, είτε επικεντρωμένου, όπως η ιατροδικαστική ή η πλαστική χειρουργική. Στο κεφάλαιο αυτό περιγράφεται η μορφολογία ενός ανθρώπινου προσώπου από την άποψη των μυών που συναντάμε σε αυτό και παρατίθενται μέθοδοι για τη συστηματική περιγραφή του μέσω αποδεκτών μετρήσεων. Η γνώση της δράσης των μυών σε ένα πρόσωπο είναι απαραίτητη για τη σωστή αναπαράσταση συναισθημάτων και ομιλίας σε ένα συνθετικό μοντέλο, ενώ η συστηματική μέθοδος μέτρησης μπορεί να μας βοηθήσει να προσαρμόσουμε ένα γενικής μορφής συνθετικό μοντέλο σε ένα συγκεκριμένο πραγματικό ή να το κατασκευάσουμε από την αρχή.

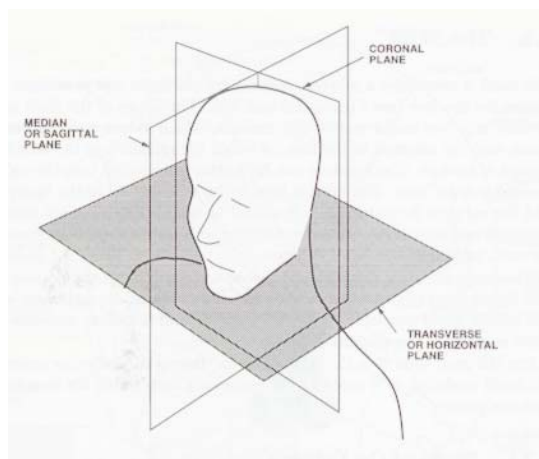
6.1 ΕΙΣΑΓΩΓΗ

Το σχήμα και η λειτουργία του ανθρώπινου σώματος μελετήθηκαν με μεγάλη λεπτομέρεια από καλλιτέχνες, κυρίως κατά την περίοδο της Αναγέννησης και με τη βοήθεια του σκίτσου, ώστε να γίνει δυνατή η κατά το δυνατόν ρεαλιστική και λεπτομερής απεικόνισή του. Ένας πρόσθετος στόχος που αφορά τη δημιουργία τρισδιάστατων μοντέλων προσώπου είναι η δημιουργία μοντέλων που θα φαίνονται ρεαλιστικά όχι μόνο στις στατικές εικόνες, αλλά και στις ακολουθίες εικόνων. Έτσι, οι μελέτες που προκύπτουν από την παρατήρηση θα πρέπει να συνδυαστούν με

ιατρικές αναλύσεις πάνω στην ανθρώπινη ανατομία με τέτοιο τρόπο ώστε το αποτέλεσμα να είναι σωστό αισθητικά, αλλά και αποδεκτό από την πλευρά της φυσιολογίας.

Ξεκινώντας μια περιγραφή της μορφολογίας της ανθρώπινης κεφαλής, ας υποθέσουμε ότι θεωρούμε τη λεγόμενη ανατομική θέση του ανθρώπινου σώματος, δηλαδή τη θέση κατά την οποία το άτομο στέκεται στητό με τα χέρια στα πλάγια και με το πρόσωπο και τις παλάμες να κοιτάζουν μπροστά. Εδώ, η θέση των μελών του σώματος περιγράφεται σε σχέση με αυτή τη θέση και τις τρεις κατηγορίες νοητών επιπέδων (Σχήμα 6-1) που περιγράφονται ως εξής:

- *ενδιάμεσο (median) επίπεδο*: πρόκειται για ένα κατακόρυφο επίπεδο που περνάει από το κέντρο του σώματος και το χωρίζει σε δεξί και αριστερό μισό.
- *στεφανιαίο (coronal) επίπεδο*: πρόκειται για ένα κατακόρυφο επίπεδο, το οποίο χωρίζει το σώμα σε μπροστά και πίσω μισό και είναι σε δεξιά γωνία σε σχέση με το ενδιάμεσο επίπεδο.
- *εγκάρσια (transverse) επίπεδα*: αυτά τα επίπεδα βρίσκονται σε δεξιά γωνία σε σχέση τόσο με το ενδιάμεσο όσο και με το στεφανιαίο επίπεδο.



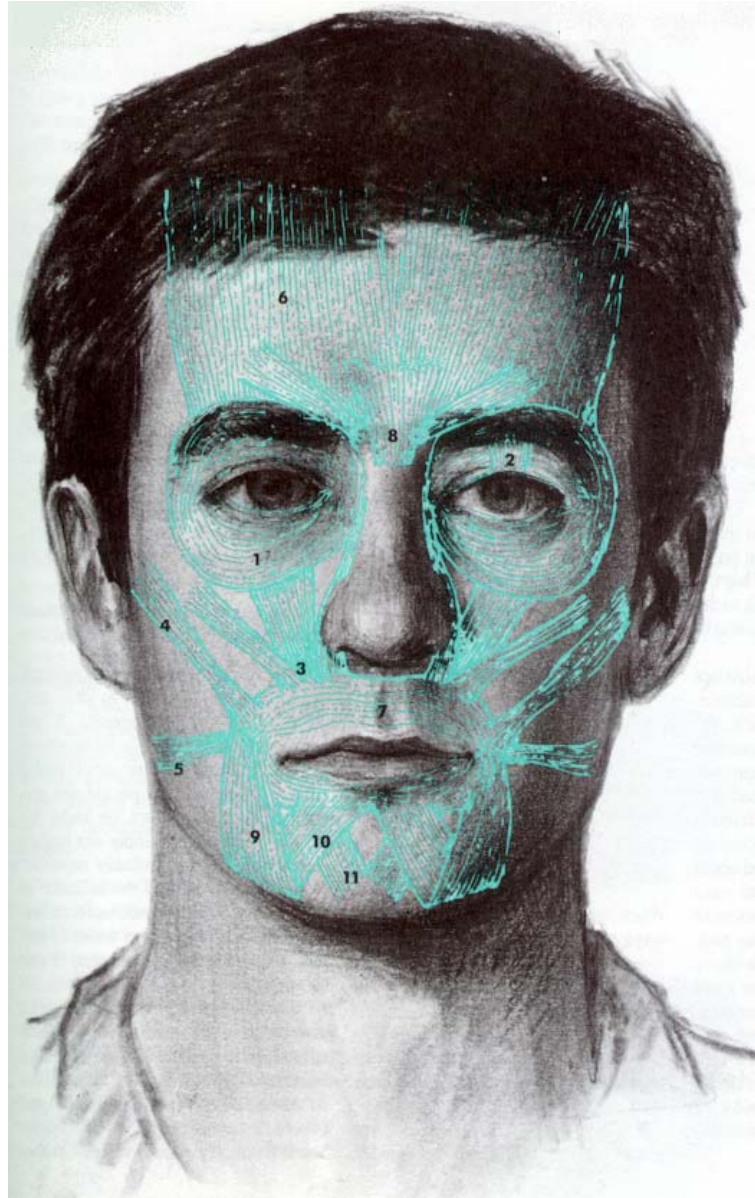
Σχήμα 6-1: Απεικόνιση των νοητών επιπέδων σε ένα ανθρώπινο κεφάλι

6.2 ΟΙ ΜΥΕΣ ΤΟΥ ΠΡΟΣΩΠΟΥ

Οι μύες είναι, γενικά, τα όργανα κίνησης για τον άνθρωπο, αφού από τη συστολή των μυών πραγματοποιούνται οι κινήσεις του ανθρώπινου σώματος. Οι μύες εκτείνονται συνήθως μεταξύ δύο κινούμενων μερών, όπως δύο οστών ή οστού και δέρματος ή μεταξύ δύο οργάνων. Σε ότι αφορά το πρόσωπο, οι σχετικοί μύες διαφέρουν από τους μύες στο υπόλοιπο σώμα στο γεγονός ότι είναι σχεδιασμένοι για να προκαλούν κίνηση δέρματος και όχι οστών. Όταν ένας μυς του προσώπου συστέλλεται, δημιουργεί ορατές αλλαγές όπως φουσκώματα ή ρυτίδες αφού, στην ουσία, μία περιοχή του προσώπου "διπλώνει".

Οι μύες του προσώπου αναφέρονται συχνά ως οι μύες των εκφράσεων. Μερικοί μύες του προσώπου πραγματοποιούν και άλλες σημαντικές λειτουργίες, όπως οι κινήσεις που πραγματοποιούνται κατά τη διάρκεια της ομιλίας ή της μάσησης της τροφής ή το άνοιγμα και το κλείσιμο των βλεφάρων. Οι μύες των εκφράσεων είναι επιφανειακοί και προσαρτημένοι στο τέλος τους σε ένα στρώμα υποδόριου λίπους και δέρματος, ενώ κάποιοι είναι προσαρτημένοι στο δέρμα στην αρχή και στο τέλος τους, όπως ο *obicularis oris*. Όταν οι μύες είναι χαλαροί, οι λιπώδεις ιστοί γεμίζουν τις κοιλότητες και αμβλύνουν τα γωνιώδη τμήματα, ώστε να είναι ορατό το σχή-

μα του κρανίου. Γενικά, το σύνολο δουλεύει σαν μία καλά οργανωμένη και συντονισμένη ομάδα, στα πλαίσια της οποίας κάθε μέλος έχει συγκεκριμένες λειτουργίες, μία εκ των οποίων είναι πρωταρχική. Έτσι, είναι σχετικά δύσκολο να διαχωριστούν ακριβώς οι λειτουργίες καθενός από αυτούς.



Σχήμα 6-2: Θέση και προσανατολισμός στους μύες του προσώπου

Στην πλειοψηφία των εκφράσεων δε συμμετέχει όλο το πρόσωπο, αλλά ένα μέρος του. Συνήθως, καθοριστικό ρόλο για την απόδοση μίας έκφρασης παίζουν τα φρύδια σε συνδυασμό με τα μάτια, καθώς και το στόμα και η περιοχή γύρω από αυτό. Κάτι τέτοιο θεωρείται λογικό αν σκεφτούμε ότι στις περιοχές αυτές συγκεντρώνονται οι περισσότεροι μύες, οπότε εκεί θα πραγματοποιούνται και οι περισσότερες αλλαγές. Άλλωστε, και η προσοχή του παρατηρητή ενός ανθρώπινου προσώπου εστιάζεται σε αυτές τις περιοχές: αν καλύψουμε την περιοχή της μύτης σε μία φωτογραφία προσώπου, η απώλεια πληροφορίας για την έκφραση δε θα είναι μεγάλη, ενώ

αν καλύψουμε την περιοχή των ματιών δε θα είναι δυνατός σε καμιά περίπτωση ο ασφαλής προσδιορισμός της έκφρασης.

Γενικά, οι μύες που συμμετέχουν στις εκφράσεις του προσώπου χωρίζονται σε άνω και κάτω πρόσωπο και μπορούν να ομαδοποιηθούν με κριτήριο τον προσανατολισμό τους, όπως φαίνεται στο Σχήμα 6-2. Τρεις τύποι μυών μπορεί να θεωρηθούν ότι πραγματοποιούν πρωταρχικές λειτουργίες: οι γραμμικοί -παράλληλοι μύες, οι οποίοι έλκουν προς τις γωνίες, όπως ο *zygomatic major* και ο *corrugator supercilii*, οι ελλειπτικοί - κυκλικοί μύες τύπου σφιγκτήρα, όπως ο *obicularis oris*, και οι επίπεδοι μύες, οι οποίοι συμπεριφέρονται σε μία σειρά γραμμικών μυών σε μία συγκεκριμένη περιοχή, όπως ο *frontalis*. Ο αριθμός των μυών που παίρνουν μέρος στις εκφράσεις του προσώπου είναι, σύμφωνα με την επιστήμη της ανατομίας είκοσι με είκοσι-έξι. Αν αγνοήσουμε τους μύες που αφορούν περιοχές του προσώπου ξένες προς τις ιδιαίτερα εκφραστικές, καταλήγουμε σε δώδεκα μύες, πέντε στην περιοχή των ματιών και επτά στην περιοχή του στόματος, οι οποίοι μπορούν να ομαδοποιηθούν σε κατηγορίες ανάλογα με την κίνηση του χείλους.

Μύες της περιοχής των ματιών

- *Orbicularis Oculi* (1): Ο μυς αυτός περιβάλλει το μάτι και λειτουργεί ως σφιγκτήρας που κλείνει το κλείνει. Προστατεύει το μάτι και κλείνει ερμητικά τα βλέφαρα, ώστε να προστατευτεί το μάτι από τη σκόνη, το έντονο ηλιακό φως ή από οποιοδήποτε αντικείμενο για να μην έρθει σε άμεση επαφή με το μάτι. Συνδέεται με το εσωτερικό της κοιλότητας του ματιού και με το δέρμα από το μάγουλο και πιέζει το μάτι όπως στον στραβισμό.
- *Corrugator Supercilii* (8): Πρόκειται για ένα μικρό μυ σχήματος πυραμίδας που είναι τοποθετημένος στο μέσο κάθε φρυδιού και χρησιμεύει στην κίνηση του φρυδιού πάνω ή κάτω. Ξεκινάει από τη ράχη της μύτης και ενώνεται με το δέρμα κάτω από το μέσο των φρυδιών. Χαμηλώνει τις εσωτερικές γωνίες των φρυδιών. Είναι γνωστός ως ο μυς της συνοφρύωσης.
- *Levator Palpebrae Superioris* (2): Ο μυς αυτός όταν συσπάται σηκώνει το άνω βλέφαρο. Ενώνεται με το άνω βλέφαρο, σηκώνει το βλέφαρο, όπως για παράδειγμα στην έκφραση της έκπληξης.

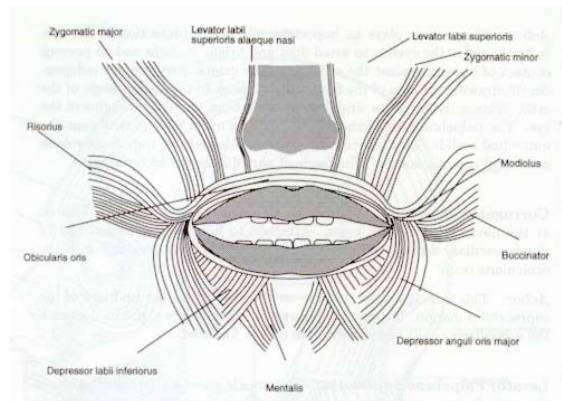
Μύες της περιοχής του στόματος

Η θέση και ο προσανατολισμός στους μύες της περιοχής του στόματος φαίνονται στο Σχήμα 6-3.

- *Orbicularis oris* (7): Ο μυς αυτός ελέγχει τα χείλη και βοηθάει στην ομιλία και στο φαγητό. Ξεκινάει από τους μύες στη γωνία του στόματος και μαζεύει και τεντώνει τα χείλη.
- *Levator labii Superiors* (3): Ο μυς αυτός σηκώνει το άνω χείλος και αποτελείται από τρεις κλάδους: ο εσωτερικός κλάδος αρχίζει από τη βάση της μύτης, ο μεσαίος κλάδος από το τέλος του κύκλου του ματιού και ο εξωτερικός κλάδος αρχίζει από το τόξο των ζυγωματικών. Όλοι οι κλάδοι καταλήγουν στην περιοχή του δέρματος πάνω από το άνω χείλος. Είναι γνωστός και ως «ο μυς του χλευασμού».
- *Zygomaticus Major* (4): Ξεκινάει από το τόξο των ζυγωματικών και καταλήγει στη γωνία του στόματος, βοηθάει στη δημιουργία του χαμόγελου και είναι γνωστός ως "ο μυς του χαμόγελου".
- *Risorius /platysma* (5): Ο *risorius* ξεκινάει από το πίσω μέρος της γνάθου και καταλήγει στη γωνία του στόματος, ενώ ο *platysma* ξεκινάει από το πάνω μέρος του θώρακα και καταλήγει και αυτό στη γωνία του στόματος. Οι δύο αυτοί μύες συνεργάζονται και τραβούν το στόμα, όπως συμβαίνει στο κλάμα. Είναι γνωστοί σαν "οι μύες που τεντώνουν τα χείλη".
- *Mentalis* (11): Ο μυς αυτός ανορθώνει το δέρμα στο πηγούνι, συμβάλλοντας στην προβολή και αναστροφή του, όπως όταν πίνουμε κάτι ξεκινάει ακριβώς κάτω από τα δόντια,

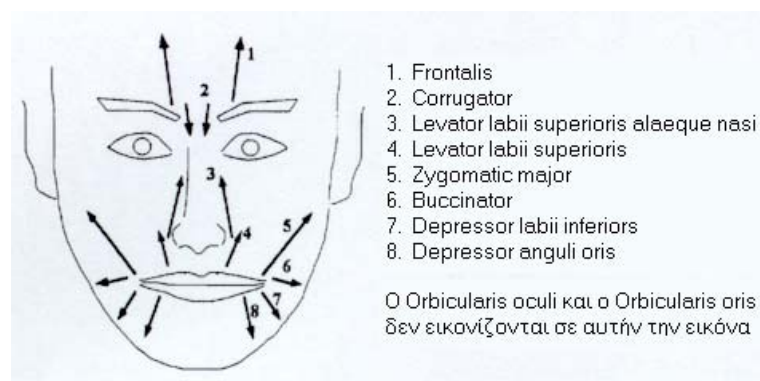
στην κάτω γνάθο και καταλήγει στο δέρμα που βρίσκεται στο κέντρο του πηγουνιού, στο οποίο και δημιουργεί ρυτίδες, ενώ πιέζει προς τα πάνω το κάτω χείλος. Είναι γνωστός ως "ο μυς του συνοφρυωμένου".

