

Τεχνολογικό Εκπαιδευτικό Ίδρυμα Κρήτης



Σχολή Τεχνολογικών εφαρμογών
Τμήμα Εφαρμοσμένης Πληροφορικής & Πολυμέσων

Πτυχιακή Εργασία
με τίτλο :

“ Μελέτη του προτύπου **MPEG-4** για την δημιουργία
δικτυακών τηλεοπτικών προγραμμάτων (**IP-TV**)”

Σπουδάστρια : **Δόγαρα Ναταλία**
Επιβλέπων Καθηγητής : **Δρ. Πάλλης Ευάγγελος**

Ακαδημαϊκό Έτος 2005-2006

Περιεχόμενα

| | |
|---|-----------|
| Πτυχιακή Εργασία | 1 |
| 1. Πρόλογος - Στόχος | 5 |
| 1.1. Περίληψη περιεχομένου εργασίας | 5 |
| 2. Η συμπίεση του βίντεο | 7 |
| 2.1. Ψηφιακές τεχνικές συμπίεσης βίντεο | 7 |
| 2.1.1. MJPEG | 8 |
| 2.1.2. ITU-T Recommendation H.261 | 10 |
| 2.1.3. CellB | 11 |
| 2.1.4. nv..... | 12 |
| 2.1.5. CUSeeMe..... | 12 |
| 2.1.6. Indeo | 13 |
| 2.1.7. Κωδικοποίηδη Run-Length..... | 13 |
| 2.1.8. Huffman και Αριθμητική κωδικοποίηση..... | 14 |
| 2.1.9. Συμπίεση αντικατάστασης (Substitutional Compression)..... | 14 |
| 2.1.10. Χωρικές Μέθοδοι Μετασχηματισμών (Spatial Transform Methods) | 14 |
| 2.1.10.1. Ο Διακριτός Συνημιτονικός Μετασχηματισμός (Discrete Cosine Transform, DCT) | 15 |
| 2.1.10.2. Ο Κυματικός Μετασχηματισμός (Wavelet Transform)..... | 17 |
| 2.2. Ελάττωση δεδομένων (data reduction) | 18 |
| 2.2.1. Κβαντοποίηση..... | 18 |
| 2.2.2. Υποδειγματοληψία (Sub-sampling)..... | 19 |
| 2.3. Δια – πλαίσια (intra – frames) σε αντιπαράθεση με τα ενδο – πλαίσια (inter – frames) | 19 |
| 2.4. Υπολογιστική ισχύς | 22 |
| 2.5. Οι τεχνικές συμπίεσης σε αντιπαράθεση με την τεχνολογία δικτύων..... | 22 |
| 3. MPEG-1 | 28 |
| 3.1. Γενικά για το πρότυπο MPEG-1 | 28 |
| 3.2. Ηχητικό MPEG-1 (Audio MPEG-1)..... | 32 |
| 4. MPEG-2 | 36 |
| 4.1. Γενικά για το πρότυπο MPEG-2 | 36 |
| 5. MPEG-4 | 42 |
| 5.1. Βασικές λειτουργίες του MPEG-4 στην έκδοση 1 | 44 |
| 5.1.1. DMIF..... | 45 |
| 5.1.2. Συστήματα | 45 |
| 5.1.3. Ήχος..... | 46 |
| 5.1.4. Εικόνα | 48 |
| 5.1.4.1. Φορμάτ που υποστηρίζονται..... | 48 |
| 5.1.4.2. Αποδοτικότητα συμπίεσης..... | 48 |
| 5.1.4.3. Λειτουργίες βασισμένες στο περιεχόμενο | 49 |

| | |
|--|-----------|
| 5.1.4.4. Κλιμάκωση των χαρακτηριστικών, των εικόνων και του βίντεο | 49 |
| 5.1.4.4.5 Μορφή και κωδικοποίηση καναλιών άλφα | 50 |
| 5.2. Βασικές λειτουργίες του MPEG-4 στην έκδοση 2 | 50 |
| 5.2.1. Συστήματα | 51 |
| 5.2.2. Εικόνα | 51 |
| 5.2.3. Body animation | 51 |
| 5.2.4. Κωδικοποίηση των τρισδιάστατων (3D) πολυγωνικών πλεγμάτων | 52 |
| 5.2.5. Ήχος | 52 |
| 5.2.6. DMIF | 52 |
| 5.3. Επεκτάσεις του MPEG-4 πέρα από την έκδοση 2 | 56 |
| 5.3.1. Εικόνα | 56 |
| 5.3.2. Συστήματα | 56 |
| 5.3.2.1 Εξελιγμένα BIFS | 56 |
| 5.3.2.2 Φορμάτ Κειμένου | 56 |
| 5.3.3. Προφίλ | 57 |
| 5.3.3.1. Οπτικά Προφίλ | 58 |
| 5.3.4. H.261 και H.263 | 60 |
| 5.4. Το MPEG-4 και η διαδραστική τηλεόραση | 63 |
| 5.5. Το μέλλον του MPEG-4 | 66 |
| 6. Internet Protocol Television (IP-TV) | 66 |
| 6.1. Εισαγωγή | 66 |
| 6.2. Το δίκτυο της IP-TV | 67 |
| 6.3. Εφαρμογές της IP-TV | 68 |
| 6.4. Τα πλεονεκτήματα της IP-TV | 68 |
| 6.5. Συμπερασματικά | 69 |
| 7. Παράδειγμα | 70 |
| 8. Πειράματα – Συμπεράσματα | 71 |
| 8.1. Mpegable Broadcaster | 71 |
| 8.2 Canopus ProCoder | 77 |
| 8.2.1. Παρατηρήσεις | 81 |
| 8.3 PCTV Vision | 81 |
| 9. Βιβλιογραφία | 83 |

Θα ήθελα να πω ένα μεγάλο ευχαριστώ στον καθηγητή μου Ευάγγελο Πάλλη για την πολύτιμη βοήθειά του καθώς και για τις εύστοχες παρατηρήσεις και διορθώσεις του.

Επίσης ευχαριστώ όλους όσους έκαναν υπομονή κοντά μου και με συμβούλευσαν όταν τους χρειάστηκα.

1. Πρόλογος - Στόχος

Το περιορισμένο εύρος ζώνης καθώς και η μη εγγυημένη ποιότητα υπηρεσιών που παρέχει το διαδίκτυο (Internet), αποτελούσαν μέχρι πρόσφατα απαγορευτικούς παράγοντες για την δημιουργία και παροχή (σε τελικούς χρήστες) δεδομένων εικόνας και ήχου σε πραγματικό χρόνο. Η μετάδοση τέτοιων ψηφιακών δεδομένων με ποιότητα όπως αυτή που προσφέρει η ψηφιακή τηλεόραση (in terms of resolution, Frame rate, picture clarity and clearness, κτλ.) προϋποθέτει την ύπαρξη ευρυζωνικών καναλιών μετάδοσης καθώς και την χρήση τεχνολογιών/πρωτοκόλλων μεταφοράς, που θα εξεσφαλίζουν την ποιότητα τέτοιων υπηρεσιών/προγραμμάτων στον τελικό χρήστη. Για παράδειγμα, ένα ψηφιακό τηλεοπτικό πρόγραμμα MPEG-2 μπορεί να ψηφιοποιηθεί και να συμπιεστεί σε ρυθμούς από 2 έως 15Mbps (ML@MP), ενώ για την διασφάλιση της ποιότητάς στον τελικό χρήστη γίνεται χρήση πολύπλοκων τεχνικών διόρθωσης λαθών (π.χ. Reed Solomon).

Η ανάπτυξη, ωστόσο, σήμερα νέων τεχνικών ψηφιοποίησης και συμπίεσης των δεδομένων εικόνας και ήχου, όπως το πρότυπο MPEG-4, επιτρέπουν την δημιουργία ψηφιακών τηλεοπτικών προγραμμάτων και την παροχή αυτών μέσα από το διαδίκτυο, προσφέροντας ρυθμούς κωδικοποίησης από μερικά kbps μέχρι 2Mbps, και συνεπώς περιορίζοντας τις δικτυακές απαιτήσεις σε εύρος ζώνης.

Στόχος αυτή της πτυχιακής εργασίας είναι η μελέτη του προτύπου MPEG-4 για την δημιουργία τηλεοπτικών προγραμμάτων που θα παρέχονται μέσα από το διαδίκτυο (IP-TV). Για εγγυημένη ποιότητα υπηρεσίας (QoS) στον τελικό χρήστη, θα δοκιμασθούν υπάρχοντα πρωτόκολλα μεταφοράς/μετάδοσης (RTP, RTSP), όπως επίσης και διάφορες τεχνικές μετάδοσης (unicast, multicast, broadcast).

1.1. Περίληψη περιεχομένου εργασίας

❖ Τί σημαίνουν τα αρχικά στο MPEG ?

Moving Pictures Experts Group.

❖ Τί είναι το MPEG ?

Το MPEG είναι μια ομάδα εργασίας, η οποία ανήκει σε μία υποεπιτροπή του ISO/IEC (the International Standards Organisation/International Electrotechnical Commission) και δημιουργεί γενικά πρότυπα (standards) για την ψηφιακή τηλεοπτική και ακουστική συμπίεση (video & audio compression). Ειδικότερα, το MPEG καθορίζει τη σύνταξη των χαμηλών ρυθμών μετάδοσης (Low bit-rate) τηλεοπτικών και ακουστικών δεδομένων και τη λειτουργία των conformant αποκωδικοποιητών. Οι αλγόριθμοι των κωδικοποιητών δεν καθορίζονται από το MPEG. Αυτό επιτρέπει τη συνεχή βελτίωση των κωδικοποιητών και την προσαρμογής τους σε συγκεκριμένες εφαρμογές, μέσα στον καθορισμό της ροής μετάδοσης (bit stream). Μαζί με την τηλεοπτική και ακουστική κωδικοποίηση (video & audio compression), το MPEG καθορίζει επίσης τα μέσα ώστε να πολυπλεχθούν διάφορες τηλεοπτικές και ακουστικές ροές (bit streams) συγχρόνως σε μία ενιαία ροή, περιγράφει τις μεθόδους για να εξεταστεί η προσαρμογή των ροών δεδομένων και των αποκωδικοποιητών στα πρότυπα, και δημοσιεύει τεχνικές αναφορές που περιέχουν λογισμικό περιγραφής της λειτουργίας των αποκωδικοποιητών και λογισμικό περιγραφής παραδειγμάτων

της λειτουργίας των κωδικοποιητών. Οι άνθρωποι του MPEG συνήθως συναντιούνται τέσσερις φορές ετησίως, για μια εβδομάδα περίπου.

❖ Έτοιμοι για συμπίεση (Dressed to compress)

Οι χειριστές των τηλεπικοινωνιών φαίνεται να επωφελούνται από τη διαθεσιμότητα της νέας τεχνολογίας σε χαμηλούς ρυθμούς μετάδοσης (new low bit-rate IP video compression technology) ώστε να τους χρησιμοποιήσουν στην IP τηλεόραση (Internet Protocol TeleVision). Αλλά πόσο σύντομα οι κατασκευαστές υλικού θα είναι σε θέση να ικανοποιήσουν τις ανάγκες της νέας τεχνολογίας σε ένα βιώσιμο οικονομικά επίπεδο τιμών; Με συνεχείς επιτυχίες η IPTV στην Ευρώπη γίνεται γρήγορα μια πραγματικότητα. Η “επανάσταση” αυτή καθιστά το τηλεοπτικό πρόγραμμα περιττό. Όμως, προς το παρόν, οι περισσότερες προσπάθειες παραμένουν σε ένα δοκιμαστικό στάδιο. Υπάρχουν διάφοροι παράγοντες που κρατούν τις μεγάλες εταιρίες τηλεπικοινωνιών πίσω. Η έλλειψη βεβαιότητας για τα πιθανά κέρδη είναι ένας από αυτούς καθώς και η αμφισβήτηση της ετοιμότητας των δικτύων τους να προσφέρουν τις τηλεοπτικές υπηρεσίες είναι σαφώς ένας ακόμα. Η έλλειψη της ικανότητας πολλαπλής διανομής στα δίκτυα και η έλλειψη εύρους ζώνης έχουν αναγνωριστεί από καιρό ως προβλήματα που αντιμετωπίζουν οι μεγάλες και επιβεβλημένες τηλεπικοινωνιακές εταιρίες. Γιατί :

“It’s all about the bandwidth...”
(actually, it’s all about the \$\$\$)

Για να ανακουφίσουν τον αντίκτυπο των τελευταίων ελλείψεων, οι τεχνικοί είναι απασχολημένοι με τις πλατφόρμες που υποστηρίζουν τις τεχνολογίες χαμηλής ροής δεδομένων.

Στην παρούσα εργασία θα αναλύσουμε την τεράστια ανάγκη που προέκυψε για συμπίεση. Συμπίεση, τόσο των δεδομένων ήχου όσο και των δεδομένων εικόνας, είτε η τελευταία είναι κινούμενη είτε όχι. Η συμπίεση που θα πραγματοποιήσουμε, θα γίνει με το πρότυπο MPEG-4 και γι’ αυτό θα αναλυθούν πλήρως οι ακρογωνιαίοι λίθοι του, καθώς και ένα προς ένα τα επακόλουθα χρήσης του. Με τη χρήση του MPEG-4 συντελούμε όλοι ώστε να κατασταθεί εφικτή η εμπειρία της ψηφιακής τηλεόρασης και να εισαχθεί η ψυχαγωγία που παρέχει αυτή στο “μοντέρνο” καθιστικό!

❖ MPEG-4: το πολυμεσικό πρότυπο του μέλλοντος

Θα επιθυμούσατε να “χωρέσετε” έναν ολόκληρο δίσκο ποιότητας DVD σε έναν δίσκο τύπου CD-R ή θα επιθυμούσατε να στείλετε (stream) μία ταινία στο κινητό τηλέφωνό σας ενώ ταξιδεύετε; Ή προτιμάτε να έχετε τη μουσική σας συμπιεσμένη με μία μορφή (format) που προσφέρει την καλύτερη ποιότητα, χρησιμοποιώντας τον λιγότερο δυνατό χώρο; Όλα αυτά μπορούν να γίνουν αληθινά. Το MPEG-4 είναι η νέα λέξη κλειδί που πρέπει να έχουμε στο μυαλό μας. Αυτό το πρότυπο θα είναι το μεγαλύτερο γεγονός στα ερχόμενα έτη.

Αρχικά, θα ήταν σημαντικό να γίνει κατανοητό ότι το MPEG-4 δεν έχει καμία σχέση με το MP3. Το MP3 είναι ένα μέρος του MPEG-1 προτύπου, ακριβώς όπως το video-CD είναι μέρος του MPEG-1 προτύπου. Ουσιαστικά το MPEG-4 δεν είναι ούτε συνέχεια του MPEG-1. Είναι ένα νέο πρότυπο που στοχεύει να γίνει το βασικότερο

πρότυπο για τα ψηφιακά μέσα ενώ το MPEG-1 στράφηκε κυρίως στο να παραδώσει “υλικό” με τον τρόπο που το ξέρουμε σήμερα.

Οι τεχνολογικές εργασίες του MPEG-4 προτύπου απλοποιούνται, με το διαχωρισμό του περιεχομένου προς επεξεργασία. Μία μικρή ταινία μπορεί, παραδείγματος χάριν, να ιδωθεί ως ήχος, βίντεο, τίτλοι και υπότιτλοι. Τέσσερα διαφορετικά αντικείμενα, που μαζί διαμορφώνουν μία πλήρη ταινία. Εάν θέλετε την καλύτερη ποιότητα χρησιμοποιώντας τον λιγότερο χώρο στο δίσκο σας, θα πρέπει να αναλύσετε κάθε ένα από αυτά τα αντικείμενα ξεχωριστά και να επιλέξετε τον κατάλληλο τρόπο συμπίεσης για το καθένα. Μπορείτε να φανταστείτε ότι εάν σε μία ταινία κάποιος κάνει μόνο ένα τηλεφώνημα, θα μπορούσατε να χρησιμοποιήσετε έναν τρόπο συμπίεσης για τον ήχο, που χρειάζεται λιγότερη ποιότητα από όταν βλέπετε μία ορχήστρα να δίνει ένα κονσέρτο μέσα σε μία όπερα, στην ίδια ταινία. Μπορείτε επίσης να φανταστείτε ότι σε περίπτωση που το πρόσωπο που κάνει το τηλεφώνημα κινεί μόνο τα χείλια του, χρειάζεστε λιγότερη ποιότητα βίντεο από όταν παρουσιάζεται ολόκληρη κινούμενη ορχήστρα παίζοντας ένα δυνατό τραγούδι.

Μεγίστης σημασίας, στο πρότυπο MPEG-4 είναι τα λεγόμενα «προφίλ». Τα προφίλ αυτά, περιγράφουν κατά κάποιο τρόπο το περιεχόμενο ενός αντικειμένου MPEG-4. Παραδείγματος χάριν, για το τηλεφώνημα θα χρησιμοποιούσαμε το ακουστικό προφίλ (Speech Audio Profile), ενώ η ορχήστρα θα απαιτούσε το υψηλής ποιότητας ακουστικό προφίλ (High Quality Audio Profile). Οι ίδιοι κανόνες ισχύουν και για το βίντεο και οι δύο ταινίες θα απαιτούσαν ένα διαφορετικό οπτικό προφίλ (visual profile). Τα προφίλ έχουν σχεδιαστεί με σκοπό να πιστοποιούν ότι μια συσκευή ξέρει εάν μπορεί να παίξει ένα συγκεκριμένο MPEG-4 αρχείο.

Αυτή η ταινία με το τηλεφώνημα και την ορχήστρα είναι μόνο ένα μικρό παράδειγμα. Τα αντικείμενα του MPEG-4, μπορούν επίσης να εμφανιστούν αυτόνομα. Για παράδειγμα, το ακουστικό φορμάτ (audio format) που χρησιμοποιήθηκε στο τηλεφώνημα θα μπορούσε επίσης να χρησιμοποιηθεί για το λογισμικό που είναι σε θέση να υποστηρίξει τα τηλεφωνήματα μέσω του Διαδικτύου. Το περιεχόμενο του MPEG-4 δεν είναι απαραίτητο να περιοριστεί σε ταινίες ή ηχητικά κομμάτια, μπορεί επίσης να χρησιμοποιηθεί για διαδραστικές εφαρμογές και παιχνίδια. Με το συνδυασμό πολλαπλών MPEG-4 τεχνολογιών, δημιουργείτε ένα .mp4 αρχείο που περιέχει όλα εσείς επιθυμείτε και που μπορεί να “παιχτεί” σε μια ευρεία σειρά συσκευών.

2. Η συμπίεση του βίντεο

2.1. Ψηφιακές τεχνικές συμπίεσης βίντεο

Το αναλογικό βίντεο ψηφιοποιείται έτσι ώστε μπορεί να επεξεργαστεί από έναν υπολογιστή. Κάθε πλαίσιο του βίντεο γίνεται ένας διδιάστατος πίνακας εικονοστοιχείων (pixels). Μια έγχρωμη εικόνα αποτελείται από τρία πλαίσια εικόνας, ένα για κάθε συστατικό χρώμα.

Οι ασυμπίεστες εικόνες και το βίντεο είναι πάρα πολύ μεγάλα για επεξεργασία και η συμπίεση απαιτείται για την αποθήκευση και τη μετάδοσή τους. Σημαντικές μετρικές της συμπίεσης είναι η αναλογία συμπίεσης (compression ratio) και τα κομμάτια (bit) ανά εικονοστοιχείο (bits per pixel) (ο αριθμός bit που απαιτείται για να αντιπροσωπευθεί ένα εικονοστοιχείο στην εικόνα).

Η συμπίεση του βίντεο είναι συνήθως απωλεστική μέθοδος, εννοώντας ότι μερικές από τις πληροφορίες χάνονται κατά τη διάρκεια του σταδίου της συμπίεσης. Αυτό

είναι αποδεκτό, επειδή οι αλγόριθμοι κωδικοποίησης έχουν ως σκοπό να απορρίψουν τις πληροφορίες που δεν είναι αντιληπτές από τους ανθρώπους ή τις πληροφορίες που είναι περιττές (πλεονάζουσες). Υπάρχουν μερικές βασικές τεχνικές, κοινές για τους περισσότερους αλγορίθμους συμπίεσης βίντεο, συμπεριλαμβανομένης της χωρικής δειγματοληψίας χρώματος (color space sampling) και της μείωσης πλεονασμού (redundancy reduction).

Η χωρική δειγματοληψία χρώματος (color space sampling) είναι μια αποτελεσματική τεχνική που χρησιμοποιείται για να μειώσει το ποσοστό των στοιχείων που πρέπει να κωδικοποιηθούν. Εάν μια εικόνα κωδικοποιείται στο διάστημα YUV, οι συνιστώσες U και V μπορούν να υποδειγματοληφτηθούν επειδή το ανθρώπινο μάτι είναι λιγότερο ευαίσθητο στις πληροφορίες χρωμικότητας (chrominance).

Η μείωση του πλεονασμού (Redundancy reduction) είναι μια άλλη τεχνική μείωσης το ποσοστού της πληροφορίας προς κωδικοποίηση. Η κωδικοποίηση Intraframe επιτυγχάνει συμπίεση με τη μείωση του χωρικού πλεονασμού (spatial redundancy) μέσα σε μια εικόνα. Αυτή η τεχνική λειτουργεί επειδή τα γειτονικά εικονοστοιχεία (pixels) είναι συνήθως παρόμοια σε μία εικόνα. Η κωδικοποίηση Interframe επιτυγχάνει τη συμπίεση με τη μείωση του χρονικού πλεονασμού (temporal redundancy) μεταξύ των εικόνων.

Υπάρχουν πολλοί διαφορετικοί τρόποι συμπίεσης δεδομένων. Αυτό που κάνουν οι τρόποι αυτοί είναι να αντικαθιστούν τις δυαδικές ακολουθίες που περιλαμβάνονται σε κάθε δέσμη δεδομένων με μια συμπιεσμένη απεικόνιση. Μερικοί τρόποι συμπίεσης οδηγούν σε απώλεια δεδομένων (οπότε καλούνται απωλεστικοί τρόποι συμπίεσης) και συνήθως χρησιμοποιούνται αρκετά για τη συμπίεση ομιλιών, ακουστικών και τηλεοπτικών δεδομένων (audio-visual data) όταν αυτά μπορούν να προσαρμοστούν σε κάποιο βαθμό απώλειας στοιχείων. Οι γενικότεροι τρόποι συμπίεσης που χρησιμοποιούνται για συμπίεση αυθαίρετων τύπων δεδομένων είναι χωρίς απώλειες, δεδομένου ότι η νέα δυαδική ακολουθία είναι ακριβώς η ίδια με την αρχική. Όπως μπορεί να αναμένεται, οι αναλογίες συμπίεσης στις απωλεστικές μεθόδους συμπίεσης είναι συνήθως πολύ υψηλότερες από εκείνες των μη απωλεστικών μεθόδων.

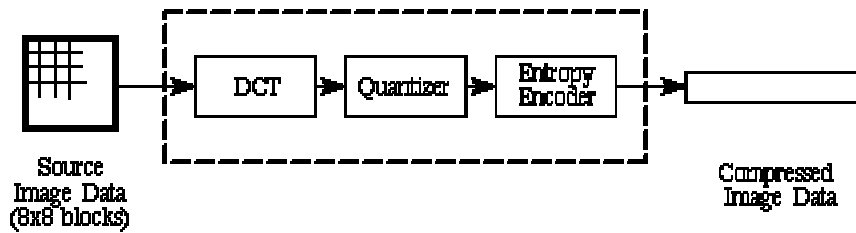
2.1.1. MJPEG

Το JPEG είναι ένα πρότυπο κωδικοποίησης για τις εικόνες που αναπτύσσονται από τους ανθρώπους του προτύπου. Αν και έχει σχεδιαστεί για τις εικόνες, με ειδικό υλικό είναι δυνατό να κωδικοποιηθεί και να αποκωδικοποιηθεί μια σειρά εικόνων JPEG σε πραγματικό χρόνο ώστε να επιτευχθεί κινούμενο βίντεο. Αυτή η χρήση της κωδικοποίησης JPEG αναφέρεται χαρακτηριστικά ως κινούμενο JPEG ή MJPEG (Moving JPEG). Παρόλα αυτά, κανένα επίσημο πρότυπο MJPEG δεν υπάρχει.

Υπάρχουν τέσσερις καθορισμένοι τρόποι λειτουργίας για το JPEG: διαδοχικός, προοδευτικός, χωρίς απώλειες, και ιεραρχικός.

Το JPEG χρησιμοποιεί μια χωρική τεχνική συμπίεσης intraframe, τον διακριτό συνημιτονικό μετασχηματισμό (DCT). Αυτή η τεχνική χρησιμοποιείται και από άλλες μεθόδους κωδικοποίησης όπως η H.261. Σε αυτήν την περίπτωση το JPEG προϋποθέτει την κωδικοποίηση μιας ενιαίας συστατικής εικόνας (grayscale). Δε θα γίνει αναφορά σε εικόνες με πολλαπλά συστατικά χρώματος. Η κωδικοποίηση JPEG είναι ανεξάρτητη του διαστήματος του χρώματος (color space independent), αν και

τα συστήματα μπορούν να επιλέξουν να μετατρέψουν τις εικόνες στο διάστημα YUV και να υποδειματοληπτήσουν τα συστατικά της χρωματικότητας (chrominance).



Σχήμα - DCT βασισμένη κωδικοποίηση

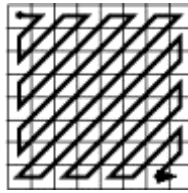
Το πρώτο βήμα της διαδικασίας κωδικοποίησης είναι να εκτελεσθεί ένας DCT μετασχηματισμός σε δεσμίδες 8x8 των δειγμάτων. Αυτό το βήμα μεταφέρει την πληροφορία στο πεδίο της συχνότητας. Η “έξοδος” αυτού του βήματος είναι 64 συνημιτονικοί συντελεστές. Ο μετασχηματισμός DCT έχει την επίδραση της συγκέντρωσης του μεγαλύτερου μέρους της πληροφορίας στην ανώτερη αριστερή γωνία της 8x8 δεσμίδας. Η μέση αξία της κάθε δεσμίδας, αποκαλούμενη “συνεχής συνιστώσα” (DC component), είναι ο συντελεστής που βρίσκεται άνω και αριστερά. Οι υπόλοιποι συντελεστές καλούνται εναλλασσόμενοι συντελεστές (AC coefficients). Καμία πληροφορία δεν χάνεται κατά τη διάρκεια του σταδίου DCT. Τα στοιχεία απλώς μετασχηματίζονται σε μια άλλη “περιοχή” και μπορούν να ανακτηθούν με την εκτέλεση ενός αντίστροφου μετασχηματισμού DCT.

Το επόμενο βήμα της διαδικασίας κωδικοποίησης είναι η κβαντοποίηση. Οι συντελεστές του DCT μετασχηματισμού διαιρούνται με μια 8x8 μήτρα κβαντοποίησης. Αυτή η μήτρα σχεδιάζεται για να μειώσει το εύρος των συντελεστών και για να αυξήσει τον αριθμό συντελεστών μηδενικής αξίας. Το βήμα της κβαντοποίησης είναι το απωλεστικό βήμα της διαδικασίας κωδικοποίησης.

Μετά από την κβαντοποίηση, μια ροή bit διαμορφώνεται από τη δεσμίδα. Η “συνεχής συνιστώσα” κωδικοποιείται ως διαφορά μεταξύ του τρέχουσας συνεχούς συνιστώσας και της συνεχούς συνιστώσας της προηγούμενης δεσμίδας. Οι εναλλασσόμενοι συντελεστές κωδικοποιούνται σε μια ακολουθία τρεκλίσματος (zig-zag) ξεκινώντας από τον ανώτερο αριστερά σε κάθε δεσμίδα και καταλήγοντας στον κατώτατο δεξιά. Οι εναλλασσόμενοι συντελεστές είναι κωδικοποιημένοι με τη μέθοδο run-length και τη μέθοδο της εντροπίας. Η κωδικοποίηση run-length αφαιρεί τις μεγάλες σειρές αποτελούμενες από μηδενικούς συντελεστές. Η μέθοδος της εντροπίας (ή κωδικοποίηση Huffman ή εναλλακτικά αριθμητική κωδικοποίηση) κωδικοποιεί την πληροφορία, η οποία είναι σίγουρα βασισμένη σε στατιστικά χαρακτηριστικά. Τα συμπλέγματα, που είναι στατιστικά πιθανότερο να εμφανιστούν κωδικοποιούνται με τις πιό σύντομες κωδικές λέξεις.

Η διαδικασία αποκωδικοποίησης είναι βασικά η αντιστροφή της διαδικασίας κωδικοποίησης. Ο αποκωδικοποιητής πρέπει να έχει πρόσβαση στους ίδιους πίνακες κβαντοποίησης και εντροπίας που χρησιμοποιήθηκαν για να κωδικοποιηθούν τα δεδομένα.

Η κωδικοποίηση JPEG επιτυγχάνει χαρακτηριστικά συμπίεση από μία αρχική κλίμακα 10:1 έως μία υψηλότερη 20:1. Εδώ πρέπει να σημειωθεί ότι όσο υψηλότερα είναι τα ποσοστά της συμπίεσης τόσο χειρότερη είναι η ποιότητα της εικόνας. Μια ποιοτική παράμετρος, η οποία διαμορφώνεται από τον χρήστη είναι συνήθως διαθέσιμη και επιτρέπει ανταλλαγές μεταξύ συμπίεσης και ποιότητας ώστε να καταφέρει ο καθένας αυτό που του χρειάζεται στην εκάστοτε εφαρμογή. Αυτή η παράμετρος προκαλεί συνήθως ενός είδους κλιμάκωση στη μήτρα κβαντοποίησης.



Σχήμα - Zig-zag

2.1.2. ITU-T Recommendation H.261

Το ITU-T H.261 είναι ένα πρότυπο συμπίεσης βίντεο που σχεδιάστηκε να λειτουργήσει στα εύρη ζώνης επικοινωνίας από 64 kbps έως 2 Mbps, μετρημένα σε διαστήματα των 64kbps. Αυτή η τεχνική αναφέρεται επίσης ως “px64” όπου το “p” κυμαίνεται από το 1 έως το 30. Το H.261 σχεδιάστηκε πρώτα για τη βίντεο - συνεδρίαση μέσω μίας ISDN σύνδεσης και διευκρινίζεται από το H.320 .

Το H.261 χρησιμοποιεί και τη χωρική καθώς και τη χρονική κωδικοποίηση intraframe. Στον τρόπο κωδικοποίησης intraframe χρησιμοποιείται η DCT βασισμένη χωρική συμπίεση. Στον τρόπο κωδικοποίησης intarframe, ο συμψηφισμός (compensation) της κίνησης εκτελείται για να υπολογίσει τις διαφορές μεταξύ των πλαισίων. Οι διαφορές, συνήθως μικρού μεγέθους, έπειτα κωδικοποιούνται με το μετασχηματισμό DCT.

Δύο μορφές εικόνων, CIF (Common Intermediate Format) και QCIF (Quarter CIF) καθορίζονται. Η λειτουργία QCIF είναι υποχρεωτική, ενώ η λειτουργία CIF είναι προαιρετική. Η QCIF χρησιμοποιείται συνήθως στους χαμηλούς ρυθμούς μετάδοσης, όπως $p < 3$. Οι εικόνες αποτελούνται από τρία τμήματα χρώματος, το Y και δύο διαφορές χρώματος, τη Cb και τη Cr (το YCbCr αντιστοιχεί σε έναν μετασχηματισμό του πεδίου YUV). Οι συνιστώσες της διαφοράς χρώματος περιέχουν το μισό ποσό πληροφορίας όπως η συνιστώσα της φωτεινότητας (για κάθε 4 δεσμίδες πληροφορίας φωτεινότητας που κωδικοποιούνται, αντιστοιχούν μόνο 2 δεσμίδες με πληροφορία χρωματικότητας). Ο πίνακας παρουσιάζει τις γραμμές ανά πλαίσιο και τα εικονοστοιχεία ανά γραμμή για τα CIF και QCIF.

Η κωδικοποίηση Intraframe δουλεύει αποτελεσματικά με το πρότυπο JPEG. Δεσμίδες 8x8 μετασχηματίζονται συνημιτονικά (DCT), κβαντοποιούνται και κωδικοποιούνται με μία από τις μεθόδους run-length ή εντροπίας. Στον τρόπο κωδικοποίησης intarframe, η πρόβλεψη για τις δεσμίδες στο τρέχον πλαίσιο γίνεται βασισμένη στο προηγούμενο πλαίσιο. Εάν η διαφορά μεταξύ της τρέχουσας δεσμίδας και της προβλεφθείσας είναι κάτω από ένα ορισμένο “κατώφλι” (threshold) τότε δεν στέλνονται καθόλου δεδομένα. Διαφορετικά, η διαφορά υπολογίζεται κανονικά, μετασχηματίζεται συνημιτονικά (DCT), κβαντοποιείται και κωδικοποιείται (με μία από τις μεθόδους εντροπίας/run-length).

| | CIF lines/frame | CIF pixels/line | QCIF lines/frame | QCIF pixels/line |
|------------------|--------------------|--------------------|---------------------|---------------------|
| Luminance | 288 | 352 | 144 | 176 |
| Chrominance (Cb) | 144 | 176 | 72 | 88 |
| Chrominance (Cr) | 144 | 176 | 72 | 88 |

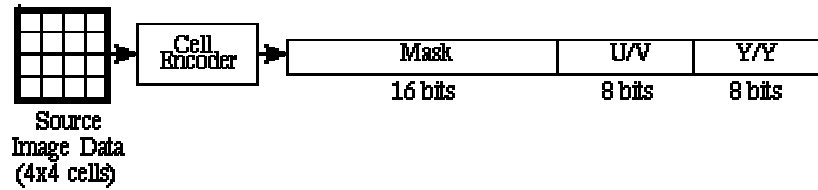
Σχήμα - CIF και QCIF μορφές εικόνων

Το βήμα κβαντοποίησης καθορίζει το ποσό πληροφορίας που στέλνεται, οπότε περισσότερη πληροφορία σημαίνει και καλύτερη ποιότητα εικόνας. Οι H.261 κωδικοποιητές ρυθμίζουν την τιμή της κβαντοποίησης ώστε να επιτύχουν ένα σταθερό ρυθμό μετάδοσης. Εάν ο καταχωρητής μετάδοσης είναι σχεδόν πλήρης, τότε αυξάνεται το βήμα του κβαντιστή, αναγκάζοντας λιγότερη πληροφορία να κωδικοποιηθεί με αποτέλεσμα να πετύχουμε χειρότερη ποιότητα εικόνας. Ομοίως, όταν ο καταχωρητής δεν είναι πλήρης, το βήμα του κβαντιστή μειώνεται, αναγκάζοντας περισσότερη πληροφορία να κωδικοποιηθεί και επιτυγχάνοντας καλύτερη ποιότητα εικόνας. Λόγω αυτής της ρύθμισης του κβαντιστή, οι γρήγορα μεταβαλλόμενες σκηνές θα έχουν χειρότερη ποιότητα από τις στατικές σκηνές.

2.1.3. CellB

Το κύτταρο είναι μια μη ειδικευμένη τεχνική συμπίεσης που αναπτύσσεται από την εταιρία Sun Microsystems. Υπάρχουν δύο τύποι συμπίεσης κυττάρων, η συμπίεση CellA και η συμπίεση CellB. Η CellA είναι μια ασύμμετρη τεχνική, που απαιτεί περισσότερο υπολογισμό για τη συμπίεση παρά για την αποσυμπίεση. Το CellB είναι παρόμοιο με το CellA αλλά είναι πιο συμμετρικό (υπολογιστικά), καθιστώντας το περισσότερο κατάλληλο για χρήση σε πραγματικό χρόνο όπως για τη βίντεο-συνεδρίαση (videoconferencing). Το CellA παίρνει ως είσοδο το RGB πρότυπο ενώ το CellB παίρνει το YUV. Και τα δύο απαιτούν το πλάτος και το ύψος της εικόνας είναι διαιρούμενα από το τέσσερα. Η εικόνα εισόδου διαιρείται σε ομάδες 4x4 εικονοστοιχείων αποκαλούμενες κύτταρα. Η τεχνική κωδικοποίησης των κυττάρων είναι βασισμένη σε μια μέθοδο αποκαλούμενη κωδικοποίηση αποκοπής δεσμίδων (Block Truncation Coding) (BTC). Τα 16 εικονοστοιχεία σε κάθε κύτταρο κωδικοποιούνται από μία μάσκα 16 bit και δύο εντάσεις (intensities) των 8 bit. Το CellA και το CellB διαφέρουν στον τρόπο που θα επιλεγθούν οι εντάσεις αυτές. Στο CellA, οι τιμές (αξίες) είναι δείκτες σε έναν χάρτη χρώματος (colormap). Στο CellB, οι τιμές είναι δείκτες στους διανυσματικούς πίνακες κβαντοποίησης. Οι προεπιλεγμένοι (default) πίνακες είναι μέρος των προδιαγραφών του CellB.

Το σχήμα παρουσιάζει ένα κύτταρο τύπου CellB και τον κυτταρικό κώδικα που διαμορφώνεται μετά από την κωδικοποίηση. Ο τομέας U/V απεικονίζει τη συνιστώσα της χρωματικότητας (chrominance). Ο τομέας Y/Y απεικονίζει δύο τιμές φωτεινότητας. Εάν το bit ενός εικονοστοιχείου τίθεται στη μάσκα των bit, το εικονοστοιχείο αποδίδεται ως < Y1, u, B > και εάν το bit δεν τεθεί στη μάσκα τότε το εικονοστοιχείο αποδίδεται ως < Y2, u, B >. Αυτό σημαίνει ότι για κάθε κωδικοποιημένο κύτταρο υπάρχουν δύο πιθανές τιμές φωτεινότητας και μια πιθανή αξία χρωματικότητας (chrominance). Το CellB κωδικοποιεί 16 εικονοστοιχεία (384 bit) χρησιμοποιώντας 32 bit. Αυτή είναι μια αναλογία συμπίεσης 12:1.



Σχήμα - Κωδικοποίηση τύπου CellB

Το CellB είναι βασικά μια intraframe τεχνική κωδικοποίησης. Εντούτοις υποστηρίζει μια απλή μορφή κωδικοποίησης intraframe καθορίζοντας έναν κώδικα (skip code), που δείχνει ότι ορισμένα κύτταρα πρέπει να προσπεραστούν, γεγονός το οποίο σημαίνει ότι η αξία τους δεν έχει αλλάξει.

Ένα πλεονέκτημα της κωδικοποίησης κυττάρων είναι ότι η διαδικασία αποκωδικοποίησης είναι παρόμοια με τη διαδικασία επίδειξης χαρακτήρα για ένα framebuffer χρώματος. Η διαδικασία επίδειξης χαρακτήρα για έναν framebuffer παίρνει ως εισαγωγή ένα χρώμα πρώτου πλάνου, το χρώμα υποβάθρου και τη μάσκα που προσδιορίζει τίνος χρώματος κάνει χρήση το εικονοστοιχείο. Δεδομένου ότι αυτό είναι μια πολύ σημαντική λειτουργία ενός παραθυρικού (windowing) συστήματος, εφαρμόζεται συχνά ως πρωτόγονη επίδειξη στους επιταχυντές γραφικής παράστασης. Επομένως, η κωδικοποίηση κυττάρων μπορεί να αξιοποιήσει τα υπάρχοντα πρωτότυπα υλικά για να επιτύχει την αποδοτική αποκωδικοποίηση.

2.1.4. nv

Το δικτυακό βίντεο (nv, Network Video) είναι ένα εργαλείο συνεδριάσεων βίντεο μέσω διαδικτύου που αναπτύχθηκε από την εταιρία Xerox//PARC. Είναι το πιο σύνθητες εργαλείο για βίντεο στο Mbone του διαδικτύου. Η σχετική με το nv τεχνική κωδικοποίησης χρησιμοποιεί τη χωρική (intraframe) και χρονική (interframe) συμπίεση. Το πρώτο βήμα του αλγορίθμου nv συγκρίνει το τρέχον πλαίσιο με το προηγούμενο πλαίσιο και επισημαίνει τις περιοχές που έχουν αλλάξει σημαντικά. Κάθε περιοχή που έχει αλλάξει συμπιέζεται χρησιμοποιώντας έναν μετασχηματισμό. Είτε έναν συνημιτονικό μετασχηματισμό (DCT) είτε τον μετασχηματισμό κυμάτων Haar (Haar wavelet transform). Ο κωδικοποιητής nv επιλέγει δυναμικά ποιος μετασχηματισμός θα χρησιμοποιηθεί βασιζόμενος στο εάν το εύρος ζώνης των δικτύων (οπότε απαιτείται χρήση του DCT) ή η τοπική υπολογιστική δύναμη (οπότε απαιτείται χρήση του μετασχηματισμού Haar) περιορίζει την απόδοση. Ο DCT προτιμάται δεδομένου ότι διπλασιάζει σχεδόν την αναλογία συμπίεσης. Η έξοδος του μετασχηματισμού κβαντοποιείται και κωδικοποιείται με τη μέθοδο run-length. Περιοδικά, τα αμετάβλητα μέρη της εικόνας στέλνονται σε υψηλότερες αναλύσεις (higher resolution), το οποίο επιτυγχάνεται με την εξάλειψη του βήματος κβαντοποίησης. Εν γένει, το nv μπορεί να επιτύχει αναλογίες συμπίεσης της κλίμακας 20:1 ή περισσότερο.

