

1. A digitális hangtechnika alapjai

1.1. Digitális hangjelek

A híradástechnikában az átviendő, vagy rögzítendő információ megjelenési formája a jel. Ezek elektromágneses jelek, mert csak ezeket lehet elektronikus úton átvinni, illetve feldolgozni.

Matematikai osztályozás alapján léteznek:

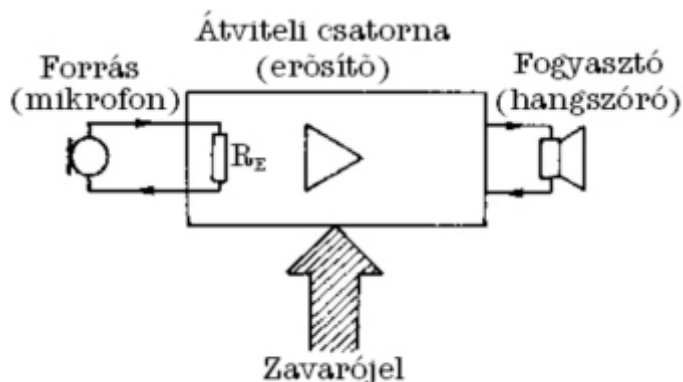
- Analóg jelek.
- Lépcsős jelek (amplitúdóban diszkrét).
- Mintavételezett jelek (időtartományban diszkrét).
- Digitális jelek.

1. táblázat: a különböző jelek tulajdonságai

Jel	Értelmezési tartomány	Értékkészlet
Analóg	folytonos	folytonos
Lépcsős	folytonos	diszkrét
Mintavételezett	diszkrét	folytonos
Digitális	diszkrét	diszkrét

A digitális hangjelek eredeti forrása mindig analóg jel. Az analóg jelből a digitalizálás során keletkezik a digitális jel. A digitális jel az analóg jelből meghatározott időközönként vett mintákból áll. Az egyes minták értékét bináris kódszavak tartalmazzák. Tehát a digitális hangjeleket feldolgozó áramkörök digitálisak, azaz logikai áramkörökből és tárolóelemekből állnak.

1.2. Az analóg és a digitális jelfeldolgozás összehasonlítása



Először vizsgáljuk meg az analóg hangátvitel folyamatát ! (1. ábra). Az erősítő a bemenetére csatlakoztatott mikrofon jelét felerősíti és a kimenetére csatlakoztatott hangszóróra továbbítja. A vázolt példában a mikrofon a jel forrása, az erősítő az átviteli csatorna, a hangszóró pedig a fogyasztó. Teljesen

1. ábra: az analóg hangátvitel

mindegy, hogy mennyire bonyolult úton jut el a jel a forrástól a fogyasztóig, ez a felbontás mindig elvégezhető. Ebből a szempontból lényegtelen, hogy analóg, vagy digitális jelfeldolgozásról van szó. Viszont fontos tudni, hogy az átviteli csatornának vannak-e olyan tulajdonságai, amelyek befolyásolják az átvitt jelet. Mielőtt tovább boncolnánk ezt a kérdést leszögezhetjük, hogy sajnos vannak ilyen tulajdonságai az átviteli csatornának. Ezekon kívül felléphetnek még különböző zavarójelek is.

1.2.1. Az átviteli csatorna jellemzői

Jel-zaj viszony

Az átvitel során a csatornán mindig fellép valamilyen zavaró hatás is, ami a fenti példában a hangszóróból áradó zajként érzékelhető. A jel-zaj viszony a hasznos és a zavaró jel (zaj) aránya dB-ben (2. ábra). A jel-zaj viszonyt lineáris frekvenciamenet és lezárt be- és kimenet mellett mérik.

A zajt az átviteli csatornát felépítő aktív és passzív alkatrészek hozzák létre. A zajfeszültség a hallható frekvenciatartomány széles sávjában jelentkezik (fehér zaj: minden frekvencia megtalálható benne csak különböző intenzitás értékekkel).

Dinamika

A kifogástalanul reprodukálható kimenő jel maximális értékének és a még zajmentesnek érzékelt kimenő jel minimális értékének aránya dB-ben. A dinamikát felülről a kivezérelhetőség, alulról a zaj korlátozza.

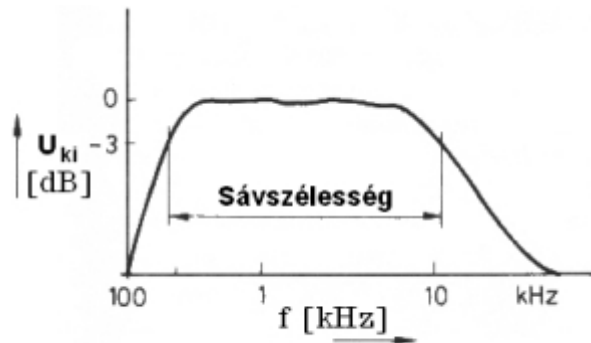
Sávszélesség

Az átviteli csatorna sávszélességét tekintve sávszűrő jellegű. A frekvencia növekedésével a hasznos jel amplitúdója általában csökken, ezenkívül egyenfeszültségű, illetve kismagasságú összetevők átvitelére általában nincs lehetőség.

A sávszélesség az a frekvenciatartomány, amelyen belül a kimenő jel értéke a vonatkoztatási frekvencián mért értékhez képest maximum ± 3 dB-lel tér el (3. ábra).



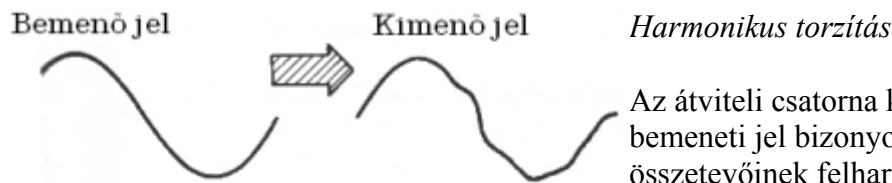
2. ábra: jel-zaj viszony



3. ábra: sávszélesség

Linearitás

Ideális átviteli csatorna esetén a bemeneti jel a kimeneti jellel lineáris viszonyban van. Ellenkező esetben torzítás keletkezik, amit nemlineáris torzításnak neveznek (4. ábra). A nemlineáris torzításokat általában az aktív elemek görbült jelleggörbéi okozzák.



4. ábra: nemlineáris torzítás

Az átviteli csatorna kimenetén általában a bemeneti jel bizonyos frekvencia összetevőinek felharmonikusai is megjelennek. A torzítás mértékét a Klirrfaktor fejezi ki: a felharmonikusok effektív

értékének és a felharmonikusokkal terhelt kimeneti jel értékének hányadosa %-ban.

1.2.2. Az analóg átvitel jellemzői

- A jel tartalmát (a jelentést) nem veszi figyelembe.
- Az átvitel során a jel csillapodik, ezért erősítőket kell alkalmazni.
- Az erősítők nem képesek a hasznos jel és a zaj szétválasztására, így a zajt is erősítik, ami a J/Z viszony folyamatos romlását idézi elő.
- Minél több erősítő fokozaton halad keresztül a jel, annál torzultabb lesz. Analóg rendszerben azonban a torzulás nem biztos, hogy információ veszteséget okoz.

1.2.3. A digitális jelfeldolgozás előnyei és hátrányai

Az egyik legnagyobb előny, hogy a zavarójelek a digitális jel durva H/L struktúráját nem tudják olyan könnyen megzavarni, mint analóg átvitel során. A zavarófeszültséggel terhelt digitális jelből egyszerű határolással visszaállítható az eredeti négyzögjel (5. ábra). Így a csatornából származó zaj kiszűrhető, emiatt pedig a J/Z viszony is jobb: általában a 90 dB-t is meghaladja. A fentiek következménye, hogy másolatok készítése esetén a másolat minősége az eredetiével megegyező.



5. ábra: a zaj hatása az analóg és a digitális jelre

Hátrányok:

- A digitális jel érzékeny az adatvesztésre és a hibák javítása nem egyszerű feladat a fogyasztói oldalon.
- A digitális jelfeldolgozó áramkörök bonyolultabbak.
- Az áramkörök túlvezérlése a digitális jel teljes összeomlásához vezet.

Fontos, hogy a 90 dB-es érték a rendszerdinamika és nem tévesztendő össze a hangfrekvenciás jel dinamika értékével, vagyis a lehangosabb és leghalkabb hangok közötti különbséggel.