- *Depressor Labii Inferioris* (10): Πιέζει το κάτω χείλος προς τα κάτω όπως γίνεται κατά τη μάσηση της τροφής. Ξεκινάει από το τέλος του πηγουνιού και καταλήγει στο κάτω χείλος, το οποίο και κατεβάζει, όπως συμβαίνει στην ομιλία.
- *Frontalis* (6): Ξεκινάει σχεδόν από την κορυφή του κρανίου, κατά μήκος της γραμμής των μαλλιών και καταλήγει στο δέρμα κάτω από τα φρύδια. Σηκώνει τα φρύδια επάνω, όπως για παράδειγμα κατά την έκπληξη. Είναι γνωστός ως "ο μυς που σηκώνει τα φρύδια".
- *Triangularis* (9): Ξεκινάει από το χαμηλότερο σημείο της γνάθου και καταλήγει στην άκρη του στόματος. Είναι γνωστός ως ο μυς "έχω άσχημη μέρα".



Σχήμα 6-3: Απεικόνιση της θέσης και του προσανατολισμού στους μύες της περιοχής του προσώπου

Μπορεί κανείς να ισχυριστεί με ασφάλεια ότι οποιαδήποτε επιφανειακή μεταβολή στην εμφάνιση του προσώπου και, ειδικότερα, η μετακίνηση των χαρακτηριστικών του οφείλεται στη δράση κάποιου από τους μύες που σχετίζονται με αυτήν την περιοχή ή, συνηθέστερα, με μια ομάδα τέτοιων μυών. Οι κινήσεις που είναι αποτέλεσμα της δράσης των μυών φαίνονται στο Σχήμα 6-4.



Σχήμα 6-4: Απεικόνιση της κίνησης των μυών στο πρόσωπο

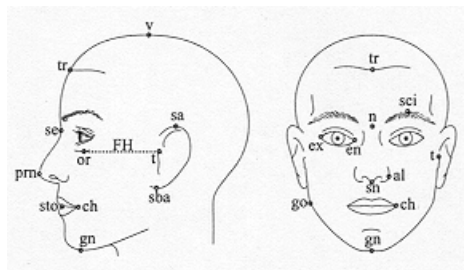
6.3 ΜΕΤΡΗΣΕΙΣ ΣΤΟ ΑΝΘΡΩΠΙΝΟ ΠΡΟΣΩΠΟ

Η ανθρωπομετρία είναι ο κλάδος της βιολογίας που ασχολείται με τη μελέτη των διαστάσεων του ανθρώπινου σώματος. Οι σχετικές συστηματικές μετρήσεις χρησιμοποιούνται σε ένα μεγάλο εύρος εφαρμογών, οι οποίες στηρίζονται στη γνώση αυτή για να προσαρμόσουν τις υπηρεσίες

τους για διάφορες ομάδες ανθρώπων. Για παράδειγμα, η γνώση των ορίων μεταβολής των μετρήσεων του ανθρώπινου σώματος μπορεί να συμβάλλει στη δημιουργία καταναλωτικών προϊόντων που θα ταιριάζουν σε ένα μεγάλο μέρος του πληθυσμού [30]. Ακόμα, στην ιατρική η σύγκριση των ανθρωπομετρικών δεδομένων με τα δεδομένα από τις μετρήσεις που πραγματοποιούνται σε ασθενείς πριν και μετά από μία εγχείρηση, βοηθούν σημαντικά κλάδους όπως η επανορθωτική ή η πλαστική χειρουργική [31] – αντίστοιχα, στην ιατροδικαστική, προβλέψεις για πιθανές μετρήσεις, οι οποίες προέρχονται από την ανθρωπομετρία συμβάλλουν αποφασιστικά στον προσδιορισμό της εμφάνισης ενός ατόμου μόνο από τα λείψανά του, ή στη δημιουργία φωτογραφιών εξαφανισμένων παιδιών, όπως θα πρέπει να είναι στην πραγματικότητα τη δεδομένη χρονική στιγμή [32].

Η διαδικασία της ανθρωπομετρικής μέτρησης αρχίζει με τη σημείωση συγκεκριμένων σημείων σε ένα αντικείμενο, τα οποία καλούνται "σημεία ορόσημο" και ορίζονται από ορατά ή απτά σημεία του αντικειμένου. Οι μετρήσεις που πραγματοποιούνται μεταξύ αυτών των σημείων, με σωστά επιλεγμένους αλγόριθμους και εργαλεία μέτρησης, γίνονται συνήθως με μεσοδιάστημα λίγων ημερών. Έτσι, οι μετρήσεις που αφορούν το ίδιο άτομο είναι εξαιρετικά αξιόπιστες, ενώ μπορούν εύκολα να γίνουν επιτυχώς συγκρίσεις μετρήσεων που προέρχονται από διαφορετικά άτομα.

Συνήθως τα αποτελέσματα προκύπτουν από ένα ευρέως χρησιμοποιούμενο σύνολο ανθρωπομετρικών μετρήσεων, το οποίο χρησιμοποιεί 47 σημεία ορόσημο για να περιγράψει το πρόσωπο [33]. Τα σημεία παίρνουν το όνομά τους από τη σύντμηση των ονομάτων των αντίστοιχων ανατομικών σημείων, για παράδειγμα η εσωτερική γωνία του ματιού συμβολίζεται με *en* από τη λέξη *endocanthion*, ενώ η κορυφή του χόνδρου μπροστά από το αυτί συμβολίζεται με *t* από το *tragion*. Ένα υποσύνολο αυτών των μετρήσεων φαίνονται στο Σχήμα 6-5.



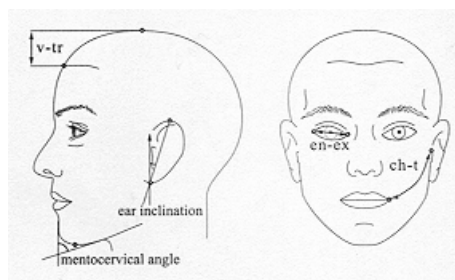
Σχήμα 6-5: Χαρακτηριστικά σημεία για τις μετρήσεις του προσώπου

Δύο από τα σημεία ορόσημο καθορίζουν την κανονική οριζόντια θέση του κεφαλιού. Το οριζόντιο επίπεδο ορίζεται από τις δύο γραμμές (σε οποιαδήποτε από τις δύο πλευρές του προσώπου), οι οποίες συνδέουν το σημείο *t* με το σημείο *or* (από το *orbitale*), το χαμηλότερο σημείο της κοιλότητας του ματιού στο κρανίο. Κατά τη μέτρηση, οι επιστήμονες της ανθρωπομετρίας ευθυγραμμίζουν το κεφάλι σύμφωνα με αυτό το οριζόντιο πεδίο, θέση η οποία είναι γνωστή ως οριζόντια θέση *Frankfurt* (FH) [34], οπότε οι μετρήσεις είναι πιο εύκολες και ακριβείς με σχέση με το συγκεκριμένο σύστημα συντεταγμένων. Επιπλέον, μία κατακόρυφη γραμμή ορίζεται από τα σημεία *n* (από το *nasion*), ένα χαρακτηριστικό μέρος του κρανίου ανάμεσα στα φρύδια, *sn* (από το *subnasale*), το κεντρικό, δηλαδή, σημείο συνάντησης της μύτης και του άνω χείλους και *gn* (*gnathion*), το χαμηλότερο σημείο του πηγουνιού.

Η εργασία [31] προτυποποιεί πέντε τύπους μετρήσεων για τα σημεία του προσώπου:

- η μικρότερη απόσταση μεταξύ δύο σημείων. Ένα παράδειγμα είναι *en-ex*, η απόσταση μεταξύ των σημείων στις γωνίες των ματιών.
- η αξονική απόσταση μεταξύ δύο σημείων. Η απόσταση που μετράται κατά μήκος ενός από τους άξονες του κανονικού συστήματος συντεταγμένων, με το κεφάλι να είναι στην FH θέση. Ένα παράδειγμα είναι η απόσταση *v-tr*, η κατακόρυφη, δηλαδή, απόσταση (διαφορά στο ύψος) μεταξύ της κορυφής του κεφαλιού (*v* από το *vertex*) και της γραμμής των μαλλιών (*tr* από το *trichion*).
- η εφαπτόμενη απόσταση μεταξύ δύο σημείων. Η απόσταση προσδιορίζεται από ένα προκαθορισμένο μονοπάτι στην επιφάνεια του προσώπου. Ένα παράδειγμα είναι η απόσταση *ch-t*, η επιφανειακή απόσταση, δηλαδή, από την άκρη του στόματος (*ch* από *cheilion*) μέχρι το *tragus*.
- η γωνία κλίσης μεταξύ δύο σημείων λαμβάνοντας υπόψη το κανονικό σύστημα συντεταγμένων. Ένα παράδειγμα είναι η κλίση του άξονα του αυτιού λαμβάνοντας υπόψη τον κατακόρυφο άξονα.
- η γωνία που δημιουργείται σε διάφορα σημεία του προσώπου, όπως η γωνία του πηγουνιού.

Οι μετρήσεις αυτές σε συγκεκριμένα σημεία του προσώπου φαίνονται στο Σχήμα 6-6.



Σχήμα 6-6: Πρακτικά παραδείγματα ανθρωπομετρικών μετρήσεων

Συνολικά, μπορούμε να περιγράψουμε 132 μετρήσεις στην περιοχή του προσώπου και του κεφαλιού. Μερικές από τις μετρήσεις αυτές σχηματίζουν ζεύγη, όταν υπάρχει αντιστοιχία μετρήσεων στο δεξί και στο αριστερό μισό του προσώπου. Γενικά, η συστηματική συλλογή ανθρωπομετρικών μετρήσεων συμβάλλει σημαντικά στη στατιστική έρευνα ομάδων εξεταζομένων ατόμων. Τα άτομα αυτά ομαδοποιούνται με βάση το γένος, τη φυλή, την ηλικία ή ακόμα και με βάση τη θελκτικότητα ή κάποιο σωματικό σύνδρομο. Η έρευνα των μέσων όρων και των διακυμάνσεων των μετρήσεων στα πλαίσια μίας κατηγορίας, παρέχει σημαντικές πληροφορίες σχετικά με τις αποκλίσεις που μπορούν να προκύψουν σε μία κατηγορία.

Εκτός από τη στατιστική των μετρήσεων, πραγματοποιούνται έρευνες σχετικά και με τις αναλογίες του σώματος, κάτι που ήδη είχε αρχίσει να ερευνάται από την εποχή του Leonardo da Vinci. Οι επιστήμονες ανακάλυψαν ότι οι αναλογίες δίνουν χρήσιμες πληροφορίες όσον αφορά στις σχέσεις μεταξύ των χαρακτηριστικών του ανθρώπου και είναι σημαντικές ενδείξεις σχετικά με την ομάδα στην οποία ανήκει το κάθε άτομο, ενώ πολλές αναλογίες προσώπου αποκαλύπτουν σημαντικές μορφολογικές διαφορές μεταξύ ομάδων πληθυσμού [34].

Παραμετρική Αναπαράσταση Ανθρώπινων Κεφαλών

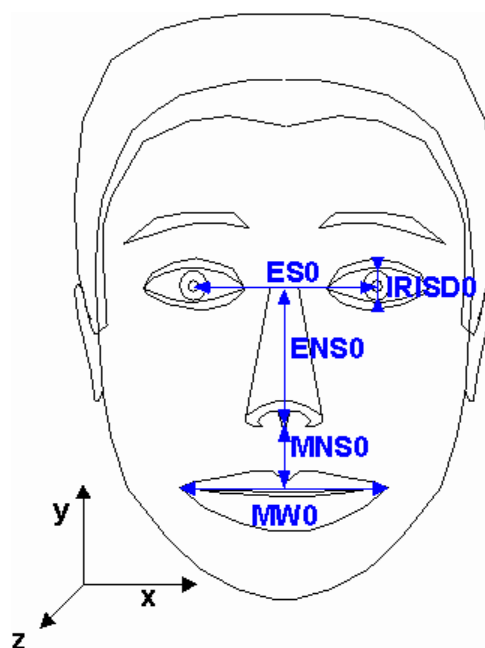
Αν και δεν αποτελεί ένα ιδιαίτερα πρόσφατο ερευνητικό αντικείμενο, η παραμετρική αναπαράσταση ανθρώπινων κεφαλιών με τη βοήθεια πολυγωνικών μοντέλων είναι ένας από τους τομείς που προσελκύουν το ενδιαφέρον τόσο από πλευράς έρευνας, όσο και από πλευράς εφαρμογών, ειδικά αφού πρόκειται για μια από τις πιο επαναστατικές καινοτομίες του προτύπου MPEG-4. Το πιο προφανές όφελος μιας τέτοιας ιδέας είναι η αφαίρεση των δεδομένων που απαιτούνται για την αναπαράσταση σε ένα σύνολο από χαρακτηριστικές παραμέτρους, οι οποίες είναι πολύ κοντά στη σημασιολογική αναπαράσταση τέτοιων αντικειμένων από τον άνθρωπο. Στην ουσία, αυτό επιτρέπει στο σχεδιαστή της εφαρμογής ή τον κωδικοποιητή να επικεντρωθεί στην απεικόνιση του περιεχομένου και όχι στην απόδοση των τεχνικών λεπτομερειών. Έτσι, γίνεται περισσότερο ορατός ο στόχος του προτύπου για αποτελεσματική υβριδική κωδικοποίηση συνθετικών και πραγματικών αντικειμένων, ενώ διευκολύνεται σημαντικά η απόδοση των αποτελεσμάτων σε δίκτυα χαμηλού εύρους ζώνης. Εδώ περιγράφεται ο τρόπος με τον οποίο το πρότυπο MPEG-4 χρησιμοποιεί τα μορφολογικά χαρακτηριστικά μιας ανθρώπινης κεφαλής για να παράγει την περιγραφή της γεωμετρίας τους και να δώσει δυνατότητες στις εφαρμογές για σωστή και αποτελεσματική εμφύχωση. Στη συνέχεια αναφέρεται μια συστηματική μέθοδος για την εκτίμηση ενός υποσυνόλου από τα μορφολογικά χαρακτηριστικά σε φωτογραφίες εμπρόσθιας και πλαϊνής όψης και προτείνεται η χρήση αυτών των μετρήσεων για την προσαρμογή ενός συνθετικού μοντέλου γενικής μορφής και η απόδοσή του με μια συνθετική εικόνα υφής.

7.1 ΕΙΣΑΓΩΓΗ

Το MPEG Group προέβλεψε έγκαιρα και με μεγάλη επιτυχία την εντεινόμενη σημασία των ομιλώντων ανθρώπινων μοντέλων κεφαλιών (talking heads) σε ένα εξαιρετικά μεγάλο εύρος εφαρμογών. Έτσι, φρόντισε να συμπεριλάβει στον πυρήνα του προτύπου MPEG-4 τις απαραίτητες δομές και λειτουργίες, ώστε να γίνει δυνατή η χρήση τέτοιων μοντέλων σε απαιτητικές, σε σχέση με την ποιότητα του αποτελέσματος, εφαρμογές, αλλά και σε τηλεπικοινωνιακές εφαρμογές point-to-point και multi-point. Ειδικά στον τελευταίο τομέα, έχουν αναπτυχθεί εμπορικές εφαρμογές που επιτρέπουν την εμφύχωση σε πραγματικό χρόνο και με ρυθμούς μετάδοσης από 300 μέχρι 2000 bits/sec. Εκτός από την απόδοση της εικόνας και της αίσθησης του animation, το

σχετικό τμήμα του προτύπου δίνει τη δυνατότητα ολοκλήρωσης με συνθετική ομιλία, η οποία μπορεί να παράγεται με βάση ένα συνοδευτικό κείμενο (text-to-speech).

Το πρότυπο περιγράφει ένα μοντέλο ανθρώπινου προσώπου με ουδέτερη έκφραση, με ένα σύνολο από σημεία αναφοράς κι ένα σύνολο από παραμέτρους εμφύχωσης (Facial Animation Parameters – FAPs), οι οποίες αντιστοιχούν σε ατομικές παραμορφώσεις του προσώπου. Η χρήση συγκεκριμένων, μη μηδενικών τιμών στις FAPs σε κάποια χρονική στιγμή θα έχει σαν αποτέλεσμα την απόδοση της εμφύχωσης στον αποκωδικοποιητή. Η ακριβής τιμή της παραμέτρου αντιστοιχεί στην ένταση της παραμόρφωσης σε επίπεδο σημασιολογικής δράσης [36], [37] κι όχι σε στοιχειώδεις μετασχηματισμούς του πολυγωνικού μοντέλου που χρησιμοποιείται. Για παράδειγμα, η έκφραση του χαμόγελου μπορεί να αποδοθεί με την ενεργοποίηση ενός κατάλληλα επιλεγμένου συνόλου από FAPs – η ένταση των επιμέρους παραμέτρων είναι αυτή που θα διακρίνει ένα αυθόρμητο, ειλικρινές χαμόγελο από ένα συγκρατημένο μειδίαμα. Αυτό ουσιαστικά σημαίνει ότι ο σχεδιαστής της εμφύχωσης ή ο κωδικοποιητής δεν είναι απαραίτητο να ασχοληθεί με την επιλογή των κορυφών του μοντέλου που θα μετασχηματιστούν, αλλά μόνο με την πραγματική ένταση του αποτελέσματος που πρέπει να αποδοθεί. Άλλωστε, αυτό είναι που ενδιαφέρει περισσότερο και γίνεται αντιληπτό από τον τελικό θεατή – χρήστη και δεν είναι εύκολο να μεταβληθεί αποτελεσματικά σε επίπεδο μετασχηματισμού πολυγώνων.