2.1.5. CUSeeMe

Το CUSeeMe είναι ένα διαδικτυακό εργαλείο βίντεο συνεδριάσεων το οποίο αναπτύχθηκε στο πανεπιστήμιο του Cornell. Χρησιμοποιεί χωρική (intraframe) και χρονική (interframe) συμπίεση, με μερικές μετατροπές για να βελτιστοποιήσει την

απόδοση του σε ένα Macintosh μηχανήμα, το οποίο είναι και η αρχική πλατφόρμα λειτουργίας του. Το CUSeeMe αντιπροσωπεύει την τηλεοπτική εισαγωγή σε 16 διαβαθμίσεις του γκρι χρησιμοποιώντας 4 bit ανά εικονοστοιχείο. Η εικόνα διαιρείται σε δεσμίδες των 8x8 εικονοστοιχείων για ανάλυση. Τα νέα πλαίσια συγκρίνονται με τα προηγούμενα πλαίσια, και εάν μία δεσμίδα έχει αλλάξει σημαντικά, τότε αναμεταδίδεται. Οι δεσμίδες αναμεταδίδονται επίσης σε περιοδική βάση για να εξουδετερώσουν τις απώλειες που μπορεί να έχουν εμφανιστεί στο δίκτυο. Τα δεδομένα προς εκπομπή συμπίεζονται από έναν μη απωλεστικό αλγόριθμο που αναπτύσσεται στο πανεπιστήμιο Cornell, ο οποίος εκμεταλλεύεται τον χωρικό πλεονασμό στην κάθετη κατεύθυνση (exploits spatial redundancy in the vertical direction). Το μέγεθος του συμπίεσμένου αρχείου είναι περίπου 60% του αρχικού (μια αναλογία συμπίεσης 1.7:1). Ο αλγόριθμος κωδικοποίησης CUSeeMe είχε ως σκοπό να λειτουργήσει σε έναν υπολογιστή Macintosh. Δουλεύει με σειρές των 4-κομματιών από 8 εικονοστοιχεία η καθεμία, ως λέξεις των 32 bit. Όλη αυτή η διαδικασία αποδίδει αρκετά καλά σε κώδικα Assembly 680x0. Το προεπιλεγμένο εύρος ζώνης για μετάδοση με το CUSeeMe είναι 80 kbps. Αυτός ο αριθμός ρυθμίζεται αυτόματα βάσει των εκθέσεων απώλειας πακέτων (packet-loss) που επιστρέφονται από κάθε πρόσωπο που λαμβάνει το βίντεο. Περίπου 100 kbps απαιτούνται για ένα τυπικό σενάριο.

2.1.6. Indeo

Το Indeo είναι μια τεχνική συμπίεσης βίντεο που σχεδιάστηκε από την Intel. Εξελίχθηκε από την τεχνολογία DVI (Ψηφιακό Διαδραστικό Βίντεο, Digital Video Interactive). Το Indeo αρχίζει με είσοδο το YUV και τις συνιστώσες U και V υποδειγματοληπτησμένες (subsampling) 4:1 και οριζόντια και κάθετα. Το Indeo υποστηρίζει την εκτίμηση κινήσεων, χρησιμοποιώντας το προηγούμενο πλαίσιο για να προβλέψει τις τιμές για το τρέχον πλαίσιο και μεταδίδει τα στοιχεία μόνο εάν η διαφορά είναι σημαντική. Η κωδικοποίηση μετασχηματισμού γίνεται χρησιμοποιώντας έναν 8x8 γρήγορο μετασχηματισμό (FST, Fast Slant Transform) στον οποίο όλες οι διαδικασίες είναι είτε μετατοπίσεις είτε προσθέσεις (κανένας πολλαπλασιασμός). Η κβαντοποίηση και η κωδικοποίηση run-length/εντροπίας χρησιμοποιείται όπως στους προηγούμενους αλγορίθμους. Το Indeo διευκρινίζει ότι η κωδικοποιημένη ροή μετάδοσης είναι το πολύ 60% των δεδομένων εισόδου, επομένως η συμπίεση είναι εγγυημένα στη χειρότερη περίπτωση 1.7:1.

2.1.7. Κωδικοποίηση Run-Length

Μια από τις απλούστερες μεθόδους συμπίεσης δεδομένων είναι γνωστή ως κωδικοποίηση Run-Length (RLE). Αυτή η μέθοδος εκμεταλλεύεται την επανάληψη κομματιών της πληροφορίας με την κωδικοποίηση τους ως tuple που αποτελούνται από τον αριθμό των επαναλήψεων και του συμβόλου ή του αριθμού που επαναλαμβάνεται. Η κωδικοποίηση Run-Length χρησιμοποιούμενη γενικά, αποδίδει καλά στα πολύ καλά συσχετισμένα στοιχεία. Ένα παράδειγμα μιας εικόνας που συμπίεστηκε καλά με τη μέθοδο Run-Length είναι μια εικόνα κινούμενων σχεδίων χαμηλού χρώματος ή ένα ασπρόμαυρο γραμμικό σχέδιο. Και τα δύο στυλ εικόνας περιέχουν μεγάλες περιοχές με τα ίδια εικονοστοιχεία χρώματος. Αυτό σημαίνει ότι πολλές πληροφορίες μπορούν να αντιπροσωπευθούν με ένα ενιαίο tuple. Εντούτοις,

οι εικόνες αληθινού χρώματος έχουν χαρακτηριστικά ανεπαίσθητες αλλαγές μεταξύ των παρακείμενων εικονοστοιχείων τους, που σημαίνει ότι τα μεγάλα τρεξίματα (Runs) είναι σπάνια σε τέτοιες εικόνες και επομένως η απόδοση συμπίεσης είναι σημαντικά λιγότερη.

2.1.8. Huffman και Αριθμητική κωδικοποίηση

Η κωδικοποίηση Huffman, χρησιμοποιεί έναν πίνακα ακολουθιών bit που αντιστοιχεί μια ακολουθία bit μεταβλητού μήκους σε ένα ιδιαίτερο σημείο. Για τη συμπίεση δεδομένων με κωδικοποίηση Huffman, οι πύ σύντομοι κώδικες μεταβλητού μήκους πρέπει να χρησιμοποιηθούν για τα σημεία με την υψηλότερη πιθανότητα της εμφάνισης σε μια ιδιαίτερη ροή δεδομένων (bit-stream). Για αυτόν τον λόγο, η κωδικοποίηση Huffman απαιτεί συνήθως γνωστές ροές δεδομένων (bit-stream) ώστε να καθορίσει τις σχετικές πιθανότητες των σημείων μέσα στην κάθε ροή. Εναλλακτικά, διάφορα δείγματα μπορούν να ανιχνευθούν για να αναπτύξουν ένα μέσο ιστόγραμμα της πιθανότητας ορισμένων σημείων που εμφανίζονται μέσα σε μία ροή. Η αριθμητική κωδικοποίηση χρησιμοποιεί μια παρόμοια αρχή με την κωδικοποίηση Huffman, αλλά δεν απαιτεί πίνακα για την κωδικοποίηση ή την αποκωδικοποίηση.

2.1.9. Συμπίεση αντικατάστασης (Substitutional Compression)

Πολλοί μη απωλεστικοί τρόποι συμπίεσης εμπίπτουν στην κατηγορία συμπίεσης αντικατάστασης, όπως ο κώδικας Ziv-Lempel και οι σχετικές μέθοδοι. Αυτές οι μορφές συμπίεσης είναι γνωστές και ως τρόποι συμπίεσης λεξικών. Ο όρος “συμπίεση λεξικών” προέρχεται αναλογικά από το ότι μια πρόταση μπορεί να γίνει πύ σύντομη με την αντικατάσταση μια ιδιαίτερης φράσης με μια κατάλληλη λέξη.

2.1.10. Χωρικές Μέθοδοι Μετασχηματισμών (Spatial Transform Methods)

Οι γενικές μέθοδοι συμπίεσης χρησιμοποιούν μια καθαρώς στατιστική μέθοδο για να μειώσουν τον αριθμό των bit ώστε να αντιπροσωπεύουν μία ιδιαίτερη δέσμη πληροφορίας. Στηρίζονται στα δεδομένα που αποτελούνται συνήθως από έναν μικρό αριθμό συμβόλων, με άλλα σύμβολα να εμφανίζονται σπάνια, για να κερδίσουν υψηλές αναλογίες συμπίεσης. Λαμβάνοντας υπόψη το γεγονός ότι το βίντεο ποικίλλει χρονικά και χωρικά, υπάρχει μια σπάνια περίπτωση κατά την οποία ένα ιδιαίτερο σημείο είναι επικρατέστερο ενός άλλου. Για αυτόν τον λόγο, οι γενικές μέθοδοι συμπίεσης, δε φημίζονται για συμπίεση βίντεο. Ως εκ τούτου, διάφορες άλλες μέθοδοι έχουν επινοηθεί για τη συμπίεση των ψηφιακών εικόνων και του βίντεο. Αυτές οι μέθοδοι περιλαμβάνουν το πρόσθετο βήμα της χωρικής μετατροπής, η οποία μετατρέπει μια εικόνα ή ένα μέρος αυτής από μια χωρική απεικόνιση σε μια απεικόνιση ψευδο-συχνότητας. Από τους πολλούς διαφορετικούς διαθέσιμους μετασχηματισμούς, μόνο δύο θα αναλύσουμε επειδή είναι οι επικρατέστεροι για τη συμπίεση ψηφιακών εικόνων και βίντεο.

2.1.10.1. Ο Διακριτός Συνημιτονικός Μετασχηματισμός (Discrete Cosine Transform, DCT)

Η διακριτός συνημιτονικός μετασχηματισμός (DCT) είναι ένα παράγωγο του διακριτού μετασχηματισμού Fourier (DFT), που χρησιμοποιείται ευρέως για τη χρονική ανάλυση πολλών διαφορετικών τύπων σημάτων ποικίλων χρόνων, σε πολλές διαστάσεις. Έχοντας, ως είσοδο, μία δέσμη των τιμών έντασης μιας εικόνας, ο DCT τη μετατρέπει σε μια ισοδύναμη ψευδο-χωρική απεικόνιση συχνότητας των ίδιων διαστάσεων. Αυτό βοηθά στη συμπίεση των εικόνων επειδή η ανταπόκριση του ανθρώπινου οπτικού συστήματος εξαρτάται πολύ από τη χωρική συχνότητα και επομένως μπορούμε να αποβάλουμε τα περιττά υψηλά χωρικά τμήματα συχνότητας, διατηρώντας μια υψηλής ποιότητας εικόνα. Υπάρχουν διάφορες μέθοδοι συμπίεσης που χρησιμοποιούν το DCT ως βάση του τρόπου συμπίεσης. Μερικά ξεχωριστά παραδείγματα περιλαμβάνουν τα JPEG, MPEG και H261/3. Το DCT, όπως χρησιμοποιείται στην εικόνα που εναλλάσσεται και στους τηλεοπτικούς τρόπους συμπίεσης, εκτελείται σε έναν 8x8 πίνακα τιμών, όπου κάθε τιμή μπορεί να είναι ένδειξη φωτεινότητας ή αξία χρώματος. Για έναν 8x8 πίνακα τιμών, ο τύπος που χρησιμοποιείται για να υπολογίσει το DCT παρουσιάζεται στην πρώτη κατα σειρά εξίσωση, με την αντίστροφη μετατροπή να παρουσιάζεται στην δεύτερη εξίσωση. Η φωτεινότητα και η αξία του χρώματος υποβάλλονται σε επεξεργασία χρησιμοποιώντας τις ίδιες εξισώσεις.

$$F[i, j] = C(j) C(i) \sum_{y=0}^7 \sum_{x=0}^7 \left(f[x, y] \cos \left(\pi j \frac{(2y+1)}{16} \right) \cos \left(\pi i \frac{(2x+1)}{16} \right) \right)$$

$$f[x, y] = \sum_{i=0}^7 \sum_{j=0}^7 \left(C(j) C(i) F[i, j] \cos \left(\pi j \frac{(2y+1)}{16} \right) \cos \left(\pi i \frac{(2x+1)}{16} \right) \right)$$

$$\text{where } C(x) = \begin{cases} \frac{1}{\sqrt{2}} & x = 0 \\ \frac{1}{2} & x = 1 \dots 7 \end{cases}$$

Η ιδιότητα του DCT που τον καθιστά εφαρμόσιμο στη συμπίεση είναι ότι παράγει

| | | | | | | | |
|----|----|----|----|-----|-----|-----|-----|
| 16 | 11 | 10 | 16 | 24 | 40 | 51 | 61 |
| 12 | 12 | 14 | 19 | 26 | 58 | 60 | 55 |
| 14 | 13 | 16 | 24 | 40 | 57 | 69 | 56 |
| 14 | 17 | 22 | 29 | 51 | 87 | 80 | 62 |
| 18 | 22 | 37 | 56 | 68 | 109 | 103 | 77 |
| 24 | 35 | 55 | 64 | 81 | 104 | 113 | 92 |
| 49 | 64 | 78 | 87 | 103 | 121 | 120 | 101 |
| 72 | 92 | 95 | 98 | 112 | 100 | 103 | 99 |

Σχήμα - Οπτικά ζυγισμένοι συντελεστές για το luna

uncorrelated συντελεστές. Στις εικόνες συνεχούς τόνου, δηλαδή εικόνες από την πραγματική ζωή, εικονοστοιχεία το ένα κοντά στο άλλο θα έχουν συχνά προβλέψιμες διαφορές μεταξύ των τιμών τους. Ο DCT μετασχηματίζει αυτό σε μια σειρά uncorrelated συντελεστών που μπορούν να συμπειστούν αποτελεσματικότερα, είτε με απωλεστικές, ή με μη απωλεστικές μεθόδους. Υπάρχουν πολλές απλές μέθοδοι που μπορούν να χρησιμοποιηθούν για να συμπίεσουν τους DCT συντελεστές με ένα ελέγξιμο ποσοστό απώλειας πληροφορίας. Για παράδειγμα, θέτοντας ως μηδενικό το

ποσοστό των συντελεστών με το χαμηλότερο πλάτος θα κρατήσει τελικά μόνο λίγους συντελεστές με πλάτη διάφορα του μηδενός και επομένως θα πραγματοποιήσει καλή συμπίεση. Εντούτοις, η πιο κοινή μέθοδος που χρησιμοποιείται για να συμπιεστούν οι DCT συντελεστές είναι η κβαντοποίηση. Η κβαντοποίηση μειώνει τον αριθμό των bit που απαιτούνται για την απεικόνιση κάθε ενός συντελεστή. Τυπικές εικόνες και βίντεο χρειάζονται 8 bit για να απεικονίσουν κάθε τμήμα luma ή χρώματος ενός εικονοστοιχείου, το οποίο σημαίνει ότι 8 bit επίσης χρησιμοποιούνται για την απεικόνιση των DCT συντελεστών. Μπορούμε να μειώσουμε τον αριθμό των bit που χρησιμοποιούνται για κάθε συντελεστή χρησιμοποιώντας μια απλή τεχνική όπως η διαίρεση κάθε ενός συντελεστή με μια δύναμη του δύο. Εντούτοις, μια σειρά οπτικά “ζυγισμένων” (visually weighted) συντελεστών έχει αναπτυχθεί για τα πρότυπα συμπίεσης JPEG που “ζυγίζουν” κάθε έναν συντελεστή σχετικά με την επίδρασή του στην ανθρώπινη αντίληψη για την εικόνα. Αυτό σημαίνει ότι ορισμένοι συντελεστές είναι λιγότερο σημαντικοί για τη γενική εικόνα και μπορούν να κβαντοποιηθούν περαιτέρω προκειμένου να επιτευχθεί μεγαλύτερη συμπίεση με τον ελάχιστο αντίκτυπο στην ποιότητα. Με την κβαντοποίηση των DCT συντελεστών, μειώνουμε αποτελεσματικά την αναλογία του σήματος ως προς το θόρυβο και επομένως δεν εισάγουμε τα λάθη στην αναδημιουργημένη εικόνα.

Το πιο σημαντικό αυτών των λαθών είναι τα αποκαλούμενα αντικείμενα δεσμίδας (blocking artifacts), τα οποία, σε κάποιες περιπτώσεις, έχουν αντίκτυπο στις 8x8 δεσμίδες των εικονοστοιχείων που έχουν ευδιάκριτα όρια (borders). Οι συντελεστές χρώματος ακολουθούν ένα παρόμοιο πρότυπο, αλλά λόγω της σχετικά φτωχής οξύτητας της ανθρώπινης όρασης, οι συντελεστές υποτονίζονται (de-emphasized) περαιτέρω. Όπως φαίνεται από αυτούς τους πίνακες, πολύ μεγαλύτερη έμφαση δίνεται στους συντελεστές πιο κοντά στη θέση [0,0]. Στην ουσία, η κβαντοποίηση έχει την επίδραση της εκτέλεσης ενός χαμηλοπερατού φίλτρου στη χωρική συχνότητα, με τη ζώνη περασμάτων να μειώνεται καθώς το επίπεδο κβαντοποίησης αυξάνεται. Στο JPEG πρότυπο, η αξία ενός χρήστη κυμαίνεται από 5 έως 95 διευκρινίζοντας την κβαντοποίηση που θα χρησιμοποιηθεί. Από την προοπτική ενός χρήστη, αυτή η αξία εμφανίζεται όπως ένας ποιοτικός παράγοντας που εκφράζεται συνήθως ως ποσοστό, δηλαδή το 100% δείχνει ότι μια μη απωλεστική συμπίεση θα έσβηνε πιθανώς όλες τις πληροφορίες από την εικόνα. Εντούτοις, αυτή η αξία δεν είναι ένα ποσοστό και συνδυάζεται με τους οπτικά ζυγισμένους συντελεστές για να λάβει τους πραγματικούς συντελεστές κβαντοποίησης για την κωδικοποίηση και την αποκωδικοποίηση της εικόνας.

Η κβαντοποίηση των συντελεστών οδηγεί στον μηδενισμό πολλών από αυτούς. Αυτό επιτρέπει μεγαλύτερα ποσοστά συμπίεσης. Το JPEG πρότυπο επανακωδικοποιεί έπειτα τους συντελεστές αυτούς σε συνδυασμό με τα Run-Length κωδικοποιημένα σημεία, τα οποία συμπίεζονται στη συνέχεια χρησιμοποιώντας είτε Huffman είτε αριθμητική κωδικοποίηση. Δεδομένου ότι η χρήση του DCT για τη συμπίεση των εικόνων και του βίντεο είναι πλέον διαδεδομένη, διάφοροι γρήγοροι αλγόριθμοι έχουν αναπτυχθεί για τον υπολογισμό του ευθύ καθώς και του αντίστροφου DCT μετασχηματισμού.

Οι γρηγορότεροι αλγόριθμοι, όπως αυτός που περιγράφεται από τους Arai, Agui, και Nakajima, κάνουν χρήση της fixed point αριθμητικής. Η αριθμητική σταθερού σημείου μπορεί να εφαρμοστεί χρησιμοποιώντας ακέραιους αριθμούς οι οποίοι βοηθούν ώστε να εκτελούνται πολύ γρηγορότερα οι υπολογισμοί από τους ισοδύναμους υπολογισμούς σημείου. Το μόνο μειονέκτημα με τη χρήση της αριθμητικής σημείου είναι ότι εισάγεται ένας βαθμός λάθους στον υπολογισμό. Για τις υψηλά κβαντοποιημένες εικόνες, αυτό είναι μόνο δευτερεύουσας σημασίας,

επειδή το λάθος που εισάγεται από την κβαντοποίηση θα είναι σημαντικά υψηλότερο από αυτό που εισάγεται από την αριθμητική σημείου. Συνολικά, οι υπολογιστικές απαιτήσεις για τον DCT μετασχηματισμό είναι αρκετά μικρές. Δεδομένου ότι ο υπολογισμός περιορίζεται σε μία 8x8 δέσμη εικονοστοιχείων ή συντελεστών, αυτό ανακουφίζει το γεγονός ότι το DCT είναι μια χρονοβόρα διαδικασία, δηλαδή ο χρόνος που λαμβάνεται για να υπολογιστεί ο DCT είναι ανάλογος προς το μέγεθος της δεσμίδας στη δύναμη του τέσσερα (4). Ως εκ τούτου, με την τήρηση του μεγέθους των δεσμίδων σε χαμηλό επίπεδο, το κόστος του υπολογισμού μειώνεται. Φυσικά, καθώς το μέγεθος της εικόνας αυξάνεται, το υπολογιστικό κόστος θα αυξηθεί στη συνέχεια. Οι προκαταρκτικές δοκιμές έδειξαν ότι οι εικόνες πλήρους χρώματος θα μπορούσαν να συμπιεστούν σε 25 εικόνες ανά δευτερόλεπτο σε μία ανάλυση 384x288 εικονοστοιχεία σχετικά με ένα Pentium III PC 600Mhz..

2.1.10.2. Ο Κυματικός Μετασχηματισμός (Wavelet Transform)

Ο διακριτός κυματικός μετασχηματισμός είναι μια ορθογώνια μετατροπή που εφαρμόζεται σε ένα σύνολο δειγματοληπτησμένων δεδομένων και είναι λειτουργικά παρόμοια με τον διακριτό μετασχηματισμό κατά Fourier. Η αρχική διαφορά μεταξύ του Fourier και του κυματικού μετασχηματισμού είναι ότι ο μετασχηματισμός Fourier χρησιμοποιεί ημιτονοειδείς συναρτήσεις ως βασικές συναρτήσεις ενώ μία βασική συνάρτηση για τον κυματικό μετασχηματισμό ορίζεται από τη διαφορική εξίσωση, που φαίνεται ακολούθως :

$$\sigma(x) = \sum_{k=0}^M c_k \sigma(2x - k)$$

Στην παραπάνω εξίσωση προσδιορίζεται ο αριθμός των συντελεστών, που είναι διάφοροι του μηδενός και αναφέρεται συνήθως ως τάξη κύματος (order of wavelet). Υπάρχει ένας συγκεκριμένος αριθμός βασικών συναρτήσεων για τα κύματα εντούτοις, οι περισσότερες από αυτές είναι περιορισμένης χρήσης. Υπάρχουν πολλές διαφορετικές οικογένειες κυμάτων, που η κάθε μια παίρνει το όνομά της από τον άνθρωπο που την ανακάλυψε, παραδείγματος χάριν, Daubechies, Haar, και Antonini. Ο διακριτός κυματικός μετασχηματισμός είναι ανάλογος με τον ταχύ μετασχηματισμό κατά Fourier δεδομένου ότι είναι μια ταχεία γραμμική λειτουργία που λειτουργεί σε ένα διάνυμα δεδομένων του οποίου το μήκος είναι μια δύναμη του δύο (2). Ο αλγόριθμος που χρησιμοποιείται για αυτόν τον υπολογισμό είναι συνήθως γνωστός ως αλγόριθμος πυραμίδων. Μια πιο νέα μέθοδος για τον ευθύ και τον αντίστροφο διακριτό κυματικό μετασχηματισμό είναι γνωστή ως σχέδιο ολίσθησης [13 ..14 ..15 ..17 ..20 ..28 ..70 ..69 ..68 ..71 ..75 ..76]. Αυτή η μέθοδος είναι πολύ απλούστερη και έχει τη δυνατότητα να είναι πολύ γρηγορότερη εφόσον μπορεί να χρησιμοποιηθεί με ακέραιους αριθμούς. Ο κυματικός μετασχηματισμός έχει την ιδιότητα να εκτελείται πολύ αποτελεσματικότερα από τον διακριτό συνημιτονικό μετασχηματισμό στη συμπίεση εικόνων. Ο κύριος λόγος για αυτό είναι ότι ο κυματικός μετασχηματισμός εκτελείται πέρα από ολόκληρη την εικόνα, που σημαίνει ότι υπάρχουν όρια πέρα από τις 8x8 δεσμίδες και επομένως, δε δημιουργούνται αντικείμενα φραγής (artifacts). Το μέγιστο μειονέκτημα του κυματικού μετασχηματισμού είναι ότι είναι υπολογιστικά ακριβότερος σε σχέση με τον DCT μετασχηματισμό.

2.2. Ελάττωση δεδομένων (data reduction)

Η ελάττωση δεδομένων αφαιρεί (“πετάει”) τα πρόσθετα στοιχεία που μπορούν να θεωρηθούν περιττά ή μη σημαντικά. Όλες οι τεχνικές ελάττωσης δεδομένων περιέρχονται σε δύο διαφορετικούς τύπους, την κβαντοποίηση (quantization) και την υποδειγματοληψία (subsampling). Με τη μέθοδο της κβαντοποίησης μειώνεται ο αριθμός των bit ανά δείγμα, μειώνοντας κατά συνέπεια το σήμα στην αναλογία θορύβου (noise ratio). Η μέθοδος της υποδειγματοληψίας μειώνει τον αριθμό δειγμάτων στο πεδίο του χρόνου ή του χώρου και αυτό οδηγεί στην απώλεια πληροφορίας για τις υψηλές συχνότητες και μπορεί ενδεχομένως να προκαλέσει aliasing.

| Colour Space | Bits per Pixel | Stereo Bandwidth (b/s) |
|---|-----------------|------------------------|
| Black & White | 1 | 20,736,000 |
| Monochrome | 8 | 165,888,000 |
| Y _{C_B} C _R 4:2:0 | 12 ¹ | 248,832,000 |
| Y _{C_B} C _R 4:2:2 | 16 ² | 331,776,000 |
| Pseudo Colour | 8 | 165,888,000 |
| Hi Colour | 16 | 331,776,000 |
| True Colour | 24 | 497,664,000 |

Σχήμα - Το εύρος ζώνης που απαιτείται από ένα βίντεο PAL με πολλαπλούς τύπους χρώματος

2.2.1. Κβαντοποίηση

Η μόνη μέθοδος κβαντοποίησης ικανή να χρησιμοποιηθεί στο βίντεο με στερεοφωνικό ήχο είναι να μειωθεί ο αριθμός των bit που απεικονίζουν το κάθε εικονοστοιχείο. Είναι γνωστό ότι τα αρχικά δεδομένα καταγράφονται (captured) χρησιμοποιώντας το DV φορμάτ που έχει 12 bit ανά εικονοστοιχείο. Εντούτοις, αυτό το σχήμα συμπίεσης μειώνει αυτήν την αναλογία σε περίπου 3 bit ανά εικονοστοιχείο. Η μόνη απεικόνιση χρώματος με χαμηλότερη αναλογία από αυτήν είναι το ασπρόμαυρο που χρησιμοποιεί μόνο 1 bit ανά εικονοστοιχείο. Μερικά πρόωρα πειράματα έγιναν, όπου μια στερεοφωνική εικόνα υποβίβασε σταδιακά την ποιότητά της, σε ένα 1 bit ανά εικονοστοιχείο, μαύρη και άσπρη εικόνα. Επίσης, διαπιστώθηκε ότι σε όλα τα επίπεδα, η στερεοσκοπική τήξη (stereoscopic fusion) θα μπορούσε να προκύψει. Παρόλα αυτά, η ζωνοποίηση (banding) αναγκάζει τις άκρες (edges) να φαίνονται, γεγονός το οποίο δε δείχνει όμορφες τις απολήξεις, και το dithering, που είναι μέχρι ενός ορισμένου βαθμού απαραίτητο για τις ασπρόμαυρες εικόνες, σβήνει ουσιαστικά τις μικρές στερεοφωνικές πληροφορίες έτσι ώστε τα απόμακρα αντικείμενα να μη φαίνονται στερεοφωνικά. Αυτό που απέδειξαν τέτοιου είδους πειράματα ήταν ότι η μείωση της εικόνας σε μονοχρωματική δεν είχε πολλές επιπτώσεις στο στερεοφωνικό εφέ, ενώ το ψευδο-χρώμα και οι ασπρόμαυρες απεικονίσεις το είχαν διαφοροποιήσει αντιληπτά. Σαφώς, η μείωση του αριθμού των bit ανά εικονοστοιχείο δεν πρόκειται από μόνη της να μας οδηγήσει σε ένα χαμηλό εύρος ζώνης. Η μείωση κβαντοποίησης στο ακατέργαστο στερεοφωνικό βίντεο οδηγεί σε άμεσες και δαπανηρές απώλειες στην ποιότητα της εικόνας, με τη μέγιστη μείωση του μεγέθους των δεδομένων να είναι 25% του αρχικού μεγέθους ώστε να διατηρηθεί μια λογικά καλής ποιότητας εικόνα.

2.2.2. Υποδειγματοληψία (Sub-sampling)

Η υποδειγματοληψία, όπως δηλώθηκε νωρίτερα, μπορεί να γίνει στο χώρο ή στο χρόνο. Στο στερεοφωνικό βίντεο η χωρική υποδειγματοληψία σημαίνει ότι ο αριθμός δειγμάτων ανά σαρώσιμη γραμμή (scanned line) και ο αριθμός σαρώσιμων γραμμών μειώνονται. Έτσι εκτιμήθηκε ότι η χρονική υποδειγματοληψία οδηγεί σε μείωση του αριθμού των πλαισίων ανά δευτερόλεπτο. Επίσης, προκειμένου να μεταβιβαστεί η κίνηση, απαιτείται ένα ελάχιστος αριθμός των 10 έως 15 πλαισίων ανά δευτερόλεπτο, περιορίζοντας έτσι το ποσό των δεδομένων που υποδειγματοληπτήθηκαν σε περίπου 50% .

Το μέγεθος της χωρικής υποδειγματοληψίας που μπορεί να γίνει περιορίζεται μέχρι ενός ορισμένου βαθμού από το μέγεθος της απεικόνισης και την απόσταση εξέτασης. Εάν το μέγεθος της απεικόνισης είναι μικρό, κατόπιν η ευκρίνεια που απαιτείται είναι μικρότερη και επομένως ένας υψηλότερος βαθμός υποδειγματοληψίας μπορεί να γίνει ανεκτός. Ομοίως, καθώς ο θεατής κινείται μακριά από την οθόνη, ένας υψηλότερος βαθμός υποδειγματοληψίας μπορεί να γίνει ανεκτός.

Μια άλλη μορφή χωρικής υποδειγματοληψίας είναι να υποδειγματοληπτηθούν το αριστερό και το δεξί κανάλι του στερεοφωνικού βίντεο με σεβασμό το ένα στο άλλο. Δηλαδή εάν η ανάλυση του αριστερού καναλιού είναι 720x576 εικονοστοιχεία, κατόπιν το δεξί κανάλι μπορεί να υποδειγματοληπτηθεί σε μία ανάλυση 360x288 εικονοστοιχεία ώστε να μειώσει το μέγεθος των δεδομένων. Μερικά πειράματα πραγματοποιήθηκαν χρησιμοποιώντας αυτήν την τεχνική, αλλά έμειναν σε αρχικό στάδιο. Διαπιστώθηκε ότι τα αποτελέσματα δεν ήταν ικανοποιητικά, δεδομένου ότι η χαμηλή ανάλυση της δεξιάς εικόνας ήταν αρκετά εμφανής. Ενώ αυτή η τεχνική δεν χρησιμοποιήθηκε σε επόμενες έρευνες, πρέπει να σημειωθεί ότι η μέθοδος που χρησιμοποιήθηκε για την απεικόνιση της υποδειγματοληπτημένης εικόνας σε μια υψηλότερη ανάλυση ήταν η zero order hold method. Αυτή είναι η απλούστερη μέθοδος παρεμβολής και επομένως παράγει μέτρια αποτελέσματα ενώ μπορεί να παραχθεί υψηλότερης ποιότητας στερεοφωνικό εφέ με τη χρήση ενός πιά σύνθετου σχεδίου παρεμβολής. Παρά ταύτα, το σημαντικότερο γεγονός που σημειώνεται είναι ότι είναι δυνατό να εφαρμοστεί σε stereoscopically fuse left and right stereo εικόνες που έχουν διαφορετικές αναλύσεις.

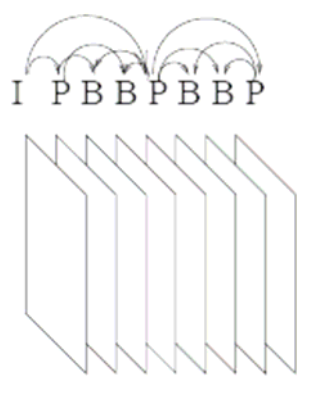
Όσο η χωρική και χρονική υποδειγματοληψία του στερεοφωνικού βίντεο μπορούν να μειώσουν το μέγεθος των δεδομένων προς μετάδοση με την διαφοροποίηση ενός μικρού παράγοντα, ένα άλλο πλεονέκτημα είναι ότι με την υποδειγματοληψία, απαιτείται χαμηλότερο υπολογιστικό κόστος σε ένα μεταγενέστερο στάδιο για τη συμπίεση του στερεοφωνικού βίντεο. Είναι σαφές εντούτοις ότι η μείωση των δεδομένων από μόνη της δεν θα είναι αποτελεσματική τόσο ώστε να επιτρέψει τη μετάδοση του στερεοφωνικού βίντεο μέσω του διαδικτύου με ταχύτητες χαμηλότερες από τα 6Mb/s χωρίς σοβαρά να αλλοιωθεί η εικόνα και η στέρεο ποιότητα.

2.3. Δια – πλαίσια (intra – frames) σε αντιπαράθεση με τα ενδο – πλαίσια (inter – frames)

Κατά τη συμπίεση των πλαισίων, πρέπει να γίνει η επιλογή εάν (μέρη ή) τα πλαίσια θα συμπεστούν ως ανεξάρτητα δια - πλαίσια ή σαν εξαρτώμενα ενδο – πλαίσια. Κατα τη δια – συμπίεση (intra – coding) πλαισίων, τα macro – blocks μεταβιβάζονται

μερικές φορές ως δια – φραγμοί (intra – blocks), επειδή ένας καλά συνδυασμένος φραγμός με το προηγούμενο πλαίσιο δεν μπορεί να βρεθεί. Ένας άλλος λόγος για τη χρήση των δια – πλαισίων (intra – frames) αντί των ενδο - πλαισίων (inter – frames) είναι η απώλεια δεδομένων κατά τη διάρκεια της μετάδοσης: εάν υπάρχουν λάθη στη συμπιεσμένη ροή, αποτρέπει όχι μόνο την αναδημιουργία από το τρέχον πλαίσιο, αλλά και από τα επόμενα ενδο – πλαίσια (τα οποία στηρίζονται σε αυτό το πλαίσιο) έως ότου διαβιβαστεί πάλι ένα δια – πλαίσιο (intra – frame). Η εισαγωγή περισσότερων δια – πλαισίων, με το κόστος μιας χαμηλότερης αναλογίας συμπίεσης, μειώνει το χρόνο κατά τον οποίο μια σειρά πλαισίων γίνεται ακατάλληλη προς χρήση.

Ο αριθμός των δια – πλαισίων (intra – frames), που πρέπει να παρεμβληθούν εξαρτάται από την ποιότητα της σύνδεσης (πόσα δεδομένα χάνονται σε μια ορισμένη χρονική περίοδο), και το χρόνο που θεωρείται ως ο λιγότερο “οδυνηρός” για τη “χαλάρωση” της σύνδεσης. Μια άλλη προσέγγιση, που θα μπορούσε να εξεταστεί είναι η αρνητική επιβεβαίωση της οντότητας λήψης, που αντιμετωπίζει τα λάθη. Εντούτοις, αυτό απαιτεί περισσότερες γενικά κεφαλίδες (overheads) για το πρωτόκολλο και βάζει περισσότερους περιορισμούς στη λανθάνουσα κατάσταση του δικτύου σε αντίθεση με ένα πρωτόκολλο που μεταφέρει χωρίς επιβεβαιώσεις.



Σχήμα – Εξαρτήσεις πλαισίων σε μια MPEG ροή bit

Λανθάνουσα κατάσταση του δικτύου

Θα πρέπει, επίσης, να εξετάσουμε έναν από τους περιοριστικούς παράγοντες για την επιλογή του δικτύου : τη λανθάνουσα κατάσταση του δικτύου. Ο χρόνος της λανθάνουσας κατάστασης ενός δικτύου σε μια τηλεδιάσκεψη καθορίζει τη χρησιμότητα του δικτύου για διαδραστική επικοινωνία. Εάν η λανθάνουσα κατάσταση είναι υψηλή (μεγαλύτερη από, για παράδειγμα, 1 δευτερόλεπτο), οι οπτικές αντιδράσεις μπορεί να μην γίνουν κατανοητές ή να παρερμηνευθούν. Από την άλλη πλευρά, μια περιττά χαμηλή λανθάνουσα κατάσταση καθιστά ένα τηλεοπτικό σύστημα διάσκεψης πάρα πολύ ακριβό για να εφαρμοστεί, καθώς απαιτεί τα καλύτερα δίκτυα και τους γρηγορότερους υπολογιστές. Για να δώσουμε μια επισκόπηση, όπου εμφανίζεται η λανθάνουσα κατάσταση ενός δικτύου, η συνολική λανθάνουσα κατάσταση θα διαιρεθεί σε διάφορα μικρότερα τμήματα λανθάνουσας κατάστασης:

1. Λανθάνουσα κατάσταση “αρπαγής” βίντεο (video grab latency): Αυτή είναι η λανθάνουσα κατάσταση μεταξύ του πραγματικού χρόνου κατά τον οποίο συμβαίνει κάτι πριν από την κάμερα, και του χρόνου, που τα δεδομένα αυτού του πλαισίου παραλαμβάνονται από τον υπολογιστή. Επειδή τα περισσότερα

ενδιάμεσα βήματα σ' αυτό το σημείο γίνονται από το υλικό (την κάμερα ή την τηλεοπτική grabber συσκευή), ο μέγιστος χρόνος είναι περίπου ο χρόνος μεταξύ δύο "αρπαγών" πλαισίων (frame grabs).

2. Λανθάνουσα κατάσταση συμπίεσης (compress latency): Αυτός είναι ο χρόνος που απαιτείται για τη συμπίεση ενός πλαισίου. Αυτή η λανθάνουσα κατάσταση εξαρτάται κυρίως από τον κωδικοποιητή που χρησιμοποιείται καθώς και από την απόδοση του συστήματος πάνω στο οποίο αυτός "τρέχει". Εάν ο κωδικοποιητής χρησιμοποιεί όλον τον υπολογιστικό χρόνο για την κωδικοποίηση (και συμπιέζει στο μέγιστο ρυθμό πλαισίων, που μπορεί να χειριστεί), η λανθάνουσα κατάσταση είναι ο χρόνος που απαιτείται για να συμπιεστεί ένα πλαίσιο. Εάν ο κωδικοποιητής λειτουργεί με ένα ρυθμό πλαισίων 10 fps (10 frames per second), η λανθάνουσα κατάσταση συμπίεσης είναι $1 / (10 \text{ fps}) = 100 \text{ ms}$.
3. Λανθάνουσα κατάσταση δικτύων (network latency): Αυτός είναι ο χρόνος που απαιτείται για να φθάσει το στοιχείο που διαβιβάζεται από τη μία άκρη του δικτύου στην άλλη. Αυτή η λανθάνουσα κατάσταση εξαρτάται βασικά από το είδος του δικτύου, που χρησιμοποιείται. Για παράδειγμα, για κάποια ATM εφαρμογή αυτή η λανθάνουσα κατάσταση είναι αμελητέα έναντι της λανθάνουσας κατάστασης συμπίεσεων. Άλλα δίκτυα, όπως το Διαδίκτυο (TCP/IP) μπορεί να έχουν λανθάνουσες καταστάσεις, που εξαρτώνται κυρίως από την κυκλοφορία του δικτύου. Μετρήσεις έχουν δείξει ότι η λανθάνουσα κατάσταση μεταξύ του δικτύου του Enschede (NL) και του δικτύου του Μίτσιγκαν (ΗΠΑ) θα μπορούσε να είναι μέχρι 10 δευτερόλεπτα.
4. Λανθάνουσα κατάσταση αποσυμπίεσης και απεικόνισης (decompress and display latency): Αν και οι δύο λανθάνουσες καταστάσεις, της αποσυμπίεσης και της απεικόνισης, θα μπορούσαν να ονομαστούν ανεξάρτητα, συνδυάζονται για ευκολία. Η λανθάνουσα κατάσταση αποσυμπίεσης απαιτεί λιγότερο χρόνο από τη λανθάνουσα κατάσταση συμπίεσης (επειδή η εκτίμηση κίνησης δεν πρέπει να γίνει κατά την αποκωδικοποίηση) και η λανθάνουσα κατάσταση απεικόνισης είναι επίσης μικρή.

Οι δύο κύριοι τελεστές στη συνολική λανθάνουσα κατάσταση είναι η λανθάνουσα κατάσταση συμπίεσης και η λανθάνουσα κατάσταση του δικτύου. Εάν τα πλαίσια κωδικοποιούνται με έναν υψηλό ρυθμό ($>10 \text{ fps}$), η λανθάνουσα κατάσταση συμπίεσης δεν θα είναι τόσο υψηλή ώστε να ενοχλήσει τη διαδραστικότητα του βίντεο. Η λανθάνουσα κατάσταση δικτύων δύσκολα θα επηρέαζε, αλλά σε μερικά δίκτυα (ATM) αυτό μπορεί να διευκρινιστεί στην κλήση της ποιότητας υπηρεσιών κατά τη διάρκεια της ίδρυσης της επικοινωνίας. Στην περίπτωση του Διαδικτύου, άλλα πρωτόκολλα πρέπει να χρησιμοποιηθούν, που χειρίζονται την απώλεια και τη αλλοίωση των πακέτων για βίντεο δεδομένα καλύτερα από το TCP/IP.