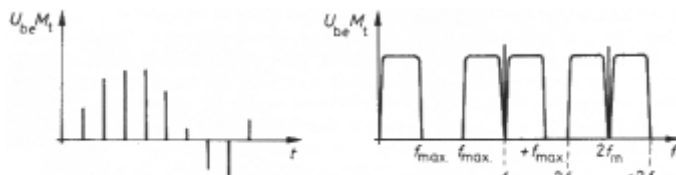
1.3. A digitalizálás folyamata

1.3.1. A mintavételi tétel

C. E. Shannon 1948-as munkájában kifejti, hogy a mintavétellel nyert diszkrét mintákból álló impulzussorozat információtartalma megegyezik az eredeti, időben folytonos analóg jel információtartalmával. Ez viszont csak bizonyos feltételek érvényesülése esetén igaz. Ezeket a feltételeket a Shannon-féle mintavételi tétel tartalmazza: a mintavételezett jelből akkor állítható vissza információ veszteség nélkül az eredeti analóg jel, ha a mintavételi frekvencia (f_m) legalább kétszerese az analóg jelben előforduló legmagasabb frekvenciának (f_{max}). A mintavételi frekvencia értékének állandónak kell lennie. Képlettel:

$$f_m \geq 2f_{max}$$

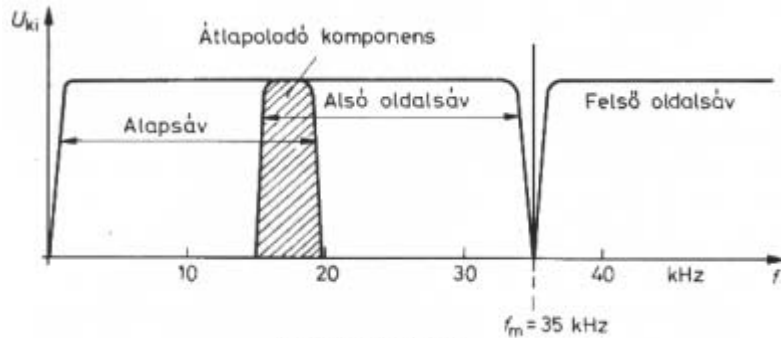
Az f_{max} frekvenciát Nyquist-frekvenciának is nevezik.



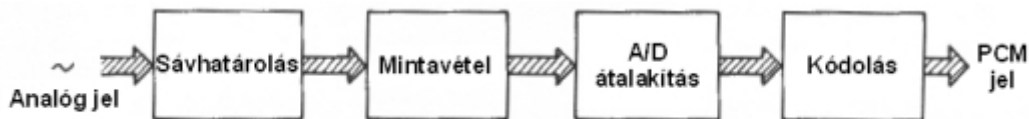
Ha nem tartjuk be a mintavételi tételt, akkor a kimeneti jelben ún. gyűrődési torzítások keletkeznek (aliasing). A 6. ábrán látható a mintavevő és -tartó áramkör kimenetén megjelenő kimeneti

6. ábra: a mintavevő és -tartó áramkör kimeneti feszültsége feszültség idő- és frekvenciatartománybeli képe.

Látható, hogy az eredeti alapsávi spektrum mellett a mintavevő frekvencia és annak felharmónikusai környezetében járulékos oldalsávok is megjelennek. Ha a mintavételi tételt nem tartjuk be, akkor az alapsáv és az f_m mellett megjelenő alsó oldalsáv között átlapolódás jön létre (7. ábra). Az ebbe a tartományba eső rész eltorzul és többé már nem rekonstruálható. A hangfrekvenciás jelek digitális feldolgozása a PCM technikán alapul (Pulse Code Modulation). Az analóg jeltől a PCM kódolású digitális jelig a 8. ábrán látható úton jutunk.



7. ábra: aliasing



8. ábra: a digitalizálás folyamata

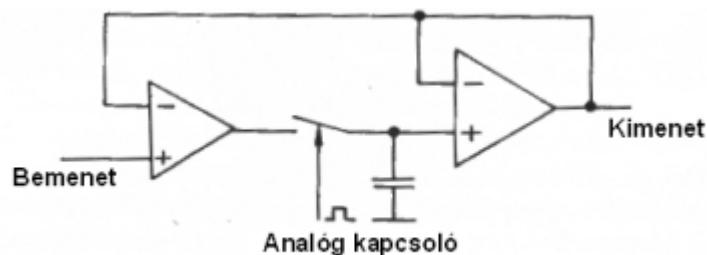
1.3.2. Bemeneti aluláteresztő szűrő

Feladata a bemenő analóg jel felső frekvenciahatárolása, hogy megfeleljen a mintavételi tételben megfogalmazott követelményeknek. Anti-aliasing szűrőnek is nevezik. Az anti-aliasing szűrővel szemben támasztott követelmények:

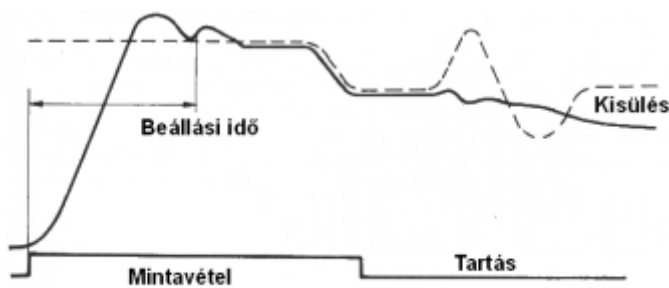
- Vízszintes átviteli karakterisztika egészen a felső határfrekvenciáig.
- Meredek átmenet az áteresztő tartományból a zárótartományba.
- Az f_{\max} frekvencia fölött min. 80 dB-es csillapítás.
- Kis harmónikus torzítás.
- Kis tranziens idő.
- Kis fázistorzítás.
- Kis csoportfutási idő.
- Az egész digitális rendszerénél kisebb saját zaj.

1.3.3. Mintavevő és -tartó áramkör

A sávhatárolt bemenő jel mintavételezését végzi. Elvi kapcsolási rajza a 9. ábrán látható. Az analóg kapcsolót a mintavevő jel vezérli. A kapcsoló zárt állapotában a kondenzátor a bemeneti feszültség pillanatnyi értékére töltődik fel, majd a kapcsoló nyit és a kondenzátorban a kapcsoló következő zárásáig tárolódik a feszültség minta.



9. ábra: mintavevő és -tartó áramkör



10. ábra: beállási idő és kisülés

Azt az időt amikor a kapcsoló zárt állapotban van nevezik mintavételi időnek, amikor pedig nyitva van az a tartási idő. A tartási idő alatt a kondenzátorban tárolt feszültség értéket a kvantáló áramkörnek fel kell dolgoznia. Ezalatt a kondenzátorban tárolt feszültség nem változhat meg. Természetesen vannak veszteségek. A kondenzátor nem képes egyik

pillanatról a másikra feltöltődni - a kondenzátor feltöltődési idejét nevezik beállási időnek. Miután a kapcsoló nyit a kondenzátor az áramkör szórt ellenállásain keresztül igen lassan kisül. A mintavevő és -tartó áramkör minőségét a beállási idő és a kisülés határozzák meg (10. ábra).

A töltőellenállás - ami a vezetékek és a kapcsoló szórt ellenállásából tevődik össze - és a tároló kondenzátor aluláteresztő szűrőt alkot, ami befolyásolja a frekvenciamenetet.

1.3.4. A/D átalakítás (kvantálás)

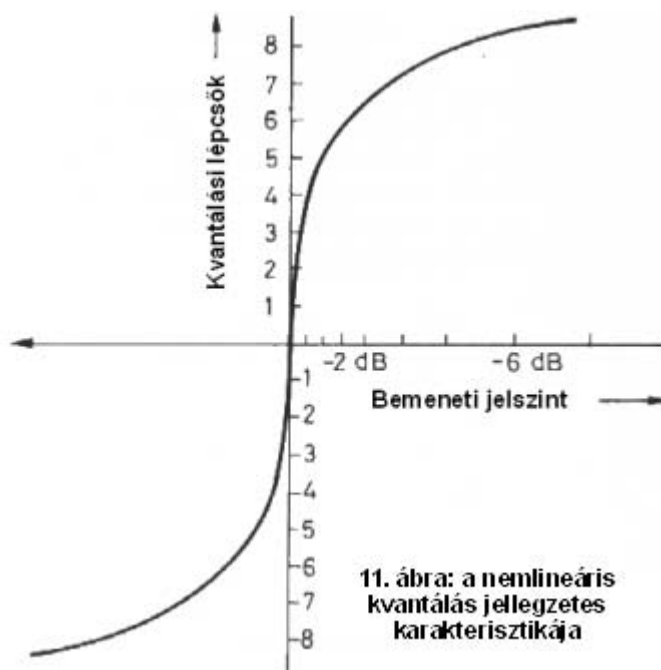
A mintavevő és -tartó kimenetén megjelenő minták még végtelen sok értéket vehetnek fel. A kvantáló feladata a jel értékkészletének diszkrétté tétele. Azaz a végtelen sok értéket véges sok érték segítségével kell ábrázolnia, ugyanis csak meghatározott számú bináris adatszó áll rendelkezésre attól függően, hogy hány bittre kvantálunk. Általában 16 bites kvantálást alkalmaznak, ami 65536 amplitúdóérték megkülönböztetését jelenti. A kvantálót is a mintavevő jel vezérli.