Σχήμα 7-1: Ένα μοντέλο ανθρώπινου προσώπου σε ουδέτερη έκφραση και οι μετρήσεις με βάση τις οποίες ορίζονται οι μονάδες FAP (FAPUs)

Για την αποκωδικοποίηση της ακολουθίας που περιγράφεται με FAPs, το τερματικό θα πρέπει να ενσωματώνει μια σειρά από κανόνες για τη "μετάφραση" των παραμέτρων εμφύχωσης, οι οποίοι εξαρτώνται στην πράξη από το συγκεκριμένο μοντέλο που χρησιμοποιείται. Αυτοί οι κανόνες μπορεί να είναι στατικά αποθηκευμένοι στον αποκωδικοποιητή, αλλά είναι δυνατόν να μεταδίδονται μαζί με το συνθετικό μοντέλο με τη μορφή Face Animation Tables (FAT), για να αποδώσουν την εμφύχωση με προσαρμοσμένο αποτέλεσμα. Εκτός από αυτό, είναι απαραίτητο να οριστούν μια σειρά από μονάδες για τις πραγματικές τιμές των FAPs, οι οποίες

να μπορούν να εφαρμοστούν σε συγκεκριμένα μοντέλα. Έτσι, το σύνολο των FAPUs (FAP Units) υπολογίζονται σαν υποδιαιρέσεις μιας σειράς από μετρήσεις ανάμεσα σε χαρακτηριστικά σημεία του προσώπου, όπως φαίνεται στο Σχήμα 7-1. Οι μονάδες που χρησιμοποιούνται για τις τιμές των FAPs υπολογίζονται με βάση τις παραπάνω μετρήσεις με βάση τις σχέσεις:

$$ES = \frac{ESo}{1000} \quad , \quad ENS = \frac{ENSo}{1000} \quad , \quad MNS = \frac{MNSo}{1000} \quad \text{και} \quad MW = \frac{MWo}{1000} \quad 7-1$$

Η σημασιολογική έννοια των αποστάσεων του Σχήματος 7-1 φαίνεται στον Πίνακα 7-1.

Μέτρηση	Σημασία	Μονάδα που προκύπτει
IRISD0	Περίμετρος της ίριδας (εξ ορισμού ίση με την κατακόρυφη απόσταση ανάμεσα στα βλέφαρα)	IRISD0 = IRISD0/1024
ES0	Απόσταση ανάμεσα στα μάτια	ES0/1024
ENS0	Απόσταση γραμμής ματιών - μύτης	ENS0/1024
MNS0	Απόσταση μύτης - στόματος	MNS0/1024
MW0	Πλάτος του στόματος	MW0/1024
AU	Μονάδα γωνιών	10 ⁻⁵ rad

Πίνακας 7-1: Ορισμός και υπολογισμός μονάδων FAPs

7.2 ΚΑΘΟΡΙΣΜΟΣ ΤΗΣ ΓΕΩΜΕΤΡΙΑΣ ΤΟΥ ΚΕΦΑΛΙΟΥ

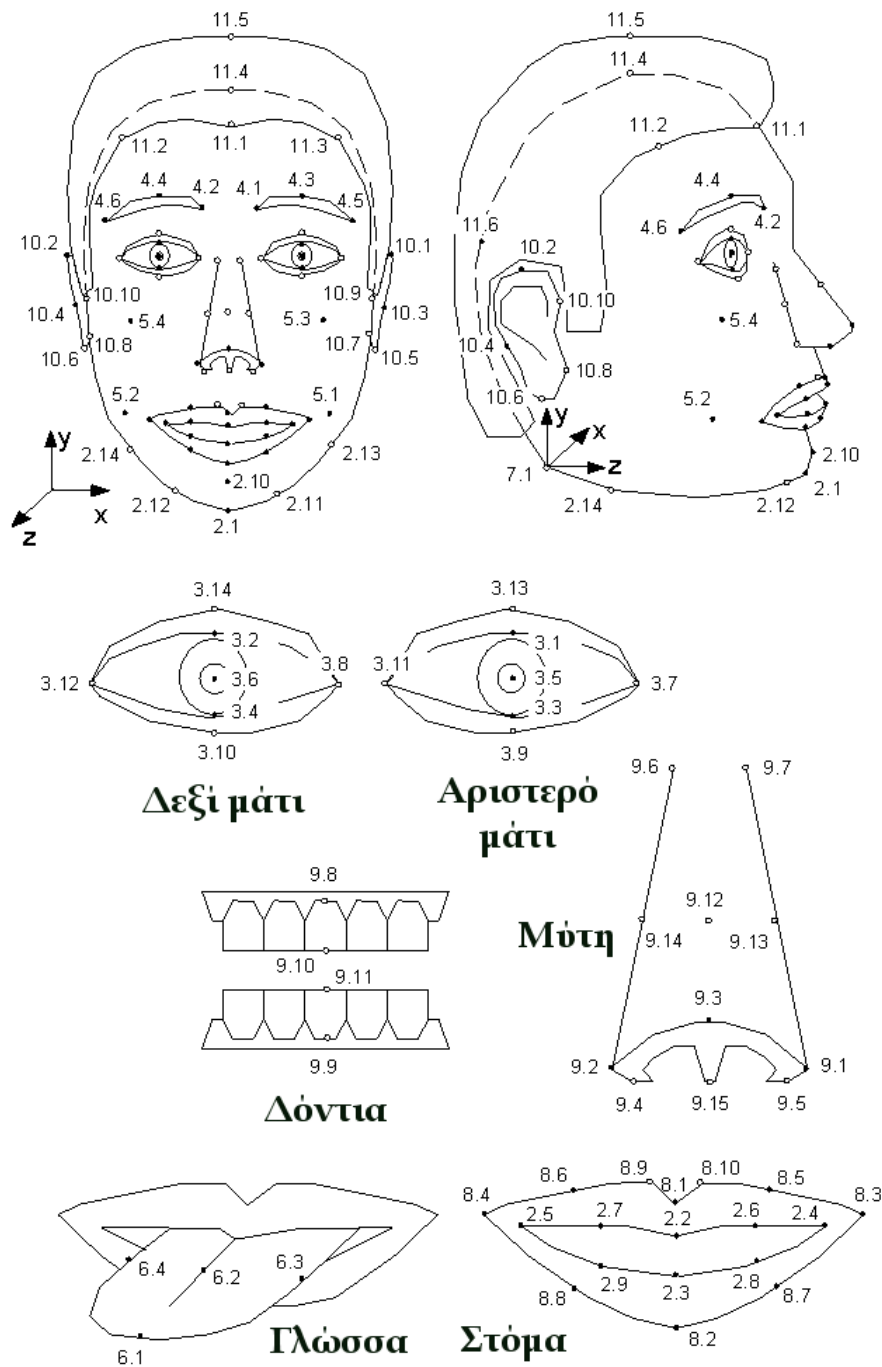
Για να μπορέσει να αποδοθεί η στατική ή χρονικά μεταβαλλόμενη έκφραση που περιγράφεται με τα FAPs, ο αποκωδικοποιητής θα πρέπει να διαθέτει τοπικά ένα συμβατό μοντέλο ανθρώπινου προσώπου. Εκτός από αυτό, το σύστημα είναι δυνατό να μεταδίδει μια σειρά από στατικές παραμέτρους που μπορούν να χρησιμοποιηθούν για την προσαρμογή του τοπικά ορισμένου μοντέλου ή ακόμα και να μεταδίδει ένα καινούριο μοντέλο από την αρχή. Σε κάθε περίπτωση, το πρότυπο εξασφαλίζει ότι η απόδοση των δράσεων που περιγράφουν τα FAPs μπορεί να γίνει με σωστό αισθητικό αποτέλεσμα.

Πάντως, το πρότυπο δεν περιγράφει με αναλυτικό τρόπο την προσαρμογή του μοντέλου σε συγκεκριμένες μετρήσεις, παρά μόνο το σύνολο των δεδομένων και των γεωμετρικών σημείων που απαιτούνται για τον πλήρη ορισμό του. Αν όλα αυτά τα δεδομένα δεν είναι διαθέσιμα, τότε τα υπάρχοντα μπορούν να χρησιμοποιηθούν σαν βάση για την προσαρμογή, με σεβασμό στους κανόνες φυσιολογίας που διέπουν ένα ανθρώπινο πρόσωπο και που περιγράφηκαν στο προηγούμενο κεφάλαιο. Οι γενικές απαιτήσεις που θέτει το πρότυπο για ένα συμβατό μοντέλο ανθρώπινου προσώπου είναι οι εξής [12]:

- το μοντέλο κοιτάζει προς τη θετική κατεύθυνση του άξονα z
- όλοι οι μύες δεν εμφανίζουν διέγερση
- τα βλέφαρα εφάπτονται στην ίριδα
- η κόρη έχει διάμετρο ίση με το ένα τρίτο της διαμέτρου της ίριδας

- το στόμα είναι κλειστό, τα χείλη εφάπτονται και η γραμμή που σχηματίζουν είναι οριζόντια και στο ίδιο ύψος με τα άκρα του στόματος
- η γλώσσα είναι επίπεδη και οριζόντια, με το άκρο της να εφάπτεται στο όριο ανάμεσα στα πάνω και κάτω δόντια.

Η πραγματική γεωμετρία του μοντέλου καθορίζεται με βάση συγκεκριμένα σημεία που περιέχονται στη ροή των FDPs (Face Definition Parameters) και ονομάζονται *feature points*. Στην πιο απλή περίπτωση, το μοντέλο που ορίζεται με αυτόν τον τρόπο αρκεί να έχει τόσες κορυφές, όσες και τα σημεία αυτά, δηλαδή σχεδόν πενήντα – ένα τέτοιο μοντέλο πάντως είναι δύσκολο να ικανοποιήσει το χρήστη της εφαρμογής από αισθητική άποψη, αν και ίσως είναι επαρκές για τηλεπικοινωνιακές εφαρμογές σε δίκτυα χαμηλού εύρους ζώνης. Τα σημεία αυτά ομαδοποιούνται με βάση την περιοχή του προσώπου που βρίσκονται και φαίνονται στο Σχήμα 7-2.



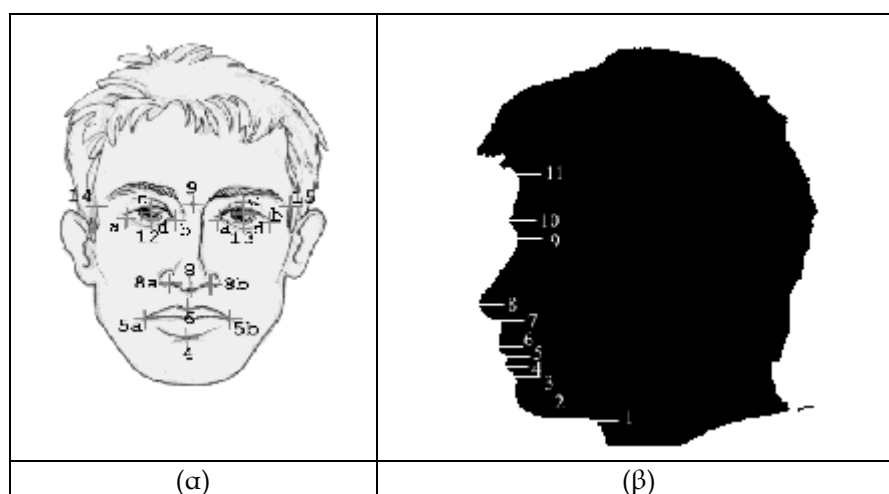
Σχήμα 7-2: Χαρακτηριστικά σημεία που εντάσσονται στο σύνολο των FDPs

Αν πρόκειται να χρησιμοποιηθεί η διαδικασία απεικόνισης υφής για την απόδοση του μοντέλου, η σχετική εικόνα θα πρέπει να μεταδοθεί και αυτή σαν τμήμα των δεδομένων για την περιγραφή του μοντέλου. Για να γίνει δυνατή η σκίαση των πολυγώνων που αποτελούν το μοντέλο με τη χρήση της εικόνας απεικόνισης υφής, θα πρέπει να καθοριστεί το σημείο της εικόνας στο οποίο θα απεικονιστεί καθεμιά από τις κορυφές του μοντέλου. Έτσι, οι συντεταγμένες απεικόνισης υφής (mapping coordinates - [38]) περιέχονται και αυτές στο σύνολο των FDPs κι επιτρέπουν τη χρήση πραγματικών εικόνων σε συνδυασμό με ένα συνθετικό μοντέλο.

7.3 ΥΠΟΛΟΓΙΣΜΟΣ ΤΩΝ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ ΣΗΜΕΙΩΝ

Όπως έχουμε αναφέρει, το πρότυπο MPEG-4 δεν προτυποποιεί τη μέθοδο υπολογισμού των παραμέτρων που απαιτούνται για την περιγραφή ενός συνθετικού κεφαλιού, παρά μόνο τη χρήση τους στον αποκωδικοποιητή. Στις εργασίες [40] και [41] προτείνεται μια μέθοδος υπολογισμού ενός υποσυνόλου των χαρακτηριστικών σημείων που περιέχονται στα FDPs και της σχετικής εικόνας απεικόνισης υψής, η οποία χρησιμοποιείται στην εργασία [39] σε συνδυασμό με τον αλγόριθμο απλοποίησης.

Η είσοδος του προτεινόμενου συστήματος είναι δύο πραγματικές φωτογραφίες ενός προσώπου, από τις οποίες μια απεικονίζει την πλάγια όψη του προσώπου και η άλλη την εμπρόσθια. Από αυτές τις όψεις δεν είναι δυνατός ο ασφαλής καθορισμός των θέσεων για όλα τα χαρακτηριστικά σημεία που περιέχονται στη ροή με τις FDPs, παρά μόνο ενός υποσυνόλου που εντοπίζεται κύρια στις εκφραστικές περιοχές του προσώπου, δηλαδή στην περιοχή των ματιών και φρυδιών και στην περιοχή γύρω από το στόμα. Αυτό συμβαίνει γιατί κάποια από τα χαρακτηριστικά του Σχήματος 7-2 είναι συνήθως κρυμμένα σε μια φυσιολογική φωτογραφία, όπως τα σημεία της περιοχής των δοντιών (9.8 μέχρι 9.11) αν το στόμα είναι κλειστό. Ένας άλλος λόγος που περιορίζει τη δυνατότητα υπολογισμού των χαρακτηριστικών σημείων είναι η έλλειψη διαχωρισμού τους από το υπόλοιπο πρόσωπο, έτσι ώστε να μπορέσουμε να τα εντοπίσουμε με μεθόδους όρασης υπολογιστών. Για παράδειγμα, τα σημεία 5.3 και 5.4 δε σχετίζονται με κάποιο από τα προφανή χαρακτηριστικά του προσώπου κι έτσι δεν είναι δυνατό να διαχωριστούν από το υπόλοιπο δέρμα της περιοχής. Πάντως, τα σύνολο των σημείων που υπολογίζεται είναι αρκετά ικανοποιητικό για εφαρμογές σε περιβάλλοντα χαμηλού εύρους ζώνης ή για εφαρμογές όπου το αισθητικό αποτέλεσμα παίζει δευτερεύοντα ρόλο σε σχέση με το υπόλοιπο υλικό.



Σχήμα 7-3: Εντοπισμός χαρακτηριστικών σημείων. (α) Εμπρόσθια όψη, (β) πλάγια όψη

Το σύνολο των σημείων που εντοπίζονται φαίνεται στο Σχήμα 7-3. Με βάση αυτά τα σημεία, μπορούμε να υπολογίσουμε μια σειρά από μετρήσεις που μας δίνουν ανθρωπομετρικές πληροφορίες για το συγκεκριμένο πρόσωπο. Οι μετρήσεις αυτές φαίνονται στον Πίνακα 7-2.

Μετρήσεις στην πλάγια όψη	Περιγραφή
Γωνία 1-2-3	Μέτρο του σαγονιού
Γωνία 7-8-9	Μέτρο της μύτης
Λόγος $d(8,10)/d(2,8)$	Κατακόρυφη αναλογία
Γωνία 5-6-7	Μέτρο του άνω χείλους
Γωνία 3-4-5	Μέτρο του κάτω χείλους
Γωνία 4-5-6	Γωνία στα χείλη
Μετρήσεις στην εμπρόσθια όψη	Περιγραφή
Λόγος $d(14,15)/d(4,9)$	Μέτρο πλάτους του προσώπου
Λόγος $d(12,13)/d(14,15)$	Απόσταση των ματιών
Λόγος $d(5a,5b)/d(14,15)$	Πλάτος του στόματος
Λόγος $d(4,6)/d(14,15)$	Ύψος του στόματος
Λόγος $d(12a,12b)/d(14,15)$	Πλάτος των ματιών
Λόγος $d(12c,12d)/d(14,15)$	Ύψος των ματιών
Λόγος $d(8,9)/d(14,15)$	Ύψος της μύτης
Λόγος $d(8a,8b)/d(14,15)$	Πλάτος της μύτης

Πίνακας 7-2: Ανθρωπομετρικές μετρήσεις στο πρόσωπο που προκύπτουν από την αντίχνευση των χαρακτηριστικών σημείων. Η έκφραση $d(i, j)$ αντιστοιχεί στην Ευκλείδεια απόσταση ανάμεσα στα σημεία i και j

Σε ότι αφορά τον εντοπισμό των χαρακτηριστικών σημείων στην εμπρόσθια όψη, η διαδικασία στηρίζεται στην ιεραρχική και αξιόπιστη επιλογή συγκεκριμένων blocks στην εικόνα και τη χρήση προτυποποιημένων τεχνικών σε καθένα από αυτά. Σε σχέση με την εφαρμογή παρόμοιων διαδικασιών σε εικόνες γενικού περιεχομένου, το ευτύχημα σε σχέση με τη χρήση τους σε φωτογραφίες ανθρώπων, είναι ότι εδώ είναι εύκολο να οριστούν οι περιοχές ενδιαφέροντος [42]. Η τεχνική που περιγράφεται εδώ αποτελεί ένα συνδυασμό των ιδεών του ταίριασματος προτύπων [43], του φιλτραρίσματος Gabor [44] και της θεωρίας των ιδιοχαρακτηριστικών (eigenfeatures) [45].