Μια τελική λύση για να μειωθεί η λανθάνουσα κατάσταση είναι να συνδυαστεί η συμπίεση και η μετάδοση του πλαισίου την ίδια στιγμή. Σε αυτήν την περίπτωση, δεν χρειάζεται να περιμένει κανείς έως ότου συμπιεστεί πλήρως ένα πλαίσιο. Στην πραγματικότητα, αυτό γίνεται από τον κωδικοποιητή H.263, ο οποίος παράγει τις ψηφιολέξεις αμέσως μόλις του γνωστοποιηθεί το υλικό.

Σύνοψη

Για να μεταφερθεί το βίντεο πάνω από ένα δίκτυο, η στρατηγική συμπίεσης καθορίζεται από τις ικανότητες του δικτύου και της υπολογιστικής ισχύος, που είναι διαθέσιμες. Όσο περισσότερο εύρος ζώνης είναι διαθέσιμο τόσο περισσότερη απώλεια εμφανίζεται, τα δια - πλαίσια σε αντιδιαστολή με τα ενδο - πλαίσια πρέπει να μεταδοθούν. Το πρωτόκολλο, επίσης, καθορίζει την επιλογή μεταξύ δια - πλαισίων ή ενδο - πλαισίων. Τέλος, το μέγεθος της λανθάνουσας κατάστασης εξαρτάται από το εάν η τηλεοπτική διάσκεψη σκεψη είναι ένας χρήσιμος τρόπος διαδραστικότητας. Για τα γρήγορα δίκτυα, η λανθάνουσα κατάσταση καθορίζεται κυρίως από τη λανθάνουσα κατάσταση συμπίεσης, η οποία είναι στη χειρότερη περίπτωση ο χρόνος για να κωδικοποιηθεί ένα πλαίσιο.

2.4. Υπολογιστική ισχύς

Εφόσον οι τρόποι συμπίεσης που συζητήθηκαν πρέπει να λειτουργήσουν σε πραγματικό χρόνο και δεν μπορούν να χρησιμοποιήσουν οποιοσδήποτε εξειδικευμένες συσκευές (hardware), η υπολογιστική ταχύτητα του αλγορίθμου είναι ένας σημαντικός παράγοντας. Για την απόφαση αυτή, υπάρχουν διάφοροι παράγοντες που εξετάζονται, πολλοί από τους οποίους δεν μπορούν να συνυπάρξουν. Οι κύριοι στόχοι της συμπίεσης είναι η ελαχιστοποίηση του εύρους ζώνης που απαιτείται για τις εκπομπές των δεδομένων, η μεγιστοποίηση της εικόνας και της στέρεο ποιότητας, η ελαχιστοποίηση οποιασδήποτε λανθάνουσας κατάστασης προκύψει και η μεγιστοποίηση του αριθμού των πλαισίων ανά δευτερόλεπτο. Όπως φαίνεται, αρκετές από αυτές τις απαιτήσεις παρεμποδίζουν η μια την άλλη, παραδείγματος χάριν, αυξάνοντας την ποιότητα μιας εικόνας γενικά, θα αυξηθεί το εύρος ζώνης που απαιτείται για την εκπομπή αυτού του σήματος. Σε αυτό το τμήμα εστιάζουμε στο υπολογιστικό κόστος, και στον αριθμό πλαισίων ανά δευτερόλεπτο, καθώς και στη λανθάνουσα κατάσταση του συστήματος μετάδοσης.

2.5. Οι τεχνικές συμπίεσης σε αντιπαράθεση με την τεχνολογία δικτύων

Εισαγωγή

Η συμπίεση του βίντεο είναι το στοιχείο - κλειδί σε ένα σύστημα τηλεδιάσκεψης μεγάλης απόστασης. Αλλά η συμπίεση δεν μπορεί να λειτουργήσει κατάλληλα χωρίς αλληλεπίδραση με το περιβάλλον της, το οποίο εκτός από κάποιον χρήστη είναι: το δίκτυο.

Το δίκτυο καθορίζει ποιου είδους συμπίεση απαιτείται. Εισάγει τη λανθάνουσα κατάσταση ενός δικτύου κοντά στη λανθάνουσα κατάσταση που προκαλείται ώσπου να συμπιεστεί ένα πλαίσιο και βάζει έναν περιορισμό και στο μέσο και μέγιστο εύρος ζώνης. Για να αντιμετωπιστούν αυτές οι συνθήκες, μια μορφή διαδραστικότητας πρέπει να υπάρξει μεταξύ του κωδικοποιητή και του πρωτοκόλλου των δικτύων. Αυτή η διαδραστικότητα εξαρτάται από τον τύπο του δικτύου που χρησιμοποιείται. Το δίκτυο καθορίζει ποιο εύρος ζώνης είναι διαθέσιμο και αυτές οι πληροφορίες

χρησιμοποιούνται για τη συμπίεση του επόμενου πλαισίου. Ο κωδικοποιητής καθορίζει το χρόνο που απαιτείται για να κωδικοποιήσει ένα πλαίσιο και αυτό έχει επιπτώσεις και στο εύρος ζώνης των δικτύων αλλά και στη συμπίεση του επόμενου πλαισίου.

Ταξινόμηση των δικτύων

Η παραδοσιακή ταξινόμηση για τα δίκτυα υπολογιστών βασίζεται στην τοπολογία:
→ ένα δίκτυο που συνδέει τους υπολογιστές σε ένα γραφείο ή ένα κτίριο καλείται δίκτυο τοπικής περιοχής (LocalAreaNetwork - LAN).

→ ενώ ένα δίκτυο που συνδέει τους υπολογιστές μεταξύ των πόλεων, των χωρών ή των ηπείρων καλείται δίκτυο ευρείας περιοχής (WideAreaNetwork - WAN).

Τα LANs είναι κυρίως γρήγορες συνδέσεις, που χωρίζονται φυσικά από τον εξωτερικό κόσμο, ενώ τα WANs είναι πιά αργά δίκτυα, τα οποία είναι προσβάσιμα παγκοσμίως. Αλλά καθώς όλο και περισσότερο τα τερματικά συνδέονται σε παγκόσμια κλίμακα, αυτή η ταξινόμηση μικραίνει, και με τις γρήγορες συνδέσεις gigabit ευρείας περιοχής που εμφανίστηκαν, αυτή η ταξινόμηση δεν θα είναι πλέον κατάλληλη.

Για να ταξινομηθεί ένα δίκτυο με βάση την κατάσταση/λειτουργία του, μπορεί να χρησιμοποιηθεί ως μέτρο η απόδοση του δικτύου. Η απόδοση καθορίζει τη μέγιστη ικανότητα του δικτύου. Εντούτοις, αυτή η ιδιότητα δεν καθορίζει πλήρως την πραγματική απόδοση, δεδομένου ότι η λανθάνουσα κατάσταση του δικτύου διευκρινίζει το χρόνο, που απαιτείται για τη μεταφορά της πληροφορίας από το ένα άκρο στο άλλο. Για παράδειγμα, οι πολυμεσικές εφαρμογές απαιτούν ένα δίκτυο, το οποίο έχει και υψηλή απόδοση, αλλά και μια χαμηλή λανθάνουσα κατάσταση για να καταστήσει τη διαδραστική (interactive communication) επικοινωνία πιθανή. Επομένως, ο συνδυασμός της απόδοσης και της λανθάνουσας κατάστασης ενός δικτύου δίνει μια καλή ένδειξη της επίδοσης ενός δικτύου. Μια άλλη ταξινόμηση που χρησιμοποιείται συχνά είναι ο διαχωρισμός μεταξύ των προσανατολισμένων προς τη σύνδεση και των χωρίς σύνδεση δικτύων. Στην πρώτη περίπτωση, πρώτα γίνεται η σύνδεση του δικτύου, και μετά διαβιβάζονται τα οποιαδήποτε δεδομένα. Στα χωρίς σύνδεση δίκτυα, καμία σύνδεση δε στήνεται, αλλά οι πληροφορίες διαβιβάζονται άμεσα στον προορισμό τους. Σε αυτά τα δίκτυα, κάθε κομμάτι πληροφορίας, που διαβιβάζεται (αποκαλούμενο πακέτο) περιέχει τη διεύθυνση του προορισμού, και επιπλέον μια διαδρομή προς αυτόν τον προορισμό πρέπει να βρεθεί για κάθε πακέτο από το δίκτυο.

Δίκτυα υπολογιστών

❖ POTS

Τα POTS είναι το παραδοσιακό αναλογικό τηλεφωνικό δίκτυο. Ένα POTS είναι ένα προσανατολισμένο ως προς τη σύνδεση δίκτυο, δεδομένου ότι η σύνδεση πρέπει πρώτα να οργανωθεί με το σχηματισμό ενός τηλεφωνικού αριθμού. Η βασική λειτουργία αυτού του δικτύου είναι μόνο η επικοινωνία φωνής καθώς το παλιομοδίτικο αυτό δίκτυο δεν προορίζεται για ψηφιακή επικοινωνία.

Εντούτοις, με τη χρήση modem οι ταχύτητες μπορούν να γίνουν περίπου 28.8 Kbps. Για τις χειρότερες ποιοτικά συνδέσεις, αυτή η ταχύτητα μειώνεται σε 14.4 Kbps ή και ακόμα λιγότερο. Η λανθάνουσα κατάσταση του δικτύου ποικίλλει από

χαμηλές τιμές για τοπικές συνδέσεις μέχρι περίπου ένα δευτερόλεπτο για τις δορυφορικές συνδέσεις μεγάλης απόστασης. Αν και η απόδοση του δικτύου δεν είναι καλή, η ευρεία, και σφαιρική εφαρμογή του, το καθιστά διαθέσιμο για όλους ανά τον κόσμο.

❖ ISDN

Η ISDN σχεδιάστηκε για τις τηλεπικοινωνίες των υπολογιστών, των fax και των τηλεφώνων. ISDN με δύο τύπους σύνδεσης: 1. Η διεπαφή βασικού ρυθμού (BasicRateInterface - BRI) προσφέρει δύο αμφίδρομα κανάλια των 64 Kbit/s και ένα κανάλι ελέγχου των 16 Kbit/s. 2. Η διεπαφή αρχικού ρυθμού (PrimaryRateInterface - PRI) η οποία προσφέρει 23 κανάλια των 64 Kbit/s και ένα κανάλι ελέγχου των 64 Kbit/s επίσης. Όπως τα POTS έτσι και η ISDN είναι προσανατολισμένη ως προς τη σύνδεση. Με βάση την απόδοση του δικτύου της ISDN, θα μπορούσαμε να την θεωρήσουμε ως μια από τις κατάλληλες υποδομές για τις εφαρμογές βίντεο διάσκεψης.

❖ Internet

Το μεγαλύτερο και ευρύτετα χρησιμοποιημένο δίκτυο για την επικοινωνία υπολογιστών είναι το Διαδίκτυο. Το Διαδίκτυο είναι πραγματικά μια συλλογή των διασυνδεδεμένων ήδη δικτύων, η οποία αναπτύσσεται στη δεκαετία του '90 από ένα ερευνητικό δίκτυο για τους επιστήμονες υπολογιστών σε ένα δίκτυο ευρέος κοινού. Στην κορυφή του χωρίς - σύνδεση IP πρωτοκόλλου του, αυτό υποστηρίζει και τις δύο αναξιόπιστες και προσανατολισμένες ως προς τη σύνδεση συνδέσεις UDP/IP όπως τις αξιόπιστες και προσανατολισμένες ως προς τη σύνδεση συνδέσεις TCP/IP. Η απόδοση και η λανθάνουσα κατάσταση του δικτύου ποικίλλουν ανάλογα με τις τοπολογίες, τη δρομολόγηση μεταξύ δύο πόλεων, καθώς και το χρόνο και την ημέρα (το φορτίο του δικτύου δεν είναι στατικό). Το αξιόπιστο TCP/IP δεν δίνει εγγυήσεις για το πότε θα φθάσει ένα πακέτο, αλλά μόνο ότι ένα πακέτο θα φθάσει και κατά σειρά με την οποία μεταδόθηκε. Αν η μέση απόδοση των δικτύων δεν βελτιώνεται σημαντικά και εισάγονται τα πρωτόκολλα που υποστηρίζουν την διαφύλαξη των πόρων, η υψηλής ποιότητας βίντεο διάσκεψη δεν είναι η καλύτερη στο διαδίκτυο.

❖ ATM

Στην επιστημονική κοινότητα, το ATM τυγχάνει μεγάλης προσοχής. Το ATM είναι βασισμένο στη μετάδοση κελιών των 53 bytes και είναι προσανατολισμένο ως προς τη σύνδεση. Το ATM μπορεί να εγγυηθεί για ορισμένο εύρος ζώνης και αυτό το καθιστά ιδανικό για έναν αριθμό εφαρμογών, που απαιτούν αυτές τις εγγυήσεις.

Στα δίκτυα ATM, οι διαφορετικές κατηγορίες υπηρεσιών είναι διαθέσιμες για τους διαφορετικούς τύπους στοιχείων: Σταθερό ρυθμός μετάδοσης (ConstantBitRate - CBR), σε πραγματικό χρόνο μεταβλητός ρυθμός μετάδοσης (real-time VariableBitRate - real-time VBR), μη-πραγματικού χρόνου μεταβλητός ρυθμός μετάδοσης (non-real-time VariableBitRate - non-real-time VBR), απροσδιόριστος ρυθμός μετάδοσης (UnspecifiedBitRate - UBR) και διαθέσιμος ρυθμός μετάδοσης (AvailableBitRate - ABR). Αυτές οι κατηγορίες υπηρεσιών επηρεάζουν και τα χαρακτηριστικά όπως το εύρος ζώνης και η λανθάνουσα κατάσταση του δικτύου, της σύνδεσης επικοινωνίας. Για τη συμπίεση του βίντεο, το σε πραγματικό-χρόνο VBR είναι το πιο κατάλληλο.

Η συμπίεση και το δίκτυο

Η μετάδοση του βίντεο εξαρτάται από την υπολογιστική δύναμη, το εύρος ζώνης και τη λανθάνουσα κατάσταση του δικτύου, από κοινού. Εάν το εύρος ζώνης του

δικτύου είναι παραπανήσιο, η συμπίεση του βίντεο δεν θα χρειαστεί καθόλου. Από την άλλη, η υψηλή δύναμη λαμβάνουσας κατάστασης του δικτύου καθιστά τη διαδραστική βίντεο διάσκεψη σχεδόν αδύνατη.

Ένας αριθμός διαφορετικών στρατηγικών συμπίεσης είναι διαθέσιμος. Ανάλογα με το μίγμα της απόδοσης του δικτύου και της υπολογιστικής δύναμης, η σωστή στρατηγική συμπίεσης είναι διαθέσιμη:

1.Καμία συμπίεση: Εάν το εύρος ζώνης του δικτύου είναι πολύ υψηλό και παραπνήσιο ή/και η υπολογιστική δύναμη είναι αραιή, το βίντεο μεταδίδεται ως ένα σύνολο ακατέργαστων δεδομένων.

2.Ενδο – συμπίεση (intra - compression): Στην ενδο - κωδικοποίηση, τα πλαίσια συμπιέζονται χωριστά. Ένα καλό παράδειγμα της ενδο - κωδικοποίησης είναι το MJPEG. Πολλά βίντεο υποστηρίζουν αυτήν την κωδικοποίηση και την αποκωδικοποίηση, η οποία μειώνει σημαντικά τη ροή του βίντεο χωρίς να απαιτεί πολλή επεξεργαστική ισχύ από την κεντρική μονάδα επεξεργασίας (CPU). Η ενδο - κωδικοποίηση είναι απωλεστική μέθοδος κωδικοποίησης και ο παράγοντας κβαντοποίησης που καθορίζει την ποιότητα και την αναλογία συμπίεσης μπορεί να ποικίλλει.

3.Διά - συμπίεση χωρίς εκτίμηση κινήσεων (inter - compression): Αυτή η στρατηγική κωδικοποίησης συγκρίνει το τρέχον πλαίσιο με το προηγούμενο πλαίσιο και τελικά συμπιέζεται η διαφορά μεταξύ αυτών των πλαισίων. Αυτός ο κωδικοποιητής μπορεί να απαιτήσει μερικές δια - συμπιεστικές κωδικοποιήσεις πλαισίων (intra - compression frame encodings), για παράδειγμα, για το πρώτο πλαίσιο ή για τα μπλόκς που έχουν αλλάξει πολύ.

4.Διά - συμπίεση με εκτίμηση κινήσεων: Η πιο δραστική στρατηγική συμπίεσης είναι η χρήση της εκτίμησης κινήσεων μέσα στη διά - συμπίεση. Τα αποτελέσματα δείχνουν καλύτερη συμπίεση για τις ακολουθίες με υψηλή κίνηση ή χαμηλούς ρυθμούς μετάδοσης πλαισίων.

Επιλογή συμπίεσης

Η εξέταση τεχνικών όπως ο συνημιτονικός μετασχηματισμός DCT, η κβαντοποίηση και η εκτίμηση κινήσεων προσφέρει μια καλή άποψη για τις διαφορετικές δυνατότητες των τεχνικών συμπίεσης. Εάν ένα QCIF φορμάτ χρησιμοποιείται ως είσοδος για ένα βίντεο, τα ακόλουθα αποτελέσματα ισχύουν για έναν κωδικοποιητή (encoder) H.263:

| Method | (Average) size of frame (bytes) | Time to encode |
|----------------------------|---------------------------------|----------------|
| No compression | 38016 | 0 |
| Hardware intra compression | 1434 | << 72 ms |
| Software intra compression | 1434 | 72 ms |
| Inter 15 fps | 260 | 98 ms |
| Inter 7.5 fps | 340 | 103 ms |

Πίνακας : Σύγκριση μεταξύ διαφορετικών μεθόδων κωδικοποίησης για ένα πλαίσιο QCIF

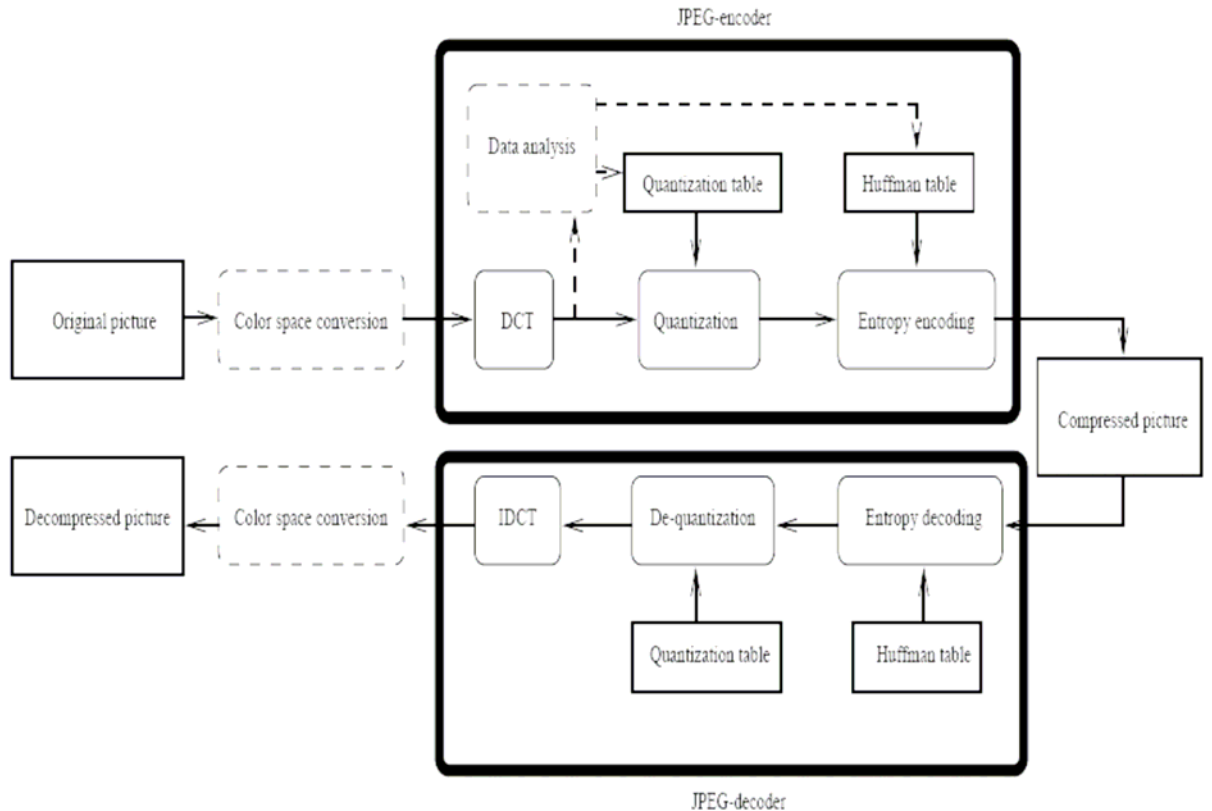
Η ποιότητα της αποκωδικοποιημένης ροής των διαφορετικών μεθόδων είναι υπό σύγκριση για τις DVC καταστάσεις. Η χρήση του ασυμπίεστου βίντεο πάνω από το δίκτυο απαιτεί μεγάλο εύρος ζώνης, το οποίο συχνά δεν είναι διαθέσιμο για μεγάλες αποστάσεις. Η χρήση της ενδο - κωδικοποίησης με τη χρήση υλικού (hardware) (όπως η εφαρμογή υλικού στο motion - JPEG) καθιστά την κωδικοποίηση εφικτή και προτιμητέα από τη δια - κωδικοποίηση λογισμικού. Η δια - συμπίεση λογισμικού περιλαμβάνει την αναδημιουργία του πλαισίου, το οποίο δεν είναι απαραίτητο, επειδή δεν υπάρχει κανένα ενδο - πλαίσιο, που να αναφέρεται στα δια - πλαίσια. Χωρίς αυτήν την αναδημιουργία, η κωδικοποίηση λογισμικού των ενδο - πλαισίων είναι στην πραγματικότητα ακόμα γρηγορότερη. Εάν θέλουμε να συμπίεσουμε τη ροή του βίντεο ακόμα περισσότερο, πρέπει να χρησιμοποιήσουμε την διά - κωδικοποίηση και δεδομένου ότι η υποστήριξη υλικού δεν είναι ακόμα διαθέσιμη, οι λύσεις λογισμικού είναι η μόνη μας επιλογή.

Οι τωρινοί κωδικοποιητές λογισμικού διαθέσιμοι για το Διαδίκτυο, όπως VIC και ο κωδικοποιητής INRIA H.261 είναι αρχικά βασισμένοι μόνο στο να δια - κωδικοποιούν και μερικές φορές συνδυάζουν και την ανίχνευση κινήσεων για να παραλείψουν τα macro - blocks που δεν έχουν αλλάξει. Οι κύριοι λόγοι για τη χρήση μόνο της δια - κωδικοποίησης είναι ότι η απώλεια πακέτων είναι υψηλή γεγονός το οποίο καθιστά τη ενδο - κωδικοποίηση λιγότερο κατάλληλη και τονίζει ότι η δια - κωδικοποίηση απαιτεί λιγότερο χρόνο στη συμπίεση.

Για τις συνδέσεις δικτύων, που έχουν χαμηλό εύρος ζώνης όπως τα POTS και η ISDN, η ενδο - κωδικοποίηση είναι πιο σημαντική και επομένως ένας γρήγορος κωδικοποιητής H.263 είναι απαραίτητος για την βίντεο διάσκεψη πάνω από τέτοια δίκτυα.

❖ JPEG

Το αρχικό υπολογιστικό κόστος για το πρότυπο συμπίεσης JPEG είναι ο ευθύς και ο αντίστροφος συνημιτονικός μετασχηματισμός (DCT). Η κβαντοποίηση που έπεται, η κωδικοποίηση εντροπίας και η χωρική μετατροπή χρώματος (colour space conversion) είναι σχετικά γρήγορες. Χρησιμοποιώντας το λογισμικό της ανεξάρτητης ομάδας JPEG, είναι δυνατό να συμπεστούν οι μονοσκοπικές (monoscopic) εικόνες σε αρκετά υψηλές ροές πλαισίων. Παραδείγματος χάριν, 25 πλαίσια ανά δευτερόλεπτο επιτεύχθηκαν για τις monoscopic εικόνες χρώματος σε μία ανάλυση 384x288 εικονοστοιχείων.



Σχήμα – Επισκόπηση της JPEG κωδικοποίησης

❖ MPEG-2

Το MPEG-2 είναι ουσιαστικά μια καθολικότερη έκδοση του MPEG-1 προτύπου με ορισμένες βελτιώσεις. Μπορεί να συμπιέσει το βίντεο σε μια ευρεία ποικιλία αναλύσεων (resolution) εξαιρετικά καλά, με πολλά διαφορετικά επίπεδα για την ποιότητα. Σε πρώιμα πειράματα, ένας κωδικοποιητής MPEG-2 είχε την ταχύτητα ενός Pentium III υπολογιστή στα 600MHz, και μια μέση ταχύτητα 4.43 πλαισίων ανά δευτερόλεπτο επιτεύχθηκε για την κωδικοποίηση του μονοσκοπικού βίντεο με μια ανάλυση 352x240 εικονοστοιχείων. Το κύριο υπολογιστικό κόστος του MPEG μετά το JPEG είναι η δημιουργία των προβλεφθέντων και bidirectionally προβλεφθέντων πλαισίων που απαιτούν σύνθετους υπολογισμούς των αλλαγών στο βίντεο. Τα προβλεφθέντα πλαίσια χρησιμοποιούνται για να συμπιέσουν το βίντεο περαιτέρω, με τη συμπίεση των αλλαγών, ή την έλλειψη αυτών, σε σκηνές μεταξύ δύο διαδοχικών πλαισίων. Ήταν σαφές, ότι η συμπίεση του στερεοφωνικού βίντεο σε ένα λογικό ποσοστό πλαισίων δεν ήταν δυνατή εάν χρησιμοποιηθεί ένας κωδικοποιητής λογισμικού. Υπολογίζεται στο εγγύς μέλλον ότι το MPEG θα έχει τις επεκτάσεις που χρειάζεται ώστε να χειριστεί το στερεοσκοπικό βίντεο, και δεδομένων των εξελίξεων απόδοσης των υπολογιστών, το MPEG θα είναι η μέθοδος συμπίεσης που θα επιλέγεται για το στερεοφωνικό βίντεο ροής μέσω του διαδικτύου.

3. MPEG-1

3.1. Γενικά για το πρότυπο MPEG-1

❖ Πώς δουλεύει το MPEG-1 βίντεο

Πρώτα από όλα, αρχίζει με μια σχετικά χαμηλής ανάλυσης βίντεο ακολουθία (η οποία ελαττώνεται σημαντικά σε σύγκριση με την αρχική) περίπου από 352 σε 240 πλαίσια, με έναν ρυθμό 30 πλαίσια ανά δευτερόλεπτο (οι ΗΠΑ έχουν διαφορετικές τιμές σχετικά με την Ευρώπη), αλλά διατηρεί τον αρχικό υψηλής ποιότητας (CD) ήχο. Οι εικόνες είναι έγχρωμες, αλλά έχουν μετατραπεί στο YUV χώρο, και τα δύο κανάλια χρωματικότητας (chrominance, το U και το V) ελαχιστοποιούνται περισσότερο από 176 σε 120 εικονοστοιχεία. Αποδεικνύεται ότι μπορούμε να χρησιμοποιήσουμε (χωρίς ιδιαίτερα κόστη) τα κανάλια με τη μικρή ανάλυση, χωρίς να γίνεται ενοχλητικά αντιληπτό, τουλάχιστον στις φυσικές (όχι σε αυτές που δημιουργούνται από υπολογιστές) εικόνες.

Το βασικό σχέδιο είναι να προβλεφθεί η κίνηση από πλαίσιο σε πλαίσιο στη διάσταση του χρόνου και έπειτα να χρησιμοποιηθεί ο DCT μετασχηματισμός (διακριτός συνημιτονικός μετασχηματισμός) για να οργανώσει τον πλεονασμό στις χωρικές κατευθύνσεις. Ο DCT γίνεται σε δέσμες 8x8, και η πρόβλεψη κινήσεων γίνεται στο κανάλι φωτεινότητας (Y) σε δέσμες 16x16. Με άλλα λόγια, λαμβάνοντας υπόψη τη δεσμίδα 16x16 στο τρέχον πλαίσιο που προσπαθούμε να κωδικοποιήσουμε, ψάχνουμε μια στενή αντιστοιχία εκείνης της δεσμίδας με ένα προηγούμενο ή μελλοντικό πλαίσιο (υπάρχουν τέτοιοι τρόποι πρόβλεψης εκτελούμενοι προς τα πίσω, όπου τα μελλοντικά πλαίσια στέλνονται πρώτα για να παρεμβληθούν στη συνέχεια μεταξύ των άλλων πλαισίων). Οι DCT συντελεστές (είτε των πραγματικών στοιχείων, είτε της διαφοράς μεταξύ αυτής της δεσμίδας και της στενής της αντιστοιχίας) κβαντοποιούνται, το οποίο σημαίνει ότι διαιρούνται με κάποια τιμή ώστε να απορριφθούν κάποια bit επειδή θα βρεθούν κάτω από ένα κατώτατο όριο. Ευτυχώς, πολλοί από τους συντελεστές θα καταλήξουν να είναι μηδενικοί. Η κβαντοποίηση μπορεί να αλλάξει για κάθε "macroblock" (ένα macroblock είναι 16x16 του Y και η αντιστοιχία 8x8 του U και του V). Τα αποτελέσματα όλου αυτού, που περιλαμβάνουν τους συντελεστές DCT, τα διανύσματα κίνησης και τις παραμέτρους κβαντοποίησης (καθώς και άλλα) υποβάλλονται σε κωδικοποίηση Huffman, η οποία κάνει χρήση σταθερών πινάκων. Επίσης, τα διανύσματα κίνησης και οι συνεχείς DCT συνιστώσες είναι DPCM κωδικοποιημένες.

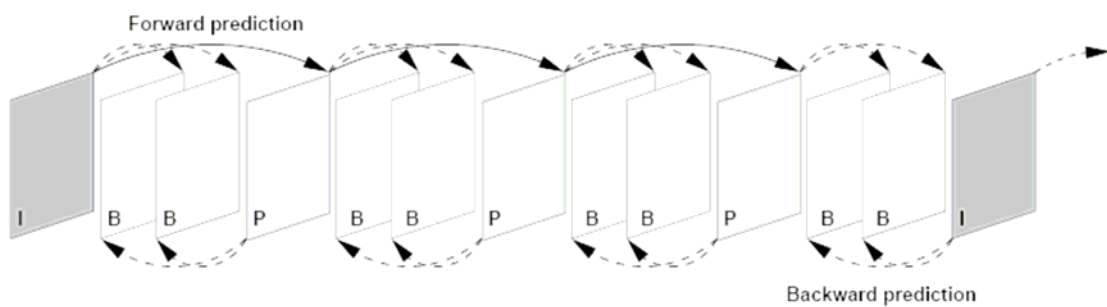
Υπάρχουν τρεις τύποι κωδικοποίησης πλαισίων. Υπάρχουν τα I ή δια- πλαίσια. Είναι απλά πλαίσια που κωδικοποιούνται ως ακίνητη εικόνα, χωρίς τη χρήση κάποιας από τις προηγούμενες μεθόδους. Κατόπιν, υπάρχουν τα P ή προβλεφθέντα πλαίσια. Υπολογίζονται από το πιο πρόσφατα αναδημιουργημένο I ή P πλαίσιο (από την άποψη του αποσυμπίεστη.) Κάθε macroblock σε ένα P πλαίσιο μπορεί να συνοδεύεται με ένα διάνυσμα και τους διαφορικούς DCT συντελεστές για μια στενή αντιστοιχία στο τελευταίο I ή P πλαίσιο, ή μπορεί ακριβώς να είναι "δια-κωδικοποιημένο" (όπως τα πλαίσια I) εάν δεν υπήρξε καμία καλή αντιστοιχία.

Τελικά, υπάρχουν τα B ή αμφίδρομα πλαίσια. Προβλέπονται από τα δύο πιο κοντινά I ή P πλαίσια, ένα στο "παρελθόν" και ένα στο "μέλλον".

Η ακολουθία των αποκωδικοποιημένων πλαισίων έχει συνήθως την εξής μορφή :

IBBPBBPBBPBBIBBPBBPB...

όπου υπάρχουν 12 πλαίσια από το I στο I (για τις ΗΠΑ και την Ιαπωνία οπωσδήποτε.) Αυτό είναι βασισμένο στην απαίτηση για τυχαία προσπέλαση ότι δηλαδή χρειάζεται μια αφετηρία (access point) τουλάχιστον μία φορά κάθε 0.4 δευτερόλεπτα περίπου. Η αναλογία των P στα B είναι βασισμένη στην εμπειρία.



Σχήμα – Ενιαίο στρώμα GOP (Group Of Pictures) πλαίσια απο 1 – 12 Single layer

Φυσικά, για τη λειτουργία ενός αποκωδικοποιητή, πρέπει να σταλεί το πρώτο P πριν από τα δύο πρώτα B, έτσι ώστε η συμπιεσμένη ροή δεδομένων να έχει κάποια τέτοια μορφή:

0xx312645...

όπου οι αριθμοί είναι αριθμοί πλαισίων, τα xx μπορεί να μην είναι κάτι (εάν αυτό είναι η αληθινή αφετηρία), ή μπορεί να είναι τα πλαίσια B -2 και -1 εάν είμαστε κάπου στη μέση της ροής δεδομένων

Το I πλαίσιο πρέπει να αποκωδικοποιηθεί πρώτα, κατόπιν το P, να κρατηθούν και τα δύο στη μνήμη, κι έπειτα να αποκωδικοποιηθούν τα δύο B. Πιθανώς, παρουσιάζετε το I ενώ αποκωδικοποιείτε το P, και τα B καθώς τα αποκωδικοποιείτε, κι έπειτα το P καθώς αποκωδικοποιείτε το επόμενο P, και τα λοιπά..

| GoP number | GoP name | Frame sequence | GoP-length |
|------------|----------|-------------------------|----------------|
| 00 | gop00 | IBBPBBPBBPBB | 12 |
| 01 | gop01 | IIIIIIII... | 1 |
| 02 | gop02 | IPIPIP... | 2 |
| 03 | gop03 | IPPIPPI... | 3 |
| 04 | gop04 | IPPPIPPI... | 4 |
| 05 | gop05 | IBIBI... | 2 |
| 06 | gop06 | IBBIBBI... | 3 |
| 07 | gop07 | IBBBIBBBI... | 4 |
| 08 | gop08 | IBBPBBPBBPBBPBBPBB... | whole sequence |
| 09 | gop09 | IBBPBBSPBBPBBSPBBPBB... | whole sequence |
| 10 | gop10 | IBBPBBSPBBPBBI... | 12 |

Σχήμα – Διαφορετικές GOP δομές

❖ Το κέρδος χρήσης πλαισίων B

Δεδομένου ότι οι αμφίδρομες προβλέψεις macroblock είναι ένας μέσος όρος δύο δεμίδων macroblock, ο θόρυβος μειώνεται στις χαμηλές ροές δεδομένων. Θεωρητικά, στις ροές MPEG-1 βίντεο (352x240x30, στο 1.15 Mbit/sec), λέγεται ότι τα πλαίσια τύπου B βελτιώνουν την αναλογία του σήματος ως προς το θόρυβο που εισάγεται

(Signal to Noise Ratio,SNR) κατά τουλάχιστον 2 DB. (το κέρδος 0.5 DB θεωρείται συνήθως σημαντικό για το πρότυπο MPEG). Εντούτοις, σε μεγαλύτερες ροές δεδομένων, τα πλαίσια B γίνονται λιγότερο χρήσιμα δεδομένου ότι από τη φύση τους δε μπορούν να συμβάλλουν στην προοδευτική “καλλιέργεια” της ακολουθίας μιάς εικόνας (παραδείγματος χάριν, δεν χρησιμοποιούνται ως πρόβλεψη από τα επόμενα κωδικοποιημένα πλαίσια). Πέρα από όλα αυτά, τα πλαίσια B είναι ακόμα υπό αμφισβήτηση για το αν θα πρέπει να χρησιμοποιούνται ή όχι.

❖ Γιατί τα πλαίσια B είναι υπό αμφισβήτηση

Η υπολογιστική πολυπλοκότητα, το εύρος ζώνης, η καθυστέρηση, και το μέγεθος των καταχωρητών που απαιτούνται για την εικόνα είναι οι τέσσερις αναβλητικοί παράγοντες για τα πλαίσια B. Η υπολογιστική πολυπλοκότητα αυξάνεται δεδομένου ότι ο υπολογισμός ενός macroblock χρησιμοποιεί τον μέσο όρο δύο άλλων macroblocks. Στην χειρότερη περίπτωση, το εύρος ζώνης για τη μνήμη αυξάνεται κατά 16 Mbytes γι’ αυτήν την επιπλέον πρόβλεψη. Χρειάζεται ένας πρόσθετος καταχωρητής εικόνας για να αποθηκεύσει τη μελλοντική αναφορά πρόβλεψης. Τέλος, εισάγεται πρόσθετη καθυστέρηση στην κωδικοποίηση δεδομένου ότι το πλαίσιο, που χρησιμοποιείται για την προς τα πίσω πρόβλεψη πρέπει να διαβιβαστεί στον αποκωδικοποιητή προτού να μπορέσουν να αποκωδικοποιηθούν και να παρουσιαστούν οι ενδιάμεσες εικόνες τύπου B.

Η καλωδιακή τηλεόραση υπήρξε ιδιαίτερος “αντίθετη” στα πλαίσια B επειδή ο πρόσθετος καταχωρητής εικόνας ωθεί τις απαιτήσεις μνήμης DRAM για τους αποκωδικοποιητές πέρα από το “μαγικό” κατώφλι των 8Mbit (1 MB) στη σφαίρα των 16 Mbits (2 MByte) για CCIR-610 πλαίσια (704x480), όμως όχι για τιμές κατώτερες του 352x480. Παρά ταύτα, το καλώδιο δεν συνειδητοποιεί ότι η μνήμη DRAM δεν κυκλοφορεί στις εξυπηρετικές συσκευασίες υψηλής απόδοσης/χαμηλού κόστους των 8Mbit αλλά υπάρχει μόνο συσκευασία των 16Mbit. Σε μερικά έτη, οι οικονομικής αξίας αποκλίσεις μεταξύ 16 Mbit και 8 Mbit θα γίνουν ασήμαντες έναντι στο κέρδος που θα απορρέει από τη συμπίεση. Προς το παρόν, τα κιβώτια καλωδίων θα αρχίσουν με 8Mbit και θα εμφανιστούν στο μέλλον βελτιώσεις ρων 16 Mbit.

❖ Μπορούν τα διανύσματα κίνησης να χρησιμοποιηθούν για να μετρήσουν την ταχύτητα αντικειμένου;

Οι διανυσματικές πληροφορίες για την κίνηση δεν μπορούν να χρησιμοποιηθούν σοβαρά ως μέσο καθορισμού της ταχύτητας ενός αντικειμένου εκτός εάν το μοντέλο του κωδικοποιητή καθορίστηκε να λειτουργεί έτσι. Αρχικά, τα μοντέλα των κωδικοποιητών, που βελτιστοποιούν την ποιότητα της εικόνας από τα διανύσματα ελαχιστοποιούν χαρακτηριστικά το λάθος πρόβλεψης και συνεπώς τα διανύσματα συχνά δεν αντιπροσωπεύουν την αληθινή απεικόνιση του αντικειμένου. Πρότυποι μετατροπείς, οι οποίοι πραγματοποιούν επαναδειγματοληψία από μία μέθοδο σε μία άλλη (όπως από NTSC σε PAL) κάνουν χρήση άλλων μεθόδων (κωδικοποίηση τομέων, ανίχνευση ορίων και άλλα) οι οποίες δεν ενδιαφέρονται για τη βελτιστοποίηση της αναλογίας του σήματος ως προς το θόρυβο (SNR) έναντι στη βελτιστοποίηση του ρυθμού μετάδοσης των δεδομένων. Αφετέρου, τα διανύσματα κίνησης δεν διαβιβάζονται για όλα τα macroblocks οπωσδήποτε.

❖ Πώς κωδικοποιείται το interlaced βίντεο με MPEG-1 σύνταξη;

Δύο μέθοδοι μπορούν να εφαρμοστούν στο interlaced βίντεο, οι οποίες διατηρούν συμβατότητα με το πρότυπο MPEG-1 (που σχεδιάστηκε αρχικά για τα προοδευτικά πλαίσια μόνο). Στη μέθοδο αλληλουχίας/συρροής (concatenation) τομέων, ο κωδικοποιητής μπορεί προσεκτικά να υπολογίσει τις προβλέψεις και τα λάθη των προβλέψεων που πραγματοποιούν την καλή συμπίεση αλλά διατηρούν την ακεραιότητα των τομέων. Μερικές τεχνικές προεπεξεργασίας μπορούν επίσης να εφαρμοστούν στο interlaced πηγαίο βίντεο που, παραδείγματος χάριν, θα ελαχιστοποιούσε τις αιχμηρές κάθετες συχνότητες. Αυτή η τεχνική δεν είναι αποδοτική, βέβαια. Από την άλλη πλευρά, εάν η πηγή ήταν προοδευτική (π.χ. μία ταινία), είναι πιά τετριμμένο να μετατραπεί η interlaced πηγή σε ένα προοδευτικό φορμάτ πριν κωδικοποιηθεί. (το MPEG-2 έπειτα θα πρόσφερε την ανώτερη απόδοση του μέσω της μεγαλύτερης ακρίβειας των συνεχών δεσμίδων, μη γραμμικού mquant, intra - VLC, κ.λπ...) Τα επαναδημιουργημένα πλαίσια επαν- interlaced στη διαδικασία απεικόνισης των αποκωδικοποιητών.