A kvantálás lehet lineáris és nemlineáris. Az előbbi esetben a kvantálási lépcsők azonos nagyságúak. A jelben azonban sokkal gyakoribbak a kis amplitúdójú minták és így a kisebb helyiértékű bitek sokkal jobban kihasználtnak. Emiatt a kvantálási bitek számát csökkentik és a bemeneti jelet egy dinamika kompresszorral vezetnek, ami a jel dinamikáját lecsökkenti, így alkalmassá válik a kevesebb bittel történő ábrázolásra. Visszaalakítás során a jelet egy dekompreszorral vezetnek keresztül, melynek karakterisztikája ellentétes a kompresszoréval, így az eredeti dinamika visszaállítható (11. ábra). Kialakításukat tekintve megkülönböztethetők integráló rendszerű és szukcesszív approximációs elven működő A/D átalakítók

Az A/D átalakítók hibái:

Kvantálási zaj: mivel a kvantáló csak véges számú értéket különböztet meg, a végtelen sok értéket felvehető minta értéke esetleg két kvantálási szint közé esik. Ebben az esetben a kvantáló dönti el, hogy a minta melyik szinthez tartozik, azaz +/-1/2 lépésközi hibát visz be a jelbe. Ezt a hibát nevezik kvantálási zajnak.

Granulációs zaj: ha a kvantálásra kerülő jel



11. ábra: a nemlineáris kvantálás jellegzetes karakterisztikája

amplitúdója olyan kicsi, hogy jellemzéséhez az LSB (Least Significant Bit = legkisebb súlyú bit) is elegendő, akkor a szinuszjel négyszögjellé torzul. A négyszögjel azonban minden frekvencia összetevőt tartalmaz, így tartalmazhat olyan összetevőket is amelyek miatt aliasing jön létre. Ez a torzítás sokkal erősebben érzékelhető, mint a nemlineáris torzítás. A granulációs torzítás értéke az amplitúdó csökkentésével növekszik.

Linearitási hiba: ideális esetben az A/D átalakító karakterisztikája lineáris, azaz lineárisan növekvő bemeneti feszültség esetén az A/D átalakítóhoz csatlakoztatott D/A átalakító kimenetén egyenletes lépcsőzésű jel jelenik meg. Az egyenletes lépcsőzéstől való eltérés az abszolút nemlinearitás, az ideális karakterisztikától való két mintavételi érték közötti eltérés pedig a differenciális nemlinearitás. Mindkettőt az LSB sokszorosaként, vagy töredékeként adják meg (12. ábra).

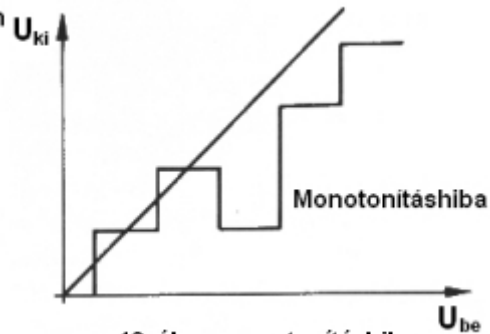
Monotonitáshiba: a lineárisan növekvő bemenő jel ellenére a D/A átalakító kimenetén a lépcsős jel folytonossága megszakad (13. ábra). Nagy torzítást okoz.

Ofszethiba: az egyenfeszültségi szint eltolódása. Hatására az A/D átalakító kimenetén is eltolódnak a bináris értékek, illetve lecsökken a kivezérelhetőségi tartomány. Az alkalmazott műveleti erősítők pontos beállításával kiküszöbölhető.

Túlvezérlési hiba: az A/D átalakító túlvezérlése esetén a kimenetén vagy a maximális érték jelenik meg, vagy annak komplemente. Ebben az esetben az A/D átalakító teljesen meghamisítja a bemeneti jelet.



12. ábra: linearitási hiba



13. ábra: monotonitáshiba

1.3.5. Kódolás

A kódolásra azért van szükség, mert az A/D átalakító kimenetén megjelenő bináris kódszavak csak a minta értékét adják meg, az előjeléről nem tartalmaznak információt. Általában kettes komplementes kódot alkalmaznak: a pozitív számokat normál bináris kóddal ábrázolják (LSB a jobb oldalon áll), a negatív számokat pedig az adott számérték kettes komplementesével. Egy szám kettes komplemente a szám bitenkénti invertálásával és +1 hozzáadásával képezhető. A következő példában 8 különböző amplitúdóértéket különböztetünk meg. Ehhez kettes komplementes kódban négy bitet alkalmaznak, mert az előjelet plusz egy biten - az MSB-n (Most Significant Bit = legnagyobb súlyú bit) ábrázolják. Nézzük pl. a 3 és a -3 kettes komplementes kódját:

$$3_{(10)} = 0011_{(2)}$$

$$-3_{(10)} = \text{inv}0011_{(2)} + 1 = 1100_{(2)} + 1 = 1101_{(2)}$$

Az előjelbit (MSB) pozitív számok esetén 0, negatív számok esetén pedig 1.

A kettes komplement kódolású jel már a PCM jelnek felel meg. Ezt a jelet nevezik forráskódnak, magát a folyamatot pedig forráskódolásnak. A PCM jel az A/D átalakító kimenetén párhuzamos, vagy soros formában jelenik meg. Az egyes bitek között a jelszint a bitek értékétől függően a 0, vagy az 1 szinten marad. Ez az NRZ kódolás (Non Return to Zero).

1.3.6. Hibafelismerés, hibajavítás

A hibák detektálásának illetve javításának nagy szerepe van a megbízható digitális átvitel kialakításában. A digitális jel az átviteli csatornán ugyanúgy ki van téve a csatornát érő zavaró hatásoknak, mint az analóg jel. A különbség az, hogy amennyiben a digitális jel eltorzul (olyan mértékben, hogy a határolással nem állítható helyre az eredeti négyzögjel), akkor megváltozik az információ tartalma. A hiba tulajdonképp egy vagy több bit invertálódását jelenti. A csatornát érő zavarimpulzusok - hosszuktól függően - több kódszót is tönkretelhetnek. Ez az ún. burst error.

Ahhoz, hogy a hiba javítható legyen először is fel kell ismerni a hiba tényét, helyét és jellegét. Ezután az alkalmazott hibajavítási eljárástól függően újrakérlik a hibás részt, vagy ha a hibafelismerő/-javító eljárás olyan, hogy önmagában is elegendő információt szolgáltat a hiba javításához, akkor a rendszer egyszerűen kijavítja a hibát.

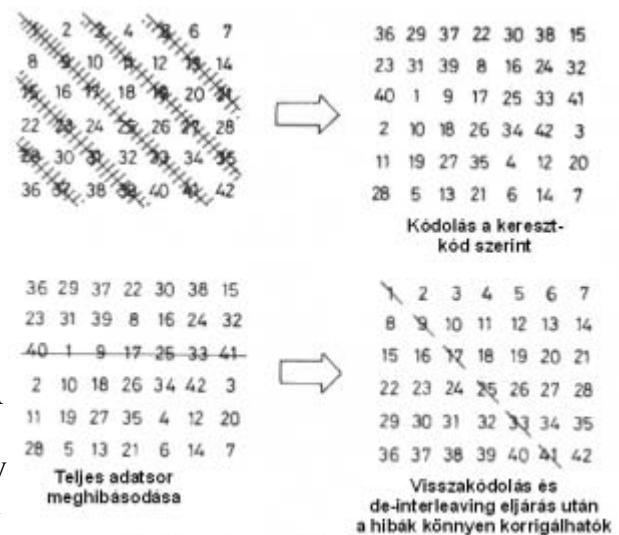
Az elterjedt hibafelismerő és -javító eljárások:

- Paritásbit.
- Hosszirányú redundancia ellenőrzés (LRC = Longitudinal Redundancy Check). Az egyszerű paritásbites eljárás továbbfejlesztése.
- Ciklikus redundancia ellenőrzés (CRC = Cyclic RC).

Ezekről az eljárásokról részletesen ír Andrew S. Tannenbaum: Számítógépes hálózatok c. művében.

1.3.7. Hibaelfedés

Egyetlen hibajavító eljárás sem képes minden hiba kijavítására. A hibaelfedés pszichoakusztikai jellemzőkre épül. A hibák miatti hiányzó adatokat a hibás kódszava(ka)t megelőző és követő kódszavak magasabb rendű interpolálásával nyerik. Az így kapott kódszóval a hibás részt elfedik (maszkolás). Ha egymás után több kódszó is megsérül, akkor a hibaelfedés nem működik helyesen. A hatékonyság a kódszavak tér- és időbeli összekeverésével növelhető (interleaving). Így a kódszavak nem a keletkezésük sorrendjében továbbítódnak. A helyes sorrendet a dekódolás során vissza kell állítani. Az eljárás következtében nagyobb lesz a valószínűsége, hogy egy hibás kódszót hibátlan kódszavak fognak közre. Az interleaving eljárás egyik speciális formája a cross interleaving (14. ábra).