Γενικά, οι ανθρωπομετρικές μετρήσεις βοηθούν στη μείωση της περιοχής της εικόνας μέσα στην οποία πρέπει να αναζητηθεί το κατάλληλο block ή το χαρακτηριστικό σημείο, αφού κανονικοποιηθεί κατάλληλα η εικόνα. Για παράδειγμα, μπορούμε να αναζητήσουμε την κατακόρυφη θέση του άκρου της μύτης (σημείο 8 στο Σχήμα 7-3) σε μια ζώνη κοντά στο μέσο της εικόνας και λίγο πάνω από αυτό, ενώ η οριζόντια θέση θα πρέπει να είναι κοντά στη μεσαία στήλη της. Αντίστοιχα, τα μάτια θα πρέπει να αναζητηθούν πάνω από το επίπεδο της μύτης και εκατέρωθεν της κατακόρυφης θέσης της, ενώ το στόμα κάτω από τη μύτη, με τα άκρα του συμμετρικά γύρω από την κατακόρυφη θέση του άκρου της. Αυτή η προεπεξεργασία είναι ιδιαίτερα σημαντική σε μεθόδους με μεγάλο υπολογιστικό βάρος, όπως το ταίριασμα προτύπων.

Αφού επιλεγεί το κατάλληλο block, θα πρέπει να εντοπίσουμε μέσα σε αυτό την ακριβή θέση του χαρακτηριστικού σημείου. Για την προσέγγιση της θέσης του κέντρου της κόρης του ματιού, αναζητούμε στον οριζόντιο και τον κατακόρυφο άξονα το σημείο με την ελάχιστη φωτεινότητα στο block του ματιού. Αυτό γίνεται με τη χρήση ολοκληρωτικών προβολών, σύμφωνα με τις σχέσεις:

$$\begin{aligned} V(x) &= \sum_{y=y_1}^{y_2} I(x, y) \\ H(y) &= \sum_{x=x_1}^{x_2} I(x, y) \end{aligned} \quad 7-2$$

Η θέση στην οποία εντοπίζουμε το ελάχιστο και στους δύο άξονες αντιστοιχεί στις συντεταγμένες του σημείου που αναζητούμε. Με την ίδια λογική, η θέση των ρουθουνιών εντοπίζεται ξεκινώντας από τα πλάγια του block της μύτης, ενώ τα φρύδια αναζητούνται με βάση τις διαδοχικές αλλαγές της φωτεινότητας προς το πάνω μέρος των ματιών.

Για να εντοπιστεί το block της εικόνας μέσα στο οποίο θα αναζητήσουμε καθένα από τα χαρακτηριστικά σημεία, χρησιμοποιούμε μια απλή διαδικασία ταιριάσματος προτύπου, η οποία ξεκινά με την επιλογή ενός πρωτότυπου για κάθε χαρακτηριστικό του προσώπου από μια σχετική βάση δεδομένων. Για τη σύγκριση με καθένα από τα blocks της εικόνας, υπολογίζουμε τη συνέλιξή του με αυτά με βάση τη σχέση:

$$r = \frac{\sum_{n_1} \sum_{n_2} A(n_1, n_2) B(n_1, n_2)}{\sqrt{\sum_{n_1} \sum_{n_2} A^2(n_1, n_2) \sum_{n_1} \sum_{n_2} B^2(n_1, n_2)}} \quad 7-3$$

όπου Α και Β είναι τα υπό έλεγχο block και πρωτότυπο. Γενικά, αυτή η διαδικασία είναι υπερβολικά χρονοβόρα σε σχέση με τα αποτελέσματα που δίνει – έτσι, χρησιμοποιούμε ένα πρόσθετο μέτρο αξιοπιστίας με τη βοήθεια ενός φίλτρο Gabor. Ένα τέτοιο φίλτρο δίνεται από τη σχέση:

$$g(x, y, u, v) = \exp \left(- \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) + 2\pi j(ux + vy) \right) \quad 7-4$$

όπου x και y αντιστοιχούν στη θέση ενός σημείου και u και v ελέγχουν την κεντρική χωρική συχνότητα του φίλτρου, μεταβάλλοντας στην ουσία τον προσανατολισμό του. Από τη μορφή της σχέσης 7-4 μπορεί να αποδειχθεί ότι βελτιστοποιείται η διαδικασία εντοπισμού, τόσο στο χώρο της εικόνας, όσο και στο χώρο της συχνότητας.

Το φίλτρο εφαρμόζεται στην επιλεγμένη περιοχή και το πρωτότυπο σε τέσσερις διαφορετικές χωρικές διευθύνσεις, με την απόκρισή του να γίνεται δεκτή μόνο αν υπερβαίνει μια δεδομένη τιμή κατωφλίου. Τελικά, η περιοχή με τη μικρότερη απόσταση από το πρωτότυπο, σε ότι αφορά τη φάση, είναι αυτή που θα επιλεγεί για την αναζήτηση του χαρακτηριστικού σημείου.

Αν το ταίριασμα προτύπων δεν επιφέρει ικανοποιητικά αποτελέσματα, μπορούμε να χρησιμοποιήσουμε την ιδέα των ιδιοδιανυσμάτων, η οποία άλλωστε επιτρέπει στα δεδομένα να έχουν υποστεί ένα μεγαλύτερο εύρος από παραμορφώσεις, όπως διαφορές στο φωτισμό, την περιστροφή ή τη μεγέθυνση. Εδώ, κάθε πρότυπο της εικόνας εισόδου συσχετίζεται με ένα σφάλμα

ανακατασκευής. Αυτό το μέγεθος, που αναφέρεται στη βιβλιογραφία σαν "distance-from-feature-space", υπολογίζεται με την προβολή του προτύπου πάνω σε ένα περιορισμένο αριθμό από ιδιοδιανύσματα και τη χρήση των συντελεστών του αποτελέσματος για την ανακατασκευή του.

Πιο συγκεκριμένα, ας θεωρήσουμε ένα σύνολο από N πρότυπα χαρακτηριστικών του προσώπου, στα οποία τα pixels ταξινομούνται με λεξικογραφική διάταξη. Έτσι, για κάθε πρότυπο I_i με $\{I_1, I_2, \dots, I_N\}$ και $I_i \in R^{n \times m}$ μπορούμε να ορίσουμε ένα σύνολο από διανύσματα $\{v_1, v_2, \dots, v_N\}$, όπου $v_i \in R^L$ και $L = n \cdot m$. Για να υπολογίσουμε τις συναρτήσεις βάσης του μετασχηματισμού Karhunen Loeve, αρκεί να επιλύσουμε το πρόβλημα ιδιοτιμών:

$$\Lambda = \Phi^T \Sigma \Phi \quad 7-5$$

όπου Σ είναι ο πίνακας συμμεταβλητότητας, Φ είναι ο πίνακας των ιδιοδιανυσμάτων του Σ και Λ είναι ο αντίστοιχος διαγώνιος πίνακας ιδιοτιμών. Για να ορίσουμε ένα διάνυσμα χαρακτηριστικών y_i για κάθε πρότυπο x_i εφαρμόζουμε ένα μερικό μετασχηματισμό Karhunen Loeve:

$$y_i = \Phi_k^T \cdot \hat{x}_i \quad 7-6$$

όπου $\hat{x}_i = x_i - \mu$, $\mu = \sum_{j=1}^N v_j$ και Φ_k αντιστοιχεί στον περικομμένο πίνακα ιδιοδιανυσμάτων

των όπου διατηρούνται μόνο τα ιδιοδιανύσματα που αντιστοιχούν στις k ιδιοτιμές με τη μεγαλύτερη τιμή. Τελικά, το σφάλμα ανακατασκευής ορίζεται σαν:

$$e_{KLT}(k, i) = (\hat{x}_i - \tilde{x}_i)^T \cdot (\hat{x}_i - \tilde{x}_i) \quad 7-7$$

όπου $\tilde{x}_i = \Phi_k \cdot y_i$ και $y_i = \Phi_k^T \cdot \hat{x}_i$. Το πρότυπο με το καλύτερο ταιρίασμα δίνεται από τη σχέση:

$$x_{opt} = \arg \min_i [e_{KLT}(k, i)] \quad 7-8$$

Στην πράξη, το σφάλμα ανακατασκευής ελαχιστοποιείται με την προβολή και ανακατασκευή των δύο πρώτων ιδιοδιανυσμάτων.

Η φιλοσοφία εντοπισμού των χαρακτηριστικών σημείων στην πλάγια όψη του προσώπου διαφέρει ριζικά από αυτή που περιγράφηκε για την εμπρόσθια. Η κύρια παραδοχή που γίνεται εδώ είναι ότι τα χαρακτηριστικά σημεία που αναζητούμε βρίσκονται πάνω στο περίγραμμα της όψης – έτσι, η διαδικασία ξεκινά με τη μετατροπή της φωτογραφίας σε μια δυαδική εικόνα, από την οποία είναι εύκολο να διακρίνουμε το περίγραμμα του κεφαλιού. Αυτή η μετατροπή γίνεται με την αύξηση της αντίθεσης στην εικόνα, ακολουθούμενη από την εφαρμογή ενός κατωφλίου, έτσι ώστε τα pixels της να αποκτήσουν μια από τις δύο ακραίες τιμές. Τελικά, μπορούμε να επεξεργαστούμε το περίγραμμα σαν ένα μονοδιάστατο σήμα, με τα χαρακτηριστικά σημεία να αντιστοιχούν σε συγκεκριμένα τοπικά ακρότατα. Η επιλογή αυτών των ακροτάτων γίνεται είτε απλά με τη δεύτερη παράγωγο, είτε με τη χρήση της μορφολογικής βαθμίδας [46]. Οι ανθρωπομετρικές μετρήσεις μπορούν να μας βοηθήσουν και σε αυτό το σημείο, αφού η γνώση της φυσιολογίας του προσώπου επιτρέπει το σαφή ορισμό των ακροτάτων που αναζητούμε. Για παράδειγμα, η θέση του άκρου της μύτης μπορεί να οριστεί σαν το σημείο της καμπύλης με την μεγαλύτερη τιμή για τη δεύτερη παράγωγο, ενώ τα μάτια θα πρέπει να αναζητηθούν σε τοπικά

ελάχιστα πάνω από τη θέση της μύτης. Το Σχήμα 7-4 απεικονίζει τη διαδικασία τοπικής κατάτμησης για τον υπολογισμό των ακροτάτων στο περίγραμμα του προσώπου.



Σχήμα 7-4: Απεικόνιση της κατάτμησης στην πλάγια όψη του προσώπου

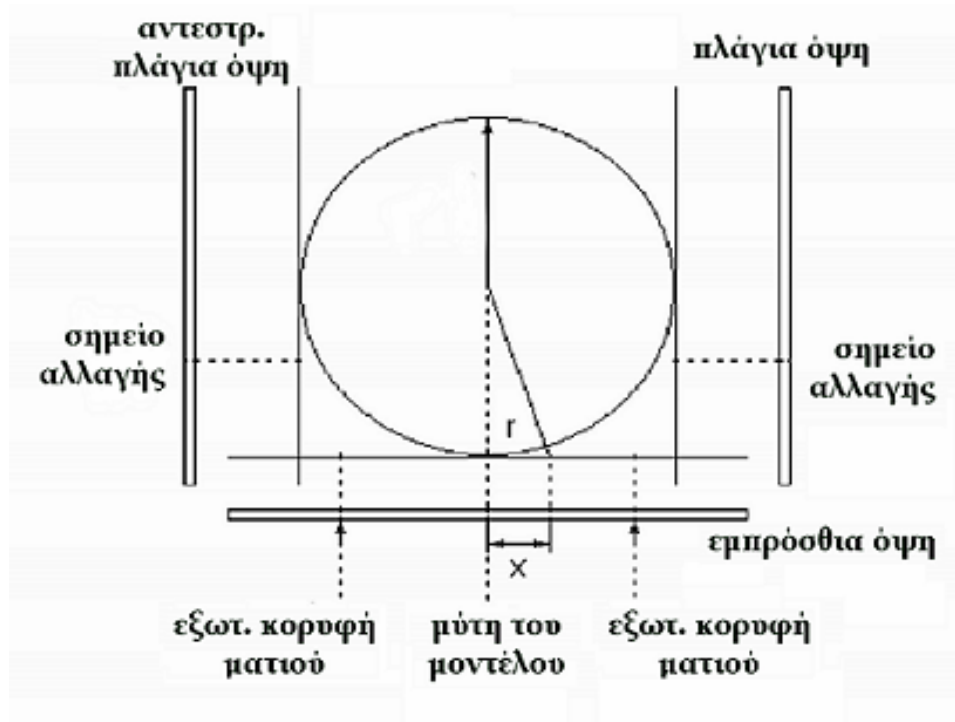
7.4 ΑΠΕΙΚΟΝΙΣΗ ΥΦΗΣ ΚΑΙ ΠΡΟΣΑΡΜΟΓΗ

Οι δύο διαθέσιμες φωτογραφίες του προσώπου μπορούν να συνδυαστούν σε μια συνολική εικόνα, η οποία χρησιμοποιείται για να αυξήσει την αποτελεσματικότητα της απόδοσης του μοντέλου. Η επιτυχία του συνδυασμού μπορεί να περιοριστεί από την απαίτηση να είναι κάθετες οι δύο όψεις, κάτι που δεν ισχύει πάντοτε αυστηρά – πάντως, για μικρές περιστροφές είναι δυνατή η διόρθωση της εικόνας [47], όπως και στις περιπτώσεις μικρής διαφοράς στις συνθήκες φωτισμού. Για να γίνει δυνατός ο συνδυασμός δύο προοπτικών προβολών, όπως είναι ουσιαστικά οι φωτογραφίες, και η απεικόνισή τους σε ένα αντικείμενο γνωστής μορφολογίας αλλά τυχαίας μορφής, όπως είναι το ανθρώπινο κεφάλι, προσεγγίζουμε το μοντέλο μας σαν ένα κύλινδρο. Έτσι, μπορούμε να μετατρέψουμε τις δύο επίπεδες προβολές σε μια συνολική κυλινδρική απεικόνιση - το μειονέκτημα αυτής της ιδέας σε ότι αφορά το μοντέλο είναι ότι σημεία της επιφάνειας τα οποία "διπλώνουν", όπως για παράδειγμα το εσωτερικό της μύτης ή του αυτιού δε θα αποδοθούν σωστά αφού η κυλινδρική προβολή θα τους αναθέσει συντεταγμένες απεικόνισης υφής ίδιες με αυτές των σημείων στην επιφάνεια του προσώπου. Πάντως, αυτά τα σημεία γίνονται ορατά σε πολύ σπάνιες περιπτώσεις, τις οποίες μπορούμε να αντιμετωπίσουμε με τη χρήση προκαθορισμένων επιφανειακών ιδιοτήτων για τις σχετικές πλευρές. Ένα άλλο πρόβλημα της κυλινδρικής προβολής είναι ότι κατά την προσαρμογή της φωτογραφίας εισάγεται παραμόρφωση η οποία μεγαλώνει καθώς πλησιάζουμε στα όρια της εικόνας. Για να αντιμετωπιστεί αυτό το πρόβλημα εναλλάσσουμε τη χρήση των δύο εικόνων, σύμφωνα με το Σχήμα 7-5. Έτσι, αφού τοποθετήσουμε τις εικόνες έτσι ώστε η θέση της μύτης που έχουμε εντοπίσει να ταιριάζει με τη σχετική κορυφή του μοντέλου μας, προβάλλουμε τις εικόνες πάνω στον ιδεατό κύλινδρο με τη σχέση:

$$r = radius \cdot \tan^{-1}\left(\frac{x}{radius}\right) \quad 7-9$$

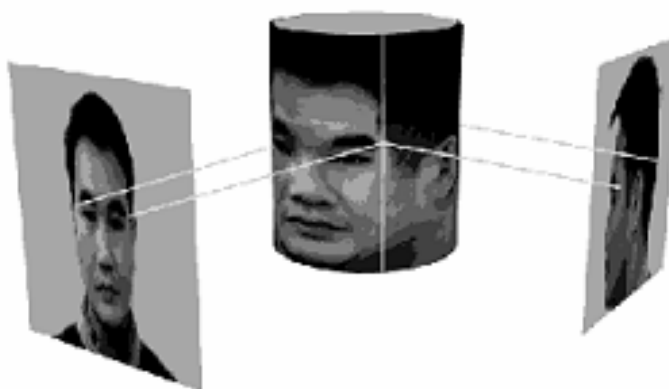
όπου x είναι η απόσταση σε pixels από τη στήλη της μύτης και $radius$ είναι η ακτίνα του κύλινδρου με τον οποίο προσεγγίζουμε το μοντέλο [41]. Στην ιδέα αυτή δεχόμαστε ότι το πρό-

σωπο και τα χαρακτηριστικά του είναι πλήρως συμμετρικά, κάτι που μπορεί να μην είναι πάντοτε σωστό. Στην πράξη όμως το αποτέλεσμα είναι επαρκές για εφαρμογές όπου ο ρεαλισμός αποτελεί δευτερεύουσα απαίτηση σε σχέση με τη συμπίεση των απαιτούμενων δεδομένων.