Η δεύτερη συντακτικά συμβατή μέθοδος κωδικοποιεί τους τομείς χωριστά. Οι τύποι εικόνων κλειδώνονται στη δραστηριότητα κινήσεων ώστε να ενισχύσουν την αποδοτικότητα της πρόβλεψης.

❖ Πως προέκυψε το 352x240 ?

Προέρχεται από το πρότυπο CCIR-610 της ψηφιακής τηλεόρασης που χρησιμοποιείται από τον επαγγελματικό ψηφιακό βίντεο εξοπλισμό. Είναι 720x243x60 τομείς (όχι πλαίσια) (στις ΗΠΑ) ανά δευτερόλεπτο, όπου απεικονίζονται οι τομείς interlaced. Τα κανάλια της χρωματικότητας (chrominance) είναι 360x243x60 τομείς το δευτερόλεπτο, τα οποία γίνονται πάλι interlaced. Αυτός ο βαθμός αποδεκατισμού της χρωματικότητας (2:1 στην οριζόντια κατεύθυνση) καλείται 4:2:2. Το φορμάτ της πηγαίας εισόδου για το MPEG-1, αποκαλούμενο SIF, είναι CCIR-610 που αποδεκατίζεται 2:1 στην οριζόντια κατεύθυνση, 2:1 στη χρονική κατεύθυνση, και επιπλέον 2:1 στην κάθετη κατεύθυνση της χρωματικότητας. Επίσης, μερικές γραμμές κόβονται για να επιβεβαιωθεί ότι τα δεδομένα διαιρούνται με το 8 ή το 16 όπου απαιτείται. Για τα 50Hz τα πρότυπα απεικόνισης (PAL, SECAM) αλλάζουν τον αριθμό γραμμών σε έναν τομέα από 243 ή 240 στα 288, και αλλάζουν και τον ρυθμό απεικόνισης σε 50 fields/s ή 25 frames/s. Ομοίως, αλλάζουν τις 120 γραμμές στα αποδεκατισμένα κανάλια της χρωματικότητας σε 144 γραμμές. Δεδομένου ότι το 288*50 είναι ακριβώς ίσο με το 240*60, τα δύο φορμάτ έχουν τον ίδιο ρυθμό πηγαίων δεδομένων.

❖ Μπορεί το MPEG-1 να κωδικοποιήσει ρυθμούς δειγματοληψίας υψηλότερους από 352x240x30 ?

Ασφαλώς και μπορεί. Η σύνταξη του προτύπου MPEG-1 επιτρέπει διαστάσεις δειγματοληψίας έως και 4095x4095x60 πλαίσια ανά δευτερόλεπτο. Το MPEG-1 είναι πραγματικά ένα υποσύνολο γνωστό ως ροή μετάδοσης περιορισμένων παραμέτρων (Constrained Parameters Bitstream).

❖ Τί είναι οι ροές μετάδοσης περιορισμένων παραμέτρων (CPB)?

Το CPB είναι ένα περιορισμένο σύνολο παραμέτρων δειγματοληψίας και ρυθμών μετάδοσης, που χρησιμοποιούνται με σκοπό να ομαλοποιηθεί η υπολογιστική πολυπλοκότητα, το μέγεθος των καταχωρητών, και το εύρος ζώνης μνήμης εξετάζοντας ακόμα και την ευρύτερη πιθανή σειρά των εφαρμογών. Το CPB περιορίζει το βίντεο σε 396 macroblocks (101.376 εικονοστοιχεία) ανά πλαίσιο εάν ο ρυθμός πλαισίων είναι μικρότερος ή ίσο προς 25 fps (πλαίσια ανά δευτερόλεπτο), και 330 macroblocks (84.480 εικονοστοιχεία) ανά πλαίσιο εάν ο ρυθμός πλαισίων είναι μικρότερος ή ίσος με 30 fps. Επομένως, το βίντεο MPEG κωδικοποιείται χαρακτηριστικά στις διαστάσεις SIF (352x240x30fps ή 352x288x25 fps).

Ο συνολικός μέγιστος ρυθμός δειγματοληψίας είναι 3.8 Ms/s (εκατομμύριο δείγματα/δευτερόλεπτο) συμπεριλαμβανομένου του χρώματος. Ο ρυθμός του κωδικοποιημένου βίντεο περιορίζεται σε 1.862 Mbit/sec. Στις βιομηχανικές εφαρμογές, ο ρυθμός μετάδοσης είναι η συχνότερα παραμερισμένη παράμετρος CPB, με ρυθμούς τόσο υψηλούς, όπως τα 6 Mbit/sec σε χρήση.

❖ Πόση συμπίεση πραγματοποιείται ?

Όπως αναφέρθηκε πριν, οι ρυθμοί δεδομένων για τους ηχητικούς δίσκους(audio CD) είναι περίπου 1.5 Mbit/sec. Το ίδιο στερεοφωνικό πρόγραμμα μπορεί να συμπιεστεί στα 256 Kbit/sec χωρίς απώλειες στην ποιότητα, που αντιλαμβάνεται ο άνθρωπος. (Βέβαια, δεν αποκλείεται, μερικές φορές να παρατηρηθεί κάποια αλλοίωση στην ποιότητα. Εντούτοις, η επίδραση αυτή είναι πολύ μικρή, και χρειάζεται έναν εκπαιδευμένο ακροατή, για να παρατηρήσει αυτούς τους ιδιαίτερους τύπους αποτελεσμάτων.) Η αναλογία αυτής της συμπίεσης είναι περίπου 6:1. Έτσι, μία ροή για ένα CD MPEG-1 θα άφηνε περίπου 1.25 Mbits/s για το βίντεο. Ο αριθμός που χρησιμοποιείται συνήθως είναι 1.15 Mbits/s (ίσως χρειάζεται το υπόλοιπο για τη ροή των δεδομένων του συστήματος). Από τους αριθμούς, που παρουσιάζονται εδώ η αναλογία συμπίεσης του βίντεο είναι 26:1. Βέβαια, αυτή η αναλογία συμπίεσης μοιάζει με θαύμα. Φυσικά, είναι μια απωλεστική συμπίεση, αλλά μπορεί γίνει αρκετά δύσκολο μερικές φορές το να παρατηρηθεί η απώλεια, εάν συγκρίνετε το αρχικό SIF με το SIF που αποσυμπιέζεται. Υπάρχει, παρόλα αυτά, μια πολύ αξιοσημείωτη απώλεια, εάν η αρχική πληροφορία είναι CCIR-610 και πρέπει να αποδεκατιστεί σε SIF, αλλά αυτό είναι ένα άλλο θέμα, το οποίο δεν υπολογίζεται στην αναλογία 26:1.

Τα πρότυπα προνόησαν, επίσης και για άλλους ρυθμούς μετάδοσης, οι οποίοι κυμαίνονται από 32Kbits/s για ένα κανάλι (mono), μέχρι 448 Kbit/s για το στερεοφωνικό κανάλι.

3.2. Ηχητικό MPEG-1 (Audio MPEG-1)

❖ Στον ήχο εφαρμόζεται η ίδια συμπίεση με το βίντεο ?

Σίγουρα όχι. Το μάτι και το αυτί, παρόλο που είναι μόνο μερικά μόνο εκατοστά μακριά, λειτουργούν πολύ διαφορετικά. Το αυτί έχει μια πολύ υψηλότερη δυναμική

περιοχή και ανάλυση. Μπορεί να κρατήσει περισσότερες λεπτομέρειες αλλά είναι πιο αργό από το μάτι.

Η επιτροπή MPEG επέλεξε να συστήσει 3 μεθόδους συμπίεσης και τις ονόμασε το ακουστικό επίπεδο I, II και III (Audio Level I, II and III). Το επίπεδο I είναι το απλούστερο, ένας κωδικοποιητής υποζωνών με ένα ψυχο-ακουστικό πρότυπο. Το στρώμα II προσθέτει τις πιο προηγμένες τεχνικές bit allocation και μεγαλύτερη ακρίβεια. Το στρώμα III προσθέτει μία υβριδική τράπεζα φίλτρων (filterbank) και μη ομοιόμορφη κβαντοποίηση. Τα στρώματα I, II και III δίνουν τις αυξανόμενες αναλογίες ποιότητας/συμπίεσης με την αυξανόμενη πολυπλοκότητα και τις απαιτήσεις σε επεξεργαστική ισχύ.

Ο λόγος για τον οποίο υποδείχθηκαν 3 μέθοδοι, είναι ότι αφ' ενός οι ελεγκτές θεώρησαν ότι κανένας από τους κωδικοποιητές δεν ήταν 100% διαφανής (transparent) σε όλο του το υλικό και αφ' ετέρου ότι ο καλύτερος κωδικοποιητής (το στρώμα III) ήταν υπολογιστικά αργός και θα προσέκρουε σοβαρά στην αποδοχή των προτύπων.

Τα specs λένε ότι ένας έγκυρος αποκωδικοποιητής στρώματος III θα είναι σε θέση να αποκωδικοποιήσει οποιαδήποτε ακουστική ροή MPEG, στρώματος I, II ή III. Ένας αποκωδικοποιητής στρώματος II θα είναι σε θέση να αποκωδικοποιήσει ροές του στρώματος I καθώς και του στρώματος II.

❖ Πώς δουλεύει το ηχητικό MPEG-1 (audio MPEG-1) ?

Αρχικά, πρέπει να ξέρετε πώς αποθηκεύεται ο ήχος σε έναν υπολογιστή. Ο ήχος είναι διαφορές πίεσης στον αέρα. Όταν λαμβάνεται από ένα μικρόφωνο και τροφοδοτείται μέσω ενός ενισχυτή αυτές οι διαφορές γίνονται επίπεδα τάσης. Η τάση δειγματοληπτείται από τον υπολογιστή σε διάφορες στιγμές ανά δευτερόλεπτο. Για την Cd-ακουστική ποιότητα χρειάζεται να δειγματοληπτείτε 44100 φορές ανά δευτερόλεπτο και κάθε δείγμα να έχει μία αναπαράσταση 16-bit. Στο στερεοφωνικό κανάλι αυτό σημαίνει 1,4Mbit ανά δευτερόλεπτο και μπορείτε προφανώς να αντιληφθείτε την ανάγκη για συμπίεση.

Για να συμπίεσει τον ήχο, το MPEG προσπαθεί να αφαιρέσει τα άσχετα και τα περιττά μέρη του σήματος. Τα μέρη του ήχου που δεν ακούμε μπορούν να απομακρυνθούν. Για να τα κάνει αυτά το ακουστικό MPEG-1 κάνει χρήση των αρχών ψυχο-ακουστικής.

❖ Πόσο καλή είναι ηχητική συμπίεση MPEG-1 (audio MPEG-1) ?

Το MPEG μπορεί να συμπίεσει μία ροή δεδομένων στα 32kbit/s και στα 384kbit/s (όταν πρόκειται για το στρώμα II). Μια ακατέργαστη ακουστική ροή μετάδοσης τύπου PCM είναι στα 705kbit/s, έτσι αυτό δίνει μια μέγιστη αναλογία συμπίεσης περίπου 22. Η αναλογία της κανονικής συμπίεσης είναι περίπου 1:6 ή 1:7. Εδώ, πρέπει να σημειώσουμε ότι αντίθετα από το βίντεο, δεν υπάρχει καμία αντιληπτή απώλεια ποιότητας. Τα 96kbit/s θεωρούνται transparent για τους περισσότερους λόγους πρακτικής. Αυτό σημαίνει ότι δεν θα παρατηρήσετε καμία διαφορά μεταξύ του αρχικού και συμπιεσμένου σήματος για τη rock'n roll ή τη μοντέρνα μουσική. Για πιο απαιτητικές ροές μετάδοσης όπως μία συναυλία πιάνο θα πρέπει να χρησιμοποιήσετε 128kbit/s.

❖ Πώς η ηχητική συμπίεση MPEG-1 (audio MPEG-1) επιτυγχάνει αυτή την αναλογία συμπίεσης ?

Με τον ήχο υπάρχουν δύο εναλλακτικές λύσεις. Είτε κάνετε λιγότερο συχνά δειγματοληψία είτε κάνετε δειγματοληψία με χαμηλότερη ανάλυση (μικρότερη από 16 bit ανά δείγμα). Εάν υπάρχει ανάγκη για ποιότητα δεν μπορείτε να κάνετε πολλά με τη συχνότητα δειγματοληψίας. Οι άνθρωποι μπορούν να ακούσουν τους ήχους με συχνότητες από 20Hz έως 20kHz. Σύμφωνα με το θεώρημα Nyquist πρέπει να επιλέξετε μία συχνότητα τουλάχιστον διπλάσια της υψηλότερης συχνότητας, που θέλετε να αναπαραγάγετε. Παίρνοντας και τα κατάλληλα μέτρα για τα μη-τέλεια φίλτρα, ένα ποσοστό δειγματοληψίας 44,1kHz είναι ένα δίκαιο ελάχιστο (minimum). Οπότε είτε προσπαθούμε να αποδείξουμε ότι το θεώρημα Nyquist κάνει λάθος είτε συνεχίζουμε την εργασία με μείωση της ανάλυσης. Η επιτροπή MPEG επέλεξε το τελευταίο.

Τώρα, ο πραγματικός λόγος για την επιλογή των 16bit είναι να αποκτηθεί μία καλή αναλογία του σήματος ως προς τον θόρυβο (s/n). Κι όταν αναφέρουμε θόρυβο, εννοούμε τον θόρυβο κβαντοποίησης, που προκύπτει από τη διαδικασία ψηφιοποίησης. Για κάθε bit που προσθέτουμε, παίρνουμε 6dB καλύτερη σηματοθορυβική αναλογία (s/n). (στο αυτί, 6dB αντιστοιχούν σε έναν διπλασιασμό του επιπέδου του ήχου.) Ο ήχος ποιότητας Cd επιτυγχάνει 90dB σηματοθορυβική αναλογία (s/n). Αυτό ταιριάζει με τη δυναμική περιοχή του αυτιού αρκετά καλά. Δηλαδή δεν θα ακούσετε οποιοδήποτε θόρυβο από το ίδιο το σύστημα. Άρα, τι συμβαίνει όταν επιλέξουμε ανάλυση των 8 bit ; Παίρνουμε ένα πολύ αξιοσημείωτο “πάτωμα” (noise floor) θορύβου στην καταγραφή μας. Μπορείτε εύκολα να το ακούσετε αυτό στις σιωπηλές στιγμές στη μουσική ή μεταξύ των λέξεων ή των προτάσεων εάν η καταγραφή μας είναι μια ανθρώπινη φωνή. Όμως, δεν παρατηρούμε κάποιο θόρυβο στις δυνατές μεταβάσεις. Αυτό είναι λόγω ενός ιδιαίτερου εφέ (masking effect) και είναι το κλειδί στην ακουστική κωδικοποίηση MPEG. Μέθοδοι όπως το masking effect ανήκουν σε μια επιστήμη ονομαζόμενη ψυχο-ακουστική, που εξετάζει τον τρόπο με τον οποίο ο ανθρώπινος εγκέφαλος αντιλαμβάνεται τον ήχο. Και το MPEG πρότυπο χρησιμοποιεί τις ψυχο-ακουστικές αρχές.

❖ Ποιές είναι οι απαιτήσεις για υλικό (hardware) ?

Το στρώμα III χρειάζεται 20MIPS ανά κανάλι για την κωδικοποίηση σε πραγματικό χρόνο. Αυτό σημαίνει ένα πραγματικά γρήγορο DSP. Το στρώμα II αφ' ετέρου χρειάζεται μόνο ένα απλό DSP όπως παραδείγματος χάριν το AD2015 που μπορεί να αποκτηθεί με μερικά δολάρια. Η διαδικασία είναι ασύμμετρη, διότι απαιτείται πολλή περισσότερη επεξεργασία από την πλευρά κωδικοποίησης. Ένας αποκωδικοποιητής θα μπορούσε να υλοποιηθεί έτσι ώστε να λειτουργεί χωρίς τη βοήθεια πρόσθετου υλικού και να λειτουργήσει μέσω ενός υπολογιστή.

❖ Ποιες συχνότητες δειγματοληψίας χρησιμοποιούνται ?

Μπορούμε να χρησιμοποιήσουμε 48kHz, (που χρησιμοποιούνται στον επαγγελματικό εξοπλισμό για τον ήχο), 44,1kHz (που χρησιμοποιούνται στον καταναλωτικό

εξοπλισμό όπως τον ήχο ποιότητας Cd) ή 32kHz (σε κάποιο εξοπλισμό επικοινωνιών).

❖ Πόσα ακουστικά κανάλια υπάρχουν/χρησιμοποιούνται ?

Το MPEG I επιτρέπει δύο ακουστικά κανάλια. Αυτά μπορούν να είναι είτε ένα ενιαίο (μονο) είτε ένα διπλό (δύο μονο κανάλια), στερεοφωνικό είτε κοινό στερεοφωνικό συγκρότημα (στερεοφωνικό συγκρότημα έντασης ή μ/s-στερεοφωνικό συγκρότημα). Στο κανονικό στερεοφωνικό συγκρότημα (l/r) ένα κανάλι φέρνει το αριστερό ακουστικό σήμα και ένα κανάλι φέρνει το σωστό ακουστικό σήμα. Στο στερεοφωνικό συγκρότημα m/s ένα κανάλι φέρνει το σήμα ποσού (l+r) και άλλο το σήμα διαφοράς (LR). Στο στερεοφωνικό συγκρότημα έντασης συνδυάζεται το μέρος υψηλής συχνότητας του σήματος (επάνω από 2kHz). Η στερεοφωνική εικόνα συντηρείται αλλά μόνο ο χρονικός φάκελος διαβιβάζεται. Το MPEG II επιτρέπει διάφορα κανάλια στο ίδιο ρεύμα.

❖ Τα συστήμα MPEG-1

Η επιτροπή του συστήματος MPEG-1 ολοκλήρωσε και ενέκρινε για κυκλοφορία τις τεχνικές προδιαγραφές για το συνδυασμό πολλαπλών κωδικοποιημένων ηχητικών και τηλεοπτικών ροών σε μία ενιαία ροή δεδομένων. Οι προδιαγραφές παρέχουν ένα πλήρως συγχρονισμένο ήχο και ένα βίντεο και διευκολύνουν την αποθήκευση καθώς και την πιθανή περαιτέρω μετάδοση των συνδυασμένων πληροφοριών μέσω ποικίλων ψηφιακών μέσων.

Αυτή η κωδικοποίηση συστημάτων περιλαμβάνει τις απαραίτητες και ικανές πληροφορίες στη ροή μετάδοσης ώστε να παρέχει στις διάφορες συναρτήσεις των επιπέδων του συστήματος, του συγχρονισμού της αποκωδικοποιημένης ακουστικής και τηλεοπτικής πληροφορίας, αρχική και συνεχή διαχείριση των καταχωρητών των κωδικοποιημένων δεδομένων για να αποτρέψει την υπερχείλιση και την υποχείλιση, το ξεκίνημα τυχαίας πρόσβασης, και τον απόλυτο χρονικό προσδιορισμό. Το στρώμα κωδικοποίησης εξηγεί ένα πολλαπλό φορμάτ δεδομένων που επιτρέπει την πολύπλεξη πολλαπλών ταυτόχρονων ακουστικών και τηλεοπτικών ροών καθώς επίσης τόσο καλά όσο αν ήταν ξεχωριστές ροές δεδομένων.

Η βασική αρχή της κωδικοποίησης συστημάτων MPEG είναι η χρήση των χρονικών γραμματοσήμων (time stamps) που διευκρινίζουν το χρόνο αποκωδικοποίησης και απεικόνισης του ήχου και του βίντεο και το χρόνο υποδοχής των πολυπλεγμένων κωδικοποιημένων δεδομένων στον αποκωδικοποιητή, όλα υπό τον περιορισμό ενός ενιαίου ρολογιού συστήματος, στα 90kHz. Αυτή η μέθοδος προσφέρει αρκετή ευελιξία σε τέτοιες περιοχές όπως τον σχεδιασμό των αποκωδικοποιητών, τον αριθμό των ροών, τα μήκη πολλαπλών πακέτων, τους ρυθμούς μετάδοσης για το βίντεο, τους ρυθμούς των κωδικοποιημένων δεδομένων μέσω ψηφιακής απομνημόνευσης ή της απόδοσης των δικτύων. Παρέχει επίσης την ευελιξία στην επιλογή της οντότητας, η οποία θα είναι η κύρια χρονική βάση, όσο εγγυείται ότι ο συγχρονισμός και η διαχείριση των καταχωρητών παραμένουν σταθερά. Ακόμα, υποστηρίζονται μεταβλητοί ρυθμοί μετάδοσης της πληροφορίας. Τελικά, προτείνεται ένα πρότυπο αναφοράς ενός συστήματος αποκωδικοποιητή, το

οποίο παρέχει τα όρια για τις σειρές των παραμέτρων, που είναι διαθέσιμες στους κωδικοποιητές και προσδιορίζει, επίσης, τις απαιτήσεις για τους αποκωδικοποιητές.

Μερικά εναλλακτικά σύνολα περιορισμών παρέχουν ένα πλαίσιο εργασίας το οποίο προορίζεται για κοινή αποδοχή από τη βιομηχανία ορισμένων βασικών παραμέτρων προς χρήση από αποκωδικοποιητές και προμηθευτές πληροφοριών.

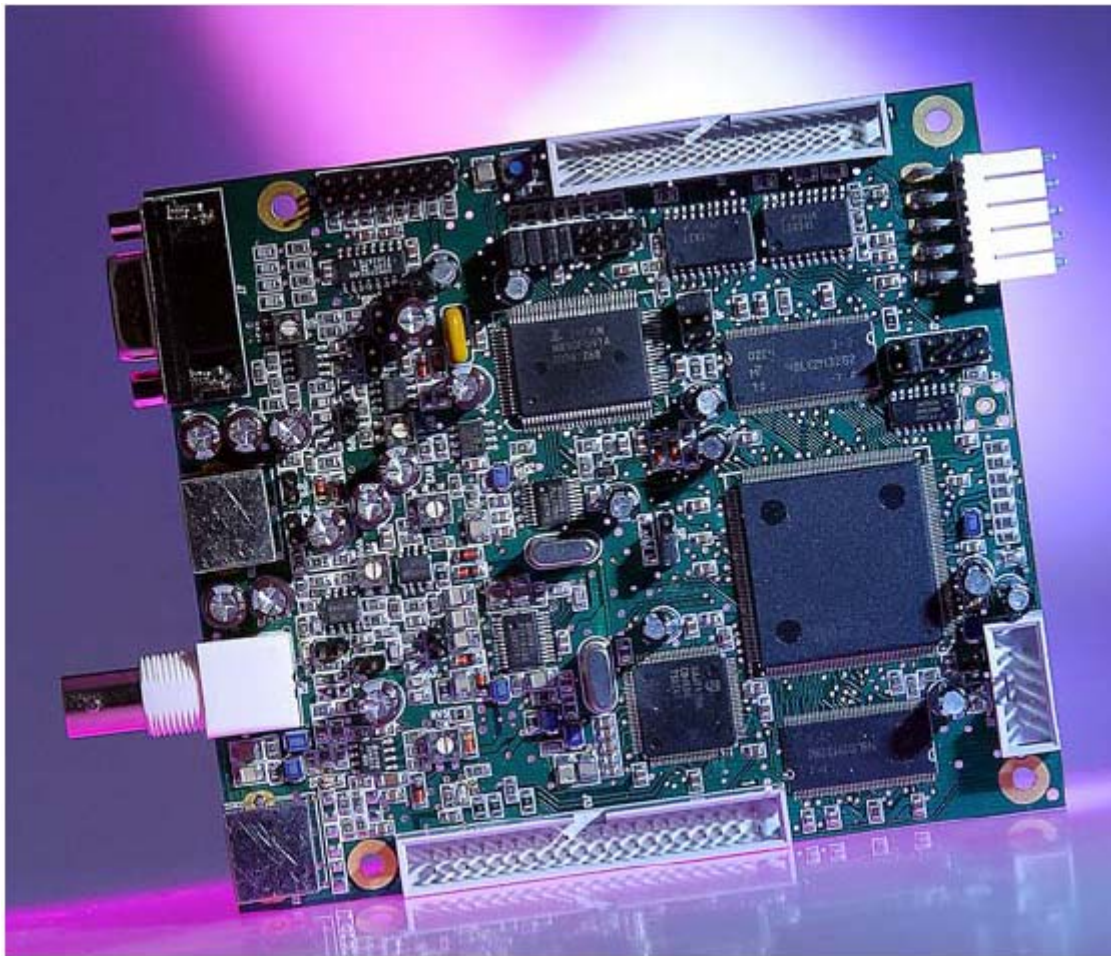
4. MPEG-2

4.1. Γενικά για το πρότυπο MPEG-2

Το MPEG-1 αναπτύσσεται παγκοσμίως σε σημαντικό και επιτυχημένο τηλεοπτικό πρότυπο κωδικοποίησης με έναν αυξανόμενο αριθμό προϊόντων που γίνονται διαθέσιμα στην αγορά. Ένας βασικός παράγοντας για αυτήν την επιτυχία είναι η γενική δομή του προτύπου, που υποστηρίζει μια ευρεία σειρά εφαρμογών και συγκεκριμένων παραμέτρων αυτών των εφαρμογών. Εντούτοις, το MPEG συνέχισε τις προσπάθειες τυποποίησής του το 1991 με μια δεύτερη φάση (το MPEG-2) ώστε να παρέχει μια λύση κωδικοποίησης για το βίντεο και τις εφαρμογές που δεν καλύφθηκαν αρχικά ή που δεν προβλέφθηκαν από το MPEG-1. Συγκεκριμένα, στον ιδρυτικό χάρτη του MPEG-2 ορίστηκε η παροχή ποιότητας βίντεο να μην είναι χαμηλότερη από NTSC/PAL και μέχρι την ποιότητα CCIR 601. Οι “επείγουσες” εφαρμογές, όπως η διανομή της ψηφιακής καλωδιακής τηλεόρασης, δικτυακές υπηρεσίες βάσεων δεδομένων μέσω ATM, ψηφιακές εφαρμογές VTR και δορυφορική και επίγεια ψηφιακή διανομή ραδιοφωνικής αναμετάδοσης, φάνηκαν να ωφελούνται από την αυξανόμενη ποιότητα που αναμενόταν να προκύψει από τη νέα φάση τυποποίησης του MPEG-2. Η εργασία εκτελέστηκε σε συνεργασία με την ITU-T SG 15 ομάδα εμπειρογνομόνων για την κωδικοποίηση του ATM βίντεο και το 1994 τα διεθνή πρότυπα MPEG-2 (που είναι ίδια με τη σύσταση του ITU-T H.262) κυκλοφόρησαν. Οι προδιαγραφές του προτύπου προορίζονται να είναι γενικές εφόσον, το πρότυπο προσπαθεί να διευκολύνει την ανταλλαγή ροών μετάδοσης πληροφορίας μέσω πολλαπλών εφαρμογών, τη μετάδοση και τα μέσα αποθήκευσης.

Βασικά, το πρότυπο MPEG-2 μπορεί να θεωρηθεί σαν ένα υπερσύνολο του MPEG-1 προτύπου κωδικοποίησης και είχε ως σκοπό να είναι συμβατό “προς τα πίσω” με το MPEG-1 και κάθε αποκωδικοποιητής MPEG-2 να μπορεί να αποκωδικοποιήσει μία MPEG-1 ροή bit. Πολλοί αλγόριθμοι κωδικοποίησης βίντεο ενσωματώθηκαν σε έναν ενιαίο για να καλύψουν τις διαφορετικές απαιτήσεις των διάφορων εφαρμογών. Νέα χαρακτηριστικά γνωρίσματα κωδικοποίησης προστέθηκαν από το πρότυπο MPEG-2 για να επιτύχουν την ικανοποιητική λειτουργία και ποιότητα, κατά συνέπεια νέοι τρόποι πρόβλεψης αναπτύχθηκαν για να υποστηρίξουν την αποδοτική κωδικοποίηση του interlaced βίντεο. Επιπλέον, εισήχθησαν επεκτάσεις για να παρέχουν τις πρόσθετες αυτές λειτουργίες, όπως η ενσωματωμένη κωδικοποίηση της ψηφιακής τηλεόρασης και της HDTV. Παρόλα αυτά, η εφαρμογή της ολοκληρωμένης μορφής του MPEG-2 προτύπου, μπορεί να μην είναι πρακτική για πολλές εφαρμογές. Το MPEG-2 έχει εισαγάγει την έννοια των “προφίλ” (profiles) και των “επίπεδων” (levels) για να “επισημάνει” τους όρους προσαρμογής του σε κάποιον εξοπλισμό που δεν υποστηρίζει την πλήρη εφαρμογή. Τα “προφίλ” (profiles) και τα “επίπεδα” (levels) παρέχουν τα μέσα στις ικανότητες των αποκωδικοποιητών ώστε να αποκωδικοποιήσουν μια ιδιαίτερη ροή μετάδοσης.

Κατά γενικό κανόνα, κάθε σχεδιάγραμμα καθορίζει ένα νέο σύνολο αλγορίθμων που προστίθενται ως υπερσύνολο στους αλγορίθμους, στο παρακάτω προφίλ. Το επίπεδο διευκρινίζει τη σειρά των παραμέτρων που υποστηρίζονται από την εφαρμογή (δηλ. μέγεθος εικόνας, ρυθμός μετάδοσης πλαισίων και ρυθμοί μετάδοσης δεδομένων). Στην ουσία του, ο αλγόριθμος MPEG-2 στο βασικό προφίλ “τονίζει” την μη-κλιμακωτή κωδικοποίηση και των προοδευτικών και των συμπελεγμένων (progressive and interlaced video) πηγαίων βίντεο. Αναμένεται ότι οι περισσότερες MPEG-2 εφαρμογές θα προσαρμοστούν τουλάχιστον στο βασικό προφίλ σε βασικό επίπεδο, το οποίο υποστηρίζει την μη-κλιμακωτή κωδικοποίηση του ψηφιακού βίντεο με περίπου τις παραμέτρους της ψηφιακής τηλεόρασης - μια μέγιστη πυκνότητα δειγμάτων 720 δείγματα ανά γραμμή και 576 γραμμές ανά πλαίσιο, ένα μέγιστο ρυθμό μετάδοσης πλαισίων 30 πλαίσια ανά δευτερόλεπτο και ένα μέγιστο ρυθμό μετάδοσης δεδομένων 15 Mbit/s.



Σχήμα - MPEG-2 κάρτα κωδικοποίησης (encoder board)

❖ Οι μη-κλιμακωτοί τρόποι κωδικοποίησης σύμφωνα με το πρότυπο MPEG-2

Ο αλγόριθμος MPEG-2 που καθορίζεται στο βασικό προφίλ είναι μια απλή επέκταση του MPEG-1 προτύπου κωδικοποίησης ώστε να επιτευχθεί η κωδικοποίηση του συμπελεγμένου (interlaced) βίντεο, διατηρώντας την πλήρη σειρά των

συναρτήσεων που παρέχεται από το MPEG-1. Ίδιος με το MPEG-1 πρότυπο, ο αλγόριθμος κωδικοποίησης MPEG-2 είναι βασισμένος στο σχέδιο της γενικής υβριδικής κωδικοποίησης DCT/DPCM, που ενσωματώνει μια δομή Macroblock, μια αποζημίωση κινήσεων και τους τρόπους κωδικοποίησης για το υπό όρους ξαναγέμισμα Macroblocks. Η έννοια/ιδέα των I-εικόνων, των P-εικόνων και των B-εικόνων διατηρείται πλήρως στο MPEG-2 για να επιτύχει την αποδοτική πρόβλεψη κινήσεων και για να βοηθήσει στην random access functionality. Παρατηρήστε, ότι ο αλγόριθμος που ορίστηκε με το απλό MPEG-2 προφίλ είναι βασικά ίδιος με αυτόν στο βασικό προφίλ, εκτός από το ότι κανένας τρόπος πρόβλεψης B-εικόνων δεν επιτρέπεται στον κωδικοποιητή. Κατά συνέπεια, η πρόσθετη πολυπλοκότητα της εφαρμογής και τα πρόσθετα αποθέματα πλαισίων για την αποκωδικοποίηση των B-εικόνων δεν είναι απαραίτητα στους MPEG-2 αποκωδικοποιητές που λειτουργούν μόνο στο απλό προφίλ.

❖ Εικόνες τομέων και πλαισίων:

Το MPEG-2 έχει εισαγάγει την έννοια των εικόνων πλαισίων και των εικόνων τομέων μαζί με τους ιδιαίτερους τρόπους πρόβλεψης πλαισίων και πρόβλεψης τομέων για να προσαρμόσουν την κωδικοποίηση του προοδευτικού και συμπεπλεγμένου (progressive and interlaced) βίντεο. Για τις interlaced ακολουθίες υποτίθεται ότι η είσοδος των κωδικοποιητών αποτελείται από μια σειρά περιέργων (κορυφών) και ακόμα και (κατώτατων σημείων) τομέων, οι οποίοι διαχωρίζονται εγκαίρως από μια περίοδο τομέων. Δύο τομείς ενός πλαισίου μπορούν να κωδικοποιηθούν χωριστά. Σε αυτήν την περίπτωση κάθε τομέας είναι χωρισμένος σε μη-επικαλυπτόμενα macroblocks και ο DCT μετασχηματισμός εφαρμόζεται σε μια βάση τομέων. Εναλλακτικά, δύο τομείς μπορούν να κωδικοποιηθούν μαζί σαν ένα πλαίσιο (εικόνες πλαισίων) παρόμοιο με τη συμβατική κωδικοποίηση των προοδευτικών ακολουθιών βίντεο. Εδώ, οι διαδοχικές γραμμές των κορυφαίων και των κατώτατων τομέων συγχωνεύονται απλά για να διαμορφώσουν ένα πλαίσιο. Παρατηρήστε, ότι και οι εικόνες πλαισίων και οι εικόνες τομέων μπορούν να χρησιμοποιηθούν σε μια ενιαία τηλεοπτική ακολουθία.

❖ Πρόβλεψη τομέων και πλαισίων:

Οι νέοι αντισταθμισμένοι τρόποι πρόβλεψης της κίνησης τομέων εισήχθησαν από το MPEG-2 για να κωδικοποιήσουν αποτελεσματικά τις εικόνες τομέων και τις εικόνες πλαισίων. Στην πρόβλεψη τομέων, οι προβλέψεις γίνονται ανεξάρτητα για κάθε τομέα με τη χρήση των στοιχείων ενός ή περισσότερων προηγούμενων αποκωδικοποιημένων τομέων, δηλαδή για έναν κορυφαίο τομέα μια πρόβλεψη μπορεί να ληφθεί είτε από έναν προηγούμενος αποκωδικοποιημένο κορυφαίο τομέα (που χρησιμοποιεί την αντισταθμισμένη πρόβλεψη κίνησης) είτε από τον προηγούμενος αποκωδικοποιημένο κατώτατο τομέα που ανήκει στην ίδια εικόνα. Γενικά η πρόβλεψη διά-τομέων από τον αποκωδικοποιημένο τομέα στην ίδια εικόνα προτιμάται εάν καμία κίνηση δεν εμφανίζεται μεταξύ των τομέων. Μια ένδειξη, που ο τομέας αναφοράς χρησιμοποιείται για την πρόβλεψη διαβιβάζεται με τη ροή μετάδοσης. Μέσα σε μια εικόνα τομέων όλες οι προβλέψεις είναι προβλέψεις τομέων. Η πρόβλεψη πλαισίων διαμορφώνει μια πρόβλεψη για μια εικόνα πλαισίων βασισμένη σε ένα ή περισσότερα προηγούμενα αποκωδικοποιημένα πλαίσια. Σε μια εικόνα πλαισίων, η πρόβλεψη τομέα ή πλαισίου μπορεί να χρησιμοποιηθεί και ο ιδιαίτερος τρόπος πρόβλεψης που προτιμάται μπορεί να επιλεγεί σε μια βάση macroblock-από-macroblock. Πρέπει να γίνει κατανοητό, ότι οι τομείς και τα πλαίσια από τα οποία γίνονται οι προβλέψεις, μπορεί να είχαν αποκωδικοποιηθεί είτε

ως εικόνες τομέων είτε ως εικόνες πλαισίων. Το MPEG-2 έχει εισαγάγει νέους τρόπους αποζημιώσεων κινήσεων για να ερευνηθούν αποτελεσματικά οι χρονικοί πλεονασμοί μεταξύ των τομέων, που καλούνται "διπλό ξεκίνημα" πρόβλεψης και η αποζημίωση κινήσεων βασισμένες στις 16x8 δεσμίδες.

❖ **Φορμάτ Χρωματικότητας:**

Το MPEG-2 έχει διευκρινίσει επιπλέον (Y:U:V) την υπο-δειγματοληψία των αναλογιών της φωτεινότητας και της χρωματικότητας που βοηθούν και ενισχύουν τις βίντεο εφαρμογές με υψηλότερες απαιτήσεις στην ποιότητα. Μετά το 4:2:0 φορμάτ, που υποστηρίχθηκε ήδη από το MPEG-1 οι προδιαγραφές του MPEG-2 επεκτείνονται στα 4:2:2 φορμάτ, τα οποία είναι κατάλληλα για τις τηλεοπτικές εφαρμογές κωδικοποίησης στα στούντιο.

❖ **MPEG-2 εξελικτικές επεκτάσεις κωδικοποίησης:**

Τα εργαλεία εξελιξιμότητας που τυποποιούνται από το MPEG-2 υποστηρίζουν εφαρμογές πέρα από εκείνες που εξετάζονται από το βασικό ΚΥΡΙΟ αλγόριθμο κωδικοποίησης σχεδιαγράμματος.

❖ **Οι κλιμακωτοί τρόποι κωδικοποίησης σύμφωνα με το πρότυπο MPEG-2**

Τα εργαλεία κλιμάκωσης που τυποποιούνται από το MPEG-2 υποστηρίζουν εφαρμογές πέρα από εκείνες που εξετάζονται από τον αλγόριθμο κωδικοποίησης του βασικού προφίλ. Ο σκοπός της κλιμακωτής κωδικοποίησης είναι να παρασχεθεί η διαλειτουργικότητα (interoperability) μεταξύ των διαφορετικών υπηρεσιών και να υποστηριχθούν ευέλικτα οι δέκτες με τις διαφορετικές δυνατότητες απεικόνισης. Οι δέκτες, που είτε δεν είναι ικανοί είτε δεν είναι πρόθυμοι να αναδημιουργήσουν το πλήρες βίντεο μπορούν να αποκωδικοποιήσουν τα υποσύνολα της διαστρωματομένης ροής μετάδοσης δεδομένων ώστε να απεικονίσουν το βίντεο στη χαμηλότερη χωρική ή χρονική ανάλυση είτε με τη χαμηλότερη ποιότητα. Ένας άλλος σημαντικός σκοπός της κλιμακωτής κωδικοποίησης είναι να παρασχεθεί μία επίσης διαστρωματομένη ροή μετάδοσης δεδομένων που συμμορφώνεται με τη μετάδοση προτεραιοτήτων. Η κύρια πρόκληση εδώ είναι να παραδοθούν αξιόπιστα τα τηλεοπτικά σήματα παρουσία των καναλιών λαθών, όπως η απώλεια κυττάρων στα ATM βασισμένα δίκτυα μετάδοσης ή η παρέμβαση co-channel στην επίγεια ψηφιακή ραδιοφωνική αναμετάδοση.

Η ευέλικτη υποστήριξη των πολλαπλών αναλύσεων είναι ιδιαίτερου ενδιαφέροντος για τη διαδίκτυωση μεταξύ της HDTV και της τυποποιημένης τηλεόρασης (SDTV), οπότε σ' αυτή την περίπτωση είναι σημαντικό για το δέκτη HDTV να είναι συμβατός με το SDTV προϊόν. Η συμβατότητα μπορεί να επιτευχθεί με τη βοήθεια της κλιμακωτής κωδικοποίησης της HDTV πηγής και η επίσημα μετάδοση δύο ανεξάρτητων ροών μετάδοσης στους HDTV και οι SDTV δέκτες μπορεί να αποφευχθεί. Άλλες σημαντικές εφαρμογές για την κλιμακωτή κωδικοποίηση περιλαμβάνουν το τηλεοπτικό ξεφύλλισμα βάσεων δεδομένων βίντεο και την αναπαραγωγή βίντεο πολλαπλών αναλύσεων σε πολυμεσικά περιβάλλοντα.