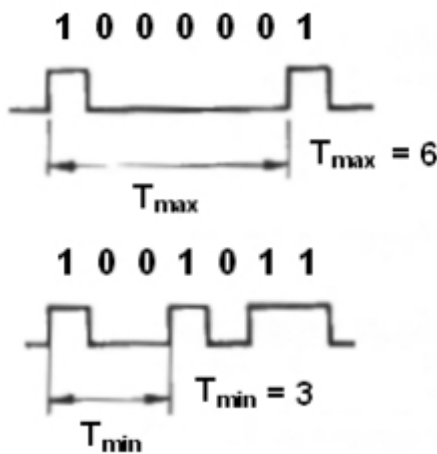
14. ábra a cross interleaving eljárás

1.3.8. Csatornakódolás, csatornamoduláció

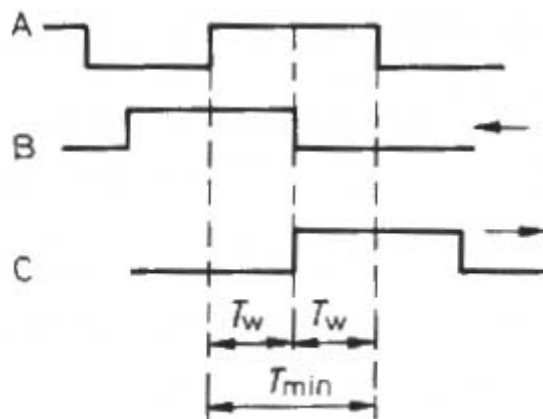
A forráskódolású jelet az átviteli csatorna jellemzőihez a csatornakódolás során illesztik. A csatornakódolásra két okból van szükség: egyrészt a csatornán felléphetnek zavarok, ami miatt meghibásodik az átvitt digitális jel. A csatornakódolás segítségével elősegíthető a hibák egyszerűbb javítása. A második ok a bitek megfelelő időben történő kiolvasása. Ha nem szinkron átvitelről van szó, akkor a csatornakódolású jelből elő kell tudni állítani az eredeti órajelet is, illetve az átvitt jel nem tartalmazhat egyenfeszültségű, vagy kisfrekvenciás összetevőket. Ezek a követelmények a forráskódolású jelbe való járulékos éltámenetek bejuttatásával elégíthetők ki.

A csatornakódok jellemzői:

- Futási hossz (T): két szintváltás közötti távolság bitekben. Az előre definiált futási hosszal rendelkező kódok az ún. RLL kódok (Run Length Limited). Az RLL kódok a két élváltás közötti minimális (T_{\min}) és maximális (T_{\max}) futási hosszal jellemezhetők (15. ábra).



15. ábra: futási hossz

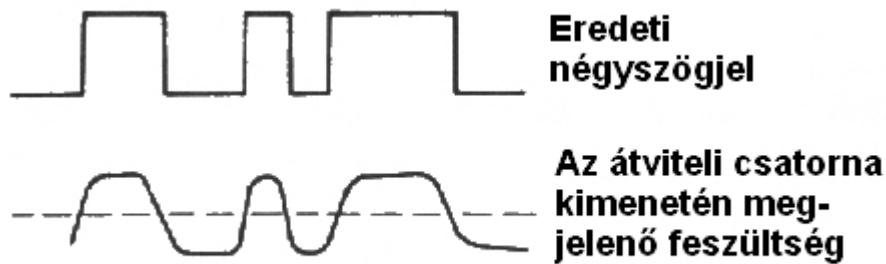


16. ábra: fázishiba

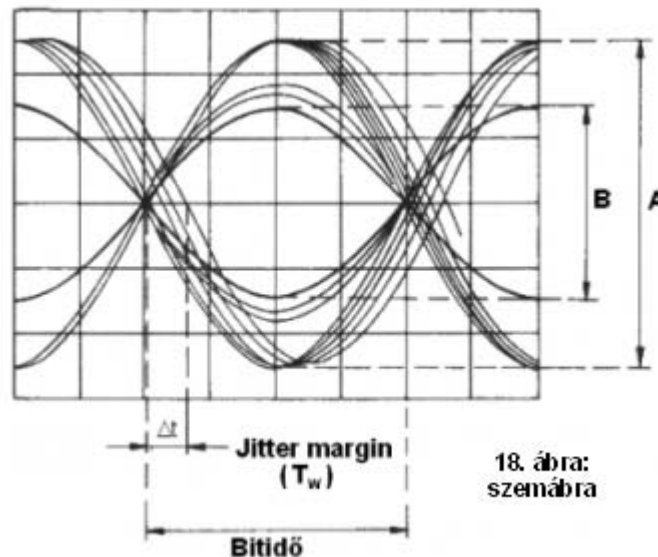
- Jitter margin (fázishiba - T_w): egy intervallumot (időablakot) jelöl ki. Soros adatjel kiértékelésénél van jelentősége. A jitteres jelet az órajelek megfelelő ütemben olvassák ki. Ha a jitter egy bizonyos szint fölé emelkedik, akkor a biteket hibásan detektálják (16. ábra). Az a tartomány, amin belül a detektálás még biztosan helyesen történik a jitter margin. Egy bit kifogástalan kiértékeléséhez a T_{\min} -től függő idő áll rendelkezésre:

$$T_w = \pm 1/2 T_{\min}$$

- DSV (Digital Sum Value): a csatornakód egyenfeszültségű komponensére jellemző összeg. A H szintű biteket +1-el, az L szintűeket -1-el helyettesítve minden bitet összeadnak - ez a DSV. Amennyiben ez 0, akkor a jel nem tartalmaz DC összetevőt.



17. ábra: az átviteli csatorna kimeneti jele



18. ábra: szemábra

- Szemábra: a sztingadozások, jitterhibák és egyéb zavarok, zajok miatt a csatornakódolt jel eltorzítva jelenik meg az átviteli csatorna kimenetén (17. ábra). Ha ezt a jelet egy oszcilloszkóp függőleges bemenetére vezetjük, és a vízszintes bemenetet az órajellel hajtjuk meg, akkor kapjuk a szemábrát (18. ábra). A szemábrán T_w -nek két jelszintátmenet közötti távolság felel meg. A maximális szemnyílás jelenti a maximális amplitúdót. Vízszintes irányban a növekvő fázishiba korlátozza a szemnyílást.

1.4. A PCM jel visszaalakítása analóg hangfrekvenciás jellé

A PCM jelet időben és értékben újra folytonossá kell tenni. Ezt a következő úton valósítják meg: először D/A átalakító segítségével PAM (Pulse Amplitude Modulation) jelet hoznak létre. Ezután a PAM impulzusok szélességét mintavevő és -tartó segítségével megnövelik, végül egy aluláteresztő szűrő segítségével integrálással nyerik az eredeti analóg jelet. Az aluláteresztő szűrőnek az integrálás mellett még egy fontos szerepe van: a PCM jel PAM jellé alakítása a D/A átalakítóval egy újabb mintavételi folyamatot jelent és így az alapsávi spektrum mellett ismét megjelenik a zavarospektrum, amit ki kell szűrni.

A bináris számsorokból álló PCM jel egyes kódszavaihoz a D/A átalakító az egyes kódszavak információtartalmával arányos feszültség értéket rendel hozzá. A legkisebb feszültség értékét az átalakító áramköri felépítése határozza meg. A maximális kimeneti feszültség:

$$U_{\text{kimax}} = U_{\text{LSB}} * 2^n$$

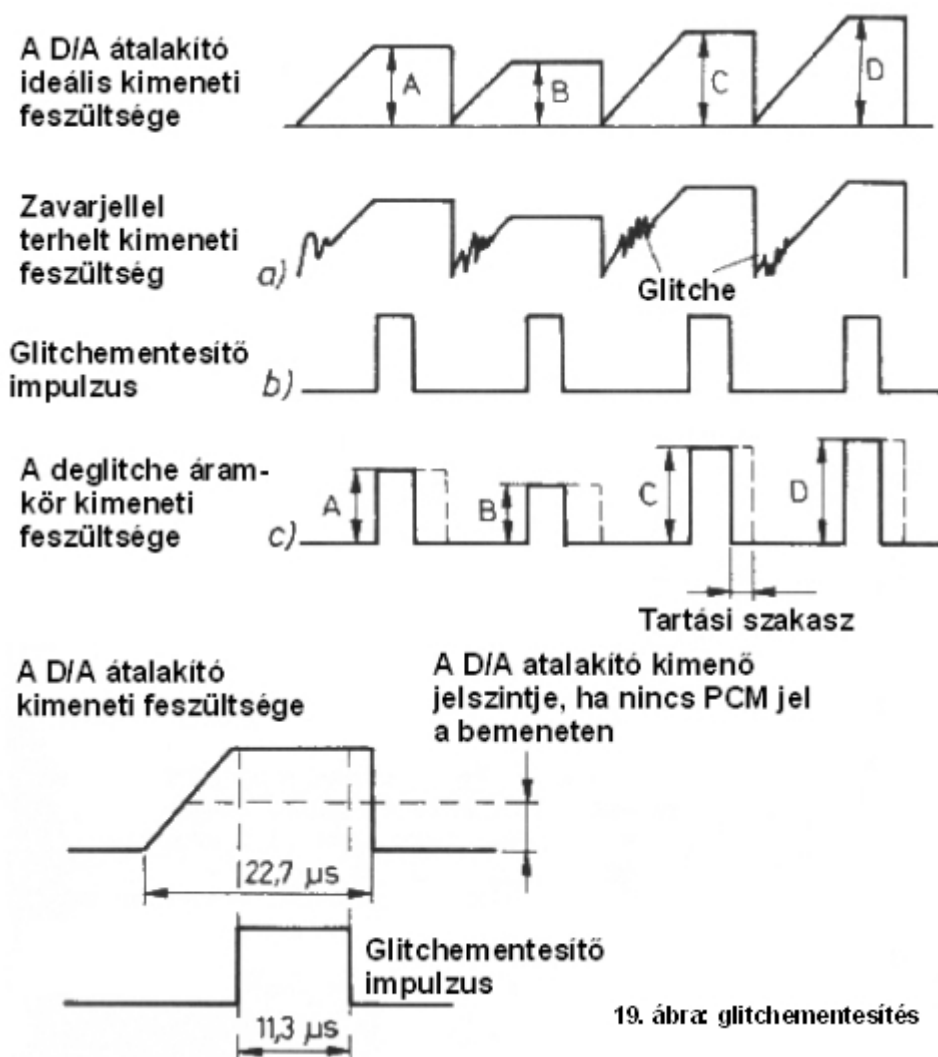
A D/A átalakítók felépítésük alapján lehetnek:

- létrahálózatot alkalmazó passzív áramosztósak.
- aktív áramosztósak
- integráló rendszerűek

A D/A átalakítók esetében is felléphetnek a fent ismertetett hibák. Ezenkívül még a beállási idő jellemzi a D/A átalakítókat, ami azt adja meg, hogy mennyi idő kell ahhoz, hogy a D/A átalakító bemenetére adott jel átalakításakor a kimeneti feszültség $\pm 1/2$ LSB pontossággal beálljon annak értékére.

A D/A átalakítóban áramgenerátor és a PCM jel által vezérelt félvezető kapcsolók segítségével történik a kódszavak feszültséggé alakítása. A kapcsolók nem egyidejű átkapcsolásának következménye az ún. glitche. Ez a beállási szakaszon látható zavarójel (19. ábra). Ezért a D/A átalakító kimeneti jelét glitchementesítő áramkörre vezetik, amelynek működése szintén a 19. ábrán látható. A glitchementesítő impulzus szélessége általában a mintavételi intervallum felével egyenlő.

A D/A átalakítók pontossága még az alkalmazott referenciafeszültség pontosságától is függ.



19. ábra: glitchementesítés

2. A DAT (Digital Audio Tape) rendszer

2.1. Bevezetés

A hagyományos, analóg mágnesszalagos hangrögzítés a CD rendszerű digitális rendszerhez képest sokkal rosszabb hangminőséget produkál. Emiatt került sor egy digitális hangrögzítési eljárást alkalmazó, mágnesszalagos rendszer kifejlesztésére. Az 1983 nyarán megrendezett Nemzetközi DAT Konferencia célja egy egységes szabvány létrehozása volt a digitális mágnesszalagos hangrögzítésre.

Két rendszer létezik: az S-Dat (Still head) és az R-DAT (Rotary head). Az előző a hagyományos audio magnóhoz hasonlóan álló fejjel, hosszanti irányú sávokba rögzített audio jelet használ. A második forgó fejjel rendelkezik és a hanganyagot ferde sávokban rögzítik a szalagra - hasonlóan a video magnóhoz. Az R-DAT rendszer bizonyult jobbnak és 1987-ben a nyilvánosság előtt bemutatták az első példányt.

2.2. Az R-DAT rendszer jellemzői

2.2.1. Általános jellemzés

A digitális jelek mágnesszalagos rögzítésekor ugyanazok a fizikai törvényszerűségek érvényesek, mint analóg rögzítés esetén, azzal a különbséggel, hogy a PCM jel sávszélesség igénye 40 - 100-szorosa az analóg jelének. Ezt a megnövekedett sávszélesség igényt csak a szalagsebesség növelésével lehet biztosítani, ami viszont drasztikusan csökkenti a műsoridőt. Emiatt a video technikában már elterjedt forgó fejdobos, ferde sávos jelrögzítési technikát alkalmazzák.

Jellemzők:

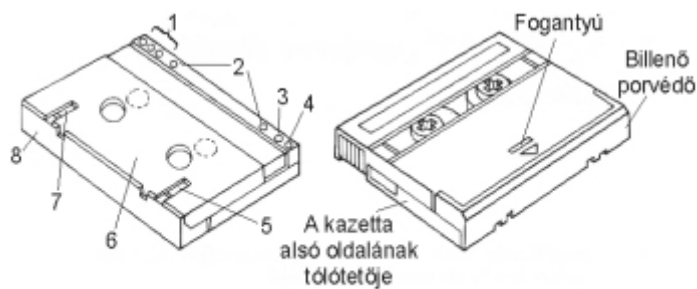
- a hagyományosnál kisebb kazetta méret
- 2-4 órás lejátszási idő
- gyorskeresési- és program funkció
- különleges funkciók
- analóg- és digitális be- és kimenetek.

A használt PCM formátumok:

- 48 kHz -es, 16 bites minták (lineáris kvantálás)
- 44,1 kHz -es, 16 bites minták (lineáris kvantálás)
- 32 kHz -es, 12 bites minták (nemlineáris kvantálás), vagy 16 bites (lineáris kvantálás).

Jogvédelmi okok miatt a kommersz készülékekből kihagyták a 44,1 kHz -es mintavételi frekvenciával történő felvétel készítés lehetőségét és egy speciális másolásvédelmi kódolást alkalmaznak (SCMS). (Természetesen léteznek stúdió célokra készült DAT magnetofonok, amelyek támogatják a 44,1 kHz -es mintavételezést is).

A DAT magnetofonok mechanikájának vezérléséről szervo áramkörök gondoskodnak.



20. ábra: az R-DAT kazetta kialakítása

2.2.2. Az R-DAT kazetta jellemzői

Az R-DAT kazetta kialakítása az 20. ábrán látható. Hasonlít a 8 mm -es videokazettára, de méretei eltérőek: 73 x 54 x 10,5 mm (hossz, szélesség, vastagság). Az alsó oldalon három azonosítási furat található a szalagfajta jelölésére és a felvétel védelmére. A szalagot a kazetta elején elhelyezett

fedél védi az érintéstől és a portól. A szalag elejét és végét egy LED -ből és fototranzisztorból álló áramkör figyeli a kazettában elhelyezett plexi prizmák segítségével, amelyek a LED fényét a fototranzisztorra vetik.

A 20. ábrán látható jelölések:

1. érzékelő nyílások a szalagfajta és a sáv szélesség felismerésére
2. illesztő furatok
3. érzékelő nyílás a törlés letiltására
4. érzékelő nyílás a gyári műsoros kazetták felismerésére
- 5., 7. vezető horony rögzítő pöccökkel
6. tolotető a szalagorsó-agyak védelmére
8. billenő porvédő

2. táblázat: az érzékelő nyílások értelmezése (1 = nyitott, 0 = zárt).

Érzékelő	1	2	3	
	0	0	0	13 mikrom-es szalagvastagság
	0	1	0	13 mikrom-es szalagvastagság
	0	0	1	13 mikrom-es szalagvastagság, megnövelt sáv szélesség
	0	1	1	13 mikrom-es szalagvastagság, megnövelt sáv szélesség
	1	x	x	Foglalt, jelenleg nem definiált

2.2.3. A mágnesszalag

Az adatsűrűség 94 kbit/mm, ezen kívül a szalagnak a nagy mechanikai igénybevételt is állnia kell, mert gyors kereséskor 200-szoros sebességgel halad el a fej előtt. Emiatt csak megfelelően jó minőségű szalag használható. Az R-DAT kazetta szalaganyaga igen egyenletes részecske szerkezetű fémporéteggel van bevonva. A szalag 3,8 mm széles, 13 mikrom vastag, remanenciája 260 mT, koercitív ereje 120 KA/m.

A gyári műsoros kazetták ún. kontaktmásolással készülnek. Szalaganyaguk az üres kezettákétól eltérő: BaFe, mert ennek az anyagnak kisebb a koercitív ereje.

2.2.4. Az R-DAT rendszer műszaki paraméterei

Négy különböző felvétel készítési üzemmód létezik:

1. Normál üzemmód: 48 kHz, 16 bites kvantálás.
2. Digitális rádióadások rögzítése: 32 kHz, 16 bites lineáris kvantálás.
3. Longplay üzemmód (felezett szalagsebesség): 32 kHz, 12 bites nemlineáris kvantálás.
4. Négy csatornás felvétel: 32 kHz, 12 bites nemlineáris kvantálás.