Σχήμα 7-5: Συνδυασμός φωτογραφιών για την εικόνα απεικόνισης υψής

Σχηματικά, το αποτέλεσμα της διαδικασίας φαίνεται στο Σχήμα 7-6. Η τελική εικόνα που προκύπτει και θα χρησιμοποιηθεί για την απεικόνιση υψής φαίνεται στο Σχήμα 7-7, μετά τη χρήση αλγορίθμων εξομάλυνσης φωτεινότητας.

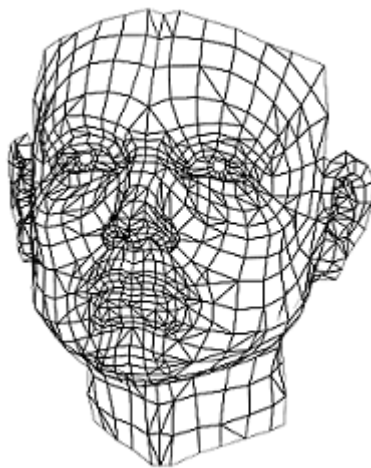


Σχήμα 7-6: Σχηματική απεικόνιση της διαδικασίας συνδυασμού δύο φωτογραφιών

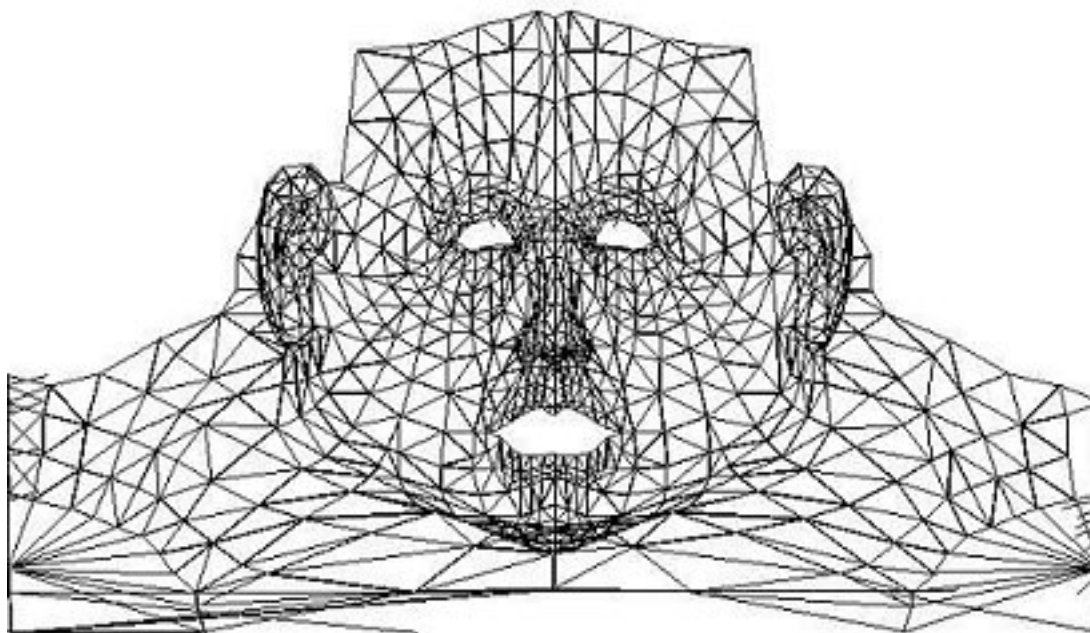


Σχήμα 7-7: Η τελική εικόνα για την απεικόνιση υφής, μετά την εφαρμογή εξομάλυνσης φωτεινότητας

Η ιδέα της κυλινδρικής προβολής χρησιμοποιείται και για την προσαρμογή του μοντέλου γενικής μορφής στις μετρήσεις που υπολογίσαμε από τις εικόνες. Πιο συγκεκριμένα, το τρισδιάστατο πλέγμα "ξετυλίγεται" σε ένα επίπεδο σύμφωνα με τη σχέση μετατροπής κυλινδρικών συντεταγμένων σε καρτεσιανές και με την απόρριψη της μεταβλητής του βάθους. Έτσι, το μοντέλο του Σχήματος 7-8 αφαιρείται σε μια επίπεδη δομή, η οποία φαίνεται στο Σχήμα 7-9.



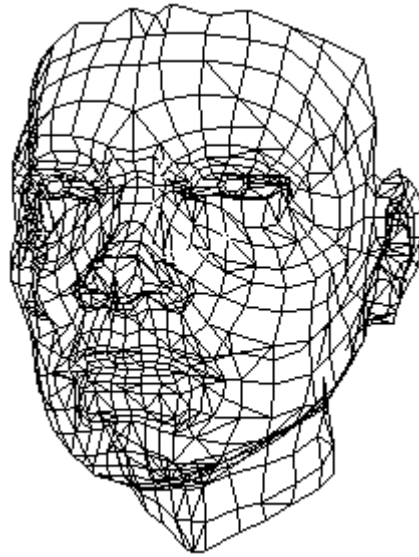
Σχήμα 7-8: Το γενικό μοντέλο πριν τη μετατροπή σε επίπεδο



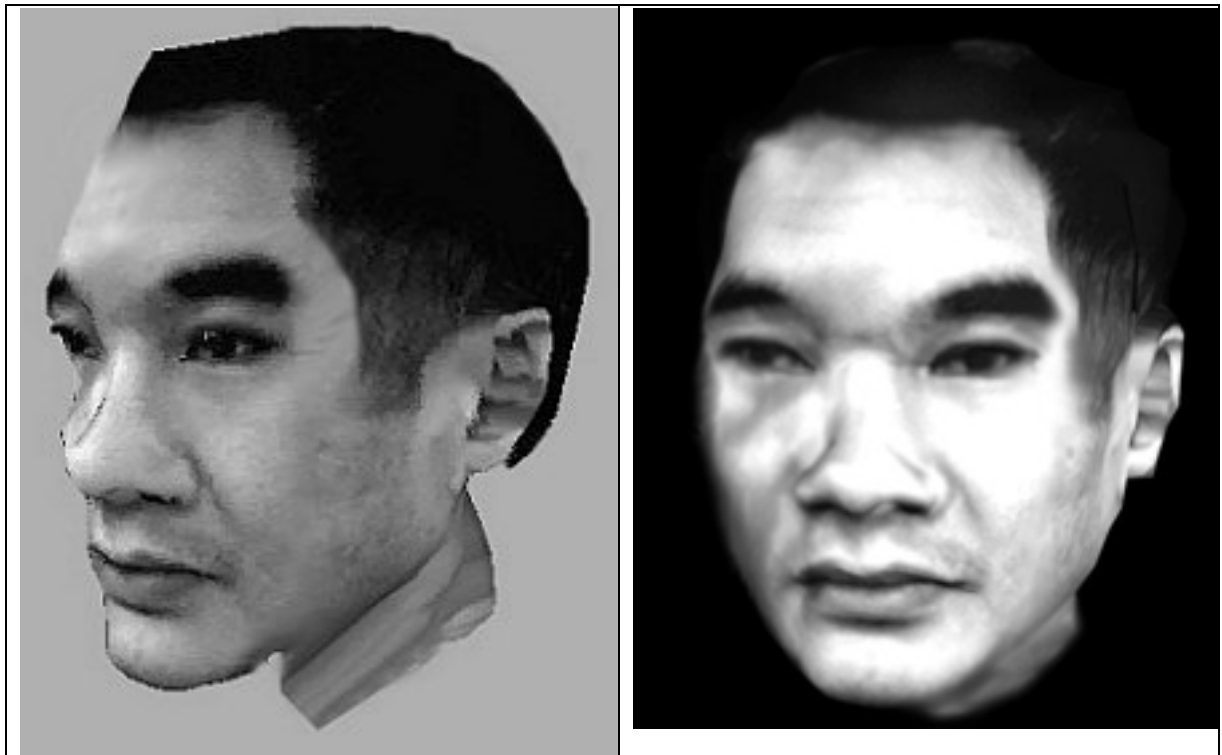
Σχήμα 7-9: Το γενικό μοντέλο μετά τη μετατροπή σε επίπεδο

Σε αυτήν τη δομή από κορυφές και ακμές είναι γνωστή η ακριβής θέση των χαρακτηριστικών σημείων που μπορούμε να εντοπίσουμε, αφού προέρχονται από τη γνωστή τοπολογία του μοντέλου. Έτσι, μπορούμε να εφαρμόσουμε μια σειρά από τοπικούς μετασχηματισμούς με περιορισμό στην τελική θέση των κορυφών που αντιστοιχούν στα χαρακτηριστικά σημεία και οι οποίες θα πρέπει να καταλήξουν σε αποστάσεις ίσες με αυτές που υπολογίστηκαν από τις φωτογραφίες. Καθένας από αυτούς τους μετασχηματισμούς επηρεάζει μια συγκεκριμένη περιοχή από πλευρές γύρω από το χαρακτηριστικό σημείο που πρέπει να προσαρμόσει, η οποία είναι προκαθορισμένη πάνω στο μοντέλο. Για να διατηρηθεί η συνέχεια της επιφάνειας στα όρια των περιοχών, η ένταση των μετασχηματισμών μειώνεται με την απόσταση από το χαρακτηριστικό σημείο και τελικά μηδενίζεται στα όρια της κάθε περιοχής. Η τελική μορφή της προσαρμογής του μοντέλου που φαίνεται στο Σχήμα 7-8 εμφανίζεται στο Σχήμα 7-10, ενώ το Σχήμα εμφανίζει δύο αποτελέσματα της απόδοσης του μοντέλου μετά τη διαδικασία απεικόνισης υφής.

Γενικά, η διαδικασία που παρουσιάστηκε δεν ανιχνεύει σημεία της φωτογραφίας τα οποία ανήκουν στο κεφάλι, αλλά όχι το πρόσωπο, όπως για παράδειγμα το περίγραμμα των μαλλιών ή σημεία που δε μπορούν να διαχωριστούν από το δέρμα. Έτσι, δεν είναι δυνατό να προσαρμόσει το μοντέλο σε όλα τα γεωμετρικά χαρακτηριστικά που μπορεί να φαίνονται στις φωτογραφίες, παρά μόνο στο υποσύνολο των FDPs που περιγράψαμε. Σε μια εφαρμογή πάνω από δικτυα χαμηλού εύρους ζώνης, αυτό δεν αποτελεί ιδιαίτερο πρόβλημα και μπορεί να αντιμετωπιστεί με έτοιμα πρότυπα αντικείμενα, τα οποία ταιριάζουν πάνω στο διαμορφωμένο μοντέλο ή με τεχνικές συστημάτων σωματιδίων (particle systems), οι οποίες πάντως είναι εξαιρετικά χρονοβόρες για να χρησιμοποιηθούν σε εφαρμογές πραγματικού χρόνου. Σε ότι αφορά τις δύο φωτογραφίες του κεφαλιού, η φαινόμενη παραμόρφωση από τη μετατροπή της προοπτικής προβολής σε κυλινδρική δεν είναι σημαντική και μπορεί να μειωθεί με τη χρήση μεγαλύτερου αριθμού από φωτογραφίες σε διαφορετικές γωνίες – σε αυτήν την περίπτωση πάντως, τίθεται το θέμα της ανίχνευσης της πραγματικής γωνίας στην οποία βρίσκεται το κεφάλι.



Σχήμα 7-10: Η τελική μορφή του προσαρμοσμένου μοντέλου



Σχήμα 7-11: Το αποτέλεσμα της προσαρμογής μετά την απεικόνιση υφής. (α) Περιστροφή 45°, (β) Περιστροφή 15°

Αναπαράσταση Συναισθημάτων σε Συνθετικά Μοντέλα

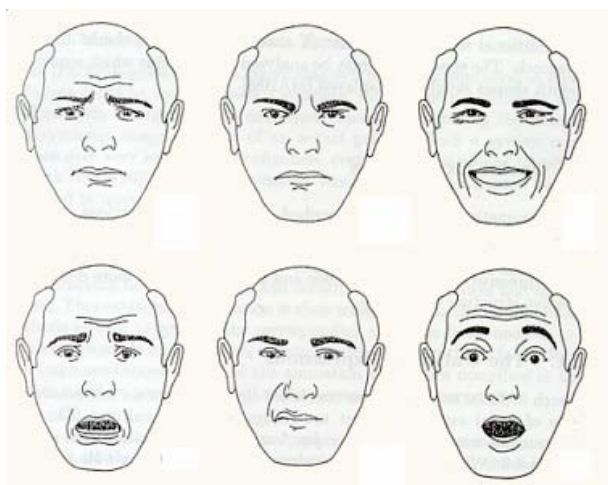
Η σωστή αισθητικά αναπαράσταση ενός συναισθήματος ή μιας έκφρασης σε ένα συνθετικό μοντέλο ανθρώπινου κεφαλιού είναι ένα πολύ σημαντικό εφόδιο για τις εφαρμογές που στηρίζονται στην επικοινωνία ανθρώπου – μηχανής. Σε ένα τέτοιο περιβάλλον, ο χρήστης μπορεί να ξεπεράσει εύκολα τις φοβίες και τους ενδοιασμούς για τη χρήση υπολογιστικών συστημάτων, ενώ τυχόν οδηγίες και πληροφορίες μπορούν να γίνουν ευκολότερα αντιληπτές, αφού παρουσιάζονται σε οικεία προς το χρήστη μορφή. Πολλά σύγχρονα προϊόντα λογισμικού περιέχουν ανθρώπινα ή ανθρωπόμορφα μοντέλα τα οποία προσεγγίζουν την κατανόηση και την αναπαραγωγή φυσικής γλώσσας, σαν υποκατάσταση των ενσωματωμένων συστημάτων βοήθειας που στηρίζονται σε επιλογή με βάση το κείμενο. Τέτοιοι πράκτορες (agents) μπορούν να χρησιμοποιηθούν σε ένα μεγάλο εύρος εφαρμογών, διευρύνοντας την αποτελεσματικότητα της επικοινωνίας. Αντίστροφα, η ανάλυση των συναισθημάτων του χρήστη ενός υπολογιστικού συστήματος μπορεί να βοηθήσει στην προσαρμογή του τρόπου εργασίας του σε αυτό ή να δώσει σε σύστημα πληροφορίες σχετικά με την επιτυχία ή την αντίδραση στις υπηρεσίες που προσφέρει στο χρήστη. Εδώ παρουσιάζουμε μια μελέτη των συναισθημάτων και εκφράσεων, η οποία μετατρέπεται σε αναπαράσταση συμβόλων με τη χρήση των εργαλείων του MPEG-4. Κατόπιν, επιχειρείται η μετατροπή του χώρου των εκφράσεων σε συνεχή, με το συνδυασμό βασικών εκφράσεων σε ένα συνθετικό μοντέλο για την παραγωγή ενδιάμεσων. Τα αποτελέσματα αυτής της διαδικασίας χρησιμοποιούνται τελικά για την ανάλυση σε μια ακολουθία εικόνων και την ανίχνευση της έκφρασης του χρήστη.

8.1 ΕΙΣΑΓΩΓΗ

Η αναγνώριση, αναπαράσταση και κατηγοριοποίηση των συναισθημάτων και των εκφράσεων που μπορούν να αποδοθούν από ένα ανθρώπινο πρόσωπο αποτελεί αντικείμενο έρευνας ακόμα και σε χώρους που είναι φαινομενικά ξένοι προς την επιστήμη των υπολογιστών, όπως η ψυχολογία ή η συμβατική ζωγραφική. Εργασίες σχετικές με το αντικείμενο, οι οποίες το προσεγγίζουν και από τις δύο όψεις, εκείνη της φυσιολογίας κι εκείνη της τέχνης μπορούν να εντοπιστούν από τις αρχές του εικοστού αιώνα. Οι περισσότερες πάντως από αυτές έχουν μικρή σχέση με τη συστηματική μελέτη και τη χρήση υπολογιστών, με πιο γνωστή και παραδεκτή το FACS (Facial Action Coding System) από τους Ekman και Friesen [48]: εδώ γίνεται μια προσπάθεια να

περιγράφουν οι εκούσιες κινήσεις που παρατηρούνται στο πρόσωπο από ένα σύνολο χαρακτηριστικών παραμέτρων, οι οποίες δρουν είτε ατομικά, είτε σε συνδυασμό. Η ανάλυση δεν επεκτείνεται ιδιαίτερα στο θέμα των συνολικών εκφράσεων, ούτε στη σημασιολογία τους και την πιθανή σχέση με μετρήσιμες μεθόδους, παρά επικεντρώνεται στη σύνδεση των στοιχειωδών αυτών δράσεων με δεδομένα φυσιολογίας του ανθρώπινου προσώπου.

Μια γενική παραδοχή που μπορούμε να κάνουμε, σε ότι αφορά τα συναισθήματα και την κατανόησή τους από τον άνθρωπο, είναι ότι έξι από αυτά μορφοποιούνται και γίνονται αντιληπτά με τον ίδιο τρόπο από άτομα ανεξάρτητα πολιτισμικού υποβάθρου. Αυτά ονομάζονται "καθολικά συναισθήματα" (universal) και είναι η λύπη, ο θυμός, η χαρά, ο φόβος, η απέχθεια και η έκπληξη [49]. Μια γενική μορφή των καθολικών συναισθημάτων φαίνεται στο Σχήμα 8-1.