Σε ένα τηλεοπτικό σχέδιο κωδικοποίησης πολλαπλών κλιμάκων παρέχονται δύο στρώματα, με κάθε στρώμα να υποστηρίζει το βίντεο σε μια διαφορετική κλίμακα, δηλαδή μια απεικόνιση πολλαπλής ανάλυσης μπορεί να επιτευχθεί με την υποκλιμάκωση του τηλεοπτικού σήματος εισόδου σε μια χαμηλότερη ανάλυση

(υποδειγματοληψία στο χώρο ή στο χρόνο). Η έκδοση με την υπο-κλίμακα κωδικοποιείται σε ένα βασικό επίπεδο με ροή μετάδοσης που έχει μειωμένο ρυθμό μετάδοσης. Το αναδημιουργημένο και υπερ-κλιμακωμένο βίντεο του βασικού στρώματος (στο χώρο ή στο χρόνο) χρησιμοποιείται ως πρόβλεψη για την κωδικοποίηση του αρχικού βίντεο εισόδου. Το λάθος πρόβλεψης κωδικοποιείται σε ροή μετάδοσης αναβαθμισμένου στρώματος. Εάν ένας δέκτης είτε δεν είναι ικανός είτε δεν είναι πρόθυμος να απεικονίσει την πλήρη ποιότητα του βίντεο, ένα υπο-κλιμακωμένο βίντεο μπορεί να αναδημιουργηθεί μόνο αν αποκωδικοποιηθεί η βασική ροή μετάδοσης. Είναι σημαντικό να παρατηρηθεί, ότι η απεικόνιση του βίντεο στην υψηλότερη ανάλυση με τη χαμηλότερη ποιότητα είναι επίσης δυνατή μόνο αν αποκωδικοποιηθεί το χαμηλότερο στρώμα της βασικής ροής μετάδοσης. Κατά συνέπεια η κλιμακωτή κωδικοποίηση μπορεί να χρησιμοποιηθεί για να κωδικοποιήσει το βίντεο με έναν κατάλληλο ρυθμό μετάδοσης που διατίθεται σε κάθε στρώμα προκειμένου να καλυφθούν οι συγκεκριμένες απαιτήσεις εύρους ζώνης των καναλιών μετάδοσης ή των μέσων αποθήκευσης. Το “ξεφύλλισμα” μέσω των βάσεων δεδομένων βίντεο και η μετάδοση του βίντεο πέρα από τα ετερογενή δίκτυα, είναι εφαρμογές που αναμένεται να επωφεληθούν από αυτήν την λειτουργία.

Κατά τη διάρκεια της φάσης της τυποποίησης του MPEG-2, φάνηκε αδύνατο να αναπτυχθεί ένα γενικό σχέδιο κλιμακωτής κωδικοποίησης ικανό να ταιριάζει με τις διαφορετικές προβλεπόμενες απαιτήσεις της κάθε εφαρμογής. Ενώ μερικές εφαρμογές είναι περιορισμένες σε χαμηλή πολυπλοκότητα, άλλες απαιτούν πολύ υψηλή αποδοτικότητα κωδικοποίησης. Κατά συνέπεια, το MPEG-2 έχει τυποποιήσει τρία κλιμακωτά σχέδια κωδικοποίησης:

- (Ποιότητα) Κλιμάκωση SNR
- Χωρική κλιμάκωση και
- Χρονική κλιμάκωση

κάθε ένα από τα οποία στοχεύει να βοηθήσει τις εφαρμογές με ιδιαίτερες απαιτήσεις.

Τα εργαλεία κλιμάκωσης παρέχουν αλγοριθμικές επεκτάσεις στο μη-κλιμακωτό σχέδιο που καθορίζεται στο βασικό προφίλ. Είναι δυνατό, να συνδυαστούν διαφορετικά εργαλεία κλιμάκωσης σε ένα υβριδικό μοντέλο κωδικοποίησης, δηλαδή η διαλειτουργικότητα μεταξύ των υπηρεσιών με διαφορετικές αναλύσεις στο χώρο και ρυθμούς πλαισίων μπορεί να υποστηριχθεί με τη βοήθεια του συνδυασμού της χωρικής κλιμάκωσης και του εργαλείου της χρονικής κλιμάκωσης σε ένα υβριδικό διαστρωματομένο σχέδιο κωδικοποίησης.

Η διαλειτουργικότητα μεταξύ των υπηρεσιών HDTV και SDTV μπορεί να παρασχεθεί μαζί με μια ορισμένη ανθεκτικότητα στα λάθη των καναλιών με το συνδυασμό των χωρικών επεκτάσεων κλιμάκωσης με το εργαλείο κλιμάκωσης SNR. Η σύνταξη του MPEG-2 υποστηρίζει μέχρι τρία διαφορετικά στρώματα κλιμάκωσης.

Η **χωρική κλιμάκωση** έχει αναπτυχθεί, για να υποστηρίξει απεικονίσεις με διαφορετικές χωρικές αναλύσεις στο δέκτη (το βίντεο χαμηλότερης χωρικής ανάλυσης μπορεί να αναδημιουργηθεί από το βασικό στρώμα). Αυτή η λειτουργία είναι χρήσιμη για πολλές εφαρμογές συμπεριλαμβανομένης και της ενσωματωμένης κωδικοποίησης για τα συστήματα HDTV/TV, που επιτρέπει την μετανάστευση από μια υπηρεσία ψηφιακής τηλεόρασης σε υψηλότερες χωρικές υπηρεσίες ανάλυσης HDTV (MPEG2, lascha). Ο αλγόριθμος είναι βασισμένος σε μια κλασσική πυραμιδική προσέγγιση για την προοδευτική κωδικοποίηση της εικόνας (purī, burt). Η χωρική κλιμάκωση μπορεί ευέλικτα να υποστηρίξει ένα ευρύ φάσμα χωρικών αναλύσεων αλλά προσθέτει μια ιδιαίτερη πολυπλοκότητα στην εφαρμογή στο βασικό προφίλ της κωδικοποίησης.

Η **κλιμάκωση SNR** είναι ένα εργαλείο, το οποίο πρώτιστα αναπτύχθηκε για να παρέχει την υποβάθμιση (ποιοτική κλιμάκωση) της τηλεοπτικής ποιότητας στα πλέον σημαντικά μέσα μετάδοσης. Εάν το βασικό στρώμα μπορεί να προστατευθεί από τα λάθη μετάδοσης, μια έκδοση του βίντεο με μειωμένη ποιότητα μπορεί να ληφθεί με την αποκωδικοποίηση του σήματος από το βασικό στρώμα μόνο. Ο αλγόριθμος που χρησιμοποιείται για να επιτύχει την υποβάθμιση είναι βασισμένος σε μια τεχνική κλιμάκωσης της συχνότητας (DCT-τομέας). Και τα δύο στρώματα κωδικοποιούν το τηλεοπτικό σήμα στην ίδια χωρική ανάλυση. Η μέθοδος εφαρμόζεται ως απλή επέκταση στο βασικό προφίλ του MPEG-2 κωδικοποιητή και επιτυγχάνει την άριστη αποδοτικότητα της κωδικοποίησης.

Στο βασικό στρώμα οι DCT συντελεστές κβαντοποιούνται σε ευμεγέθη κομμάτια και μεταδίδονται για να επιτύχουν μία μέτρια ποιότητα εικόνας μέσω ενός χαμηλού ρυθμού μετάδοσης. Το στρώμα ενίσχυσης (enhancement layer) κωδικοποιεί και μεταδίδει τη διαφορά μεταξύ των μη-κβαντοποιημένων DCT συντελεστών και των κβαντοποιημένων συντελεστών από το βασικό στρώμα με το καλύτερο βήμα κβαντοποίησης. Στον αποκωδικοποιητή, το σήμα του βίντεο με την υψηλότερη ποιότητα αναδημιουργείται με την αποκωδικοποίηση και του χαμηλότερου αλλά και του υψηλότερου στρώματος της ροής μετάδοσης. Είναι επίσης δυνατό, να χρησιμοποιηθεί αυτή η μέθοδος για να αποκτηθεί το βίντεο με τη χαμηλότερη χωρική ανάλυση από τον δέκτη. Εάν ο αποκωδικοποιητής επιλέγει τους χαμηλότερους NxN DCT συντελεστές από τη ροή μετάδοσης βασικού στρώματος, μεταβλητές μεγέθους NxN, του αντίστροφου DCT μετασχηματισμού, μπορούν να χρησιμοποιηθούν για να αναδημιουργηθεί το βίντεο στη μειωμένη χωρική ανάλυση. Εντούτοις, ανάλογα με τις εφαρμογές, που υποστηρίζουν οι κωδικοποιητές και οι αποκωδικοποιητές το υποκλιμακωμένο βίντεο του χαμηλότερου στρώματος μπορεί να υποβληθεί σε ολίσθηση.

Το εργαλείο της **χρονικής κλιμάκωσης** αναπτύχθηκε με έναν στόχο παρόμοιο με αυτόν της χωρικής κλιμάκωσης, το στερεοσκοπικό βίντεο (stereoscopic video) να μπορεί να υποστηριχθεί από μία διαστρωματομένη ροή μετάδοσης (layered bit stream) κατάλληλη για τους δέκτες με τις στερεοσκοπικές ικανότητες απεικόνισης. Η διάταξη σε στρώματα επιτυγχάνεται με την παροχή μιας πρόβλεψης για μια από τις εικόνες του στερεοσκοπικού βίντεο (για παράδειγμα, αριστερή όψη) στο στρώμα ενίσχυσης, βασισμένη στις κωδικοποιημένες εικόνες από την αντίθετη όψη, που μεταδίδεται στο βασικό στρώμα.

Ο **διαχωρισμός των δεδομένων** προορίζεται να βοηθήσει με την απόκρυψη των λαθών στη μετάδοση ή στα κανάλια ATM, στην επίγεια ραδιοφωνική μετάδοση ή στα περιβάλλοντα μαγνητικής καταγραφής. Επειδή το εργαλείο μπορεί να χρησιμοποιηθεί εξ ολοκλήρου ως εργαλείο μετα-επεξεργασίας και προ-επεξεργασίας σε οποιοδήποτε απλό στρώμα του σχεδίου κωδικοποίησης, δεν έχει τυποποιηθεί τυπικά με το πρότυπο MPEG-2, αλλά αναφέρεται στο πληροφοριακό παράρτημα του εγγράφου MPEG-2 DIS. Ο αλγόριθμος είναι, παρεμφερής με το εργαλείο κλιμάκωσης SNR, βασισμένος στον διαχωρισμό των DCT συντελεστών και εφαρμόζεται με την πολύ χαμηλή πολυπλοκότητα έναντι στα άλλα εξελικτικά σχέδια κωδικοποίησης. Για να παρέχουν την προστασία λάθους, οι κωδικοποιημένοι DFT-συντελεστές στο ρεύμα κομματιών είναι απλά χωρισμένοι και διαβιβασμένοι σε δύο στρώματα με τη διαφορετική πιθανότητα λάθους.

5. MPEG-4

Τα πρότυπα MPEG-4 παρέχουν ένα σύνολο τεχνολογιών για να ικανοποιήσουν τις ανάγκες των συντακτών/δημιουργών, των προμηθευτών διαφόρων υπηρεσιών και των τελικών χρηστών, επίσης.

- ❖ Για τους συντάκτες, το MPEG-4 επιτρέπει την παραγωγή υλικού που έχει την μεγαλύτερη δυνατότητα επαναχρησιμοποίησης, έχει τη μεγαλύτερη ευελιξία από όσο είναι δυνατόν σήμερα με τις μεμονωμένες τεχνολογίες όπως την ψηφιακή τηλεόραση, animated γραφικά, ιστοσελίδες World Wide Web (WWW) και τις επεκτάσεις τους. Επιπλέον, είναι δυνατό να διαχειριστεί καλύτερα και να προστατεύσει δικαιώματα των ιδιοκτητών.
- ❖ Για τους φορείς παροχής υπηρεσιών δικτύου, το MPEG-4 προσφέρει σαφείς πληροφορίες, οι οποίες μπορούν να είναι ερμηνευμένες και μεταφρασμένες στα κατάλληλα μηνύματα κάθε δικτύου με τη βοήθεια των σχετικών οργανισμών προτύπων. Τα παραπάνω, όμως, αποκλείουν οποιαδήποτε εκτίμηση για Ποιότητα Υπηρεσιών, για τις οποίες το MPEG-4 παρέχει έναν γενικό περιγραφέα για κάθε διαφορετικό μέσο(media) MPEG-4. Οι ακριβείς μεταφράσεις από τις παραμέτρους της QoS για κάθε μέσο στο δίκτυο είναι πέρα από το πεδίο του MPEG-4 και αφήνεται στους προμηθευτές δικτύων. (Signaling of the MPEG-4 media QoS descriptors end-to-end enables transport optimization in heterogeneous networks.)
- ❖ Για τους τελικούς χρήστες, το MPEG-4 φέρνει τα πιο υψηλά επίπεδα αλληλεπίδρασης με το “υλικό”(content), μέσα στο σύνολο των ορίων που επιβάλλουν οι συντάκτες. Φέρνει επίσης τα πολυμέσα στα “νέα” δίκτυα, συμπεριλαμβανομένων και εκείνων που κάνουν χρήση σχετικά χαμηλών ρυθμών μετάδοσης. Ένα έγγραφο εφαρμογών MPEG-4 υπάρχει στη σελίδα του MPEG (www.cselt.it/mpeg), που περιγράφει πολλές εφαρμογές για τελικούς χρήστες, συμπεριλαμβανομένης της αλληλεπιδραστικής μετάδοσης πολυμέσων και της κινητής τηλεπικοινωνίας. Για όλα τα εμπλεκόμενα μέρη, το MPEG επιδιώκει να αποφύγει ένα πλήθος ιδιωτικών μη-αλληλεπιδραστικών formats και players.

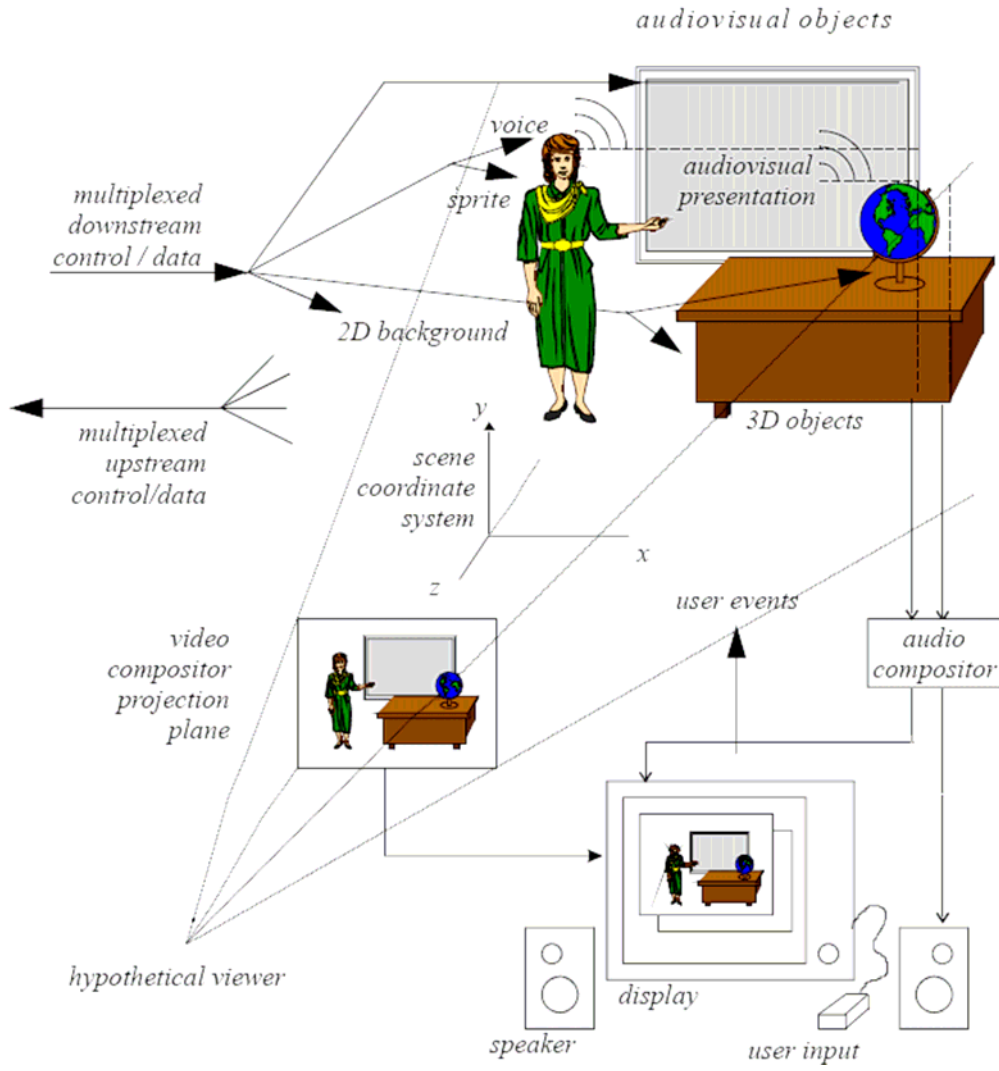
Το MPEG-4 επιτυγχάνει αυτούς τους στόχους, με την παροχή τυποποιημένων τρόπων:

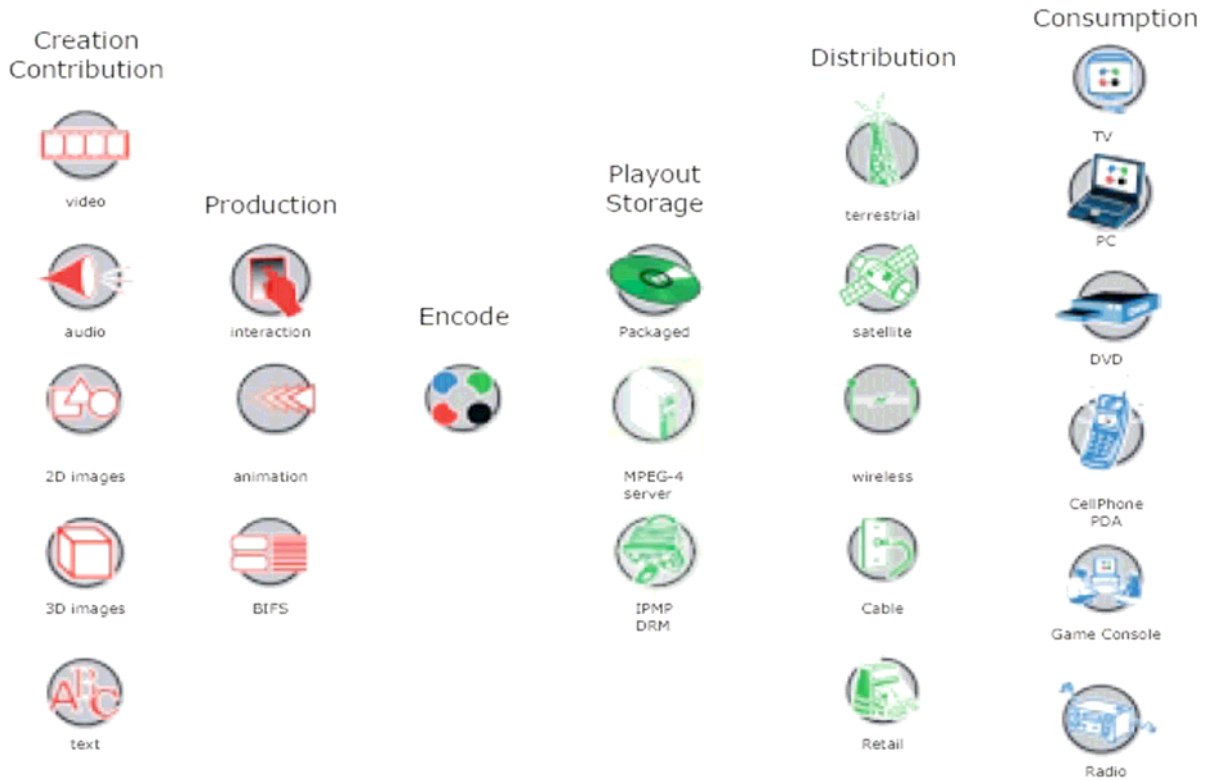
1. Να απεικονίσει μονάδες ακουστικού, οπτικού ή οπτικοακουστικού περιεχομένου, αποκαλούμενες “πολυμεσικά αντικείμενα”. Αυτά τα πολυμεσικά αντικείμενα μπορούν να είναι φυσικής ή ψηφιακής προέλευσης, πράγμα που σημαίνει ότι θα μπορούσαν να καταγραφούν με μια φωτογραφική μηχανή ή ένα μικρόφωνο, ή να έχουν παραχθεί από έναν υπολογιστή.

2. Να περιγράψει τη σύνθεση αυτών των αντικειμένων ώστε να δημιουργηθούν τα σύνθετα πολυμεσικά αντικείμενα, που διαμορφώνουν οπτικοακουστικές σκηνές.

3. Να πολυπλέξει και να συγχρονίσει τα δεδομένα που συνδέονται με τα πολυμεσικά αντικείμενα, έτσι ώστε να μπορούν να μεταφερθούν από τα κανάλια δικτύων που παρέχουν μία QoS κατάλληλη για τη φύση του εκάστοτε πολυμεσικού αντικειμένου.

4. Να αλληλεπιδράσει με την οπτικοακουστική σκηνή που παράγεται στον τελικό δέκτη.

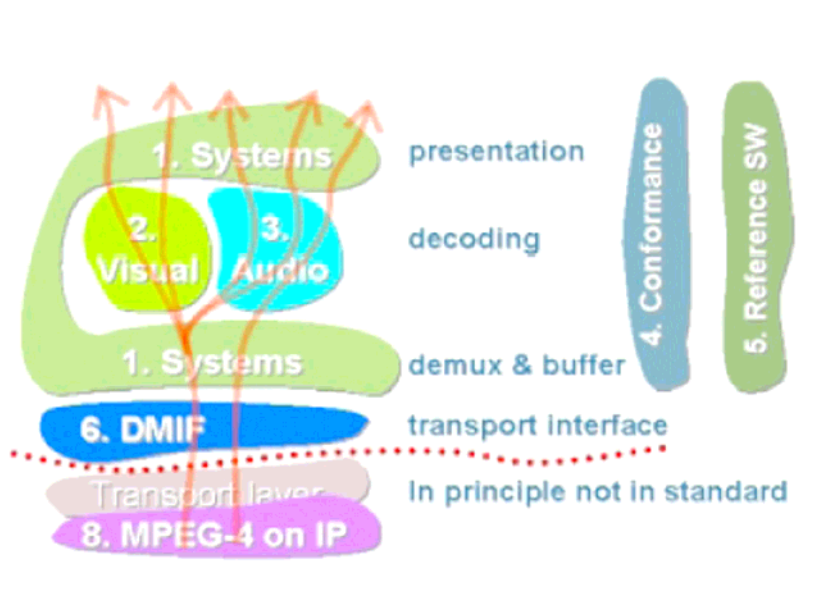




Σχήμα - Το MPEG-4 εξασφαλίζει τη δημιουργία πολυμεσικού υλικού το οποίο μεταδίδεται μέσω οποιουδήποτε δικτύου, σε οποιονδήποτε χρήστη με οποιαδήποτε συσκευή.

5.1. Βασικές λειτουργίες του MPEG-4 στην έκδοση 1

Αυτό το τμήμα περιέχει, με έναν ιδιαίτερα αναλυμένο τρόπο, τις σημαντικότερες λειτουργίες του MPEG-4 για την οριστικοποιημένη έκδοση 1.

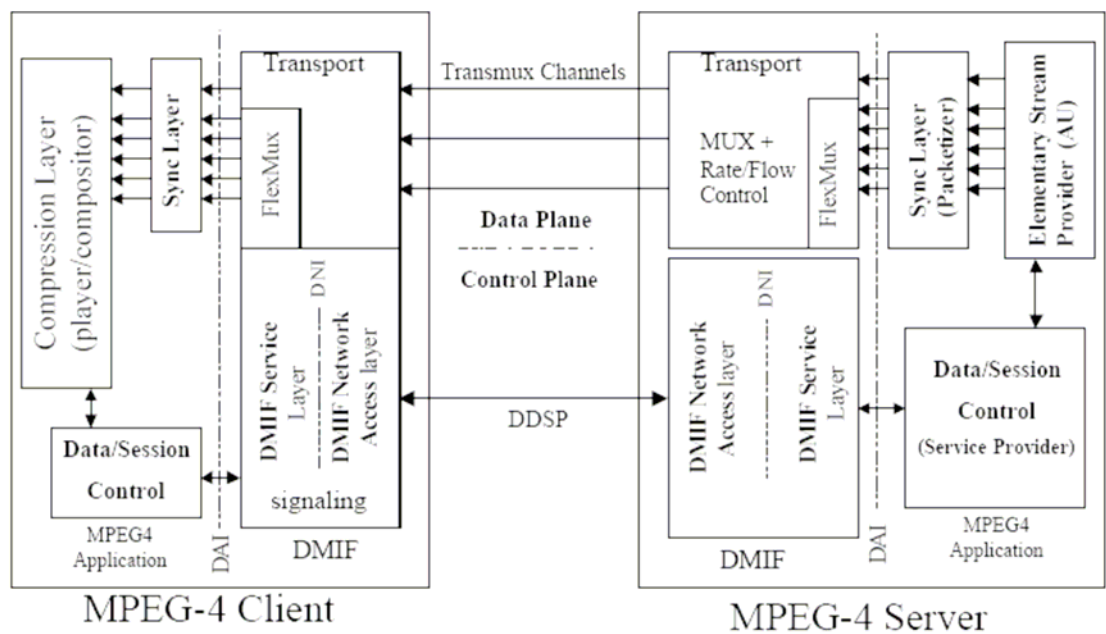


Σχήμα – Τα μέρη του MPEG-4. Τα βέλη δείχνουν τη ροή των bit μέσα σε ένα σύστημα MPEG-4

5.1.1. DMIF

Το DMIF υποστηρίζει τις ακόλουθες λειτουργίες:

- ❖ Σαφής διεπαφή MPEG-4 DMIF-εφαρμογής, ανεξάρτητα από το εάν το peer είναι απομακρυσμένο αλληλεπιδραστικό peer, εκπομπής ή τοπικής αποθήκευσης.
- ❖ Έλεγχος της καθιέρωσης των καναλιών FlexMux
- ❖ Χρήση των ομοιογενών δικτύων μεταξύ των αλληλεπιδραστικών peers: IP, ATM, κινητή τηλεφωνία, PSTN
- ❖ Περιορισμένου εύρους ζώνης ISDN.



Σχήμα – Η αρχιτεκτονική του συστήματος

5.1.2. Συστήματα

Όπως εξηγείται παραπάνω, το MPEG-4 συνιστά μια εργαλειοθήκη προηγμένων αλγορίθμων συμπίεσης για τον ήχο και το βίντεο. Οι ροές δεδομένων (στοιχειώδεις ροές, Elementary Streams ES) που προκύπτουν από τη διαδικασία κωδικοποίησης μπορούν να εκπεμπθούν ή να αποθηκευτούν χωριστά, και να πρέπει να συντεθούν ώστε να δημιουργηθεί η πραγματική πολυμεσική απεικόνιση στην πλευρά τωνδεκτών. Τα συστήματα – μέρη του MPEG-4 αναλαμβάνουν την περιγραφή της σχέσης μεταξύ των οπτικοακουστικών συστατικών που αποτελούν μια σκηνή. Η σχέση περιγράφεται σε δύο κύρια επίπεδα.

- ❖ Το δυαδικό φoρμάτ για τις σκηνές (BIFS, BInary Format for Scenes) περιγράφει τις χωρο-χρονικές θέσεις των αντικειμένων στη σκηνή. Οι θεατές μπορεί να έχουν τη δυνατότητα αλληλεπίδρασης με τα αντικείμενα, παραδείγματος χάριν, να επαναπροσδιορίζουν τη θέση τους στη σκηνή ή να αλλάζουν την οπτική τους γωνία μέσω ενός τρισδιάστατου εικονικό περιβάλλοντος. Η περιγραφή της σκηνής παρέχει ένα πλούσιο σύνολο κόμβων για τη δισδιάστατη 2-D και την τρισδιάστατη 3-D απεικόνιση, στους χειριστές των συνθέσεων καθώς και τους γραφίστες.

❖ Σε ένα χαμηλότερο επίπεδο, οι περιγραφείς αντικειμένου (Ods, Object Descriptors) καθορίζουν τη σχέση μεταξύ των στοιχειωδών ροών με κάθε αντικείμενο (π.χ. η ροή ήχου και βίντεο ενός συμμετέχοντος σε μια τηλεδιάσκεψη). Οι περιγραφείς αντικειμένου, παρέχουν επίσης πρόσθετες πληροφορίες, όπως το URL, που απαιτείται για να υπάρχει πρόσβαση στις στοιχειώδεις ροές, τα χαρακτηριστικά των αποκωδικοποιητών, πνευματική ιδιοκτησία και άλλα.

❖ Άλλα ζητήματα που αντιμετωπίζονται από τα συστήματα MPEG-4

- ❖ Διαδραστικότητα, που περιλαμβάνει: διαδραστικότητα βασισμένη σε σχέση πελάτη – εξυπηρετητή (client - server), ένα γενικό μοντέλο γεγονότων για να προκαλεί γεγονότα ή να δρομολογεί τις ενέργειες των χρηστών, το γενικό χειρισμό των γεγονότων και τη δρομολόγηση μεταξύ των αντικειμένων στη σκηνή, επάνω στο χρήστη ή στα προκαλούμενα στη σκηνή γεγονότα.
- ❖ Ένα εργαλείο για την παρεμβολή πολλαπλών ροών σε μία ενιαία ροή, συμπεριλαμβανομένων των πληροφοριών χρονισμού (Εργαλείο FlexMux).
- ❖ Ένα εργαλείο για την αποθήκευση των MPEG-4 δεδομένων σε ένα αρχείο (το μορμάτ των MPEG-4 αρχείων είναι .mp4).
- ❖ Διεπαφές στις διάφορες πλευρές των τερματικών και των δικτύων, υπό τη μορφή Java API's (MPEG-J)
- ❖ Ανεξαρτησία από το στρώμα μεταφορών. Αναφορές στις σχετικές λίστες του πρωτοκόλλου μεταφοράς συσσωρεύονται, όπως (RTP)/UDP/IP ή MPEG-2 TS μπορούν να καθοριστούν ή καθορίζονται από τους αρμόδιους οργανισμούς τυποποίησης.
- ❖ Απεικόνιση κειμένων με την υποστήριξη διεθνών γλωσσών.
- ❖ Η έναρξη και η συνεχής διαχείριση των καταχωρητών του λαμβάνοντος τερματικού.
- ❖ Μηχανισμοί προσδιορισμού, συγχρονισμού και αποκατάστασης χρονισμού.
- ❖ Σύνολα δεδομένων που καλύπτουν τον προσδιορισμό των δικαιωμάτων πνευματικής ιδιοκτησίας σχετικά με τα πολυμεσικά αντικείμενα.

5.1.3. Ήχος

Ο MPEG-4 ήχος διευκολύνει πολλές εφαρμογές, οι οποίες ποικίλλουν από μία απλή ομιλία έως και έναν υψηλής ποιότητας πολυκάναλο ήχο, και από τους φυσικούς ήχους έως τους ψηφιακούς ήχους. Ειδικότερα, υποστηρίζει την ιδιαίτερα αποδοτική περιγραφή και παράσταση των ακουστικών αντικειμένων που συνιστώνται στα εξής:

- ❖ Σήματα Λόγου: Η κωδικοποίηση του λόγου μπορεί να γίνει χρησιμοποιώντας ρυθμούς μετάδοσης από 2 Kbit/s μέχρι 24 Kbit/s με τη χρήση των εργαλείων κωδικοποίησης του λόγου. Χαμηλότεροι ρυθμοί μετάδοσης, όπως ένας μέσος όρος 1.2 Kbit/s, είναι επίσης δυνατοί όταν επιτρέπονται οι μεταβλητοί ρυθμοί κωδικοποίησης. Η χαμηλή καθυστέρηση είναι δυνατή στις εφαρμογές των τηλεπικοινωνιών. Με τη χρήση των εργαλείων HVXC, η ταχύτητα και το βήμα μπορούν να τροποποιηθούν υπό τον έλεγχο των χρηστών κατά τη διάρκεια αναπαραγωγής του ήχου. Εάν γίνεται χρήση των εργαλείων CELP, τότε μια αλλαγή της ταχύτητας αναπαραγωγής του ήχου μπορεί να επιτευχθεί χρησιμοποιώντας ένα επιπλέον εργαλείο για την επεξεργασία των εφέ.

- ❖ Συνθετιμένη ομιλία: Για τους κλιμακωτούς κωδικοποιητές TTS, οι ρυθμοί μετάδοσης κυμαίνονται από 200 bit/s έως 1.2 Kbit/s, γεγονός το οποίο επιτρέπει σε ένα απλό κείμενο, ή ένα κείμενο με προσωδιακές παραμέτρους (περίγραμμα βημάτων, διάρκεια φωνήματος, κτλ), να θεωρηθούν είσοδοί του για να παραγάγει τη συνθετική ομιλία. Περιλαμβάνει τις ακόλουθες λειτουργίες :
- ❖ Σύνθεση λόγου, που χρησιμοποιεί την προσωδία της αρχικής ομιλίας
- ❖ Έλεγχος του χειλικού συγχρονισμού με τις πληροφορίες φωνήματος.
- ❖ Τρόποι λειτουργίας : μικρή διακοπή (pause), επαναλαμβάνει (resume), πηδά προς τα εμπρός/προς τα πίσω (forward/backward).
- ❖ Υποστήριξη διεθνών γλωσσών και διαλέκτων για το κείμενο. (δηλ. μπορεί να επισημανθεί ποιες γλώσσες και διάλεκτοι πρέπει να χρησιμοποιηθούν στη ροή μετάδοσης)
- ❖ Υποστήριξη διεθνών συμβόλων για τα φωνήματα.
- ❖ Υποστήριξη για τη διευκρίνιση της ηλικίας, του γένους, και του ρυθμού ομιλίας του ομιλητή.
- ❖ Υποστήριξη για τη μετάδοση των σελιδοδεικτών της παραμέτρου facial animation (FAP).
- ❖ Γενικά ηχητικά σήματα: Από τις τεχνικές κωδικοποίησης με μετασχηματισμούς παρέχεται υποστήριξη για την κωδικοποίηση ηχητικών σημάτων, από τους πολύ χαμηλούς ρυθμούς μετάδοσης μέχρι και υψηλές ποιότητες. Με αυτήν την λειτουργία καλύπτεται ένα ευρύ φάσμα ρυθμών μετάδοσης καθώς πολλά εύρη ζώνης. Ξεκινάει από έναν ρυθμό μετάδοσης της τάξης των 6 Kbit/s και ένα εύρος ζώνης υπό τα 4 KHz αλλά περιλαμβάνει ήχο με ποιότητα μετάδοσης από μόνο μέχρι πολυκάναλο.
- ❖ Συνθετιμένος ήχος: Η συνθετική ακουστική υποστήριξη παρέχεται από έναν αποκωδικοποιητή ήχου, εφαρμογή, που επιτρέπει τον έλεγχο των μουσικών οργάνων βασισμένο σε αποτελέσματα που περιγράφονται από μια ειδική γλώσσα.
- ❖ Φραγμένης-πολυπλοκότητας συνθετικός ήχος: Αυτό παρέχεται από έναν αποκωδικοποιητή ήχου, εφαρμογή, που επιτρέπει την επεξεργασία ενός τυποποιημένου κυματικού (wavetable) φορμάτ. Παραδείγματα πρόσθετων λειτουργιών είναι ο έλεγχος της ταχύτητας και η αλλαγή του βήματος για τα σήματα ομιλίας και η κλιμάκωση στα πλαίσια του ρυθμού μετάδοσης, του εύρους ζώνης, των λαθών, της πολυπλοκότητας κ.λπ. όπως καθορίζονται παρακάτω.
- ❖ Η λειτουργία αλλαγής ταχύτητας επιτρέπει την αλλαγή της κλίμακας του χρόνου χωρίς αλλαγή του βήματος κατά τη διάρκεια της αποκωδικοποίησης. Αυτό μπορεί, παραδείγματος χάριν, να χρησιμοποιηθεί για μια λειτουργία τύπου “fast forward” (στην περίπτωση αναζήτησης σε μία βάση δεδομένων) ή για να προσαρμόσει το μήκος μιας ακολουθίας ήχου σε μία δεδομένη βίντεο ακολουθία.
- ❖ Η λειτουργία αλλαγής βημάτων επιτρέπει την αλλαγή των βημάτων χωρίς αλλαγή της χρονικής κλίμακας κατά τη διάρκεια της διαδικασίας κωδικοποίησης ή αποκωδικοποίησης. Αυτό μπορεί να χρησιμοποιηθεί, παραδείγματος χάριν, για την αλλαγή φωνής ή εφαρμογές τύπου Karaoke. Αυτή η τεχνική εφαρμόζεται μόνο στις μεθόδους κωδικοποίησης παραμετρικού και δομημένου ήχου.
- ❖ Η κλιμάκωση του ρυθμού μετάδοσης επιτρέπει σε μια ροή δεδομένων να αναλυθεί σε ροή του χαμηλότερου ρυθμού μετάδοσης ώστε ο συνδυασμός να μπορεί ακόμα να αποκωδικοποιηθεί σε ένα σημαντικό σήμα. Η ανάλυση της ροής μπορεί να πραγματοποιηθεί είτε κατά τη διάρκεια της μετάδοσης είτε στον αποκωδικοποιητή.
- ❖ Η κλιμάκωση του εύρους ζώνης είναι μια ιδιαίτερη περίπτωση κλιμάκωσης του ρυθμού μετάδοσης, όπου μέρος μιάς ροής, που αντιπροσωπεύει μέρος του

φάσματος συχνότητας μπορεί να απορριφθεί κατά τη διάρκεια της μετάδοσης ή της αποκωδικοποίησης.

- ❖ Η κλιμάκωση πολυπλοκότητας των κωδικοποιητών επιτρέπει στους κωδικοποιητές διαφορετικής πολυπλοκότητας να παράγουν έγκυρες και σημαντικές ροές δεδομένων.
- ❖ Η κλιμάκωση πολυπλοκότητας των αποκωδικοποιητών επιτρέπει ροές δεδομένων να αποκωδικοποιηθούν από τους αποκωδικοποιητές σε διαφορετικά επίπεδα πολυπλοκότητας. Τα αποτελέσματα στην ποιότητα του ήχου, γενικότερα, εξαρτώνται από την πολυπλοκότητα.
- ❖ Τα ηχητικά εφέ παρέχουν τη δυνατότητα επεξεργασίας των αποκωδικοποιημένων ηχητικών σήματα με την πλήρη ακρίβεια συγχρονισμού για να επιτευχθούν οι λειτουργίες για τη μίξη, την αντήχηση, το spatialization, κ.λπ.

5.1.4. Εικόνα

Το πρότυπο εικόνας του MPEG-4 θα επιτρέψει την υβριδική κωδικοποίηση των φυσικών εικόνων (βασισμένη σε εικονοστοιχεία) και του βίντεο μαζί με τις συνθετικές σκηνές (αυτές που παράγονται από ηλεκτρονικούς υπολογιστές). Αυτό, παραδείγματος χάριν, θα επιτρέψει εικονική παρουσία συμμετεχόντων σε μία βίντεο - συνεδρία. Για αυτόν τον λόγο, τα πρότυπα εικόνας θα συμπεριλάβουν εργαλεία και αλγόριθμους, που υποστηρίζουν την κωδικοποίηση των φυσικών εικόνων (βασισμένη σε εικονοστοιχεία) και των ακολουθιών βίντεο καθώς επίσης και εργαλεία για να υποστηρίξουν τη συμπίεση των συνθετικών 2D δισδιάστατων και 3D τρισδιάστατων γεωμετρικών παραμέτρων της γραφικής (δηλαδή συμπίεση των παραμέτρων πλέγματος καλωδίων, συνθετικό κείμενο). Παρακάτω δίνεται μια αναλυμένη επισκόπηση των λειτουργιών που τα εργαλεία και οι αλγόριθμοι του MPEG-4 προτύπου θα υποστηρίζουν.

5.1.4.1. Φορμάτ που υποστηρίζονται

Τα ακόλουθα φορμάτ και οι ρυθμοί μετάδοσης θα υποστηριχθούν από την έκδοση 1 του MPEG-4:

- ❖ Ρυθμοί μετάδοσης: χαρακτηριστικά μεταξύ 5 Kbit/s και 10 Mbit/s
- ❖ Φορμάτ: προοδευτικό καθώς επίσης και συμπλεγμένο βίντεο
- ❖ Ανάλυση: χαρακτηριστικά από sub-QCIF μέχρι και πέρα από HDTV

5.1.4.2. Αποδοτικότητα συμπίεσης

- ❖ Η αποδοτική συμπίεση του βίντεο θα υποστηριχθεί για όλους τους ρυθμούς συμπίεσης που εξετάζονται. Αυτό περιλαμβάνει την συμπαγή κωδικοποίηση του κάθε σήματος με μια ποιότητα ρυθμιζόμενη μεταξύ “αποδεκτού” για πολύ υψηλές αναλογίες συμπίεσης έως και “περίπου μη απωλεστικές” κωδικοποιήσεις.
- ❖ Αποδοτική συμπίεση των σημάτων για τη χαρτογράφηση σύστασης σε 2D δισδιάστατα και 3D τρισδιάστατα πλέγματα.

- ❖ Τυχαία πρόσβαση στο βίντεο για να επιτραπούν λειτουργίες όπως η μικρή διακοπή (pause), γρήγορα μπροστά (fast forward) και γρήγορη αντιστροφή (fast reverse) στο αποθηκευμένο βίντεο.

5.1.4.3. Λειτουργίες βασισμένες στο περιεχόμενο

- ❖ Βασισμένη στο περιεχόμενο, κωδικοποίηση των εικόνων και του βίντεο για να επιτραπεί η χωριστή αποκωδικοποίηση και αναδημιουργία των αυθαίρετα διαμορφωμένων βίντεο αντικειμένων.
- ❖ Τυχαία πρόσβαση στο περιεχόμενο των βίντεο ακολουθιών για να επιτραπούν λειτουργίες όπως η μικρή διακοπή (pause), γρήγορα μπροστά (fast forward) και γρήγορη αντιστροφή (fast reverse) των αποθηκευμένων βίντεο αντικειμένων.
- ❖ Παρατεταμένος χειρισμός του περιεχόμενου των βίντεο ακολουθιών για να επιτραπούν λειτουργίες όπως η παραμόρφωση από το συνθετικό ή φυσικό κείμενο, επικαλύψεις εικόνας και βίντεο στο αναδημιουργημένο βίντεο. Ένα παράδειγμα είναι η χαρτογράφηση του κειμένου μπροστά από ένα βίντεο αντικείμενο όπου το κείμενο κινείται λογικά σε σχέση με το αντικείμενο.