A stúdió célú DAT magnók esetén megtalálható még a 44,1 kHz -es, 16 bites lineáris kvantálású felvétel készítési lehetőség is.

További üzemmódot jelent a 44,1 kHz -es, 16 bites lineáris kvantálású hanganyagok lejátszása.

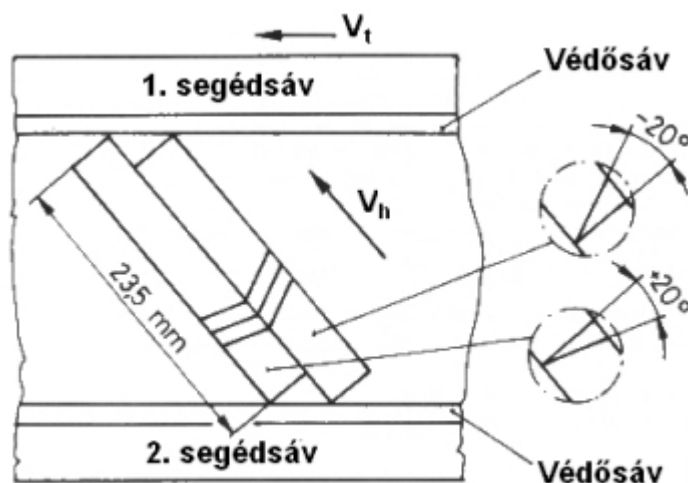
A rendszer további paramétereit az 3. és a 4. táblázat tartalmazza.

3. táblázat: általános adatok.	
Jellemzők	
Frekvenciamenet	5 Hz - 22 kHz + 0,5 dB (normál üzemmód esetén) 5 Hz - 14 kHz + 0,5 dB (long-play üzemmód esetén)
Jel-zaj viszony	90 dB (előkiemelés - utóelnyomás van; normál üzemmód) 88 dB (előkiemelés - utóelnyomás van; long-play üzemmód)
Rendszerdinamika	90 dB (előkiemelés - utóelnyomás van; normál üzemmód) 88 dB (előkiemelés - utóelnyomás van; long-play üzemmód)
Harmónikus torzítás	0,005 % - 1 kHz, normál üzemmód 0,05 % - 1 kHz, long-play üzemmód
Szinkronizálási hiba	nem mérhető
Hozzáférési idő keresési funkcióban	15 s (2 órás kazettánál)
Analóg bemenet	640 mV, 47 ohm (teljes kivezérlésnél)
Analóg kimenet	3 V, 2 ohm (teljes kivezérlésnél)

4. táblázat: az egyes üzemmódok adatai.					
Üzemmódok	1.mód	2.mód	3.mód	4. mód	Műsoros kazetták
Hangcsatornák száma	2	2	2	4	2
Mintavételi frekv.(kHz)	48	32	32	32	44,1
Kvantálás	16 bites, lineáris		12 bites, nem lin.		16 bites, lineáris
Kódolás	2 -es komplementes kód				
Szalagsebesség (mm/s)	8,15	8,15	4,08	8,15	8,15
Adatátviteli seb. (Mbps)	2,46	2,46	1,23	2,46	2,46
Hibajavítás	Kettős Reed - Solomon kód				
Redundanciatényező (%)	37,5	58,3	37,5	37,5	42,6
Csatornakódolás	MNRZ-I m=8 n=10 (Eight-to-Ten Modulation)				
Max. felvételi idő (min)	120	120	240	120	120

2.2.5. Sávfekvés

A szalag futási irányához képest a forgó fejdob ferdén áll. A fejek a fejdob ellentétes oldalán helyezkednek el, tehát 180 fokra egymástól. A fejdob 30 mm átmérőjű és 2000 1/min fordulatszámmal forog. A szalag a fejdob előtt 8,15 mm/s -os sebességgel halad el. A szalag futási irányához képest a fejdob 6 fokkal dől, így a fej és a szalag között 3,133 m/s -os relatív sebesség jön létre. A 7,5 Mbps -es adatátviteli sebesség elérése csak így módon lehetséges.



21. ábra: a sávok elhelyezkedése a szalagon

A ferde sávok szélessége 13,5 mikrom, hossza 23,5 mm és közvetlenül egymás mellé veszik fel őket (nincsenek védősávok). Az áthallások elkerülése végett a két fej között ún. azimut eltolást alkalmaznak: a fejrések egymáshoz képest +/- 20 fokkal elforgatott helyzetűek. Az

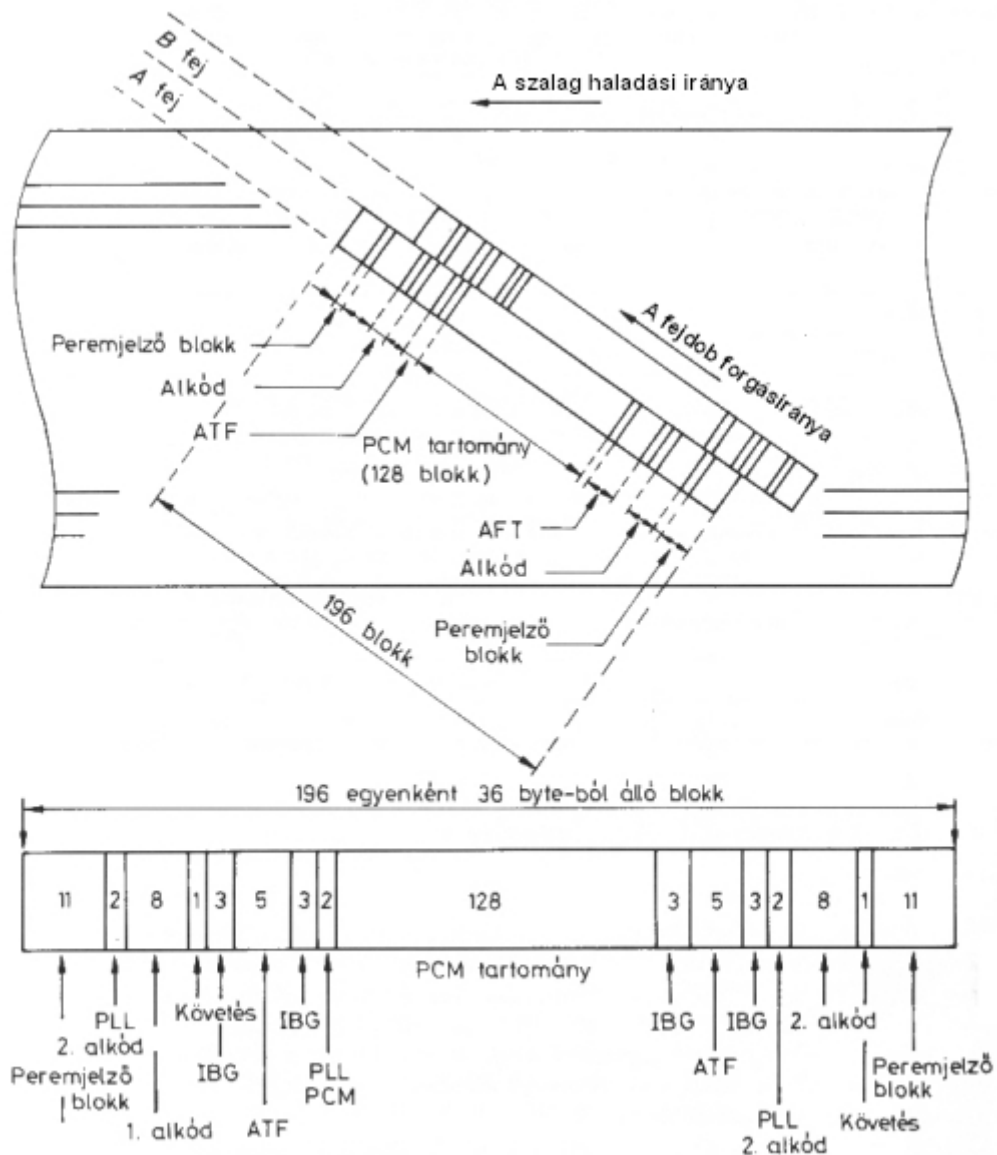
adott sávot csak a megfelelő résferdeségű fej képes olvasni. A két fej közötti +/-20°-os azimuteltolás tűrése +/-15'. A ferde sávokat a szalagon hosszanti irányban két segédsáv határolja, melyeknek különleges funkcióik vannak. A fent leírtakat illusztrálja a 21. ábra.

2.2.6. Felvételi formátum

A rögzített adatok blokkokra oszlanak. Sávonként 196 adatblokk található. Az egyes adatblokkok 36, egyenként 8 bites szimbólumból állnak. A PCM információt hordozó keretek mellett léteznek az ATF rendszernek, alkód információknak, szinkronizálási célokra és a rendszeradatok elválasztására fenntartott keretek.