Σχήμα 8-1: Αναπαράσταση των καθολικών συναισθημάτων. Πάνω σειρά λύπη, θυμός, χαρά. Κάτω σειρά φόβος, απέχθεια, έκπληξη

Μπορούμε να θεωρήσουμε αυτές τις αναπαραστάσεις σαν "οικογένειες" συναισθημάτων, μέσα στις οποίες ανήκει ακόμα ένα πλήθος διαφορετικών εκφράσεων. Οι διαφοροποιήσεις μπορεί να είναι αποτέλεσμα της αλλαγής στην ένταση, όπως από τη λύπη στην απόγνωση, ή να βρίσκονται στα σημασιολογικά όρια δύο οικογενειών, όπως η ευχάριστη ή αμήχανη έκπληξη. Φαίνεται λοιπόν ότι υπάρχει η ανάγκη για τη μετατροπή αυτής της διακριτής κατηγοριοποίησης σε όσο το δυνατόν συνεχή. Η έννοια του "χώρου συναισθημάτων" εισάγεται από το Whissel στην εργασία [50], και περιγράφεται με δύο άξονες: την ενεργοποίηση - activation και την επαλήθευση - evaluation. Στην πράξη, η επαλήθευση φαίνεται να εκφράζει την εσωτερική συναισθηματική κατάσταση ή τις σκέψεις του ανθρώπου, οι οποίες δεν μπορούν να ανιχνευθούν από τη μορφοποίηση του προσώπου. Από τη άλλη, η ενεργοποίηση σχετίζεται άμεσα με την ολική κίνηση του κεφαλιού, καθώς και με τη δραστηριοποίηση των μυών του προσώπου. Ο Πίνακας 8-1 περιέχει μια σειρά από χαρακτηριστικές τιμές για τη θέση κάποιων καθολικών και ενδιάμεσων συναισθημάτων, σύμφωνα με την εργασία του Whissel.

	Ενεργοποίηση	Επαλήθευση		Ενεργοποίηση	Επαλήθευση
Φόβος	4.9	3.4	Θυμός	4.2	2.7
Ντροπή	2	2.7	Ικανοποίηση	4.2	6.4
Αηδία	5	3.2	Προσμονή	5	5.1
Ενοχή	4	1.1	Χαρά	5.4	6.1
Υπομονή	3.3	3.8	Λύπη	3.8	2.4
Έκπληξη	6.5	5.2			

Πίνακας 8-1: Οι τιμές για μερικά συναισθήματα, σύμφωνα με το Whissel

Από την πλευρά της σύνθεσης, το πρότυπο MPEG-4 υποστηρίζει την περιγραφή συναισθημάτων και εκφράσεων σαν ένα σύνολο από FAPs. Κάθε τέτοιο σύνολο μπορεί να αποθηκευτεί τοπικά στον αποκωδικοποιητή ή να μεταδοθεί σαν τμήμα της ροής - στη συνέχεια, τα σύνολα αυτά μπορούν να συνδυαστούν ανά δύο, για να δώσουν μια ενδιάμεση έκφραση. Όπως συμβαίνει σε όλους τους τομείς του προτύπου, δε γίνεται κάποια περιγραφή για το πώς θα εξαχθούν τα δεδομένα, παρά μόνο η χρήση τους για την αναπαραγωγή. Έτσι, μένει στην πλευρά του κωδικοποιητή να αποφασίσει για τα FAPs που θα χρησιμοποιήσει για να περιγράψει κάθε έκφραση. Εξάλλου, η απλή παρεμβολή ανάμεσα στις παραμέτρους που ορίζουν δύο εκφράσεις δεν εξασφαλίζει ότι το αποτέλεσμα θα είναι σημασιολογικά και αισθητικά σωστό. Για αυτό το λόγο, θα πρέπει να ορίσουμε κάποιους κανόνες για τη σύνθεση ενδιάμεσων εκφράσεων με βάση τις καθολικές.

8.2 ΚΑΘΟΛΙΚΑ ΣΥΝΑΙΣΘΗΜΑΤΑ

Η αναπαράσταση των καθολικών συναισθημάτων μπορεί να χρησιμεύσει σαν βάση για τη σύνθεση των ενδιάμεσων εκφράσεων. Ας δούμε μια αφαιρετική περιγραφή των χαρακτηριστικών που εκφράζουν καθένα από αυτά και που τα διακρίνουν από τα υπόλοιπα:

Λύπη

Είναι το πλέον έντονο συναισθημα και είναι δυνατό να αφήσει μια σκιά στο πρόσωπο, όταν άλλες εκφράσεις θα είχαν ατονήσει. Τα φρύδια παίρνουν μέρος σε κάθε παραλλαγή αποτύπωσης της λύπης. Στην πιο συνηθισμένη περίπτωση, κυρτώνουν προς τα πάνω, ενώ στο κλάμα πιέζονται προς τα κάτω. Στην περίπτωση, όμως, έντονης λύπης, το σημείο κλειδί του προσώπου είναι τα μάτια, αφού το πόσο σφιχτά είναι κλεισμένα μαρτυρά την ένταση της λύπης. Ακόμα, στην "απλή" λύπη οι εσωτερικές γωνίες των φρυδιών κυρτώνουν προς τα πάνω. Στο δέρμα και στο μαλακό ιστό κάτω από το φρύδι δημιουργούνται πτυχές πάνω από το άνω βλέφαρο. Τα μάτια είναι ελαφρώς κλειστά εξαιτίας της πίεσης από τον ιστό πάνω από το βλέφαρο και της κίνησης του κάτω βλεφάρου προς τα πάνω.

Οι ρυτίδες που συνδέονται με την έκφραση της λύπης είναι είτε οριζόντιες πτυχές κατά μήκος των φρυδιών είτε κατακόρυφες γραμμές μεταξύ των φρυδιών. Επιπλέον, δημιουργούνται πλάγιες γραμμές πάνω από το άνω βλέφαρο. Τέλος, υπάρχουν και ρυτίδες σε σχήμα χαμόγελου κάτω από το κάτω χείλος.

Η λύπη έχει πολλές διαβαθμίσεις και παραλλαγές, στις οποίες περιλαμβάνονται κλάμα με ανοιχτό ή κλειστό το στόμα και η συγκαλυμμένη λύπη. Παράλληλα, λυπημένος θεωρείται και αυτός που είναι έτοιμος να κλάψει, αλλά και ο δυστυχισμένος. Τέτοιες παραλλαγές μπορεί να περιλαμβάνουν εντελώς κατεβασμένα φρύδια, ερμητικά κλειστά μάτια, ανοιχτό τετραγωνισμένο στόμα, εξόγκωμα στο πηγούνι ή και έντονο ρινοχειλικό δίπλωμα.

Θυμός

Το κοινό σημείο σε όλα σχεδόν τα θυμωμένα πρόσωπα είναι τα κατεβασμένα φρύδια, ενώ τα μεταβαλλόμενα χαρακτηριστικά είναι το άνοιγμα των ματιών και του στόματος. Όσο πιο ανοικτό είναι το μάτι, τόσο πιο έντονος είναι ο θυμός, αλλά δεδομένων των ματιών, κάποιος φαίνεται πιο θυμωμένος όταν φωνάζει και το στόμα του είναι ανοικτό, από κάποιον που κρατάει τα χείλη σφιχτά κλεισμένα.

Γενικά, στο θυμό τα σημεία κλειδιά είναι τα κατεβασμένα φρύδια, τα οποία πρέπει να βρίσκονται κάτω από το ανώτερο μέρος των πάνω βλέφαρων, τα ανοιχτά μάτια, στα οποία συνήθως η ίριδα δεν σκεπάζεται από το άνω βλέφαρο σε ποσοστό μεγαλύτερο του 25%, και το στόμα, το οποίο είναι είτε ανοικτό αφήνοντας να φανούν τα δόντια είτε σφιχτά κλειστό. Στον απλό θυμό οι εσωτερικές γωνίες των φρυδιών σμίγουν προς τα κάτω και η χαμηλότερη άκρη είναι στο ίδιο επίπεδο με το άνω βλέφαρο. Τα μάτια είναι ορθάνοιχτα, αλλά τα κατεβασμένα φρύδια εμποδίζουν την εμφάνιση του άσπρου των ματιών, πάνω από την ίριδα. Το στόμα είναι κλειστό και το άνω χείλος είναι ελαφρώς συμπιεσμένο ή τετραγωνισμένο.

Οι διαβαθμίσεις του θυμού περιλαμβάνουν τον κραυγάζοντα θυμό, τη λυσσαλέα οργή και τη βλοσυρότητα. Οι διαβαθμίσεις αυτές μπορεί να περιλαμβάνουν σφιχτά συμπιεσμένα χείλη και εξόγκωμα στο πηγούνι ή ανοικτό στόμα που σχηματίζεται από το άνω χείλος σε σχήμα περιπαικτικού χαμόγελου και το ίδιο κάτω χείλος, αφήνοντας να φανούν και η άνω και η κάτω οδοντοστοιχία.

Χαρά

Η έκφραση της χαράς είναι, με μεγάλη διαφορά, η πιο σύνθετη από όλες τις άλλες εκφράσεις και η πιο δύσκολη στην ανάλυσή της. Το χαμόγελο μπορεί να συνδέεται με ένα πλήθος εκφράσεων και συναισθημάτων, από την ευφορία μέχρι τη μελαγχολία ή απλά να χρησιμοποιείται για καθαρά κοινωνικούς σκοπούς. Από την πλευρά της φυσιολογίας, το χαμόγελο είναι αποτέλεσμα της αλληλεπίδρασης δύο μυών, των *zygomatic major* και *orbicularis oris*.

Στην απλή έκφραση της χαράς τα φρύδια είναι χαλαρωμένα. Το άνω βλέφαρο είναι ελαφρά χαμηλωμένο και το κάτω βλέφαρο πιέζεται προς τα πάνω από το πάνω μέρος από το μάγουλο. Ακόμα, το στόμα είναι ορθάνοιχτο και οι άκρες των χειλιών είναι τραβηγμένες προς τα αυτιά. Στην περίπτωση που το στόμα είναι κλειστό, τα χείλη είναι λεπτά και σφιγμένα ως προς το κόκαλο που βρίσκεται κάτω από το στόμα. Αν το στόμα είναι ανοιχτό, το άνω χείλος είναι σε ευθεία γραμμή, επιτρέποντας στην άνω οδοντοστοιχία να φανεί, ενώ το κάτω χείλος είναι ίσιο στο κέντρο και σχηματίζει γωνίες στις άκρες.

Στη χαρά οι ρυτίδες σχηματίζουν ένα 'δίχτυ' στην εξωτερική γωνία των ματιών (*crow's feet*), ενώ δημιουργείται μία πτυχή σε σχήμα χαμόγελου κάτω από το κάτω βλέφαρο, λακκάκια στα μάγουλα και μία βαθιά πτυχή από τη μύτη προς το πηγούνι. Οι τρόποι εκφράσεις της χαράς

ποικίλουν και περιλαμβάνουν θορυβώδη ή όχι γέλια, χαμόγελο με ανοιχτό ή κλειστό στόμα, χαμόγελο που έχουμε καταπνίξει, ή άλλου είδους χαμόγελα, μελαγχολικά, ενθουσιώδη, πρόθυμα, ύπουλα, ένα χαμόγελο μπορεί επίσης να ξελογιάζει ή να είναι διεφθαρμένο και αυτός που χαμογελάει μπορεί να έχει ανοικτά ή κλειστά τα μάτια. Το χαμόγελο ή το γέλιο μπορεί να είναι και ψεύτικο, γεγονός που μπορεί να γίνει αντιληπτό από το μειωμένο αριθμό ρυτίδων στις εξωτερικές γωνίες των ματιών και από τις ελάχιστες ή ανύπαρκτες ρυτίδες κάτω από τα κάτω βλέφαρα.

Φόβος

Η διάκριση της έκφρασης ενός προσώπου σε τρόμο, φόβο ή απλή ανησυχία εξαρτάται ουσιαστικά από το πόσο ανοιχτά είναι τα μάτια. Αν τα μάτια είναι αρκετά ανοικτά, αμέσως συμπεραίνουμε ότι πρόκειται για τρόμο, ενώ τα πρόσωπα που εκφράζουν απλά φόβο έχουν τα μάτια λιγότερο ανοιχτά από το κανονικό και τα φρύδια ελαφρώς υψωμένα. Για να δείξει ανησυχία ένα πρόσωπο πρέπει να συνδυάζει τα υψωμένα φρύδια με την κίνηση τουλάχιστον ενός ακόμα χαρακτηριστικού του προσώπου, όπως ελαφρώς γουρλωμένα μάτια, ανοικτό στόμα ή στόμα σφιγμένο.

Γενικά, ο φόβος μπορεί να ποικίλλει από απλή ανησυχία μέχρι τρόμο. Για να εκφράσει κάποιος φόβο σηκώνει και πιέζει μαζί τα φρύδια. Οι εσωτερικές γωνίες των φρυδιών διπλώνουν προς τα πάνω. Τα μάτια είναι σε επαγρύπνηση, ενώ το στόμα μπορεί να αφήνεται ελαφρώς ανοικτό και οι άκρες του τεντώνονται. Η έκφραση του φόβου δημιουργεί οριζόντιες πτυχές κατά μήκος των φρυδιών και κατακόρυφες γραμμές μεταξύ των φρυδιών, λακκάκια πάνω από τα φρύδια και πλάγιες γραμμές πάνω από το άνω βλέφαρο.

Στην ανησυχία, τα χείλη είναι σφιγμένα γερά και τα περιφερειακά όρια των χειλιών εξαφανίζονται, ενώ συγχρόνως εμφανίζεται ένα εξόγκωμα κάτω από το κάτω χείλος και πάνω από το πηγούνι. Στον τρόμο τα μάτια και το στόμα είναι ορθάνοικτα, το πάνω χείλος είναι χαλαρωμένο, ενώ το κάτω είναι εντελώς τεντωμένο και σφιχτό αφήνοντας να φανούν τα κάτω δόντια. Η ρινοχειλική πτυχή γίνεται ευθεία και ρηχή και ρυτίδες σε σχήμα παρένθεσης εμφανίζονται στα πλάγια του κάτω χείλους.

Απέχθεια

Αηδία, αποστροφή, περιφρόνηση, σαρκασμός είναι τα διαφορετικά ονόματα μίας κατηγορίας εκφράσεων, με το κοινό τους σημείο είναι η ενεργοποίηση του χλευαστικού μυός, του *levator labii superioris*. Μόνο σε πολύ ακραίες περιπτώσεις συμμετέχει στην έκφραση ο εσωτερικός κλάδος του μυός, στις περισσότερες περιπτώσεις συμμετέχει ο κεντρικός κλάδος του μυός. Όσο μειώνεται η ένταση του συναισθήματος, η έκφραση εκφυλίζεται από έκφραση του συνόλου του προσώπου με αλλαγές ορατές από το ανώτερο μέρος από το μάγουλο μέχρι το πηγούνι, σε μία έκφραση που είναι κυρίως ορατή λόγω του επιπέδου σχήματος του άνω χείλους.

Η απέχθεια κυμαίνεται από περιφρόνηση μέχρι φυσική αποστροφή. Στην απέχθεια τα φρύδια είναι χαλαρωμένα και τα βλέφαρα είναι επίσης χαλαρωμένα ή ελαφρώς κλειστά. Το άνω χείλος είναι υψωμένο σε περιπαικτικό, συχνά ασύμμετρο, χαμόγελο. Το κάτω χείλος είναι χαλαρωμένο. Αντίστοιχα, στην περιφρόνηση τα βλέφαρα μπορεί να είναι μερικώς κλειστά με τα μάτια να κοιτάζουν προς τα κάτω. Στη φυσική αποστροφή τα φρύδια, κυρίως οι εσωτερικές τους

γωνίες, είναι χαμηλωμένα, τα μάτια μπορεί να είναι σχεδόν κλειστά και να δίνουν την εντύπωση ότι “στραβοκοιτάζουν”. Το άνω χείλος υψώνεται σε έντονο περιπαικτικό χαμόγελο που μπορεί να κάνει ορατή την άνω οδοντοστοιχία, το κάτω χείλος είναι ελαφρώς υψωμένο.