5.1.4.4. Κλιμάκωση των χαρακτηριστικών, των εικόνων και του βίντεο

- ❖ Η κλιμάκωση πολυπλοκότητας των κωδικοποιητών επιτρέπει στους κωδικοποιητές διαφορετικής πολυπλοκότητας να παράγουν έγκυρες και σημαντικές ροές δεδομένων.
- ❖ Η κλιμάκωση της πολυπλοκότητας στον αποκωδικοποιητή επιτρέπει σε ένα δεδομένο χαρακτηριστικό, μια εικόνα ή μια ροή βίντεο να αποκωδικοποιούνται από τους αποκωδικοποιητές διαφορετικών επιπέδων πολυπλοκότητας. Η αναδημιουργημένη ποιότητα, γενικά, συσχετίζεται με την πολυπλοκότητα του αποκωδικοποιητή, που χρησιμοποιήθηκε. Αυτό μπορεί να συνεπάγεται ότι λιγότερο ισχυροί αποκωδικοποιητές, αποκωδικοποιούν μόνο ένα μέρος της ροής μετάδοσης (bitstream).
- ❖ Η χωρική κλιμάκωση επιτρέπει στους αποκωδικοποιητές να αποκωδικοποιήσουν ένα υποσύνολο της ροής μετάδοσης που παράγεται από τον κωδικοποιητή ώστε να αναδημιουργήσουν και να απεικονίσουν τα χαρακτηριστικά, τις εικόνες και το βίντεο σε μειωμένη χωρική ανάλυση. Γίνεται υποστήριξη το περισσότερο έντεκα (11) επιπέδων χωρικής κλιμάκωσης για τα χαρακτηριστικά και τις εικόνες.
- ❖ Η χρονική κλιμάκωση επιτρέπει στους αποκωδικοποιητές να αποκωδικοποιήσουν ένα υποσύνολο της ροής μετάδοσης που παράγεται από τον κωδικοποιητή ώστε να αναδημιουργήσουν και να απεικονίσουν το βίντεο σε μειωμένη χρονική ανάλυση. Ένα μέγιστο τριών επιπέδων χωρικής κλιμάκωσης υποστηρίζεται.
- ❖ Η κλιμάκωση της ποιότητας επιτρέπει σε μια ροή μετάδοσης να αναλυθεί σε διάφορα επίπεδα ροών μετάδοσης διαφορετικού ρυθμού μετάδοσης έτσι ώστε ο συνδυασμός ενός υποσυνόλου των επιπέδων να μπορεί ακόμα και να αποκωδικοποιηθεί σε ένα σήμα με νόημα. Η ροή μετάδοσης μπορεί να ανακτηθεί είτε κατά τη διάρκεια της μετάδοσης είτε κατά την αποκωδικοποίηση. Η

αναδημιουργημένη ποιότητα, συνήθως σχετίζεται με τον αριθμό επιπέδων που χρησιμοποιούνται για αποκωδικοποίηση και αναδημιουργία.

5.1.4.4.5 Μορφή και κωδικοποίηση καναλιών άλφα

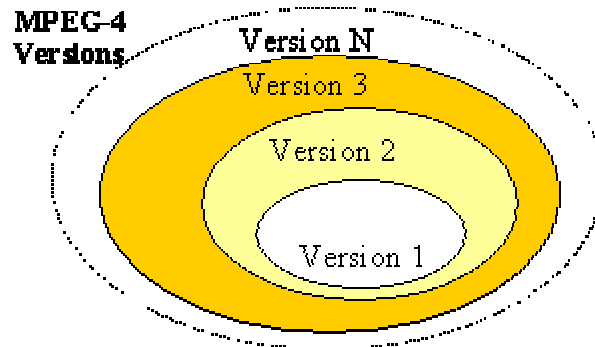
Η κωδικοποίηση μορφής υποστηρίζεται για να βοηθήσει στην περιγραφή και τη σύνθεση συμβατικών εικόνων και βίντεο καθώς επίσης και αυθαίρετα διαμορφωμένων τηλεοπτικών αντικειμένων. Εφαρμογές που επωφελούνται από τους δυαδικούς χάρτες μορφής με εικόνες είναι απεικονίσεις βασισμένες στο περιεχόμενο για βάσεις δεδομένων εικόνων, αλληλεπιδραστικά παιχνίδια και animation. Οι αποδοτικές τεχνικές παρέχονται και επιτρέπουν αποδοτική κωδικοποίηση της δυαδικής μορφής. Ένας δυαδικός χάρτης άλφα καθορίζει εάν ένα εικονοστοιχείο ανήκει ή όχι σε ένα αντικείμενο. Μπορεί να είναι “on” ή “off”.

Κλίμακα του γκρι ή άλφα κωδικοποίηση μορφής

Ένα επίπεδο άλφα καθορίζει τη διαφάνεια/σαφήνεια ενός αντικειμένου, το οποίο δεν είναι απαραίτητος ενιαίο. Οι χάρτες άλφα, πολλαπλών επιπέδων, χρησιμοποιούνται συχνά για να συνδυάσουν διαφορετικά επίπεδα/στρώματα ακολουθιών εικόνας. Άλλες εφαρμογές που επωφελούνται από τους σχετικούς δυαδικούς χάρτες άλφα με τις εικόνες είναι απεικονίσεις βασισμένες στο περιεχόμενο για βάσεις δεδομένων εικόνων, αλληλεπιδραστικά παιχνίδια και animation.. Αποδοτικές τεχνικές παρέχονται, οι οποίες επιτρέπουν την αποδοτική κωδικοποίηση του δυαδικού καθώς επίσης και της κλίμακας του γκρι άλφα επιπέδου. Ένας δυαδικός χάρτης άλφα καθορίζει εάν ένα εικονοστοιχείο ανήκει ή όχι σε ένα αντικείμενο. Μπορεί να είναι “on” ή “off”. Ένας χάρτης κλίμακας του γκρι προσφέρει τη δυνατότητα να καθοριστεί η ακριβής transparency κάθε εικονοστοιχείου.

5.2. Βασικές λειτουργίες του MPEG-4 στην έκδοση 2

Η έκδοση 2 πάγωσε το Δεκέμβρη του 1999. Τα υπάρχοντα εργαλεία και τα υπάρχοντα προφίλ, από την έκδοση 1, δεν αντικατάθηκαν από κάτι νεότερο στην έκδοση 2. Η τεχνολογία θα προστεθεί στο MPEG-4 υπό τη μορφή νέων προφίλ. Το παρακάτω σχήμα απεικονίζει τη σχέση μεταξύ των δύο εκδόσεων. Η έκδοση 2 είναι προς τα πίσω συμβατή επέκταση της έκδοσης 1. Η έκδοση 2 στηρίζεται στην έκδοση 1 του MPEG-4. Το στρώμα συστημάτων της έκδοσης 2 είναι επίσης προς τα πίσω συμβατό με αυτό της έκδοσης 1. Στον τομέα του ήχου και της εικόνας η έκδοση 2 θα προσθέσει νέα προφίλ στην έκδοση 1. Θα προστεθούν περισσότερες λειτουργίες, σε καθορισμένες με σαφήνεια περιοχές. Η ίδια αρχή ισχύει και για τα νέα εργαλεία ώστε να βρεθεί τρόπος να συμπεριληφθούν στα νέα προφίλ του προτύπου. Αυτό σημαίνει ότι το υπάρχον σύστημα θα παραμένει συμβατό, επειδή τα προφίλ δεν θα διαφοροποιούνται κατά την αναδρομική εξέταση.



Σχήμα – Η σχέση των εκδόσεων του MPEG-4

Η έκδοση 2 των mpeg-4 συστημάτων επεκτείνει την έκδοση 1 για να καλύψει τα ζητήματα όπως εκτεταμένο BIFS λειτουργίες, και υποστήριξη της Java (mpeg-j). Η έκδοση 2 επίσης διευκρινίζει ένα σχήμα αρχείων στο κατάστημα Mpeg-4 περιεχόμενο.

5.2.1. Συστήματα

Η έκδοση 2 των MPEG-4 συστημάτων επεκτείνει την έκδοση 1 για να καλύψει τα ζητήματα όπως οι λειτουργίες του εκτεταμένου BIFS, και η υποστήριξη της Java (MPEG-J). Η έκδοση 2 επίσης διευκρινίζει ένα φορμάτ αρχείων για να αποθηκεύσει υλικό MPEG-4 .

5.2.2. Εικόνα

Η έκδοση 2 του MPEG-4 προσθέτει τεχνολογία στην εικόνα στις ακόλουθες περιοχές:

- ❖ Αυξανόμενη ευελιξία στη βασισμένη σε αντικείμενα κλιμακωτή κωδικοποίηση.
- ❖ Βελτιωμένη αποδοτικότητα κωδικοποίησης.
- ❖ Βελτιωμένη σταθερότητα στη χρονική ανάλυση με χαμηλή καθυστέρηση των καταχωρητών.
- ❖ Βελτιωμένη ευρωστία(robustness) λάθους.
- ❖ Κωδικοποίηση πολλαπλών απόψεων: Οι ενδιάμεσες απόψεις ή οι στερεοσκοπικές απόψεις θα υποστηριχθούν βασισμένες στην αποδοτική κωδικοποίηση των πολλαπλών εικόνων ή των ακολουθιών βίντεο. Ένα ιδιαίτερο παράδειγμα είναι η κωδικοποίηση των στερεοσκοπικών εικόνων ή του βίντεο με μείωση του πλεονασμού των πληροφοριών που περιλαμβάνονται μεταξύ των εικόνων των διαφορετικών απόψεων.

5.2.3. Body animation

Η έκδοση 2 προσθέτει το Body Animation στο ήδη υπάρχον Face Animation της έκδοσης 1.

5.2.4. Κωδικοποίηση των τρισδιάστατων (3D) πολυγωνικών πλεγμάτων

Η έκδοση 2 του MPEG-4 παρέχει μια ακολουθία εργαλείων για τα τρισδιάστατα πολυγωνικά πλέγματα. Τα πολυγωνικά πλέγματα χρησιμοποιούνται ευρέως ως γενική απεικόνιση των τρισδιάστατων αντικειμένων. Οι ελλοχεύουσες (underlying) τεχνολογίες συμπίεζουν τη συνδετικότητα, τη γεωμετρία, και τις ιδιότητες όπως η σκίαση των χρωμάτων και των συντεταγμένων των τρισδιάστατων πολυγωνικών πλεγμάτων.

5.2.5. Ήχος

Η ηχητική έκδοση 2 του MPEG-4 είναι μια επέκταση της ηχητικής έκδοσης 1. Προσθέτει νέα εργαλεία και λειτουργίες στο πρότυπο MPEG-4, ενώ κανένα από τα υπάρχοντα εργαλεία της έκδοσης 1 δεν αντικαθίσταται. Οι ακόλουθες πρόσθετες λειτουργίες παρέχονται από την ηχητική έκδοση 2 του MPEG-4:

- ❖ Αυξανόμενη ευρωστία (robustness) λάθους.
- ❖ Κωδικοποίηση ήχου που συνδυάζει την υψηλή ποιότητα με τη χαμηλή καθυστέρηση.
- ❖ Fine Grain κλιμάκωση (ανάλυση κλιμάκωσης μέχρι 1 Kbit/s ανά κανάλι).
- ❖ Παραμετρική κωδικοποίηση ήχου για να επιτραπεί “υγιής” χειρισμός στις χαμηλές ταχύτητες.
- ❖ Συμπύεση σιωπής (CELP), σε περαιτέρω χαμηλούς ρυθμούς μετάδοσης για την κωδικοποίηση της ομιλίας.
- ❖ Ελαστική στα λάθη, παραμετρική κωδικοποίηση φωνής.
- ❖ Περιβαλλοντικό spatialization - η δυνατότητα να αναδημιουργηθεί το περιβάλλον του ήχου με τη χρήση ψυχο-ακουστικών ή/και φυσικών τεχνικών διαμόρφωσης.
- ❖ Ένα οπίσθιο κανάλι που είναι χρήσιμο για να ρυθμίσει τη κωδικοποίηση ή την κλιμακωτή αναπαραγωγή σε πραγματικό χρόνο.
- ❖ Χαμηλά overhead, συγκεκριμένος μηχανισμός μετάδοσης για τον MPEG-4 ήχο.

5.2.6. DMIF

Τα σημαντικότερα χαρακτηριστικά γνωρίσματα που εισάγονται από την έκδοση 2 του DMIF καλύπτουν (την περιορισμένη) υποστήριξη για δίκτυα κινητής τηλεφωνίας και εποπτεία της QoS. Επίσης, εισήχθησαν μερικές δευτερεύουσες προσθήκες.

5.2.6.1. Υποστήριξη δικτύων κινητής τηλεφωνίας

Από κοινού με το ITU-T, η προδιαγραφή H.245 έχει επεκταθεί (H.245v6) για να συμπεριλάβει και την υποστήριξη των MPEG-4 συστημάτων. Η προδιαγραφή DMIF παρέχει το κατάλληλο πέρασμα και χαρτογράφηση στα H.245 σήματα. Τα κινητά τερματικά μπορούν τώρα να χρησιμοποιήσουν τα χαρακτηριστικά γνωρίσματα συστημάτων MPEG-4 όπως ροές BIFS και OD, αν και με κάποιο περιορισμό (η παρουσίαση MPEG-4 είναι μεμονωμένα επιλεγμένα από τον στόχο-peer)

5.2.6.2. Εποπτεία της QoS

Η έκδοση 2 του DMIF εισάγει την έννοια του ελέγχου της ποιότητας της υπηρεσίας που παραδίδεται πραγματικά από το δίκτυο. Η διεπαφή της DMIF-εφαρμογής έχει επεκταθεί αναλόγως. Το πρότυπο επιτρέπει τρεις διαφορετικούς τρόπους ελέγχου της QoS : συνεχής έλεγχος, συγκεκριμένα ερωτήματα, και ανακοίνωση “παραβίασης” της QoS.

| Parameter | Description |
|-----------------------------------|--|
| Throughput | The maximum number of bytes, contained in Service Data Units (SDUs) that may be successfully transferred in unit time by the service provider over the connection on a sustained basis. |
| Transit delay | The time delay between the issuing of a <i>data.request</i> and the corresponding <i>data.indication</i> . The parameter is usually specified as a pair of values, a statistical average and a maximum. Those data transfers where a receiving service user exercises flow control and excluded. The computations are all based on SDUs of a fixed size. |
| Residual error rate | The probability that an SDU is transferred with error or that is lost, or that a duplicated copy is transferred. |
| Establishment delay | The delay between the issuing <i>connect.request</i> and the corresponding <i>connect.confirm</i> . |
| Establishment failure probability | The probability that a requested connection is not established within the specified maximum acceptable establishment delay as a consequence of actions that are solely attributable to the service provider. |
| Transfer failure probability | The probability that the observed performance with respect to transit delay, residual error rate or throughput will be worse than the specified level of performance. The failure probability is, as such, specified for each measure of performance of data transfer discussed above. |
| Resilience | The probability that a service provider will, on its own, release the connection, or reset it, within a specified interval of time. |
| Release delay | The maximum delay between the issuing of a <i>disconnect.request</i> primitive by the service user and a corresponding <i>disconnect.indication</i> primitive issued by the service provider. |
| Release failure probability | The probability that the service provider is unable to release the connection within a specified maximum release delay. |

Πίνακας – Οι παράμετροι της QoS βασισμένες στην απόδοση (OSI)

| Parameter | Description |
|-------------------|--|
| Protection | The extend to which a service provider attempts to prevent unauthorised monitoring or manipulation of a user data. The level of protection is specified qualitatively by selecting either (i) no protection, (ii) protection against passive monitoring, (iii) protection against modification, addition or deletion, or (iv) a combination of (i) and (ii). |
| Priority | High-priority connections are serviced before lower ones. Lower-priority connection packets will be dropped before high-priority packets should the network become congested. |
| Cost determinants | A parameter to define the maximum acceptable cost for a network connection. It may be stated in relative or absolute terms. Final actions on this parameter are to the specific network providers. |

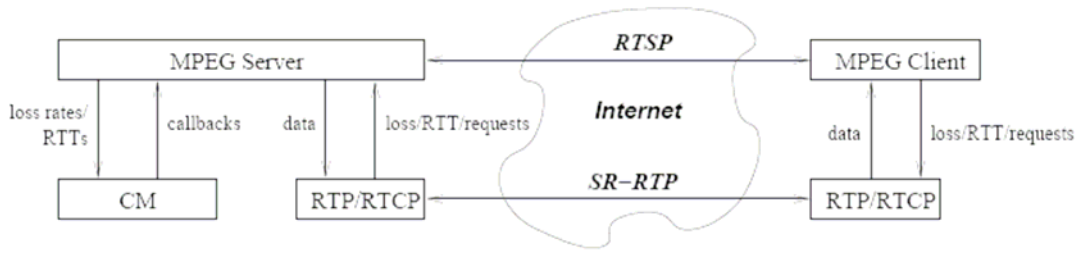
Πίνακας - Οι παράμετροι της QoS μη - βασισμένες στην απόδοση (OSI)

5.2.6.3. Εντολές χρηστών με Ack

Το πρότυπο DMIF επιτρέπει στις peer εφαρμογές να ανταλλάσουν μηνύματα οποιουδήποτε είδους (συμπεριλαμβανόμενων και μηνυμάτων ελέγχου των ροών). Η έκδοση 2 του DMIF προσθέτει την υποστήριξη για τα μηνύματα αναγνώρισης (ACK) στην έκδοση 1.

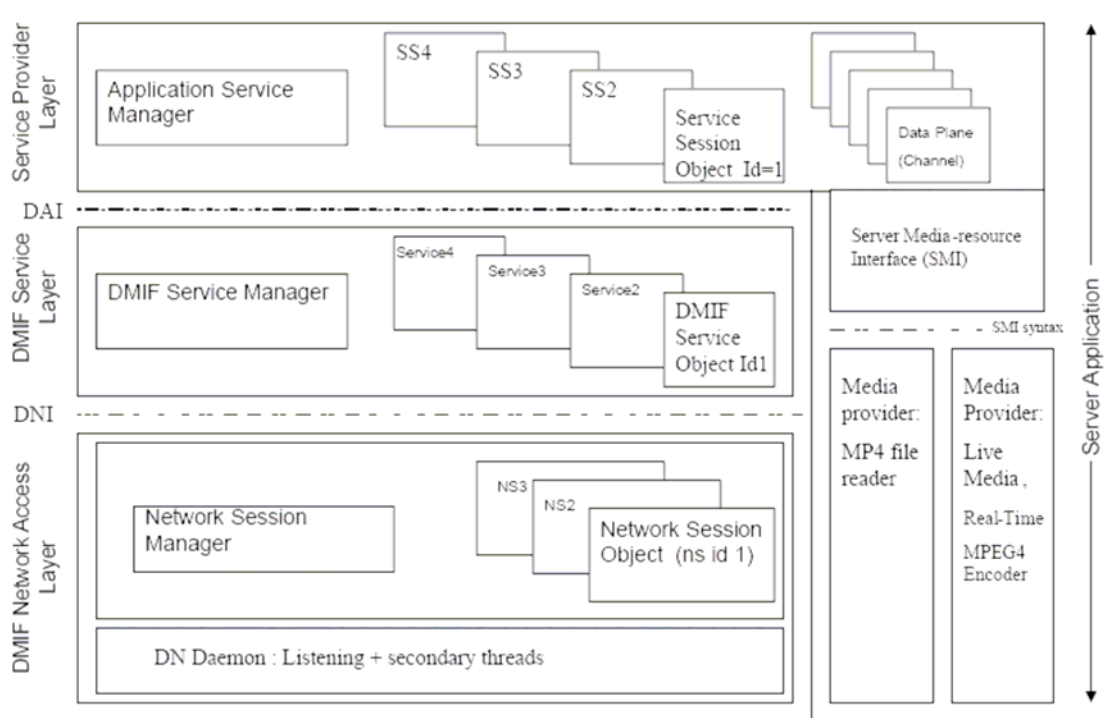
5.2.6.4. Διαχείριση των πληροφοριών στρώματος Sync του MPEG-4

Η έκδοση 2 ενισχύει το πρότυπο DMIF ώστε να επιτρέψει στις εφαρμογές να ανταλλάξουν οριζόμενα από την εφαρμογή δεδομένα με το στρώμα DMIF. Αυτή η προσθήκη εισήχθη για να επιτρέψει, μέσα στο πρότυπο, την ανταλλαγή Μονάδων Δεδομένων του πρωτοκόλλου του στρώματος Sync ως συνδυασμό καθαρών δεδομένων media (PDU, pure media data) και πληροφοριών του λογικού στρώματος Sync. Το πρότυπο αναγνωρίζει ότι ανάμεσα στις υπάρχουσες στοίβες μεταφοράς υπάρχουν χαρακτηριστικά γνωρίσματα τα οποία επικαλύπτονται με το στρώμα συστημάτων Sync του MPEG-4. Αυτή είναι η περίπτωση του RTP και του MPEG-2 PES (Packetized Elementary Streams) καθώς επίσης και MP4 άτομα (atoms) στο φορμάτ του αρχείου. Σε όλες αυτές τις περιπτώσεις μια DMIF εφαρμογή χαρτογραφεί (map) τις πληροφορίες στρώματος Sync που εξάγονται από εκείνες τις δομές, ομοίως με ένα αληθινό SL-PDU, σε μια ομοιόμορφη λογική απεικόνιση της κεφαλίδας του στρώματος Sync. Κατά συνέπεια, οι κατάλληλες παράμετροι έχουν εισαχθεί στο DAI, προσέχοντας, όπως συνήθως, η σημασία τους να είναι ανεξάρτητη της στοίβας μεταφοράς καθώς και της εφαρμογής.



Σχήμα – Η αρχιτεκτονική του συστήματος: Η ανάδραση στέλνεται στην εφαρμογή ροής δεδομένων μέσω RTCP Feedback, το οποίο χρησιμοποιείται για να ρυθμίσει όπως πρέπει το ρυθμό μεταφοράς.

Η επιλεκτική αξιοπιστία καθίσταται εφικτή με τη χρήση του SR-RTP σε συνδυασμό με το RTCP κανάλι ανάδρασης. Με το SR-RTP, τις “προς τα πίσω” συμβατές επεκτάσεις με το RTP, ο πελάτης μπορεί επιλεκτικά να ζητήσει να μεταδοθούν σε επανάληψη ορισμένα πακέτα, που έχουν χαθεί.



Σχήμα - MPEG-4 (streaming server application)

5.2.6.5. Η σύνταξη του DAI σε γλώσσα C

Η έκδοση 2 του DMIF περιλαμβάνει ένα παράρτημα πληροφοριών που δίνει τη σύνταξη της διεπαφής της DMIF εφαρμογής είτε σε γλώσσα C είτε σε γλώσσα C++, ως μία συνιστώμενη σύνταξη API.

5.3. Επεκτάσεις του MPEG-4 πέρα από την έκδοση 2

Το MPEG ασχολείται αυτήν την περίοδο με διάφορες επεκτάσεις της έκδοσης 2, στον οπτικό τομέα και στα συστήματα . Δεν καταβάλλεται καμία προσπάθεια για την επέκταση του MPEG-4 DMIF ή του ήχου πέρα από την έκδοση 2.

5.3.1. Εικόνα

Στον τομέα της εικόνας, οι ακόλουθες τεχνολογίες είναι στο στάδιο της προσθήκης:

- ❖ Η κλιμάκωση “Fine Grain” είναι στη φάση της ψήφισης, με προτεινόμενα τα προφίλ ροής βίντεο (“Advanced Simple” και “Fine Grain Scalability”). Η κλιμάκωση “Fine Grain” είναι ένα εργαλείο που επιτρέπει μικρά ποιοτικά βήματα με την προσθήκη ή τη διαγραφή στρωμάτων πρόσθετης πληροφορίας. Είναι χρήσιμο, σε έναν αριθμό από περιβάλλοντα, ειδικότερα για λόγους ροής αλλά και για δυναμική (“στατιστική”) πολύπλεξη του προ-κωδικοποιημένου περιεχομένου στα περιβάλλοντα εκπομπής (broadcasting).
- ❖ Εργαλεία για τη χρήση του MPEG-4 στο στούντιο. Για αυτά τα εργαλεία, έχει μεριμνηθεί ώστε να διατηρηθεί κάποια μορφή συμβατότητας με τα προφίλ MPEG-2. Αυτήν την περίοδο, το απλό στούντιο προφίλ είναι σε μια φάση ψήφισης, το οποίο είναι ένα προφίλ μόνο με I-πλαίσια κωδικοποιούμενο σε πολύ υψηλούς ρυθμούς μετάδοσης (πολλές εκατοντάδες Mbit/s) το οποίο υιοθετεί την κωδικοποίηση μορφής (shape coding). Αναμένεται η προσθήκη ενός προφίλ πυρήνων στούντιο (με I και P πλαίσια).
- ❖ Ο ψηφιακός κινηματογράφος είναι ακόμα υπό μελέτη. Αυτή η εφαρμογή θα απαιτήσει την πραγματική χωρίς απώλειες κωδικοποίηση, και όχι μόνο την κωδικοποίηση εικόνας χωρίς απώλειες που το MPEG-4 έχει παράσχει μέχρι τώρα.

5.3.2. Συστήματα

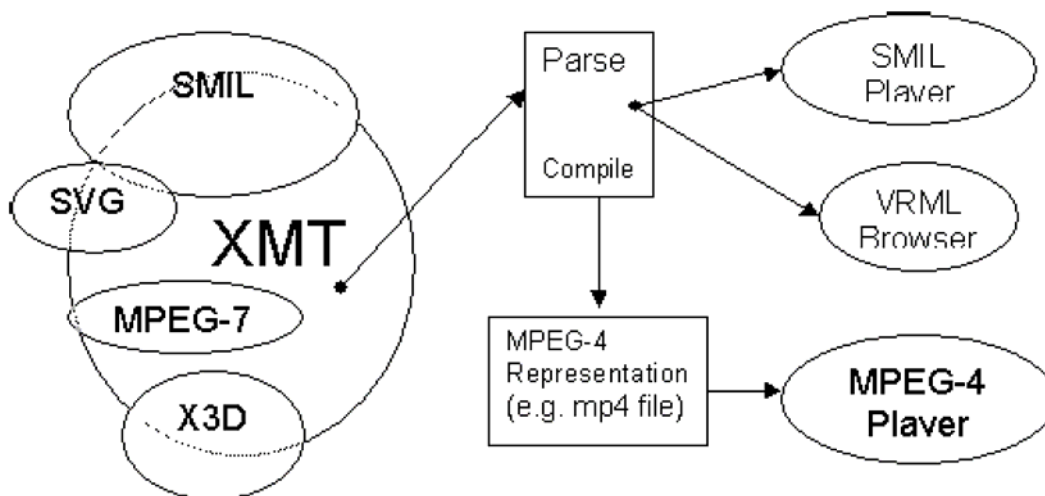
5.3.2.1 Εξελιγμένα BIFS

Παρέχουν νέους κόμβους που χρησιμοποιούνται στη γραφική αναπαράσταση της σκηνής για τον έλεγχο των διαθέσιμων media και τη διαχείριση των media, όπως η αποστολή εντολών σε έναν κεντρικό υπολογιστή (server), ο προηγμένος έλεγχος της αναπαραγωγής των media, και το αποκαλούμενο EXTERNPROTO, ένας κόμβος που παρέχει περαιτέρω συμβατότητα με την VRML, και που επιτρέπει τη σύνταξη μακροεντολών που καθορίζουν τη συμπεριφορά των αντικειμένων. Επίσης, η προηγμένη συμπίεση των BIFS δεδομένων είναι “καλυμμένη” (covered), και ιδιαίτερα καλή συμπίεση για πλέγματα και για πίνακες δεδομένων.

5.3.2.2 Φορμάτ Κειμένου

Το επιδεκτικό επέκτασης φορμάτ κειμένου του MPEG-4 (XMT) είναι ένα πλαίσιο για την περιγραφή της απεικόνισης της σκηνής του MPEG-4 που χρησιμοποιεί μια σύνταξη κειμένου. Το XMT επιτρέπει στους συντάκτες περιεχομένου να

ανταλλάξουν το περιεχόμενό τους με άλλους συντάκτες, εργαλεία ή φορείς παροχής υπηρεσιών, και διευκολύνει τη διαλειτουργικότητα και με τα δύο, το επιδεικτικό επέκτασης τρισδιάστατο 3D (X3D), που αναπτύσσεται από την κοινοπραξία Web3D, και τη συγχρονισμένη Γλώσσα Ολοκλήρωσης Πολυμέσων (SMIL) από τη W3C κοινοπραξία. Το φορμάτ XMT μπορεί να ανταλλαχθεί μεταξύ των φορέων SMIL, των φορέων VRML, και των MPEG-4 φορέων. Το φορμάτ μπορεί να αναλυθεί και να αναπαρασταθεί άμεσα από έναν W3C φορέα SMIL, που προεπεξεργάζεται Web3D X3D και να αναπαραχθεί από έναν φορέα VRML, ή να συνταχθεί σε μια απεικόνιση MPEG-4 όπως το .mp4, το οποίο μπορεί έπειτα να παιχτεί από έναν φορέα MPEG-4. Δείτε παρακάτω το διάγραμμα περιγραφής της διαλειτουργικότητας του XMT.



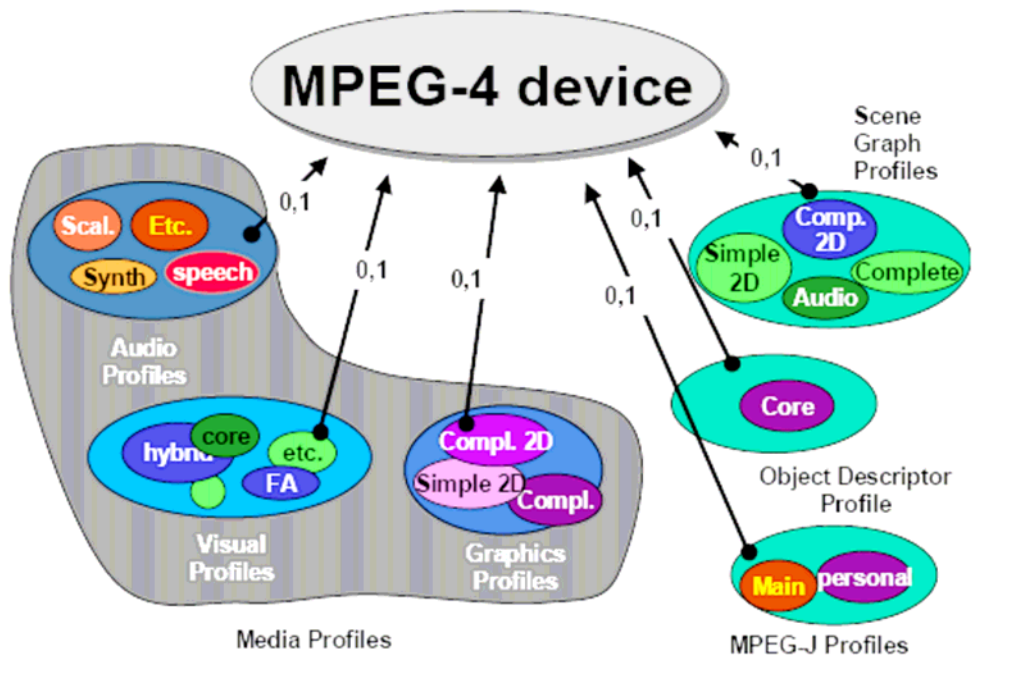
Σχήμα - Η σχέση του MPEG-4 XMT με τις υπόλοιπες προδιαγραφές

5.3.3. Προφίλ

Το MPEG-4 παρέχει ένα μεγάλο και πλούσιο σύνολο εργαλείων για την κωδικοποίηση των οπτικοακουστικών αντικειμένων. Προκειμένου να επιτραπούν οι αποτελεσματικές εφαρμογές των προτύπων έχουν προσδιοριστεί, υποσύνολα των σετ εργαλείων των MPEG-4 συστημάτων, εικόνων, και ήχων, τα οποία μπορούν να χρησιμοποιηθούν για τις συγκεκριμένες εφαρμογές. Αυτά τα υποσύνολα, αποκαλούμενα “προφίλ”, περιορίζουν το σύνολο εργαλείων που ένας αποκωδικοποιητής πρέπει να εφαρμόσει. Για κάθε ένα από αυτά τα προφίλ, ένα ή περισσότερα επίπεδα έχουν τεθεί, περιορίζοντας την υπολογιστική πολυπλοκότητα. Η προσέγγιση είναι παρεμφερής με το MPEG-2, όπου ο πιο γνωστός συνδυασμός προφίλ/επίπεδου είναι “Βασικό προφίλ σε Βασικό Επίπεδο” (Main Profile @ Main Level). Ένας συνδυασμός προφίλ σε επίπεδο(Profile@Level) επιτρέπει:

- ❖ Ένας σχεδιαστής codec να εφαρμόσει μόνο το υποσύνολο του προτύπου που χρειάζεται, διατηρώντας τις αλληλεπιδράσεις με άλλες συσκευές MPEG-4, που χρησιμοποιούνται στον ίδιο συνδυασμό, και
- ❖ Να ελέγξει εάν οι συσκευές MPEG-4 συμμορφώνονται με τα πρότυπα (“conformance testing”). Τα προφίλ υφίστανται για διάφορους τύπους πολυμεσικού περιεχομένου (ακουστικό, οπτικό, και γραφική αναπαράσταση) και για περιγραφές της σκηνής. Το MPEG δεν ορίζει ούτε προτείνει τους

συνδυασμούς αυτών των προφίλ, αλλά έχει σημειωθεί ότι οι καλές αντιστοιχίες υπάρχουν μεταξύ διαφορετικών περιοχών.



Σχήμα - προφίλ και επίπεδα

5.3.3.1. Οπτικά Προφίλ

Το οπτικό μέρος του προτύπου παρέχει τα προφίλ για την κωδικοποίηση φυσικού, συνθετικού, και συνθετικού/φυσικού υβριδικού οπτικού περιεχομένου. Υπάρχουν πέντε προφίλ για το φυσικό βίντεο περιεχόμενο:

- ❖ Το απλό οπτικό προφίλ παρέχει την αποδοτική, ελαστική στα λάθη ορθογώνια κωδικοποίηση βίντεο αντικειμένων, κατάλληλη για εφαρμογές κινητών δικτύων (mobile networks), όπως τα PCS και τα IMT2000.
- ❖ Το απλό οπτικό προφίλ κλιμάκωσης προσθέτει την υποστήριξη για την κωδικοποίηση των αντικειμένων με χωρική ή χρονική κλιμάκωση στο απλό οπτικό προφίλ. Αυτό είναι χρήσιμο για τις εφαρμογές που παρέχουν τις υπηρεσίες τους σε περισσότερα από ένα επίπεδα ποιότητας λόγω των περιορισμών των ρυθμών μετάδοσης ή των αποκωδικοποιητών, όπως η χρήση του διαδικτύου και η αποκωδικοποίηση λογισμικού.
- ❖ Το οπτικό προφίλ πυρήνων προσθέτει την υποστήριξη για την κωδικοποίηση αυθαίρετα διαμορφωμένων και χρονικά κλιμακωτών αντικειμένων στο απλό οπτικό προφίλ. Είναι χρήσιμο για τις εφαρμογές όπως αυτές που παρέχουν σχετικά απλή διαδραστικότητα περιεχομένου (εφαρμογές πολυμέσων Διαδικτύου).
- ❖ Το κύριο οπτικό προφίλ προσθέτει την υποστήριξη για την κωδικοποίηση συμπλεγμένων (interlaced) ημιδιάφανων (semi-transparent), και sprite αντικειμένων στο οπτικό προφίλ πυρήνων. Είναι χρήσιμο για διαδραστική και ψυχαγωγική ποιότητα εκπομπής καθώς και εφαρμογές DVD.

- ❖ Το οπτικό προφίλ N-bit προσθέτει την υποστήριξη κωδικοποίησης βίντεο αντικειμένων, που το “βάθος” των εικονοστοιχείων τους κυμαίνεται από 4 έως 12 bit στο οπτικό προφίλ πυρήνων. Είναι κατάλληλο για χρήση σε εφαρμογές “επιτήρησης” (surveillance).

Τα προφίλ για το συνθετικό και συνθετικό/φυσικό υβριδικό οπτικό περιεχόμενο είναι:

- Το απλό facial (του προσώπου) οπτικό προφίλ animation παρέχει απλά μέσα για να animate (ζωντανεψει) ένα πρόσωπο-μοντέλο, κατάλληλο για τις εφαρμογές όπως η ακουστική/βίντεο απεικόνιση για τους κουφούς.
- Το οπτικό προφίλ κλιμακωτής σύστασης παρέχει τη χωρική κλιμακωτή κωδικοποίηση της εικόνας (σύσταση) χρήσιμη για τις εφαρμογές που χρειάζονται τα πολλαπλά επίπεδα κλιμάκωσης, όπως η χαρτογράφηση της ιδιοσυστασίας (texture) επάνω στα αντικείμενα σε παιχνίδια, και τις υψηλής ευκρίνειας ψηφιακές φωτογραφικές μηχανές.
- Το βασικό animated 2D οπτικό προφίλ σύστασης παρέχει τη χωρική κλιμάκωση, την κλιμάκωση SNR, και το animation βασισμένο σε πλέγματα για τα αντικείμενα της εικόνας (συστάσεις) και επίσης απλό animation προσώπου.
- Το υβριδικό οπτικό προφίλ συνδυάζει τη δυνατότητα να αποκωδικοποιηθούν τα αυθαίρετα-διαμορφωμένα και χρονικά κλιμακώσιμα φυσικά βίντεο αντικείμενα (όπως στο οπτικό προφίλ πυρήνων) με τη δυνατότητα να αποκωδικοποιηθούν αρκετά τα συνθετικά και υβριδικά αντικείμενα, συμπεριλαμβανομένων και του απλού προσώπου και της animated ακόμα εικόνας. Είναι κατάλληλο για τις διάφορες εφαρμογές πολυμέσων πλούσιου-περιεχομένου.

Η έκδοση 2 προσθέτει τα ακόλουθα προφίλ για το φυσικό βίντεο:

- Το προηγμένο σε πραγματικό χρόνο απλό προφίλ (Advanced Real-Time Simple (ARTS)) παρέχει προηγμένες τεχνικές κωδικοποίησης, ελαστικές σε λάθη, ορθογώνιων βίντεο αντικειμένων που χρησιμοποιούν ένα οπίσθιο κανάλι και βελτιωμένης σταθερότητας της χρονικής ανάλυσης με τη χαμηλή καθυστέρηση των καταχωρητών. Ενδείκνυται για κωδικοποιήσεις εφαρμογών πραγματικού χρόνου όπως η βίντεο-κλήση, η τηλε-διάσκεψη και η μακρινή παρατήρηση.
- Το κλιμακωτό προφίλ πυρήνων προσθέτει την υποστήριξη για την κωδικοποίηση χρονικά και χωρικά κλιμακωτών, αυθαίρετα διαμορφωμένων αντικειμένων στο προφίλ πυρήνων. Η κύρια λειτουργία αυτού του προφίλ είναι αντικειμενικά βασισμένη στο SNR και στη χωρική/χρονική κλιμάκωση για τις περιοχές ή τα αντικείμενα ενδιαφέροντος. Είναι χρήσιμο σε εφαρμογές όπως το Διαδίκτυο, η κινητή τηλεφωνία και η ραδιοφωνική μετάδοση.
- Το προφίλ προηγμένης κωδικοποίησης αποδοτικότητας (Advanced Coding Efficiency) βελτιώνει την κωδικοποίηση αποδοτικότητας και για τα ορθογώνια και για τα αυθαίρετα διαμορφωμένα αντικείμενα. Είναι κατάλληλο για εφαρμογές όπως κινητή υποδοχή εκπομπών, απόκτηση των ακολουθιών εικόνας (camcorders) και άλλες εφαρμογές όπου απαιτείται η υψηλή κωδικοποίηση αποδοτικότητας και το μικρό ίχνος δεν είναι η πρωταρχική ανησυχία.

Τα προφίλ της έκδοσης 2 για το συνθετικό και συνθετικό/φυσικό υβριδικό οπτικό περιεχόμενο είναι:

- Το προηγμένο προφίλ κλιμακωτής σύστασης υποστηρίζει την αποκωδικοποίηση της αυθαίρετα διαμορφωμένης σύστασης και ακόμα εικόνες που περιλαμβάνουν κλιμακωτή κωδικοποίηση μορφής, επικεράμωσης κυματακιών(wavelet tiling) και της λάθος-ανθεκτικότητας. Είναι χρήσιμο για εφαρμογές όπου απαιτείται γρήγορη τυχαία προσπέλαση καθώς και τα πολλαπλά επίπεδα κλιμάκωσης και η αυθαίρετα-διαμορφωμένη κωδικοποίηση των αντικειμένων.
- Το εξελιγμένο προφίλ πυρήνων συνδυάζει τη δυνατότητα να αποκωδικοποιηθούν τα αυθαίρετα-διαμορφωμένα τηλεοπτικά αντικείμενα (όπως στο οπτικό προφίλ πυρήνων) με τη δυνατότητα να αποκωδικοποιηθεί η αυθαίρετα-διαμορφωμένη κλιμακωτή εικόνα αντικειμένων (όπως στο προηγμένο προηγμένο προφίλ κλιμακωτής σύστασης). Είναι κατάλληλο για τους διάφορες εφαρμογές πολυμέσων πλούσιες σε περιεχόμενο όπως τα διαδραστικά πολυμέσα που μεταδίδονται μέσω Διαδικτύου.
- Το απλό προφίλ animation προσώπου και σώματος είναι υπερσύνολο του απλού προφίλ animation προσώπου, προσθέτωντας, προφανώς και το animation σώματος.