A sávok kezdetét és végét 11 peremjelző blokk határolja. A PCM kódolású hangadatok a redundancia- és a hibajavító bitekkel együtt 128 blokkot foglalnak el. A különböző formátumú keretek (ATF, alkód, PCM adat) ún. IBG blokkokkal vannak elválasztva (Inter Block Gap), amelyeknek a különböző formátumú keretek közötti áthallás kiküszöbölése a feladatuk. Az IBG-ben $f_{ch}/6$ frekvenciát rögzítenek ($f_{ch} = 9.408$ MHz - ez a csatornafrekvencia). Így egyes blokkok szükség szerint törölhetők, vagy felülírhatók.

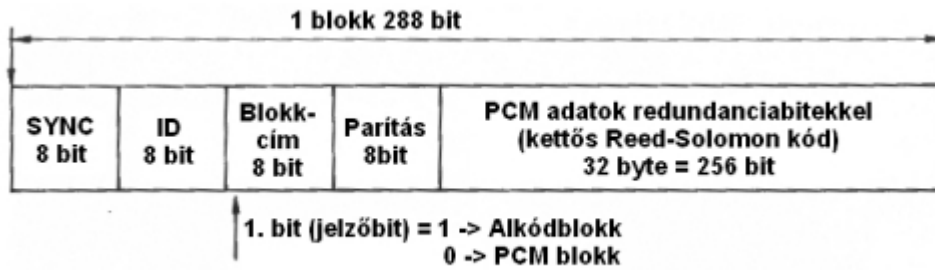
Az alkód és a PCM tartományban mindig egy PLL jelet helyeznek el az információ elé, aminek a frekvenciája $f_{ch}/2$. Ez segíti a lejátszóoszillátor szinkronizálását. A PLL a preamble (előjelző) jel. A szubkód adatokat egy postamble követi, ami szintén $f_{ch}/2$ frekvenciájú jelet tartalmaz, de hossza kb. a fele a PLL hosszának. A sávok blokkfelosztása a 22. ábrán látható.



22. ábra: a sávok blokkfelosztása

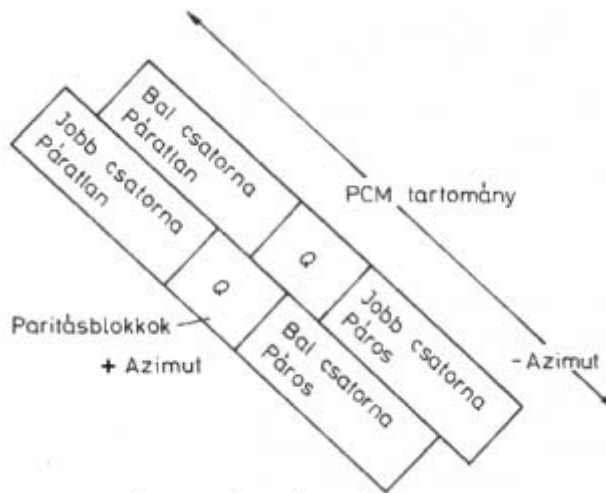
2.2.7. PCM formátum

A PCM kódolású, redundancia bitekkel ellátott hangjel 128 db., egyenként 36 szimbólumot tartalmazó blokkot foglal el egy ferde sávon. Egy blokkból a PCM adatok 32 db. 8 bites szimbólumot foglalnak le. A maradék 4 szimbólum szolgál a szinkronizálás, az azonosítás, a paritásvizsgálat (Q) és a blokkcímezés adatainak tárolására. A PCM formátum a 23. ábrán látható.



23. ábra: a PCM formátum

A megfelelő üzemmód kiválasztása az ID információk alapján történik. Az A/D átalakító által előállított forráskódolású jelet a rögzítés előtt csatornakódolásnak kell alávetni. Ez a megbízható rögzítés és hibafelismerés ill. -javítás miatt szükséges, valamint a forráskódolású jelet még a mágneses jelrögzítés sajátosságaihoz is illeszteni kell. A normál összekeverés mellett (interleaving) a keresztirányú összekeverést is alkalmazzák (cross interleaving). Így a PCM blokkokat páros és páratlan blokkcímek szerint különböztetik meg és így rögzítik a szalagra. (24. ábra) Így nagyobb kiterjedésű hiba is kijavítható.



24. ábra: a páros és a páratlan blokkok

2.2.8. ATF formátum

ATF = Automatic Track Following. Az automatikus sáv letapogatást biztosítja. Mivel a ferde sávok között nincs védősáv, így az ATF szervó áramkörökhöz szükséges vonatkoztatási jelek rögzítésére sem használnak külön sávot. Ehelyett a szükséges jelet a ferde sávokból kell előállítani. A ferde sávon két tartományt alakítanak ki, amikre azonosítás céljából időben elkülönített frekvenciacsomagokat rögzítenek (ATF frekvenciák), így pontos követés érhető el még sávtorzulás esetén is. Ezek azt adják meg, hogy páros, vagy páratlan

adatblokk-sorozatot tartalmaz a letapogatott adatsáv. Az egyes zónákban különböző ATF frekvenciák találhatók:

- f1: pilotjel (130,67 kHz)
- f2: szinkronjel (522,67 kHz)
- f3: szinkronjel (784 kHz)
- f4: törlőjel (1,56 MHz)

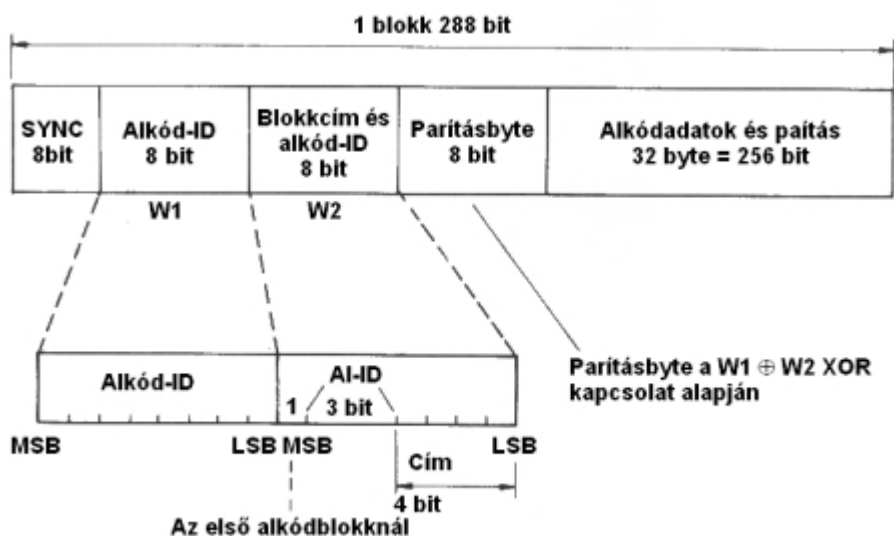
A frekvenciacsomagok sávonként más elrendezésűek, a séma négy sáv után ismétlődik.

2.2.9. Alkód formátum

A szalagon az audio adatok mellett kiegészítő információk is tárolhatók alkód formájában. Alapvetően kétféle alkód létezik:

1. a felhasználó által definiált alkód. A ferde sáv mindkét alkód zónájában megtalálható (8-as blokkok a 22. ábrán). A PCM adatoktól függetlenül változtatható, adatkódra és jelzőkódra osztható (Lsd. 5. táblázat).
2. a blokkokon belüli alkódot az ID szimbólumok képezik. A tartalma a PCM adatoktól függetlenül nem változtatható meg.

A felhasználói alkód blokk szintén 36 szimbólumból áll, melyből 32 ténylegesen az alkód adat, 4 pedig a szinkronizáció, címzés, ID - felismerés, paritásvizsgálat céljait szolgálja (25. ábra).



25. ábra: a belső alkódformátum

A belső (felhasználó által nem módosítható) azonosítók a PCM tartományban a PCM adatokkal együtt kerülnek rögzítésre. 8 különböző PCM azonosító értelmezett (ID 0-7). Mindegyik 2 bites, a rögzítés együtt történik a PCM azonosító keret (PCM-ID-frame) címével 8 bites szavakban (W1). 8 PCM azonosítóból áll egy PCM azonosító keret. Minden W1 szóban a cím 4 bitet foglal el, ezért 4 bit marad az azonosítók számára, vagyis szavanként 2 azonosító rögzíthető. Egy sávon belül a 128 PCM blokkból minden második tartalmaz PCM azonosítókat, így egy sávon 16 PCM azonosító rögzíthető. A hozzátartozó blokkcím a W2 adatszóban kerül átadásra. A belső azonosítók jelentését a 6. táblázat tartalmazza.

5. táblázat: példa a felhasználói alkód alkalmazására.		
Abszolút idő	Az eltelt műsoridő kijelzése a kazetta elejétől számítva.	Jelzőkódok
Program idő	Az eltelt műsoridő kijelzése a zeneszám kezdetétől számítva.	
Index	Az eltelt műsoridő kijelzése az utolsó indextől számítva.	
Tartalom	A tárolt zeneszám címek, a teljes műsoridő kijelzése.	
Programsorszám	Az aktuális program sorszámának kijelzése.	
Start	A program startpozíciójának kijelzése.	Adatkódok
Skip	Ugrás utasítás a következő start pozícióra.	
End	Stop utasítás gyors előre tekerésnél, vagy lejátszásnál.	
TOC	Tartalomjegyzék kiírása.	