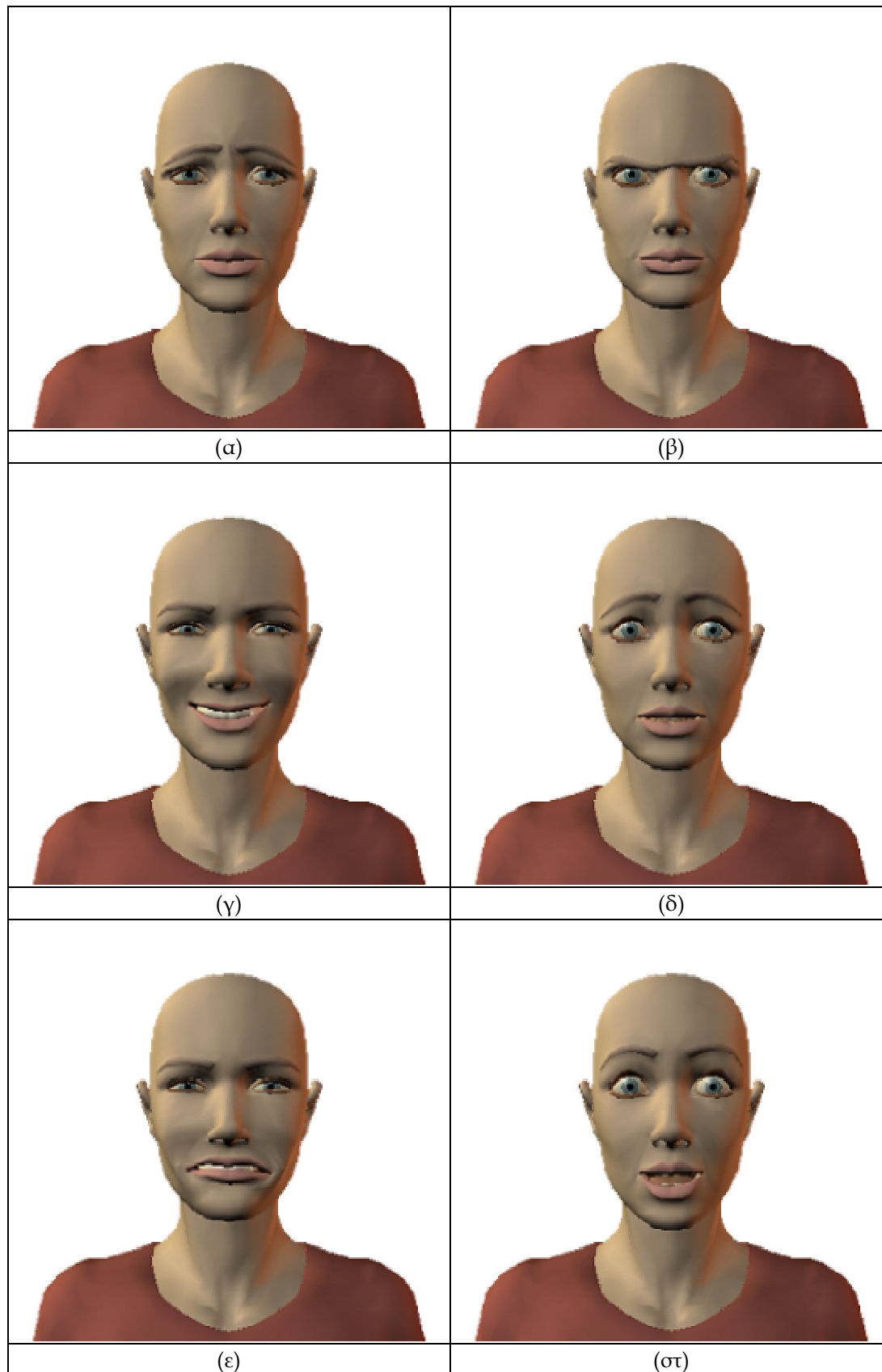
Έκπληξη

Στην έκπληξη τα φρύδια είναι υψωμένα όσο το δυνατόν περισσότερο. Τα άνω βλέφαρα είναι ορθάνοιχτα και τα κάτω είναι χαλαρωμένα, το στόμα αφήνεται ανοικτό σε σχήμα οβάλ, χωρίς να τεντωθούν οι μύες. Ένα συχνό χαρακτηριστικό στην έκφραση της έκπληξης είναι ότι δημιουργούνται οριζόντιες πτυχές κατά μήκος των φρυδιών.

8.3 ΠΕΡΙΓΡΑΦΗ ΜΕ FAPs

Ο Πίνακας 8-2 εμφανίζει το σύνολο των παραμέτρων FAP που μπορούν να χρησιμοποιηθούν για την αναπαράσταση στοιχειωδών ή οργανωμένων παραμορφώσεων σε ένα μοντέλο ανθρώπινου προσώπου. Με βάση την περιγραφή των καθολικών συναισθημάτων και με δεδομένες τις παραμορφώσεις των FAPs, μπορούμε να κατασκευάσουμε την αναπαράστασή τους με σύμβολα ανώτερου επιπέδου. Τα FAPs που χρησιμοποιούνται για κάθε οικογένεια καθολικών εκφράσεων φαίνονται στον Πίνακα 8-3.

Ο καθορισμός των FAPs που εμπλέκονται σε κάθε έκφραση, είτε πρόκειται για καθολική, είτε για κάποια ενδιάμεση μπορεί να υποστηριχθεί και από μετρήσεις σε πραγματικά δεδομένα. Για να γίνει αυτό, θα πρέπει να συσχετίσουμε την κίνηση χαρακτηριστικών σημείων στο ανθρώπινο πρόσωπο με συγκεκριμένα FAPs, κάνοντας έτσι την παραδοχή ότι η κίνηση προκύπτει σαν αποτέλεσμα μιας στοιχειώδους δράσης, αντίστοιχης της συνθετικής παραμόρφωσης που περιγράφει η παράμετρος εμπύχωσης. Για παράδειγμα, η χαμηλού επιπέδου παράμετρος `open_jaw` μπορεί να προσεγγιστεί από τη μεταβολή της απόστασης ανάμεσα στο άνω άκρο του άνω χείλους και το κάτω άκρο του κάτω χείλους σε μια πραγματική εικόνα ή ακολουθία εικόνων. Η εφαρμογή των αλγορίθμων που παρουσιάστηκαν στο προηγούμενο κεφάλαιο, σε συνδυασμό με έναν απλό αλγόριθμο εκτίμησης κίνησης, μπορεί να μας δώσει μια αρκετά καλή αρχική τιμή για την αναπαράσταση. Τα αποτελέσματα των μετρήσεων σε στατικές εικόνες από το σύνολο του Ekman, αλλά και σε ακολουθίες εικόνων από το Medialab φαίνονται στον Πίνακα 8-4. Το Σχήμα 8-2 εμφανίζει τα αποτελέσματα της σύνθεσης με βάση τα βεβαιωμένα FAPs.



Σχήμα 8-2: Αναπαράσταση των καθολικών συναισθημάτων με βάση τα πειραματικά βεβαιωμένα FAPs. (α) Λόπη, (β) Θυμός, (γ) Χαρά, (δ) Φόβος, (ε) Αλέχθεια και (στ) Έκπληξη

Ονομασία	Περιγραφή	Uni - Bi-directional	Κατεύθυνση θετικών τιμών	Μονάδα
"move_h_l_eyeball"	Οριζόντια κίνηση του βολβού του αριστερού ματιού.	B	δεξιά	AU
"move_h_r_eyeball"	Οριζόντια κίνηση του βολβού του δεξιού ματιού.	B	δεξιά	AU
"move_v_l_eyeball"	Κατακόρυφη κίνηση του βολβού του αριστερού ματιού.	B	κάτω	AU
"move_v_r_eyeball"	Κατακόρυφη κίνηση του βολβού του δεξιού ματιού.	B	κάτω	AU
"enlarge_l_pupil"	Μεγέθυνση της κόρης του αριστερού ματιού.	B	έξω	ES
"enlarge_r_pupil"	Μεγέθυνση της κόρης του δεξιού ματιού.	B	έξω	ES
"close_upper_l_eyelid"	Κατακόρυφη κίνηση του άνω αριστερού βλεφάρου.	B	κάτω	ENS
"close_upper_r_eyelid"	Κατακόρυφη κίνηση του άνω δεξιού βλεφάρου.	B	κάτω	ENS
"close_lower_l_eyelid"	Κατακόρυφη κίνηση του κάτω αριστερού βλεφάρου.	B	πάνω	ENS
"close_lower_r_eyelid"	Κατακόρυφη κίνηση του κάτω δεξιού βλεφάρου.	B	πάνω	ENS
"move_o_l_eyeball"	Μέσα/έξω κίνηση του αριστερού βολβού.	B	μπροστά	ES
"move_o_r_eyeball"	Μέσα/έξω κίνηση του δεξιού βολβού.	B	μπροστά	ES
"raise_l_o_eyebrow"	Κατακόρυφη κίνηση του εξωτερικού μέρους του αριστερού φρυδιού.	B	πάνω	ES
"raise_r_o_eyebrow"	Κατακόρυφη κίνηση του εξωτερικού μέρους του δεξιού φρυδιού.	B	πάνω	ENS
"raise_l_m_eyebrow"	Κατακόρυφη κίνηση του μέσου του αριστερού φρυδιού.	B	πάνω	ENS
"raise_r_m_eyebrow_m"	Κατακόρυφη κίνηση του μέσου του δεξιού φρυδιού.	B	πάνω	ENS
"raise_l_i_eyebrow"	Κατακόρυφη κίνηση του εσωτερικού μέρους του αριστερού φρυδιού.	B	πάνω	ENS
"raise_r_i_eyebrow"	Κατακόρυφη κίνηση του εσωτερικού μέρους του δεξιού φρυδιού.	B	πάνω	ENS
"squeeze_l_eyebrow"	Οριζόντια κίνηση του αριστερού φρυδιού.	U	μέσα	ES
"squeeze_r_eyebrow"	Οριζόντια κίνηση του δεξιού φρυδιού.	U	μέσα	ES
"stretch_l_nose"	Τέντωμα/συμπίεση του αριστερού	B	προς το άκρο	ENS

Ονομασία	Περιγραφή	Uni - Bi-directional	Κατεύθυνση θετικών τιμών	Μονάδα
	ρουθουνιού.			
"stretch_r_nose"	Τέντωμα/συμπίεση του δεξιού ρουθουνιού.	B	προς το άκρο	ENS
"raise_nose"	Κατακόρυφη κίνηση της μύτης.	U	πάνω	ENS
"open_jaw"	Περιστροφική κίνηση της γνάθου.	U	κάτω	MNS
"move_hori_jaw"	Αριστερή/δεξιά κίνηση της γνάθου.	B	δεξιά	MNS
"move_fwd_jaw"	Προς τα μπροστά κίνηση της γνάθου.	B	μπροστά	MNS
"depress_chin"	Ανύψωση και συμπίεση του σαγονιού (όπως στην έκφραση της λύπης).	U	πάνω	0 - 10
"puff_l_cheek"	Ρούφηγμα του αριστερού μάγουλου.	B	αριστερά	ES
"puff_r_cheek"	Ρούφηγμα του δεξιού μάγουλου.	B	δεξιά	ES
"lift_l_cheek"	Ανύψωση του αριστερού μάγουλου.	U	πάνω	ENS
"lift_r_cheek"	Ανύψωση του δεξιού μάγουλου.	U	πάνω	ENS
"lower_lowerlip"	Κατακόρυφη κίνηση του κάτω χείλους.	U	κάτω	MNS
"lower_l_cornerlip"	Κατακόρυφη κίνηση του αριστερού άκρου των χειλιών.	U	κάτω	MNS
"lower_r_cornerlip"	Κατακόρυφη κίνηση του δεξιού άκρου των χειλιών.	U	κάτω	MNS
"raise_l_cornerlip"	Κατακόρυφη κίνηση (προς τα πάνω) του αριστερού άκρου των χειλιών.	U	πάνω	MNS
"raise_r_cornerlip"	Κατακόρυφη κίνηση (προς τα πάνω) του δεξιού άκρου των χειλιών.	U	πάνω	MNS
"raise_u_midlip"	Κατακόρυφη κίνηση του μέσου του άνω χείλους.	U	πάνω	MNS
"raise_l_midlip"	Κατακόρυφη κίνηση του μέσου του κάτω χείλους.	U	πάνω	MNS
"push_pull_upperlip"	Προβολή ή ρούφηγμα του άνω χείλους.	U	μπροστά	MNS
"push_pull_lowerlip"	Προβολή ή ρούφηγμα του κάτω χείλους.	U	πίσω	MNS
"stretch_l_cornerlip"	Τέντωμα του αριστερού άκρου των χειλιών.	U	αριστερά	MW
"stretch_r_cornerlip"	Τέντωμα του δεξιού άκρου των χειλιών.	U	δεξιά	MW
"squeeze_l_cornerlip"	Σφίξιμο του αριστερού άκρου των χειλιών.	U	δεξιά	MW
"squeeze_r_cornerlip"	Σφίξιμο του δεξιού άκρου των χειλιών.	U	αριστερά	MW

Ονομασία	Περιγραφή	Uni - Bi-directional	Κατεύθυνση θετικών τιμών	Μονάδα
"move_h_tongue_tip"	Αριστερή/δεξιά κίνηση της άκρης της γλώσσας.	B	δεξιά	MW
"raise_tongue_tip"	Κατακόρυφη κίνηση της άκρης της γλώσσας.	B	πάνω	MW
"proturd_tongue"	Οριζόντια κίνηση, προβολή της γλώσσας.	B	μπροστά	MW
"raise_tongue"	Κατακόρυφη κίνηση της γλώσσας.	B	πάνω	MW
"tongue_roll"	Τύλιγμα της γλώσσας σε σχήμα U.	U	προς το κέντρο	AU
"raise_left_ear"	Ανύψωση του αριστερού αυτιού.	B	πάνω	ENS
"raise_right_ear"	Ανύψωση του δεξιού αυτιού.	B	πάνω	ENS
"pull_left_ear"	Τράβηγμα του αριστερού αυτιού προς τα μπροστά με πιθανή περιστροφή.	U	μπροστά	ENS
"pull_right_ear"	Τράβηγμα του δεξιού αυτιού προς τα μπροστά με πιθανή περιστροφή.	U	μπροστά	ENS
FAPs υψηλού επιπέδου				
"viseme"	Οπτικοποίηση φωνήματος			
"expression"	Μεταβλητή που προσδιορίζει την έκφραση του προσώπου.			

Πίνακας 8-2: Τα FAPs για την αναπαράσταση της εμφύχωσης στο πρότυπο MPEG-4

οργή	squeeze_l_eyebrow (+) lower_t_midlip (-) raise_l_i_eyebrow (+) close_t_r_eyelid (-) close_b_r_eyelid (-)	squeeze_r_eyebrow (+) raise_b_midlip (+) raise_r_i_eyebrow (+) close_t_l_eyelid (-) close_b_l_eyelid (-)
λόπη	raise_l_i_eyebrow (+) close_t_l_eyelid (+) raise_l_m_eyebrow (-) raise_l_o_eyebrow (-) close_b_l_eyelid (+)	raise_r_i_eyebrow (+) close_t_r_eyelid (+) raise_r_m_eyebrow (-) raise_r_o_eyebrow (-) close_b_r_eyelid (+)
Έκπληξη	raise_l_o_eyebrow (+) raise_l_i_eyebrow (+) raise_l_m_eyebrow (+) squeeze_l_eyebrow (-) open_jaw (+)	raise_r_o_eyebrow (+) raise_r_i_eyebrow (+) raise_r_m_eyebrow (+) squeeze_r_eyebrow (-)
χαρά	close_t_l_eyelid (+) close_b_l_eyelid (+) stretch_l_cornerlip (+) raise_l_m_eyebrow (+) lift_r_cheek (+) lower_t_midlip (-) H open_jaw (+)	close_t_r_eyelid (+) close_b_r_eyelid (+) stretch_r_cornerlip (+) raise_r_m_eyebrow (+) lift_l_cheek (+) raise_b_midlip (-)
απέχθεια	close_t_l_eyelid (+) close_t_r_eyelid (+) lower_t_midlip (-) squeeze_l_cornerlip (+) KAI/H squeeze_r_cornerlip (+)	close_b_l_eyelid (+) close_b_r_eyelid (+) open_jaw (+)
φόβος	raise_l_o_eyebrow (+) raise_l_m_eyebrow (+) (+) raise_l_i_eyebrow (+) squeeze_l_eyebrow (+) open_jaw (+) H close_t_l_eyelid (-) lower_t_midlip (-) H lower_t_midlip (+)	raise_r_o_eyebrow (+) raise_r_m_eyebrow (+) raise_r_i_eyebrow (+) squeeze_r_eyebrow (+) close_t_r_eyelid (-)

Πίνακας 8-3: Αναπαράσταση των οικογενειών καθολικών συναισθημάτων με FAPs

			Θ	Λ	Χ	Α	Φ	Ε
F_1 (ES)	squeeze_l_eyebrow	Mean	-57	*	*	*	*	*
		StD	28	*	*	*	*	*
F_2 (ES)	squeeze_r_eyebrow	Mean	-58	*	*	*	*	*
		StD	31	*	*	*	*	*
F_3 (MNS)	lower_t_midlip	Mean	-73	*	-271	-234	*	*
		StD	51	*	110	109	*	*
F_4 (MNS)	raise_b_midlip	Mean	*	*	*	-177	218	543
		StD	*	*	*	108	135	203
F_5 (ENS)	raise_l_i_eyebrow	Mean	-83	85	*	*	104	224
		StD	48	55	*	*	69	103
F_6 (ENS)	raise_r_i_eyebrow	Mean	-85	80	*	*	111	211
		StD	51	54	*	*	72	97
F_7 (ENS)	raise_l_o_eyebrow	Mean	-66	*	*	*	*	54
		StD	35	*	*	*	*	31
F_8 (ENS)	raise_r_o_eyebrow	Mean	-70	*	*	*	*	55
		StD	38	*	*	*	*	31
F_9 (ENS)	raise_l_m_eyebrow	Mean	-149	*	24	-80	72	144
		StD	40	*	22	53	58	64
F_{10} (ENS)	raise_r_m_eyebrow	Mean	-144	*	25	-82	75	142
		StD	39	*	22	54	60	62
F_{11} (MNS)	open_jaw	Mean	*	*	*	*	291	885
		StD	*	*	*	*	189	316
F_{12} (IrisD)	close_t_l_eyelid - close_b_l_eyelid	Mean	*	-153	-254	-203	244	254
		StD	*	112	133	148	126	83
F_{13} (IrisD)	close_t_r_eyelid - close_b_r_eyelid	Mean	*	-161	-242	-211	249	252
		StD	*	109	122	145	128	81
F_{14} (MW)	stretch_l_cornerlip - stretch_r_cornerlip	Mean	*	*	234	*	*	-82
		StD	*	*	98	*	*	39
F_{15} (ES)	squeeze_l_eyebrow & squeeze_r_eyebrow	Mean	-69	-56	*	-52	*	86
		StD	51	35	*	34	*	60

Πίνακας 8-4: Πειραματικές τιμές για FAPs και συνδυασμούς τους από πραγματικά δεδομένα στα καθολικά συναισθήματα. Το σύμβολο * αντιστοιχεί σε αμελητέα συνεισφορά.