Στις επόμενες εκδόσεις, τα ακόλουθα σχεδιαγράμματα προστέθηκαν :

- Το προηγμένο απλό προφίλ μοιάζει σαν απλό δεδομένου ότι έχει μόνο τα ορθογώνια αντικείμενα, αλλά έχει μερικά πρόσθετα εργαλεία που το καθιστούν αποδοτικότερο: Β-πλαίσια, αποζημίωση κινήσεων $\frac{1}{4}$ pel και σφαιρική αποζημίωση κινήσεων.
- Το απλό στούντιο προφίλ είναι ένα προφίλ με πολύ υψηλής ποιότητας εφαρμογές για τη χρήση στο στούντιο. Έχει μόνο τα πλαίσια I, αλλά υποστηρίζει την αυθαίρετη μορφή και στην πραγματικότητα πολλαπλά κανάλια άλφα. Οι ρυθμοί μετάδοσης ανεβαίνουν σε σχεδόν 2 Gigabit ανά δευτερόλεπτο.
- Το προφίλ στούντιο πυρήνων προσθέτει τα πλαίσια P στο απλό στούντιο προφίλ, που το καθιστούν αποδοτικότερο αλλά και απαίτηση των πύθ σύνθετων εφαρμογών.

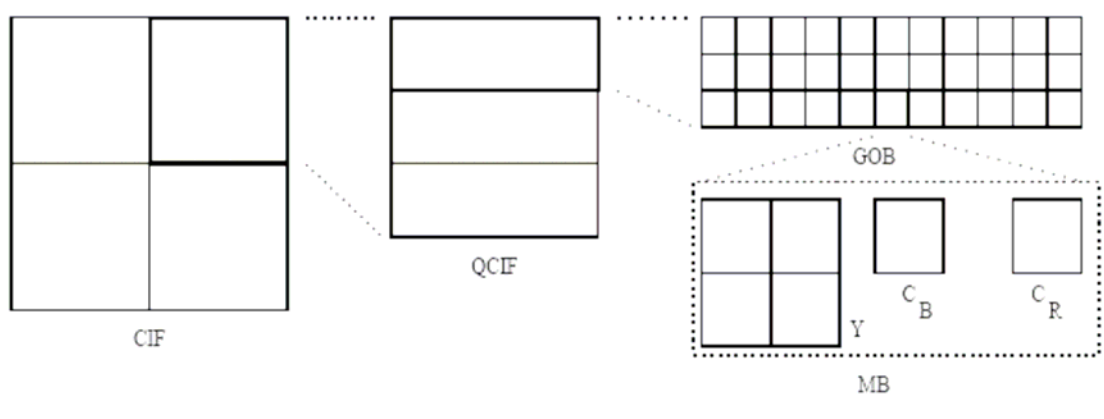
5.3.4. H.261 και H.263

Η CCITT ανέπτυξε τα τηλεοπτικά πρότυπα συμπίεσης H.261 που σχεδιάζονται για τις τηλεοπτικές επικοινωνίες πάνω από δίκτυα ISDN. Τα H.261 μπορούν να χειριστούν $\pi * 64$ Kbps (όπου $\pi = 1, 2, 3, \dots, 30$) τηλεοπτικές ροές και αυτό είναι ίσο με τα πιθανά εύρη ζώνης στην ISDN. Το πρότυπο H.261 υποστηρίζει τα δύο ακόλουθα βίντεο φορμάτ:

1. Κοινό ενδιάμεσο φορμάτ (CommonIntermediateFormat CIF): Αυτό το φορμάτ έχει μία ανάλυση 360 X 288 εικονοστοιχεία για τη συνιστώσα της φωτεινότητας (Luminance, Y) της ροής του βίντεο και μια ανάλυση 180 X 144 εικονοστοιχεία για τα δύο μέρη της χρωματικότητας (Chrominance, $C_R C_B$) της ροής του βίντεο.

2. Τέταρτο - CIF (QuarterCIF): Αυτό το φορμάτ περιέχει ένα τέταρτο της

πληροφορίας μιάς CIF βίντεο ροής. Αυτό σημαίνει ότι η ανάλυση της φωτεινότητας είναι 180 X 144 εικονοστοιχεία και η ανάλυση των δύο συνιστωσών της χρωματικότητας είναι 90 X 72 εικονοστοιχεία.



Σχήμα – Η σύνθεση ενός H.261 CIF.

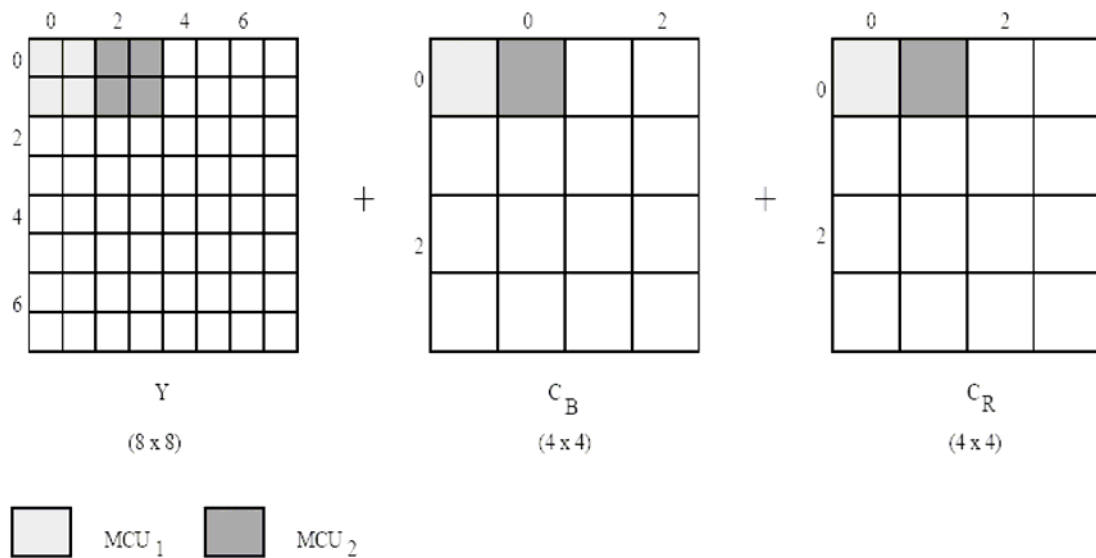
Ο μέγιστος ρυθμός πλαισίων για μία ροή βίντεο H.261 είναι περίπου 30 πλαίσια ανά δευτερόλεπτο. Το CIF και το QCIF αποτελούνται από εικόνες για κάθε πλαίσιο, και μέσα σε κάθε εικόνα της ομάδας των blocks (Group of blocks ή GOBs). Ένα QCIF έχει 3 GOBs, ενώ ένα CIF έχει 12 GOBs. Κάθε GOB αποτελείται από 3 X 11 macro – blocks (MB). Ένα macro – block αποτελείται από 4 8x8 blocks φωτεινότητας και 2 8x8 blocks χρωματικότητας (C_R C_B). Ένα macro – block μπορεί να συγκριθεί με ένα MCU στο JPEG.

Ο κωδικοποιητής H.261 μπορεί να λειτουργήσει σε δύο καταστάσεις. Στη λειτουργία δια – πλαισίων, όπου DCT – μετασχηματίζεται κάθε 8x8 block, κβαντοποιείται γραμμικά, και αποστέλεται στον πολυπλέχτη του βίντεο. Στη λειτουργία των ενδο – πλαισίων, κάθε 8x8 block DCT – μετασχηματίζεται επίσης και κβαντοποιείται γραμμικά, αλλά το αποτέλεσμα αποστέλεται αρχικά σε έναν compensator - κίνησης προτού αποσταλεί στον πολυπλέχτη βίντεο. Ο compensator - κίνησης χρησιμοποιείται για να συγκρίνει το macro – block του τρέχοντος πλαισίου με τα blocks δεδομένων του προηγούμενου πλαισίου. Εάν η διαφορά είναι κάτω από ένα προκαθορισμένο κατώτατο όριο (κατώφλι - threshold), καθόλου δεδομένα δεν αποστέλονται για αυτό το block. Διαφορετικά, η διαφορά DCT – μετασχηματίζεται, και στη συνέχεια κβαντοποιείται γραμμικά.

Το τελικό βήμα κωδικοποίησης είναι ο βίντεο πολυπλέχτης, που χρησιμοποιεί έναν κωδικοποιητή μεταβλητού μήκους λέξης για να μειώσει ακόμα περισσότερο τη ροή των bit. Μετά από τον βίντεο πολυπλέχτη, το αποτέλεσμα εισέρχεται σε έναν καταχωρητή μετάδοσης, ο οποίος ελέγχει το γραμμικό κβαντιστή προκειμένου να ρυθμίσει την εξερχόμενη ροή bit.

Το H.261 είναι παρόμοιο με το πρότυπο MPEG όσον αφορά την DCT - κωδικοποίηση και την κβαντοποίηση. Κατά τη διάρκεια τυποποίησης του MPEG-1, αυτό έγινε επίτηδες ώστε να απλοποιήσει τις εφαρμογές που κωδικοποιούν ή αποκωδικοποιούν και H.261 και MPEG. Η κύρια διαφορά μεταξύ H.261 και MPEG-1 είναι ότι τα διανύσματα κίνησης στο H.261 είναι περιορισμένα κατά 15 εικονοστοιχεία μακρύτερα από την αρχική τους θέση στην εικόνα. Επιπλέον, καμία μελλοντική – πρόβλεψη δεν πραγματοποιήθηκε για το H.261, που σημαίνει ότι το H.261 δεν έχει κανένα αντίτιμο των πλαισίων – B όπως το MPEG. Το σύστημα βίντεο διάσκεψης INRIA (IVS) είναι μια εφαρμογή ενός εργαλείου διάσκεψης βίντεο,

που χρησιμοποιεί το H.261 για συμπίεση. Το Vic είναι επίσης ένα εργαλείο διάσκεψης βίντεο, που υποστηρίζει το H.261.



Σχήμα - Component interleaving(παρεμβολή λευκών σελίδων) με χρήση MCU

H.263

Το προσχέδιο ITU είναι μια βελτίωση πάνω στο H.261. Το H.263 αναπτύχθηκε για τις επικοινωνίες με χαμηλά εύρη ζώνης, πάνω από παλιά τηλεφωνικά συστήματα (POTS) και ιδιαίτερα modems στα 28.8 Kbps. Σε σύγκριση με το H.261, αυξήθηκε ο αριθμός των διαθέσιμων φορμάτ για τις εικόνες, βελτιώθηκε ο αλγόριθμος αποζημίωσης – κίνησης (motion – compensation), χρησιμοποιείται μια καλύτερη κωδικοποίηση, αυτή της εντροπίας και εισάγεται ένα νέο πλαίσιο, που επιτρέπει μια απλή μορφή πρόβλεψης προς τα μπροστά (forward prediction).

Τρία νέα βίντεο φορμάτ προστέθηκαν στο H.263: ένα υπο – QCIF φορμάτ προστέθηκε (128 X 96), ένα 4CIF φορμάτ προστέθηκε (704 X 786), και ένα 16CIF φορμάτ προστέθηκε (1408 X 1152). Όπως και στο H.261, ο αριθμός των εικονοστοιχείων της χρωματικότητας είναι πάντα ο μισός από τον αριθμό των συνολικών εικονοστοιχείων της φωτεινότητας. Αυτό σημαίνει ότι για 2x2 εικονοστοιχεία φωτεινότητας, χρησιμοποιείται ένα C_B και ένα C_R εικονοστοιχείο.

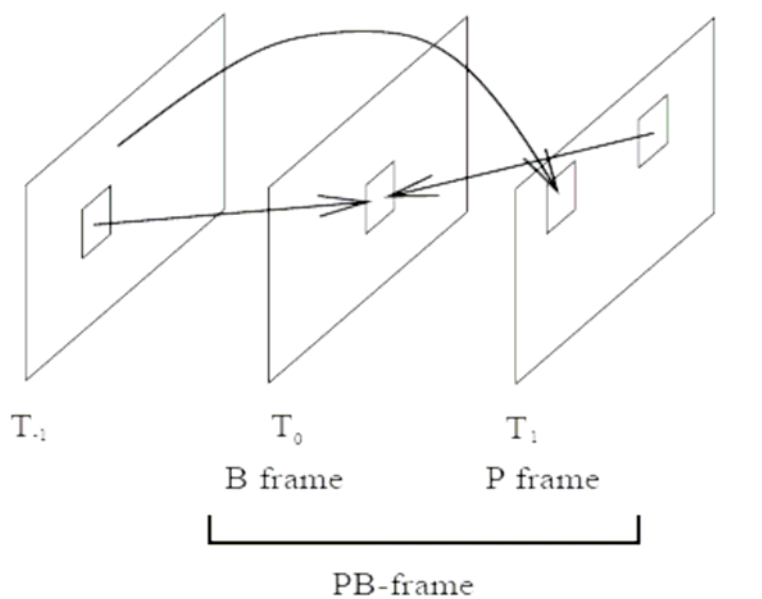
Το H.263 υποστηρίζει επίσης τον απεριόριστο τρόπο κίνησης των διανυσμάτων. Στον προεπιλεγμένο τρόπο κίνησης των διανυσμάτων (που είναι περιορισμένος), το block στο οποίο αναφερόμαστε πρέπει να είναι πλήρως μέσα στην εικόνα. Στον μη καταγεγραμμένο διανυσματικό τρόπο κίνησης, ένας αυθαίρετος αριθμός εικονοστοιχείων μπορεί να εκτός της εικόνας. Για κάθε αναφορά σε αυτά τα εικονοστοιχεία, χρησιμοποιείται το πιό κοντινό εικονοστοιχείο αντ' αυτών.

Ο προηγμένος τρόπος πρόβλεψης είναι επίσης μια καινοτομία στο H.263. Αντί του ενός διανύσματος κίνησης των 16x16 macro – blocks, τέσσερα διανύσματα κίνησης 8x8 blocks χρησιμοποιούνται για την πρόβλεψη. Αν και αυτή η κωδικοποίηση χρησιμοποιεί περισσότερα bits απ' ότι θα γινόταν με ένα διάνυσμα κίνησης, η ποιότητα της πρόβλεψης βελτιώνεται σημαντικά.

Μια άλλη βελτίωση του H.263 είναι η χρήση των διανυσμάτων κίνησης, που αναφέρονται στις μετατοπίσεις μισού εικονοστοιχείου αντί των μετατοπίσεων με έναν ακέραιο αριθμό εικονοστοιχείων. Για να υπολογιστούν τα αναφερόμενα υπο – εικονοστοιχεία, παραποιείται με διαφόρων ειδών προσθήκες η αξία των περιβαλλόντων εικονοστοιχείων.

Εκτός από τα δια - και τα ενδο - κωδικοποιημένα πλαίσια, το H.263 εισάγει τα PB – πλαίσια. Το όνομα προέρχεται από τα πλαίσια P και B του MPEG. Ένα PB – πλαίσιο αποτελείται από δύο μικτά πλαίσια: ένα “κανονικό” P – πλαίσιο και ένα αμφίδρομο πλαίσιο πρόβλεψης. Όταν αποκωδικοποιείται ένα PB – πλαίσιο, πρώτα παρουσιάζεται το B – πλαίσιο και μετά από αυτό το P – πλαίσιο. Το P – πλαίσιο μπορεί να χρησιμοποιήσει τους δια ή ενδο – τρόπους κωδικοποίησης, αλλά το B – πλαίσιο μπορεί να χρησιμοποιήσει μόνο το νέο προς τα μπροστά ή τον παλιό προς τα πίσω τρόπο πρόβλεψης και όχι τον δια – τρόπο κωδικοποίησης. Το B – πλαίσιο μπορεί να αναφερθεί στο σχετικό P – πλαίσιο μέσα στο PB – πλαίσιο, καθώς και σε έναν μέσο όρο του σχετικού P – πλαισίου και του προηγούμενως κωδικοποιημένου P – πλαισίου.

Το Test Model Near (TMN) είναι μια εφαρμογή του ανερχόμενου H.263 προτύπου και χρησιμοποιείται ως μοντέλο δοκιμής για αυτό το πρότυπο. Υποστηρίζει ότι έχει έναν παράγοντα του δύο, καλύτερης συμπίεσης από το H.261.



Σχήμα – πλαίσιο-B του H.263

5.4. Το MPEG-4 και η διαδραστική τηλεόραση

Παράλληλα με το σημαντικά λιγότερο εύρος ζώνης για την ίδια ποιότητα, η εγγενής υποστήριξη για την αλληλεπίδραση είναι μια βασική διαφορά μεταξύ του MPEG-4 και του MPEG-2, που επεκτείνεται ευρέως στα ψηφιακά τηλεοπτικά συστήματα. Σε κάθε περίπτωση, η λειτουργία της διαδραστικής τηλεόρασης σε ένα MPEG-2 βασισμένο περιβάλλον, σημαίνει ότι οι χειριστές πρέπει να υιοθετήσουν μια ή περισσότερες ικανές λύσεις, ή λύσεις βασισμένες σε τεχνολογίες μη εγγενείς με το MPEG, και να τις προσθέσουν παραγωγικά σε ένα MPEG-2 περιβάλλον. Αυτό έχει οδηγήσει στην εμφάνιση ανταγωνισμού στις τεχνολογίες προσθήκης που ανταγωνίζονται για την “εξέλιξη” των ITV χειριστών.

Κάθε χειριστής διαθέτει μια μοναδική σύνθετη λύση των τεχνολογιών, καθορισμένη συνήθως από την πλατφόρμα συμπίεσης MPEG, το σύστημα πρόσβασης τους και την πλατφόρμα Middleware, για παράδειγμα η OpenTV. Αυτό έχει οδηγήσει στην εμφάνιση διάφορων ασυμβίβαστων λύσεων και αγορών. Το πρόβλημα με τις “κάθετες” (vertical) αγορές, όχι μόνο υπό την επιχειρησιακή έννοια αλλά και από την πλευρά της τεχνολογίας, είναι ότι στο τέλος της ημέρας, οι τελικοί χρήστες δεν επωφελούνται από αυτές και η επέκταση υπηρεσιών επιβραδύνεται.

Η “οικονομική βαρύτητα” του παρόντος, καθιστά τις “κάθετες” αγορές αναπόφευκτες χωρίς ανοικτά πρότυπα. Παρά τις πολλαπλές προσπάθειες που έγιναν και γίνονται ώστε να εξαλειφθούν όλες αυτές οι ασυμβατότητες. Μια πολλά υποσχόμενη τεχνολογία είναι το πρότυπο DVB για τη διαδραστική τηλεόραση και την πλατφόρμα των πολυμέσων (MultimediaHomePlatform).

Θα δώσουμε μερικά παραδείγματα για το πώς η “δύναμη” του MPEG-4 μπορεί να προστεθεί εύκολα ώστε να συμπληρώσει τα σημαντικότερα πρότυπα σε αυτήν την περιοχή. Για παράδειγμα, σε μια πλατφόρμα με βάση το διαδικαστικό περιεχόμενο, εξετάζουμε τους τρόπους με τους οποίους το MPEG-4 μπορεί να προσθέσει αξία στο MHP. Στη συνέχεια, στην άλλη διαδρομή της διαδραστικής τηλεόρασης, μακριά από την Java, υπάρχει σημαντικό ενδιαφέρον για W3C εργασία και τις νεότερες γενεές των μεταγλωσσών του βασισμένες στην HTML, π.χ. XML και XHTML. Στην πραγματικότητα η πλατφόρμα MHP υποστηρίζει και τις δύο.

❖ MPEG-4 και MHP

Μια συχνά διατυπωμένη ερώτηση για το MPEG-4 είναι πώς σχετίζεται με την προδιαγραφή του DVB (DigitalVideoBroadcasting) για την πλατφόρμα πολυμέσων. Το πρώτο πράγμα που πρέπει να εξηγήσουμε είναι ότι υπάρχουν δύο σχετικές ομάδες προδιαγραφών της DVB. Αρχικά, η ομάδα DVB 1.0, η οποία είναι η βάση των μεταφορών της οικογένειας των DVB προτύπων. Αυτή η προδιαγραφή εξηγεί πώς θα εφαρμοστούν οι DVB-συμβατές ροές MPEG. Έτσι, το MPEG-4 φαίνεται ως μια λογική εξέλιξη, που θα είναι αποδοτικότερη όταν οι υπηρεσίες DVB πρόκειται να παραδοθούν πάνω από IP διαυθύνσεις.

❖ MPEG-4 : το πρότυπο των πολυμέσων

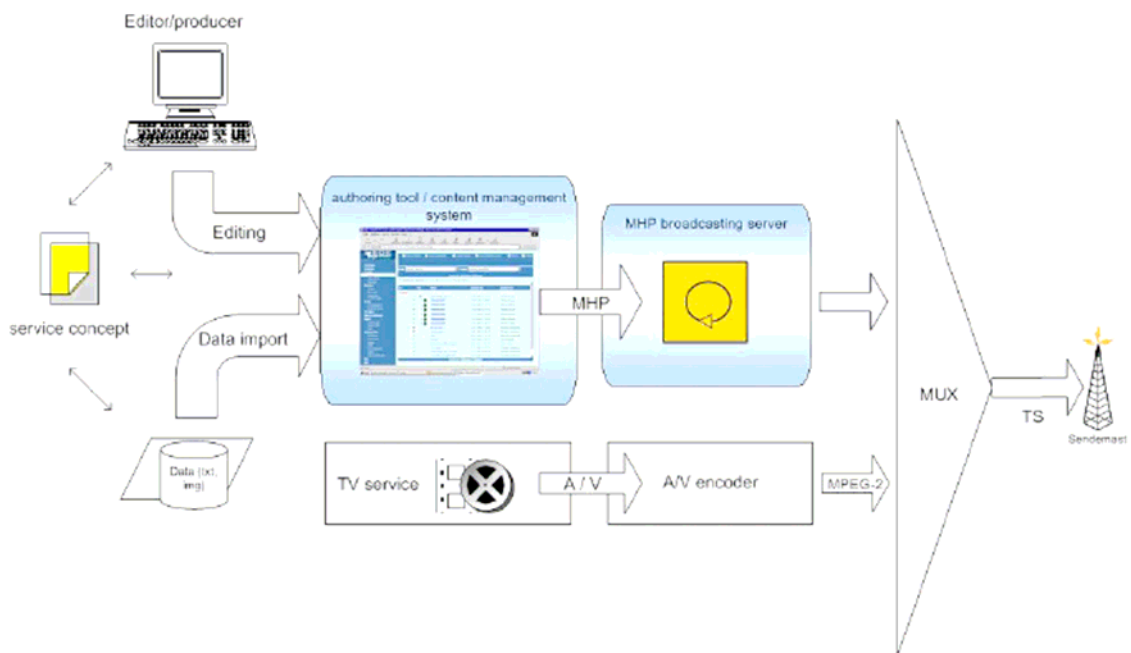
Η δεύτερη ομάδα προδιαγραφών του DVB, μερικές φορές καλούμενη και ως DVB 2.0, “καταπιάνεται” με την πλατφόρμα πολυμέσων (MHP) και ποικίλες εφαρμογές παράδοσης επόμενης γενεάς, συμπεριλαμβανομένης της προστασίας αντιγράφων (Copy Protection) και της διαχείρισης αντιγράφων (Copy Management) και της παράδοσης των υπηρεσιών του DVB πάνω από IP διαυθύνσεις. Η πλατφόρμα πολυμέσων (MHP) καθορίζει μια γενική διεπαφή μεταξύ των διαδραστικών ψηφιακών εφαρμογών και των τερματικών στα οποία εκείνες οι εφαρμογές εκτελούνται. Η προδιαγραφή της πλατφόρμας MHP διευκρινίζει πώς να μεταφορτωθούν οι εφαρμογές και το πολυμεσικό υλικό, τα οποία τυπικά θα παραδοθούν με μια DVB-συμβατή ροή μετάδοσης και εναλλακτικά μέσω ενός καναλιού επιστροφής.

❖ Το MPEG-4 είναι φυσικός “συνοδός” των MHP εφαρμογών, με χαμηλούς ρυθμούς μετάδοσης και φορμάτ απεικόνισης σκηνών που παραδίδονται μέσω της IP στους μετασηματιστές (set top boxes). Τα μοντέλα εφαρμογής, διαδραστικότητας και συγχρονισμού του MPEG-4 επιτρέπουν στο δυναμικότερο περιεχόμενο να προστεθεί σε εφαρμογές τύπου MHP. Επειδή το MPEG-4 μπορεί να μεταφερθεί από MPEG-2 μέσα μπορούμε να επιτύχουμε ένα πολύ μικρό ίχνος συγχρονισμού μεταξύ του προγράμματος εκπομπής και του MPEG-4 πολυμεσικού περιεχομένου. Η ενσωμάτωση MHP μέσα στο MPEG-4 μπορεί να επιτρέψει την αντικείμενο-βασισμένη (object-based) διαδραστική ψηφιακή τηλεόραση. Ο συνδυασμός της MHP

πλατφόρμας και του MPEG-4 παρέχει τη δυνατότητα να αναπτυχθούν πολύ ευέλικτες και πλούσιες διαδραστικές εφαρμογές για τη διαδραστική περιοχή εκπομπών (interactive broadcast domain). Τα χαρακτηριστικά γνωρίσματα του MPEG-4 μπορούν να εισαχθούν ομαλά και βαθμιαία, κατά τρόπο συμβατό προς τα πίσω (backwards-compatible).

❖ MPEG-J και DVB-J

Η αρχιτεκτονική της πλατφόρμας MHP καθορίζεται αναφορικά με τρία στρώματα: πόροι (resources), λογισμικό συστημάτων (system software) και εφαρμογές (applications). Οι χαρακτηριστικοί MHP πόροι είναι εκείνα τα στοιχεία που μπορούν να κληθούν από εφαρμογές για να εκτελεστούν ορισμένες λειτουργίες, παραδείγματος χάριν MPEG επεξεργασία, I/O, CPU, χειρισμός μνήμης και γραφικών. Το λογισμικό συστημάτων παρουσιάζει μια τυποποιημένη περιληπτική άποψη των πόρων της πλατφόρμας στις εφαρμογές, επιτρέποντας κατά συνέπεια τη “ανεξαρτησία πλατφορμών” (platform independence). Ένας “διαχειριστής εφαρμογών” (application manager) παρέχεται ώστε να διαχειριστεί την αλληλεπίδραση μεταξύ όλων αυτών των στοιχείων. Γενικές διεπαφές εφαρμογών προγραμμάτων (ApplicationProgramInterfaces APIs) διευκρινίζονται από το DVB-MHP, το οποίο βασίζεται γύρω από το DVB-J και περιλαμβάνει την εικονική μηχανή της Java (VirtualMachine VM) όπως με πρωτοτυπία διευκρινίζεται από τη Sun Microsystems. Οι εφαρμογές MHP μπορούν να έχουν πρόσβαση στους πόρους της πλατφόρμας μόνο μέσω αυτών των διευκρινισμένων APIs, ένα χαρακτηριστικό γνώρισμα το οποίο εγγυάται τη σταθερότητα της πλατφόρμας και την ευρωστία της ενάντια σε εφαρμογές “απάτες” (rogue). Αυτά τα APIs διευκρινίζονται από την τεχνική ενότητα της DVB (TAM) και εξετάζονται για προσαρμογή με την προδιαγραφή MHP μέσω της χρήσης συμφωνημένων εφαρμογών δοκιμής. Το MPEG-J(ava) προσφέρει ένα σύνολο λειτουργιών συμπληρωματικών σε εκείνες που προσφέρονται από το DVB-J. Το MPEG-J στο MPEG-4 είναι ένα σύνολο από APIs σε Java, τα οποία μπορούν να είναι παρόντα σε ένα MPEG-4 τερματικό. Οι MPEG-J εφαρμογές (MPEGlets), που στέλνονται ως τμήμα της παρουσίασης, χρησιμοποιούν τα MPEG-J APIs ώστε να ελέγξουν τις ικανότητες του MPEG-J τερματικού.



Σχήμα – Ροή εργασιών από τις εφαρμογές του MHP

5.5. Το μέλλον του MPEG-4

Τελικά, μέχρι πού θα επεκταθούν τα πρότυπα MPEG-4;

Τα πρώτα συστήματα θα χρησιμοποιηθούν για την παράδοση πολυμεσικού υλικού πάνω από το διαδίκτυο. Ο αποκωδικοποιητής βίντεο στον player της Microsoft λειτουργεί τώρα σύμφωνα με το κείμενο των τελικών προτύπων MPEG-4. Οι κατασκευαστές εξοπλισμού της κινητής τηλεφωνίας και οι παροχείς υπηρεσιών της κινητής τηλεφωνίας θα είναι πιθανώς οι επόμενοι, που θα υιοθετήσουν το πρότυπο. Σ' αυτό το σημείο, φρόνιμο θα ήταν να σημειώσουμε ότι η εφαρμογή ολόκληρου του προτύπου θα ήταν υπερβολή για μια κινητή συσκευή. Αντ' αυτού, διάφορα προφίλ, που καθορίζουν υποσύνολα του προτύπων, έχουν περιληφθεί ώστε οι σχεδιαστές να επιλέξουν από αυτά. Τα προφίλ αυτά είναι διαθέσιμα για απλές περιπτώσεις για την κινητή χρήση ή για σύνθετες με την προηγμένη τρισδιάστατη γραφική παράσταση. (Ομοίως, τα MPEG-2 προφίλ υπάρχουν για την ψηφιακή τηλεόραση.)

Μερικοί αναμεταδότες ήχου έχουν εκφράσει ήδη το ενδιαφέρον τους για το MPEG-4, των οποίων η ποιότητα του ήχου έχει κριθεί από την Ευρωπαϊκή Περιορισμένης Ζώνης Ψηφιακή Ακουστική Ομάδα Αναμετάδοσης (European Narrowband Digital Audio Broadcasting Group) καλύτερη από την αναλογική αναμετάδοση AM για το ίδιο εύρος ζώνης.

Μερικές αντιδράσεις είναι ακόμη και εχθρικές, πιστεύοντας ότι το πρότυπο προορίζεται να εκτοπίσει το MPEG-2. Προφανώς, στη δαπανηρή διαδικασία από αναλογική τηλεόραση στην ψηφιακή υπάρχουν άτομα, που δε θα θεωρήσουν τα MPEG πρότυπα ιδιαίτερα ελκυστικά. Αλλά το MPEG-4 δεν προορίστηκε να αντικαταστήσουν το MPEG-2 και αντ' αυτού, επιτρέπει νέες εφαρμογές και τους νέους τύπους υλικού (content), καθώς επίσης και περισσότερους τύπους συνδέσεων. Στην πραγματικότητα, η ομάδα εμπειρογνομόνων κινούμενης εικόνας (Moving Picture Experts Group) σχεδιάζει τρόπους ώστε το MPEG-4 και το MPEG-2 να συνεργαστούν. Δεδομένου ότι η τηλεοπτική μετάδοση γίνεται ψηφιακή παγκοσμίως και ότι οι τηλεοπτικοί δέκτες τείνουν να μοιάσουν με τα διαδραστικά τερματικά (interactive terminals) έχει αρχίσει μια νέα φάση στην εξέλιξη των πολυμέσων.

6. Internet Protocol Television (IP-TV)

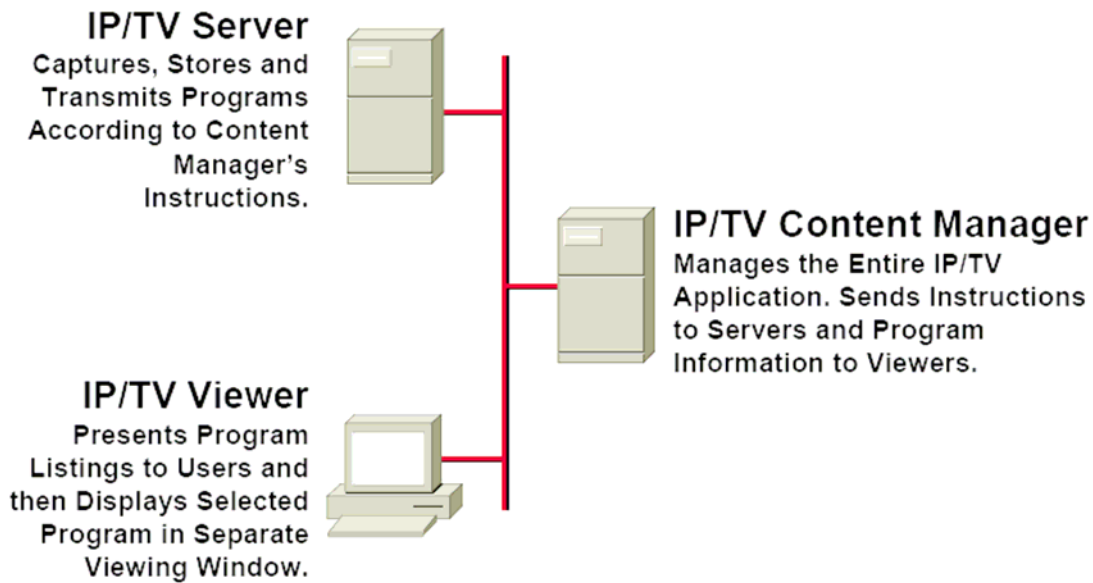
6.1. Εισαγωγή

Ένα σημαντικό ζήτημα που αντιμετωπίζεται στις επιχειρήσεις σήμερα είναι πώς να μεταφερθούν οι πληροφορίες στους υπαλλήλους με τρόπο ουσιαστικό και "πρόθυμο να εξυπηρετήσει". Ένας αποτελεσματικός τρόπος είναι να διανεμηθεί το πολυμεσικό υλικό μέσω ενός δικτύου Ethernet. Ένα converged δίκτυο επιτρέπει την παράδοση ζωντανού ή καταγεγραμμένου βίντεο και ήχου, συνδιάσκεψης, από απόσταση εκμάθησης, και παρουσιάσεων ζωντανής κίνησης (animated) σε οποιοδήποτε υπολογιστή που συνδέεται με το δίκτυο. Η ιδέα εξεταζόταν, αλλά η τεχνολογία προσφάτως ωρίμασε ώστε να την κάνει πραγματικότητα και να την υλοποιήσει στις εφαρμογές. Η IP-TV είναι μια πρόσφατη σύλληψη, η οποία προωθεί μεθόδους κατανεμημένου βίντεο και ήχου, και άλλων πολυμεσικών υλικών μέσω ενός δικτύου το οποίο αναπτύσσεται όπως η τηλεόραση. Στο πρότυπο της IP-TV, πολλαπλά κανάλια διαφορετικού περιεχομένου θα είναι διαθέσιμα και εύκολα προς χρήση από έναν πελάτη που εργάζεται σε έναν υπολογιστή γραφείου ή κάποια

εξειδικευμένη συσκευή. Για την επέκταση της IP-TV, διάφορες τεχνολογίες εμπλέκονται, όπως η ποιότητα υπηρεσιών (QoS), πρωτόκολλα πολλαπλής διανομής, εξειδικευμένες μορφές κωδικοποίησης βίντεο (MPEG-2, Microsoft Windows Media Video 9), και ειδικευμένοι κεντρικοί υπολογιστές καθώς και πελάτες.

6.2. Το δίκτυο της IP-TV

Η IP-TV χρησιμοποιεί ένα μοντέλο server-client, προστατεύοντας το σχεδιασμό της απέναντι στην προσέγγιση τύπου ένας-προς-πολλούς (one-to-many). Οι κεντρικοί υπολογιστές (servers) που χρησιμοποιούνται για την IP-TV έχουν δύο σημαντικές δομές. Σε καθεμία περίπτωση, οι κεντρικοί υπολογιστές (servers) είναι μεγάλης χωρητικότητας μηχανήματα, τα οποία έχουν πολύ γρήγορες συνδέσεις στο δίκτυο. Εντούτοις, ο ένας τύπος κεντρικού υπολογιστή σχεδιάζεται για την καταγραφή (capture) ζωντανών δεδομένων βίντεο και ήχου, ώστε να κωδικοποιηθούν αποτελεσματικά και να μεταδοθούν μέσω του δικτύου. Ο άλλος είναι ένας κεντρικός υπολογιστής, ο οποίος αποθηκεύει δεδομένα στη μνήμη cache για γρήγορη προσπέλαση (cached server) και κρατά το ηχογραφημένο εκ των προτέρων περιεχόμενο ώστε να το παραδίδει καθ' απαίτηση (on demand). Οι πελάτες της IP-TV μπορούν να είναι βασισμένοι είτε στο υλικό είτε στο λογισμικό (hardware or software based). Οι περισσότεροι από αυτούς υποστηρίζουν συγκεκριμένα φορμάτ υψηλής απόδοσης κωδικοποίησης βίντεο όπως το MPEG-2, αλλά υπάρχουν και μερικοί που “εργάζονται” βάσει της μονοπωλιακής κωδικοποίησης βίντεο με το φορμάτ Microsoft Windows Media Player 9. Ο κεντρικός υπολογιστής και οι πελάτες διαμορφώνονται έτσι ώστε να επικοινωνούν ο ένας με τον άλλον, γεγονός το οποίο μειώνει τις γενικότερες δαπάνες για την ενσωμάτωση μηχανισμών ανακαλύψεων και συγχρονισμού και επομένως επιτρέπεται γρηγορότερη και αποδοτικότερη επικοινωνία. Μεταξύ του πελάτη και του κεντρικού υπολογιστή, το δίκτυο πρέπει να έχει ενεργοποιημένη την ποιότητα υπηρεσιών τόσο για τις ροές βίντεο όσο και για τις ροές ήχου ώστε να διατηρήσει την ακεραιότητά τους και να μην τις αλλοιώνει από άλλους τύπους “κίνησης” (traffic). Όπως με οποιοδήποτε δίκτυο που παραδίδει πολυμεσικό υλικό, η υποδομή πρέπει να υποστηρίζει τα πρωτόκολλα πολλαπλής διανομής. Οι IP-TV εφαρμογές χρησιμοποιούν τα πρωτόκολλα πολλαπλής διανομής αραιού-τρόπου (sparse-mode) για να παραδώσουν το υλικό επειδή έχουν σχεδιαστεί με σκοπό τη διανομή από μια πηγή προς πολλούς συνδρομητές και αποδίδουν καλύτερα σε δίκτυα όπου το εύρος ζώνης έχει μεγάλη αξία. Αυτήν την περίοδο, η κύρια τάση στη βιομηχανία τηλεπικοινωνιών είναι να παρασχεθούν οι υπηρεσίες της IP-TV στο μέσο καταναλωτή. Εντούτοις, ο μέσος καταναλωτής δεν έχει πάντα μια αρκετά γρήγορη σύνδεση στο διαδίκτυο για να εκμεταλλευθεί πλήρως τις υπηρεσίες αυτές. Τελικά, καθώς το Gigabit Ethernet (GigE) και το 10 GigE γίνονται κομμάτι της καθημερινότητας και η πρόσβαση μεγάλων ταχυτήτων στο διαδίκτυο είναι ευρέως διαθέσιμη, η IP-TV θα είναι μια πραγματικότητα. Εντούτοις, μέχρι τότε, οι παροχείς υπηρεσιών IP-TV θα πρέπει να χρησιμοποιήσουν τις τρέχουσες μεθόδους όπως η πρόσβαση στο δίκτυο μέσω του τηλεοπτικού καλωδίου, της τηλεφωνικής γραμμής, ή της ασύρματης πρόσβασης. Ευτυχώς, το δίκτυο μιας επιχείρησης έχει αρκετό εύρος ζώνης και είναι προσπελάσιμο από όλους τους υπαλλήλους, γεγονός το οποίο ανοίγει έναν ολόκληρο νέο κόσμο δυνατοτήτων για τις εφαρμογές IP-TV.



Σχήμα – Τα συνιστώσα μέρη της IPTV

6.3. Εφαρμογές της IP-TV

Οι επιχειρήσεις τείνουν να αποκτήσουν πολλά διασκορπισμένα γραφεία σε ολόκληρη τη χώρα τους ή ακόμα και τον κόσμο, τα οποία είναι περιοριστικά στην επικοινωνία στο επίπεδο του ηλεκτρονικού ταχυδρομείου και της τηλεφωνίας. Ενώ αυτές οι μέθοδοι είναι επαρκείς για πολλούς σκοπούς, υπάρχουν φορές κατά τις οποίες δε μπορούν να μεταφέρουν αρκετή πληροφορία ώστε να είναι αποτελεσματικές. Για παράδειγμα, εάν προσπαθούσατε να εκπαιδεύσετε μια ομάδα υπαλλήλων στο πώς να χειρίζεται ένα πρόγραμμα λογισμικού χρησιμοποιώντας ένα ενδιάμεσο γραφικό (user interface), θα ήταν εξαιρετικά δύσκολο να ολοκληρώσετε το έργο αυτό μέσω τηλεφώνου. Ένα οπτικό μέσο (visual medium) μεταβιβάζει περισσότερες πληροφορίες και αυξάνει το ενδιαφέρον του ακροατηρίου ενώ επιτρέπει θέματα, που είναι πιά περίπλοκα να καλυφθούν ευκολότερα. Η IP-TV εξυπηρετεί πολύ καλά τέτοιες περιπτώσεις και μπορεί να εφαρμοστεί με τους ακόλουθους τρόπους:

- Από απόσταση εκπαίδευση (distance learning)
- Επιχειρησιακά σεμινάρια (company – wide seminars)
- Τηλεοπτική διάσκεψη (video conferencing)
- Τηλεοπτική και ακουστική αρχειοθέτηση (video and audio archiving)
- Τηλεόραση εσωτερικού κλειστού κυκλώματος (internal closed circuit television)

6.4. Τα πλεονεκτήματα της IP-TV

Αυτό που είναι μοναδικό στην IP-TV είναι ότι έχει τα πολλαπλά κανάλια περιεχομένου υψηλής ποιότητας, αλλά οι τεχνικές συμπίεσης που χρησιμοποιούνται είναι αποδοτικότερες και απαιτούν λιγότερο εύρος ζώνης από άλλες συγκρίσιμες τεχνολογίες. Στο παρελθόν, για την πλήρη εκμετάλλευση διαφορετικών ροών πολυμεσικού περιεχομένου θα χρειαζόσαστε πολλαπλούς κεντρικούς υπολογιστές (servers). Σαν χρήστης, που θέλει να βρει αυτές τις ροές πολλαπλής διανομής, έπρεπε

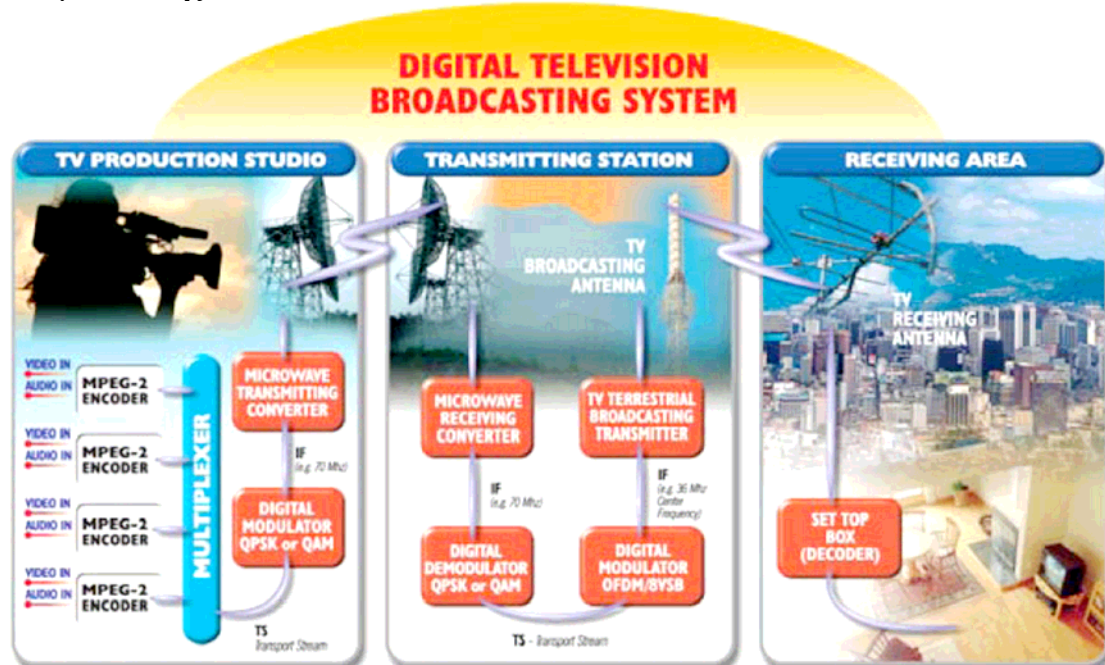
να λειτουργήσετε “με το χέρι”, εάν ξέρατε πού να κοιτάξετε στο δίκτυο. Η IP-TV λύνει τα προβλήματα με την ανάπτυξη κεντρικών υπολογιστών, που μπορούν να χειριστούν τις πολλαπλές αυτές ροές αμέσως και να τις παρέχουν στους πελάτες - χρήστες για να τις “περιεργαστούν” (browse) μέσω των διαθέσιμων ροών IP-TV. Ένας πελάτης IP-TV ή ένας θεατής βασισμένος σε ένα πρόγραμμα λογισμικού ή μια συσκευή IP-TV δίνει στο χρήστη έναν τρόπο να “κινηθεί” (channel surf) μέσα στο διαθέσιμο περιεχόμενο για να βρεί τι θέλει, ή για να δει απλά εάν συμβαίνει κάτι σε αυτό που θέλει να δει. Θα ήταν δυνατό να αναπτυχθούν πολλαπλοί τύποι κατάρτισης, οι οποίοι θα ήταν προσιτοί από σχεδόν οπουδήποτε μέσα στο εταιρικό δίκτυο. Οι υπάλληλοι θα μπορούσαν να συνεργαστούν μέσω της τηλεοπτικής διάσκεψης και θα ήταν δυνατό να εκμεταλλευτούν τα νέα και άλλους τύπους πληροφοριών, επίσης. Για παράδειγμα, μια αεροπορική εταιρία θα μπορούσε να χρησιμοποιήσει την IP-TV για να παρέχει διαφορετικά κανάλια ψυχαγωγίας στους επιβάτες της μέσω ενός ασύρματου δικτύου. Οποιοσδήποτε επιβάτης με ένα lap-top, έναν ασύρματο μετασχηματιστή και έναν viewer (Windows Media Player) θα μπορούσε να χρησιμοποιήσει τα διαφορετικά κανάλια ενώ περίμεναν την πτήση τους. Ένα άλλο παράδειγμα θα ήταν να δημιουργηθεί ένα τηλεοπτικό αρχείο, που θα περιήχε πολλούς διαφορετικούς τύπους ταινιών και θα τους διένιμε μέσω της IP-TV. Θα ήταν δυνατό για ένα πανεπιστήμιο να δημιουργήσει μια τηλεοπτική βιβλιοθήκη (video library), που θα μπορούσε να προσεγγιστεί από οπουδήποτε στην πανεπιστημιούπολη. Ένα τέτοιο εργαλείο θα ήταν απίστευτα χρήσιμο σχεδόν σε κάθε ακαδημαϊκό τμήμα.

6.5. Συμπερασματικά

Η IP-TV είναι μια τεχνολογία, η οποία τώρα αναφαίνεται και μία μέρα θα είναι τόσο κοινή όσο η συμβατική τηλεόραση. Στο μεταξύ, έχει μερικές πολύ αναγκαίες εφαρμογές, που μπορούν να βοηθήσουν τις επιχειρήσεις και τις οργανώσεις σήμερα. Η IP-TV έχει αναπτύξει πολλαπλά κανάλια ενεργού ή αρχειοθετημένου πολυμεσικού περιεχομένου για να παρέχει κατάρτιση, βελτιωμένη συνεργασία, ένα μέσο για διάσκεψη, και να παρέχει ακόμη και ψυχαγωγία στους πελάτες της. Επιπλέον, τα οφέλη της IP-TV είναι πολύ απτά επειδή η καθ'απαίτηση πρόσβαση στην πληροφορία μπορεί να αυξήσει την παραγωγικότητα, να βελτιώσει το επίπεδο επικοινωνίας, και να παρέχει έναν τρόπο στους υπαλλήλους να προωθήσουν τις καριέρες τους. Οπουδήποτε συναντάται η καθ'απαίτηση πρόσβαση σε πολυμεσικό περιεχόμενο, η IP-TV μπορεί να την παραδώσει μέσω ενός υπάρχοντος δικτύου Ethernet.

7. Παράδειγμα

Παράδειγμα : Δομή ενός ψηφιακού multiprogram συστήματος τηλεοπτικής αναμετάδοσης



Στούντιο παραγωγής της τηλεόρασης :

Το στούντιο παραγωγής της τηλεόρασης παράγει περισσότερα ηχητικά/βίντεο προγράμματα (στο παράδειγμα είναι 4) που κωδικοποιούνται ψηφιακά σύμφωνα με το πρότυπο MPEG-2 και πολυπλέκονται (αυτό ενοποιείται ώστε να αποτελέσει μία ενιαία ροή ψηφιακών δεδομένων, η οποία καλείται Ροή μεταφοράς – Transport Stream, TS). Η ροή μεταφοράς διαμορφώνει ψηφιακά μία φέρουσα μέσης συχνότητας (συνήθως 70MHz), σύμφωνα με τη διαμόρφωση QPSK ή QAM. Η φέρουσα μέσης συχνότητας μετατρέπεται σε μικροκυματικές συχνότητες και διαβιβάζεται άμεσα στο σταθμό εκπομπής (επίγεια μικροκυματική σύνδεση) ή μέσω ενός δορυφόρου ή ενός επίγειου αναμεταδότη.

Σταθμός εκπομπής :

Το λαμβανόμενο μικροκυματικό σήμα μετατρέπεται στη μέση συχνότητα - 70MHz και αποδιαμορφώνεται ψηφιακά ώστε να παράγει τη ροή μεταφοράς, που περιέχει τα τέσσερα προγράμματα. Ο αποδιαμορφωτής μπορεί επίσης να αποκωδικοποιήσει τα τέσσερα προγράμματα, ώστε αυτά να είναι διαθέσιμα ξεχωριστά σε αναλογική ή/και σε ψηφιακή μορφή, για άλλους λόγους (π.χ. να εκπεμφθούν μέσω πομπών αναλογικής τηλεόρασης). Στο σημείο αυτό, η ροή μεταφοράς διαμορφώνει ψηφιακά μία φέρουσα (συνήθως στα 36 ή 44MHz) σύμφωνα με τα πρότυπα της επίγειας ψηφιακής αναμετάδοσης, OFDM (DVB-T) ή 8VSB (H1A- ATSC). Στη συνέχεια η φέρουσα μέσης συχνότητας μετατρέπεται στη VHF ή UHF ζώνη, όπου ενισχύεται και εκπέμπεται μέσω της κεραίας αναμετάδοσης, για να είναι διαθέσιμη στην περιοχή λήψης.

Περιοχή λήψης :

Το σήμα της ψηφιακής εκπομπής λαμβάνεται μέσω της κεραίας του δέκτη και τροφοδοτείται σε έναν κατάλληλο δέκτη/ αποκωδικοποιητή (συνήθως αποκαλούμενος settop-box ή IRD) που συνδέεται στην τηλεόραση(και λειτουργεί ως video/audio monitor).

8. Πειράματα – Συμπεράσματα

Στην εργασία αυτή μελετήθηκαν διάφοροι τύποι λογισμικών. Αρχικά έγινε εγκατάσταση αυτών των προγραμμάτων και για αρκετό διάστημα επεξεργασία πολυμεσικού υλικού στα πλαίσια του εκάστοτε τρόπου κωδικοποίησης.

- ❖ Mpegable Broadcaster Version 2.1
- ❖ Canopus ProCoder 2.0
- ❖ PCTV Vision Recordings

8.1. Mpegable Broadcaster

- ❖ Περιγραφή του Mpegable Broadcaster:

Το Mpegable Broadcaster είναι το πιο εξελιγμένο λογισμικό κωδικοποίησης βίντεο, διαθέσιμο σήμερα, που προσφέρει μια εύχρηστη, γρήγορη και επαγγελματική λύση για τη μετατροπή πολυμεσικών δεδομένων σε αρχεία τύπου MPEG-4.

Το Mpegable Broadcaster, που ενσωματώνει και που επεκτείνει όλα τα χαρακτηριστικά γνωρίσματα του Mpegable X4 Live, επιτρέπει τη διακωδικοποίηση ποικίλων αρχείων πολυμέσων σε αρχεία mpeg-4. Προαιρετικά, μπορούν να προστεθούν Hint tracks για ροές συμβατές με το ISMA, καθώς επίσης και live κωδικοποίηση ήχου και βίντεο. Το Mpegable Broadcaster επιτρέπει με ιδιαίτερα περίπλοκες ρυθμίσεις παραμέτρων να ρυθμιστεί η ποιότητα παραγωγής του βίντεο και όσον αφορά στο εύρος ζώνης.

Σε συνεργασία με οποιοδήποτε κεντρικό υπολογιστή ISMA (streaming server) το πολυμεσικό υλικό, που δημιουργείται μπορεί να χρησιμοποιηθεί από το ευρύ κοινό αποδοχής του MPEG-4 ISO, που χρησιμοποιεί συμβατούς players (όπως αυτοί της Apple, το QuickTime, Real Network, RealOne)

Και επαγγελματίες και ερασιτέχνες στη βιομηχανία, ικανοποιούν τις απαιτήσεις τους σχετικά με την κωδικοποίηση του ήχου καθώς και του βίντεο με την υψηλότερη πιθανή ποιότητα.

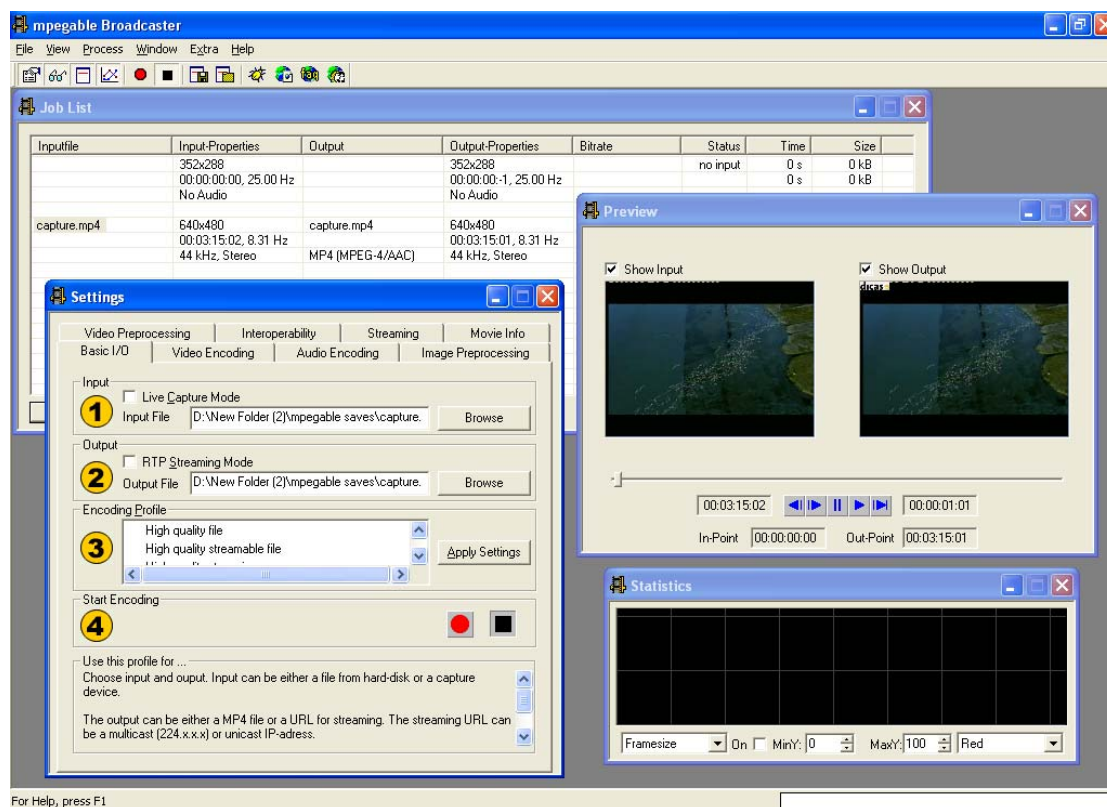
Χαρακτηριστικά γνωρίσματα : Πλήρης συμβατότητα με το MPEG-4 (ISO/IEC 14496), απλό και προηγμένο απλό προφίλ για το βίντεο, AACLC για τον ήχο. Εύχρηστο με το παράθυρο προβολής του αποτελέσματος. Πραγματοποιεί προεπεξεργασία για το εργαλείο του cropping, την υποδειγματοληψία και το de-interlacing. Αναλύσεις καθορισμένες από το χρήστη από 16 x 16 μέχρι 2048 x 2048. Κωδικοποίηση πραγματικού χρόνου για Pentium III. Κλιμακωτό από 5 Kbit/s σε 35 Mbit/s σταθερό και μεταβλητό bit rate. Επεξεργασία με συγκεκριμένη σειρά με το διαμορφώσιμο κατάλογο εργασιών. Αυτόματη ανίχνευση των σκηνών, που πρέπει να διαγραφούν. Παράθυρο στατιστικής για τον έλεγχο των παραμέτρων και της ποιότητας. Και τέλος περιλαμβάνει τον mpegable player

Format εισόδου:

- AVI, MPEG-1
- Live capture από αναλογικές πηγές
- οτιδήποτε πολυμεσικό υλικό, που υποστηρίζεται από DirectShow (MPEG-2).

Format εξόδου:

- .MP4 αρχεία



Σχήμα – Διεπαφή του Mpegable Broadcaster

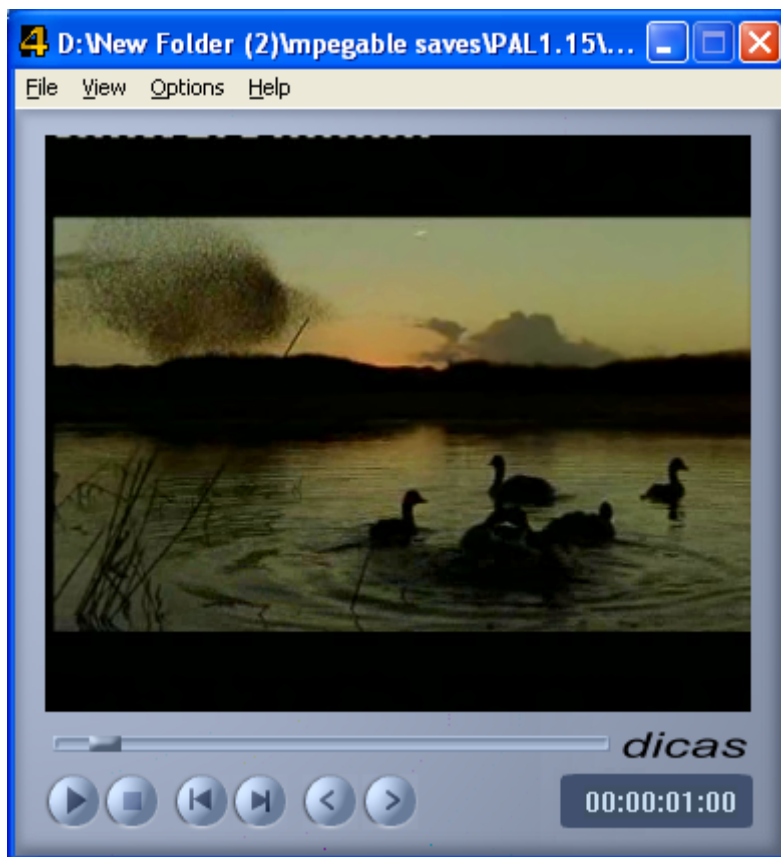
❖ Παραδείγματα καταγραφής με το Mpegable Broadcaster



Σχήμα – Advanced profile



Σχήμα – NTSC1.15 Mbps (absolute size 704x480)



Σχήμα – PAL 1.15 Mbps (absolute size 704x576)



Σχήμα – QCIF 1.15 Mbps (absolute size 176x144)



Σχήμα –SIF 1.15 Mbps (absolute size 352x288)



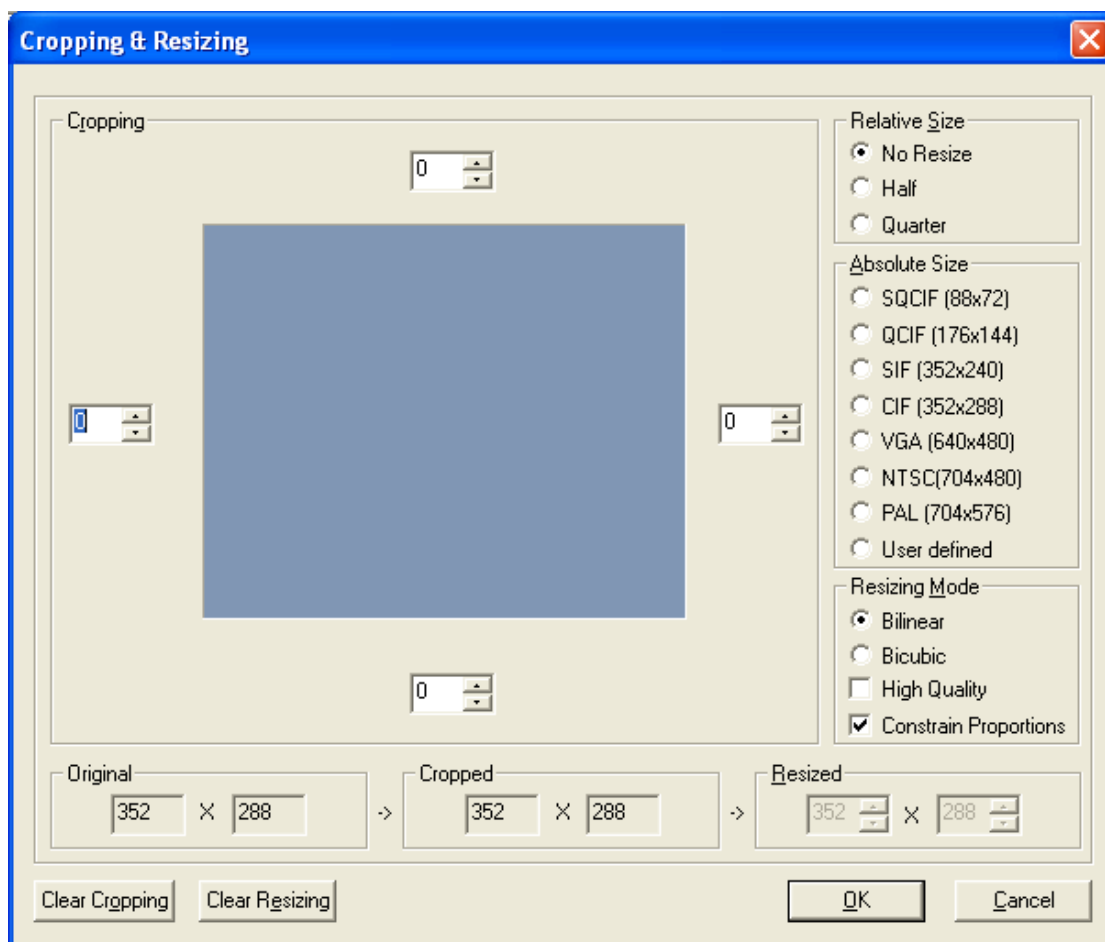
Σχήμα –SQCIF 1.15Mbps (absolute size 88x72)



Σχήμα –QCIF 3Mbps (absolute size 88x72)



Σχήμα –VGA 1.15Mbps (absolute size 640x480)



Σχήμα – Διεπαφή ρυθμίσεων

8.2 Canopus ProCoder

❖ Ο τελικός μετατροπέας format

Το λογισμικό Canopus ProCoder 2.0 συνδυάζει την ταχύτητα και την ευελιξία σε ένα βελτιωμένο εργαλείο μετατροπής βίντεο για επαγγελματίες. Με χαρακτηριστικά γνωρίσματα της ευρείας κλίμακας επιλογές για το υλικό εισόδου/εξόδου (I/O), το προηγμένο φιλτράρισμα (advanced filtering), την επεξεργασία κατά δεσμίδες με τη βοήθεια μιας εύχρηστης διεπαφής το ProCoder 2.0 είναι η ιδανική λύση για την υψηλής ποιότητας δημιουργία πολυ-μορφικών βίντεο (multi-format video).

Το ProCoder 2.0 μπορεί να χρησιμοποιηθεί με ποικίλους τρόπους: σαν ένας wizard, σαν μια αυτόνομη εφαρμογή, ή ως σύνδεση (plug-in) για τα μη γραμμικά πακέτα συμπεριλαμβανομένου του Canopus EDIUS και του Adobe Premiere Pro 1.x.

Με την κωδικοποίηση MPEG βίντεο για την παραγωγή DVD, που παράγει πολυμεσικό υλικό Windows (Windows media) για streaming ή για διακωδικοποίηση από NTSC σε PAL, το ProCoder 2.0 καθιστά τη μετατροπή του format του βίντεο γρήγορη και εύκολη. Μπορείτε να κωδικοποιήσετε μια είσοδο (single source) σε πολλαπλούς “στόχους”/εξόδους (targets) ταυτόχρονα, να εκτελέσετε πολλαπλές εργασίες μετατροπής back-to-back κατά δεσμίδες, ή να χρησιμοποιήσετε τα εκ των προτέρων ρυθμισμένα εικονίδια τύπου drag-and-drop για να ξεκινήσετε την κωδικοποίηση με ένα απλό click. Το ProCoder 2.0 υποστηρίζει χαρακτηριστικά

γνωρίσματα πολλών βίντεο codecs συμπεριλαμβανομένης της καινοτόμου υψηλής ποιότητας, της μεγάλης ταχύτητας, των DV codecs από το ίδιο το Canopus, καθώς επίσης και MPEG-2 codecs, και υποστηρίζει επίσης επιπλέον codecs εγκατεστημένους ήδη στο σύστημα.

Κύρια χαρακτηριστικά :

- Επαγγελματική κωδικοποίηση και μεταροπή βίντεο σε όλα τα δημοφιλή format με μεγάλη ταχύτητα και υψηλή ποιότητα (MPEG-1, MPEG-2, Windows Media, QuickTime, DivX)
- Ειδικές συναρτήσεις για αυτοματοποιημένη κωδικοποίηση
- Manager εργασιών με συναρτήσεις προτεραιότητας
- Συρραφή πολλαπλών πηγαίων αρχείων ώστε να δημιουργηθεί ένα ενιαίο αρχείο εξόδου
- Ταυτόχρονη δημιουργία πολλαπλών αρχείων εξόδου
- Μετατροπή του aspect ratio
- Μετατροπή από NTSC σε PAL
- Παρεμβολή και μετατροπή του ρυθμού πλαισίων
- Automatic adaptive deinterlacing
- 3:2 pulldown
- Πολυεπίπεδη επεξεργασία του VBR (VariableBitRate)
- Πλήρη υποστήριξη HD για MPEG και το Windows Media
- Αρχείο τύπου VOB και δημιουργία DVD-video image με κεφάλαια
- Αποθήκευση droplets για τη γρήγορη μετατροπή τύπου drag-and-drop
- Χωριστές επιλογές για τον ήχο εισόδου/εξόδου συμπεριλαμβανομένων και των πολυπλέξεων
- Πολυεπεξεργασία και υποστήριξη Hyper-Threading για γρηγορότερη κωδικοποίηση
- Πολλαπλή υποστήριξη colorspace συμπεριλαμβανομένου YUV και RGB

Τεχνικές προδιαγραφές :

Format εισόδου:

- Canopus DV
- HQ Canopus
- Microsoft DV
- Οποιοδήποτε βίντεο για τον windows codec
- Οποιοσδήποτε DirectShow codec
- Οποιοσδήποτε QuickTime codec
- MPEG-1, MPEG-2
- Macromedia Flash (βίντεο μόνο)
- Windows Media
- DivX
- MP3 και αρχεία PCM WAV
- AVISynth script αρχεία
- Ακόμα και εικόνες συμπεριλαμβανομένων : BMP, TGA, του TIF, του PNG και JPG

Format εξόδου:

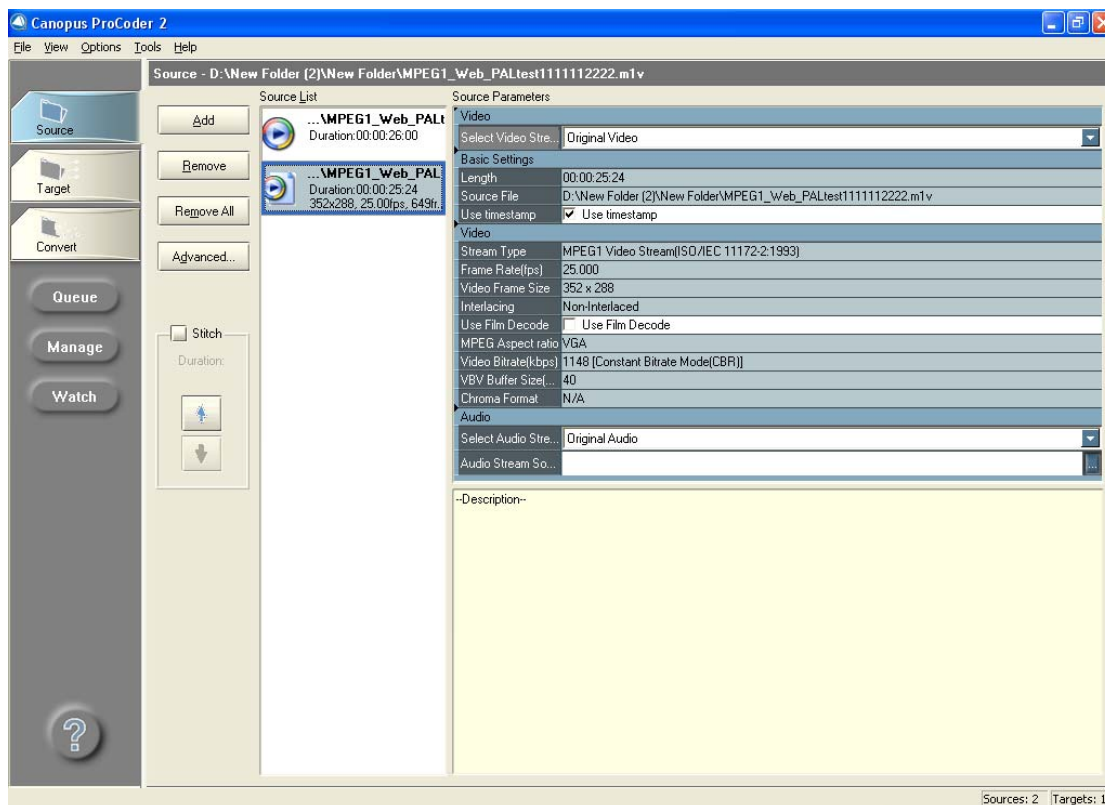
- Canopus DV
- Microsoft DV
- Οποιοδήποτε βίντεο για τον windows codec
- Οποιοσδήποτε DirectShow codec
- Οποιοσδήποτε QuickTime codec
- MPEG-1, MPEG-2 (που χρησιμοποιούν τον Canopus codec)
- VOB (DVD-video image)
- ACM codecs
- Macromedia Flash βίντεο (με το Flash MX εγκατεστημένο)
- Windows media
- RealVideo
- DivX
- MP3 και αρχεία PCM WAV

Φίλτρα για την επεξεργασία του βίντεο:

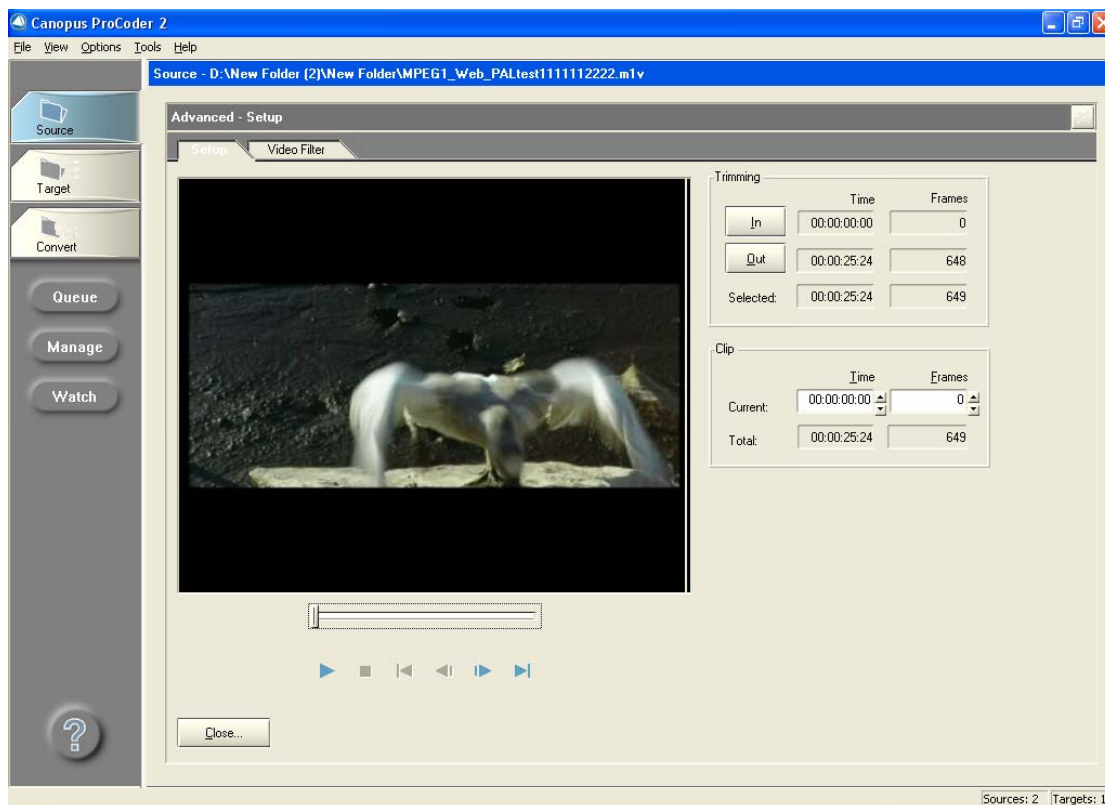
- Διόρθωση χρώματος 601 έως 709 (ITU-ρ BT 601 SD στο ITU-ρ BT 709-5 HD)
- Διόρθωση χρώματος 709 έως 601 (ITU-ρ BT 709-5 HD στο ITU-ρ BT 601 SD)
- Διόρθωση 601 – επέκταση του διαστήματος χρώματος (Expand Color Space)
- Διόρθωση 601 – συρρίκνωση του διαστήματος χρώματος (Shrink Color Space)
- Προσαρμοστικό de – interlace (Adaptive de – interlace)
- Διαμόρφωση δυαδικών αρχείων εικόνας (Bitmap Keying)
- Ασπρόμαυρη διόρθωση (Black/white correction)
- Blur
- Circular Blur
- Διόρθωση χρώματος
- Color Safe
- Crop
- Fade In/Out
- Gamma correction
- Gaussian Blur
- Median
- Pulldown
- Rotate
- Sharpen
- Προσωρινός μειωτής θορύβου (Temporal Noise Reducer)

Φίλτρα για την επεξεργασία του ήχου:

- Χαμηλής διέλευσης (LowPass)
- Κανονικοποίησης (Normalize)
- Volume



Σχήμα – Διεπαφή Canopus ProCoder



Σχήμα – Διεπαφή Canopus ProCoder

8.2.1. Παρατηρήσεις

Κατα τη δημιουργία αρχείων με το MPEG – εργαλείο Canopus ProCoder, σημειώθηκαν τα εξής συμπεράσματα:

- ❖ .m1a : αρχείο τύπου MPEG-1 αλλά ενώ είναι Quick Time application δεν μπορεί να αναπαραχθεί από τον Quick Time Player.
- ❖ .mp4 :
 - 1024 PAL Streaming : Η κωδικοποίηση δεν πραγματοποιείται επιτυχώς.
 - 256 PAL Streaming : Η κωδικοποίηση δεν πραγματοποιείται επιτυχώς.
 - 1024 PAL Download : Δεν αναπαράγεται από τον MPEG-4 player.
 - 256 PAL Download : Επιτυχής κωδικοποίηση και αναπαραγωγή αλλά δεν μπορούμε να προσθέσουμε φίλτρα επεξεργασίας του υλικού.
 - 512 PAL Download : Καμία δυνατότητα αναπαραγωγής και προσθήκης φίλτρων.
- ❖ .mpg : Είναι όλα τα MPEG-2 αρχεία, τα οποία κωδικοποιήθηκαν και αναπαράχθηκαν χωρίς καμία δυσκολία και χωρίς κανέναν περιορισμό. Στις δοκιμαστικές περιπτώσεις χρησιμοποιήθηκαν όλα τα φίλτρα, που παρέχει το Canopus ProCoder αλλά καταλήξαμε ότι τα πιο χρήσιμα είναι: Color Correction, Color Safe και Sharpen.

Τέλος, πρέπει να σημειωθεί ότι έγιναν πολλά πειράματα με το ίδιο βίντεο ως είσοδο και κωδικοποίηση με διαφορετικό bit rate κάθε φορά ώστε να διαπιστωθούν οι διαφοροποιήσεις της ποιότητας.

8.3 PCTV Vision

❖ Το λογισμικό PCTV Vision

Το PCTV Vision είναι ένα πακέτο λογισμικού για το βίντεο capture και την παρακολούθηση τηλεόρασης. Αποτελείται από ένα απλό παράθυρο, το οποίο προβάλλει το αποτέλεσμα και κάποιες ρυθμίσεις για το VCR στο κατώτατο σημείο του παραθύρου. Έχει μερικές επιλογές διαμόρφωσης για το capture του βίντεο και μπορείτε να αλλάξετε ρυθμίσεις όπως το χρώμα, τον ήχο, και το interlacing. Κατά το capture, κανονικά, εμφανίζεται το παράθυρο gallery, δίπλα στο παράθυρο του vision και δείχνει σε μορφή thumbnails τα καταγεγραμμένα clips.

Μετατρέπει το pc σε τηλεόραση και προσωπικό τηλεοπτικό εργαλείο εγγραφής βίντεο (PVR Personal Video Recorder). Το PCTV Vision συνδυάζει μια υψηλής ποιότητας λήψη τηλεοπτικού καθώς και ραδιοφωνικού σήματος με προηγμένα χαρακτηριστικά γνωρίσματα καταγραφής βίντεο.

Format εισόδου:

- S-video
- Composite-video

Format εξόδου:

- 2.5 mm header (internal)

Stereo υποστήριξη ήχου:

- Ακουστικό καλώδιο για τη σύνδεση σε ένα soundboard ή μέσω της PCI ροής ήχου

Σήμα τηλεόρασης (TV signal):

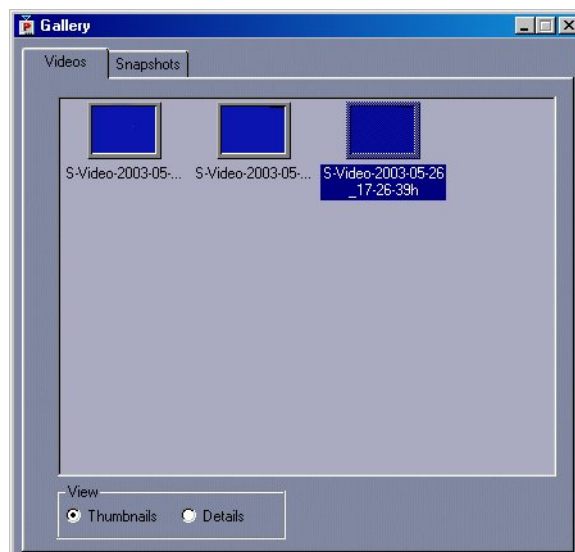
- F-connector (NTSC, PAL M/N) μέσω ειδικού προσαρμοστή (adapter)

Format capture:

- Κάνει capture και αποθήκευση στο σκληρό δίσκο με MPEG-1 πραγματικού χρόνου.



Σχήμα – Το παράθυρο vision



Σχήμα – Το παράθυρο gallery

9. Βιβλιογραφία

Links

- ❖ <http://www.chiariglione.org/mpeg>
- ❖ <http://www.mpeg4.net/>
- ❖ <http://mpeg4ip.sourceforge.net/documentation/index.php>
- ❖ <http://www.skystream.com/press/2005/>
- ❖ http://www.tribecaexpress.com/cisco_IPTV.htm
- ❖ <http://www.matrixstream.com/IMX%20solution%20overview.cfm>
- ❖ <http://informatv.com/articles/2004/10/21/pacedevelopsmpeg4/>
- ❖ <http://www.redorbit.com/news/display/?id=53196>
- ❖ <http://www.isp-planet.com/technology/2005/envivio.html>
- ❖ <http://www.envivio.com/products/4front.html>
- ❖ <http://citeseer.ist.psu.edu/>

Papers - Δημοσιεύσεις

- ❖ Paper[1] : “Adaptive Delivery of Real-Time Streaming Video” by Nicholas G. Feamster, Electrical Engineering and Computer Science, Massachusetts Institute of Technology
- ❖ Paper[2] : “Low Bit Rate IP Video” by Stuart Thomson
- ❖ Paper[3] : “Internet Protocol Television IP-TV” an Alcatel executive brief
- ❖ Paper[4] : “Realizing MPEG-4 Streaming Over the Internet: A Client/Server Architecture using DMIF” by K. Asrar Haghighi, IEEE Student Member, Y. Pourmohammadi, IEEE Student Member, H. M. Alnuweiri, IEEE Member, Department of Electrical & Computer Engineering University of British Columbia.
- ❖ Paper[5] : “Statistical properties of MPEG video traffic and their impact on traffic modelling in ATM systems” by O. Rose, Institute of Computer Science University of Würzburg.
- ❖ Paper[6] : “Mpeg-4 - The media standard” by the MPEG-4 Industry Forum
- ❖ Paper[7] : “Video Compression Techniques over low-bandwidth lines” by Roalt Aalmoes, Twente University