Τα συναισθήματα είναι Θ (Θυμός), Λ (Λύπη), Χ (Χαρά), Α (Αηδία),

Φ (Φόβος) και Ε (Εκπληξη)

8.4 ΣΥΝΘΕΣΗ ΕΝΔΙΑΜΕΣΩΝ ΕΚΦΡΑΣΕΩΝ

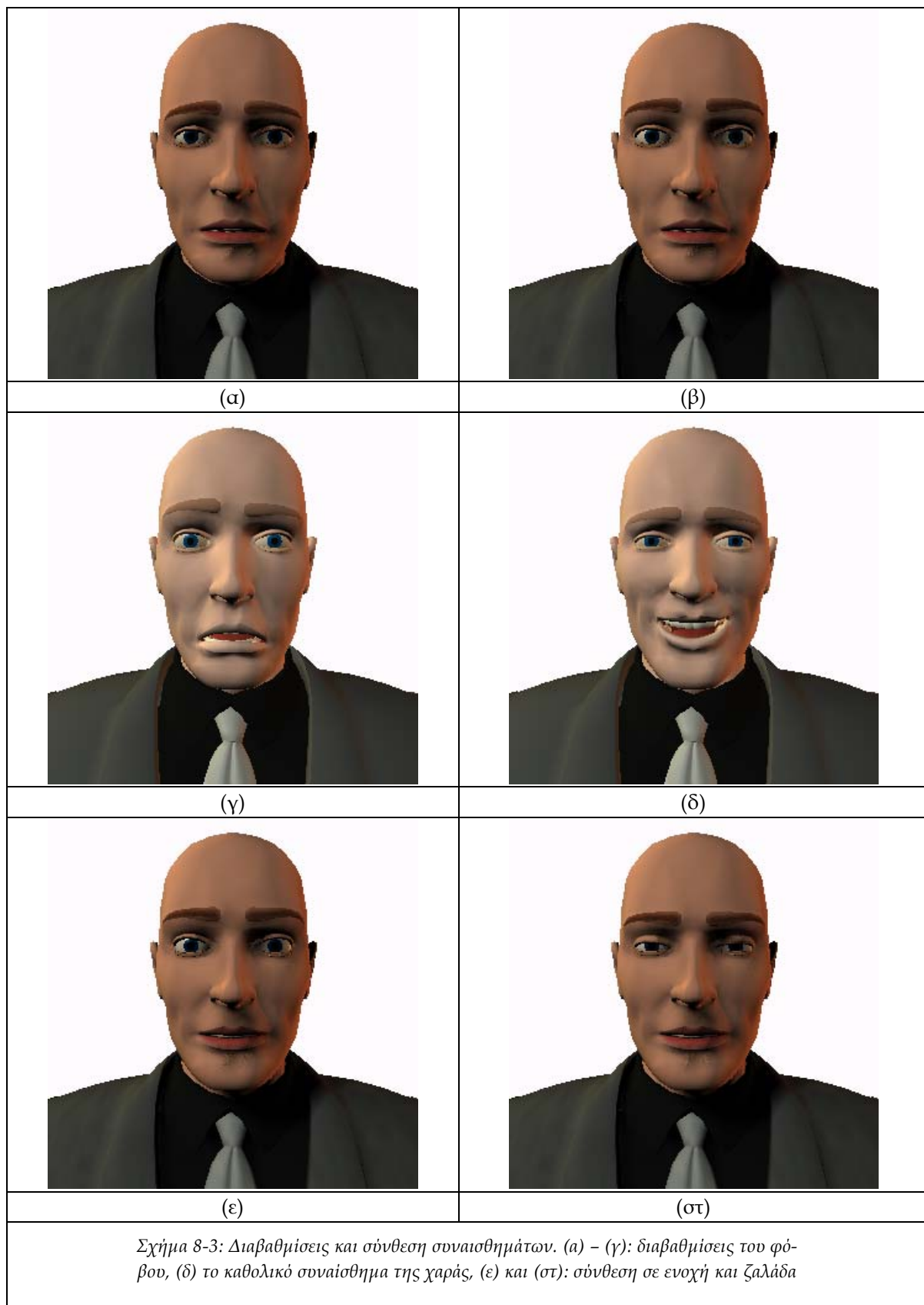
Οι καθολικές εκφράσεις μπορούν να χρησιμοποιηθούν σαν βάση για τη σύνθεση των ενδιάμεσων εκφράσεων. Για να γίνει όμως αυτό θα πρέπει να ορίσουμε μια σημασιολογική βάση μέσα στην οποία να έχουν νόημα οι σχέσεις ανάμεσα σε διαφορετικά συναισθήματα. Επειδή η κατανόηση ενός φαινόμενου ενδιάμεσου συναισθήματος ή έκφρασης είναι έτσι κι αλλιώς υποκειμενική, η βάση αυτή ορίζεται από αντικειμενικούς παρατηρητές στην εργασία [51] σαν το "Feeltrace Plane". Εδώ, τα διάφορα συναισθήματα παρουσιάζονται σαν σημεία σε ένα επίπεδο με άξονες την ενεργοποίηση και την επαλήθευση. Ακολουθώντας, τα συμπεράσματα και οι σχέσεις

που ορίζονται χρησιμοποιούνται στην εργασία [52] για να ορίσουν τις συνεχείς σχέσεις ανάμεσα στα σημεία.

Συγκεκριμένα, αν οι εκφράσεις που πρέπει να συνδυάσουμε ανήκουν στην ίδια σημασιολογική κατηγορία, είναι δηλαδή διαβαθμίσεις του ίδιου καθολικού συναισθήματος, τότε αρκεί να παρεμβάλλουμε τις τιμές των FAPs που ενεργοποιούνται και στις δύο περιπτώσεις για να υπολογίσουμε τις τελικές τιμές. Αν υπάρχουν FAPs που απαντώνται σε μια μόνο από τις συνδυαζόμενες εκφράσεις, τότε θα πρέπει να ελέγξουμε τη συνάφειά του με την κεντρική τιμή του καθολικού συναισθήματος της συγκεκριμένης οικογένειας: αν είναι κοντά σε αυτήν, τότε το FAP διατηρείται με μειωμένη ένταση, ενώ στην αντίθετη περίπτωση απορρίπτεται.

Στις περιπτώσεις που οι εκφράσεις ανήκουν σε διαφορετική σημασιολογική κατηγορία, τότε θα πρέπει να "επικρατήσει" εκείνο το συναισθηματικό για το οποίο η παράμετρος της ενεργοποίησης [50] έχει μεγαλύτερη τιμή. Αυτό συμβαίνει γιατί η συγκεκριμένη παράμετρος σχετίζεται περισσότερο με την ένταση της εκδήλωσης μιας έκφρασης και καλύπτει σε μεγάλο βαθμό το σημασιολογικό της υπόβαθρο. Το τελικό αποτέλεσμα είναι συνήθως μια διαφορετική εκδοχή της κυρίαρχης έκφρασης με τη μικρή παραλλαγή που εισάγεται από τη μειωμένη συνεισφορά των FAPs της δεύτερης που δεν υπάρχουν στην κυρίαρχη. Οι εργασίες [52] και [53] παρουσιάζουν μια πιο προχωρημένη μέθοδο συνδυασμού, βασισμένη σε νευρωνικά δίκτυα και ασαφή λογική.

Το Σχήμα εμφανίζει διαφορετικές διαβαθμίσεις στην περιοχή συναισθημάτων του φόβου, όπως τον τρόπο ή την ανησυχία. Αντίστοιχα, στο Σχήμα γίνεται συνδυασμός της συγκεκριμένης οικογένειας με αυτήν της χαράς για να προκύψουν εκφράσεις ενοχής ή ακόμα και αστάθειας – ζαλάδας.



Bιβλιογραφία

- [1] Garland M. and Heckbert P.S., "Simplifying Surfaces with Color and Texture Using Quadric Error Metrics", Proc. SIGGRAPH 97, pp. 209-216, August 1997.
- [2] Hoppe H., DeRose T., Duchamp T., McDonald J., Stuetzle W., "Mesh Optimization", Proc. SIGGRAPH 93, pp. 19-26, August 1993.
- [3] Clark J., "Hierarchical Geometric Models for Visible Surface Algorithms", Communications of the ACM, Vol. 19, No 10, pp. 547-554.
- [4] Hoppe H., "View-Dependent Refinement of Progressive Meshes", Proc. SIGGRAPH 97, pp. 189-198, August 1997
- [5] Schroeder W., Zarge J., Lorensen W., "Decimation of Triangle Meshes", Proc. SIGGRAPH 92, August 1996.
- [6] Turk G., "Re-tiling Polygonal Surfaces", Proc. SIGGRAPH 92, August 1992.
- [7] Low K. L., Tan T. S., "Model Simplification Using Vertex Clustering", Proc. ACM Symposium on Interactive 3D Graphics, pp. 75-82, 1995.
- [8] Cohen J., Varshney A., Manocha D., Turk G., Weber H., Agarwal P., Brooks F., Wright W., "Simplification Envelopes", Proc. SIGGRAPH 96, August 1996.
- [9] Hoppe H., "Progressive Meshes", Proc. SIGGRAPH 1996, August 1996.
- [10] Stolnitz E., DeRose T., Salesin D., "Wavelets for Computer Graphics: Theory and Applications", Morgan-Kaufmann, San Francisco, 1996, pp. 141-167
- [11] Eck M., DeRose T., Duchamp T., Hoppe H., Lounsbery M., Stuetzle W., "Multiresolution Analysis of Arbitrary Meshes", Proc. SIGGRAPH 95, pp. 173-182, August 1995.
- [12] Tekalp M. and Ostermann J., "Face and 2-D Mesh Animation in MPEG-4", Image Communication Journal, Tutorial Issue on the MPEG-4 Standard, Elsevier Publishers, 2000.
- [13] Tao H., Chen H.H., Wu W., Huang T.S., "Compression of Facial Animation Parameters for Transmission of Talking Heads", IEEE Circuits and Systems for Video Technology, vol. 9, No. 2, pp. 264-276, 1999
- [14] Koenen R., Pereira F., Chiariglione L., "MPEG-4: Context and Objectives", Image Communication Journal, vol. 9, No. 4, May 1997.
- [15] Koenen R., "MPEG-4: Multimedia for our Time", IEEE Spectrum, Vol. 36, No 2, February 1999.
- [16] Wang Y. and Ostermann J., "Evaluation of Mesh-based Motion Estimation in H.263-like Coders", IEEE Transactions on Circuits and Systems on Video Technology, pp. 243 – 252, 1998.
- [17] Nakaya Y. and Harashima H., "Motion Compensation based on Spatial Transformations", IEEE Transactions on Circuits and Systems on Video Technology, Vol. 4, No. 3, pp. 339 – 356, June 1994.
- [18] Signes J., Fisher Y., Eleftheriadis A., "MPEG-4's Binary Format for Scene Description", Image Communication Journal, Tutorial Issue on the MPEG-4 Standard, Elsevier Publishers, 2000.
- [19] ISO-IEC 14772-1, "The Virtual Reality Modeling Language", 1997, <http://www.vrml.org/Specifications/VRML97>.

- [20] ISO/IEC JTC1/SC29/WG11, "Multi-users Technology (Requirements and Applications)", Maui, December 1999.
- [21] VRML Group, "H-Anim Version 1.0 Specification", <http://ece.waterloo.ca/~h-anim/spec.html>
- [22] Witten I. H., Neal R. M., Cleary J. G., "Arithmetic Coding for Data Compression", Communications of the ACM, Vol. 30, No. 6, pp. 520 – 540, June 1987.
- [23] Badler N. I., Manoochchri K., Walters G., "Articulated Figure Positioning by Multiple Constraints", IEEE Computer Graphics & Applications, pp. 28-38, June 1987.
- [24] Maestri G., "Digital Character Animation", New Riders Publishing, Indianapolis, IN, USA, 1996.
- [25] Reese S., "Character Animation with 3D Studio MAX", Coriolis Group Books, Scottsdale, AZ, USA, pp. 29-30, 1996.
- [26] Hamill J., Knutzen K. M., "Biomechanical Basis of Human Movement", Williams & Wilkins, 1995.
- [27] Tzavidas S., "Parametric Representation of Human Gait", Diploma Thesis, Image, Video and Multimedia Systems Laboratory, NTUA, Athens, Greece, 1998.
- [28] Maestri G., "Digital Character Animation", New Riders Publishing, Indianapolis, IN, USA, 1996.
- [29] Karpouzis K., Tzavidas S., and Kollias S., "Parametric Representation of Basic Human Locomotion", Journal of Systems Analysis, Modeling and Simulation, Vol. 38, pp. 281-293, 2000.
- [30] Dooley M., "Anthropometric Modeling Programs - a Survey", IEEE Computer Graphics and Applications, vol. 2, pp. 17-25, November 1982.
- [31] Farkas L., "Anthropometry of the Head and Face", Raven Press, 1994.
- [32] Rogers S., "Personal Identification from Human Remains", Charles C. Thomas Publisher, LTD, 1984.
- [33] Farkas L., "Anthropometric Facial Proportions in Medicine", Thomas Books, 1987.
- [34] Kolar J. and Salter E., "Craniofacial Anthropometry: Practical Measurement of the Head and Face for Clinical, Surgical and Research Use", Charles C. Thomas Publisher, LTD, 1996.
- [35] Pandzic I., Ostermann J., Millen D., "User Evaluation: Synthetic Talking Faces for Interactive Services", The Visual Computer, Special Issue on Real-time Virtual Worlds, 1999.
- [36] Kalra P., Mangili A., Magnenat - Thalmann N., Thalmann D., "Simulation of Facial Muscle Actions Based on Rational Free-form Deformations", Proc. of Eurographics 92, pp. 59-69, 1992.
- [37] Terzopoulos D. and Waters K., "Physically-based Facial Modeling, Analysis and Animation", Journal of Visualization and Computer Animation, Vol. 1, pp. 73-80, 1990.
- [38] Foley J., van Dam A., Feiner S., Hughes J., "Computer Graphics: Principles and Practice", Addison-Wesley, 1990.
- [39] Karpouzis K., Votsis Y., Tsapatsoulis N., Kollias S., "Compact 3D Model Generation based on 2D Views of Human Faces: Application to Face Recognition", Machine Graphics and Vision, vol. 7, Nos. 1/2, pp. 75-85, 1998.
- [40] Votsis Y., Tsapatsoulis N., Karpouzis K., Kollias S., "A Simplified Representation of 3D Human Faces Adapted from 2D Images", Noblesse Workshop on Non-Linear Model Based Image Analysis, Glasgow, Scotland, 1998.
- [41] Tsapatsoulis N., Karpouzis K., Votsis Y., Kollias S., "Analysis by Synthesis of Facial Images Based on Frontal and Profile Views", Proc. of International Workshop on Synthetic Natural Hybrid Coding 97, Rhodes, Greece.
- [42] Samal A. and Iyengar P., "Automatic Recognition and Analysis of Human Faces and Facial Expressions: A Survey", Pattern Recognition, vol. 25, no. 1, pp. 65-77, 1992.

- [43] Brunelli R. and Poggio T., "Face Recognition: Features versus Templates", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 15, no. 10, pp. 1042-1052, Oct. 1993.
- [44] McKenna S. J., Gong S., Wurtz R. P., Tanner J. and Banin D., "Tracking Facial Feature Points with Gabor Wavelets and Shape Models", Lecture Notes in Computer Science, Springer Verlag, 1997.
- [45] Moghaddam B., and Pentland A. "Probabilistic Visual Learning for Object Representation," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, pp. 696-710, 1997.
- [46] Maragos P., "Differential Morphology and Image Processing", IEEE Transactions on Image Processing, vol. 5, pp. 922-937, June 1996.
- [47] Sengupta K. and Ohya J., "An Affine Coordinate based Algorithm for Reprojecting the Human Face for Identification Tasks", Proc. of International Conference on Image Processing 97, Santa Barbara, USA, Oct. 1997.
- [48] Ekman P. and Friesen W., "The Facial Action Coding System", Consulting Psychologist Press, San Francisco, CA, 1978.
- [49] Faigin G., "The Artist's Complete Guide to Facial Expression", Watson - Guptill Publications, New York, 1990.
- [50] Whissel C. M., "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.
- [51] Cowie R., Douglas-Cowie E., Tsapatsoulis N., Votsis G., Kollias S., Fellenz W. and Taylor J., "Emotion Recognition in Human-Computer Interaction," to appear in the IEEE Signal Processing Magazine, 2001.
- [52] Karpouzis K., Tsapatsoulis N. and Kollias S., "Moving to Continuous Facial Expression Space using the MPEG-4 Facial Definition Parameter (FDP) Set", Proc. of Electronic Imaging 2000, San Jose, CA, USA, Jan. 2000.
- [53] Tsapatsoulis N., Karpouzis K., Stamou G., Piat F. and Kollias S., "A Fuzzy System For Emotion Classification Based On The MPEG-4 Facial Definition Parameter Set", Proc. of EUSIPCO 2000, Tampere, Finland.