6. táblázat: a blokkon belüli alkódok.	
ID0	Formátumbesorolás (osztályozás). Audiojelek esetén 00. Lehetséges másfajta jelzés is pl. adatrögzítő rendszerekben. A 00-tól eltérő jelzés audio rendszerekben a némítást aktiválja.
ID1	A felvétel során alkalmazott kiemelés: 00 - kikapcsolva 01 - bekapcsolva, 50/15 mikros
ID2	Mintavevő frekvencia 00 - 48 kHz 01 - 44,1 kHz 10 - 32 kHz
ID3	Csatornák száma 00 - 2 csatorna 01 - 4 csatorna
ID4	A felbontás 00 - 16 bites, lineáris 01 - 12 bites, nemlineáris
ID5	Az alkalmazott sáv szélesség 00 - 13,591 mikrom 01 - 20,410 mikrom (műsoros, kontaktmásolással készült szalagok esetén)
ID6	A másolásvédelmi rendszer jelzései 00 - digitális másolás megengedett 10 - digitális másolás tiltott 11 - egyszeri digitális másolás megengedett
ID7	ez az ún. PACK-ID (csomag azonosító). 32 PACK ID 64 bitet tesz ki, és ez egy csomag (PACK). 16 különböző csomag definiálható, amelyből egyidejűleg 7 jelenhet meg, egy 8. lesz a paritáscsomag (Parity PACK), ami az előző 7 adatai alapján képződik.

Rögzíthetők még keresési- (Search Code -> SC) és segéd adatok (Auxiliary Code -> AC). Egy adott sáv minden 128 blokkjából minden második W1 adatszóban kap helyet, szavanként egy SC vagy AC. A PCM adatokkal együtt kerülnek rögzítésre, így a PCM adatok megváltoztatása nélkül nem változtathatók meg.

Az SC tartalmazza a felvétel számát, indexét, abszolút idejét, a szalag kezdet- és végfelismeréshez és a számok kezdetének felismeréséhez szükséges adatokat. 4 részre oszlik (SC1 - SC4).

Az AC is 4 részre oszlik. Az AC1 tartalmazza a 4. - 7. biteken a funkcióhozrendelés item kódját. Tárolható abszolút-, program-, és index idő, illetve jelzők (flagek), továbbá tartalomjegyzék (TOC), katalógusszám, nemzetközileg szabványosított felvételkedok és szoftverködok.

2.2.10. Az SCMS rendszer

Serial Copy Management System (soros másolás kezelő rendszer). A zenei adatok és a DAT felvételek veszteségmentes, digitális másolását szabályozza. Jelforrás lehet CD lejátszó, digitális rádió, vagy másik DAT magnetofon. A rendszer meggátolja a profi kalózmásolatok előállítását.

A kommersz készülékek digitális audio felülete mindig az SPDIF-nek megfelelő a kompatibilitás érdekében.

Az audio adatok mellett átmásolódnak kategória- és másolásvédelmi ismertetőjegyek, illetve az audio jel lejátszásához szükséges adatok. DAT felvételkor a 6-os PCM azonosító beállítása a kategória- és a másolásvédelmi ismertetőjegyek függvényében történik. Ezeket a funkciókat a felhasználó nem tudja befolyásolni, nem is kerülnek kijelzésre.

A szubkód azonosítók jelforrástól függően kerülnek átvitelre:

CD-ről DAT-ra nem viszik át a szubkódokat.

DAT-ról DAT-ra átvitelre kerülnek a START és SKIP azonosítók is.

DSR-ről nem viszik át a szubkódokat.

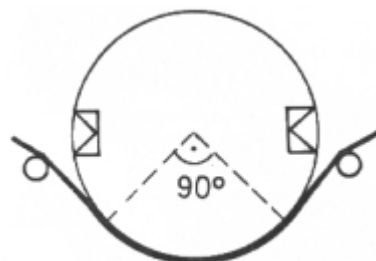
Általában egy programról egy db digitális másolat készítése megengedett. Ez a másolat az SCMS rendszer jelzőbitjein már az 10 ismertetőjegyet fogja hordozni. Ez vonatkozik a saját felvételekre is. Nincs lehetőség a digitális jelnek több készüléken keresztül történő végigláncolására sem. Forráskódolás segítségével akarják megoldani, hogy a saját felvétel esetén a másolásvédelem megszüntethető legyen.

Az 1. generációs DAT magnók nem tartalmazzák az SCMS rendszert, de ezekkel a készülékekkel gyakran nem is lehetséges közvetlen digitális másolat készítése CD-ről.

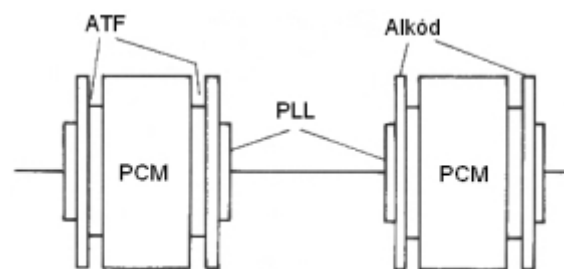
2.3. Az R-DAT magnetofon felépítése

A mechanikai felépítés hasonló a képmagnókéhoz, de lényegesen precízebb, a méretek kisebbek, a szalagfutás és -feszítés stabilabb. Nincs szükség törlőfejre, mert a régi adatok törlése felülírással történik.

Az alkódokkal megvalósított különleges funkciók megvalósításának feltétele, hogy a szalag még 100-szoros sebesség esetén is körülfogja a fejdobot. A nagy súrlódás csökkentése és a szalag élettartamának megnövelése érdekében a szalag csak 90°-os szögtartományban fogja át a fejdobot. (26.ábra). Vagyis a fejek a körülfordulási időnek csak az 50 %-ban érintkeznek a szalaggal. Emiatt a rögzítésre kerülő adatfolyamot időben komprimálni kell. Lejátszás folyamán a komprimálást megszüntetik és újra folyamatossá alakítják az adatfolyamot. A 27. ábrán látható a rögzítésre kerülő komprimált jel. Néhány hordozható készülékben a fejdob átmérőjét 15 mm-re csökkentették, ebben az esetben a fejdob körülfogási szög 180°. A szalagvezérlést két további, egymáshoz képest 90°-ra elhelyezett fejjel oldják meg.



26. ábra a szalag fejtőfogása



27. ábra az időben komprimált adatformátum

2.4. A jelfeldolgozás folyamata

2.4.1. A felvétel

Az analóg bemenetről történő felvétel készítésekor a csatornaelválasztás jobb, mint 80

dB 8 kHz-es frekvencián. Az analóg hangfrekvenciás jel sávhatárolását az anti-aliasing szűrő végzi. A vágás 15, vagy 20 kHz -nél történik meg a választott üzemmódtól függően. Ezután a mintavevő és -tartó áramkör segítségével létrehozzák a PAM impulzusokat. Az A/D átalakító ezeket kvantálja és az értékeket kettes komplementes kódban ábrázolja. A mintavevő és -tartó és az A/D átalakító fizikailag egy egységet alkotnak. Az A/D átalakító 32, 44,1 és 48 kHz -es mintavételi frekvenciájú jeleket is képes kvantálni. Az átkapcsolásról a folyamatvezérlő áramkör gondoskodik. Professzionális készülékek esetén lehetőség van a mintavételi frekvencia és az előkiemelés (pre-emphasis) szabad megválasztására. Kómersz készülékek esetén az előkiemelés mindig be van kapcsolva.

A jobb és bal csatorna jelét ezután egy TDM multiplexerre vezetik. 48 kHz -es mintavételi frekvencia esetén az adatátviteli sebesség:

$$48 \text{ kHz} * 2 * 16 \text{ bit} = 1,53 \text{ Mbps}$$

Ezután a multiplexelt jelet egy interfészre vezetik. Ide csatlakozik a digitális bemenet is. A digitális be-/kimenetek általában SPDIF előírás szerinti (Sony/Philips Digital Interface Format), amelyek az IEC 958-2 előírásainak megfelelőek. A digitális be-/kimenetek koax kábelesek, vagy fénykábelesek.

Ezután történik meg a blokkosítás és a hibajavítást szolgáló redundancia szimbólumok beiktatása. Az R-DAT rendszer egy speciális hibafelismerő és -javító eljárást használ: a kettős Reed-Solomon kódot, amelyet I. S. Reed és G. Solomon 1960-ban hoztak nyilvánosságra. A hibajavító fokozatok neve: C1- és C2-kódoló. A feldolgozásra kerülő szóhossz nem lehet nagyobb 8 bitnél. Minden hibajavító szón egy ECC-interleaving (Error Correction Code) műveletet hajtanak végre, majd az adatok még egy Cross-interleavingen mennek át.

A C1 kódolóban minden 28 szimbólumhoz 4 hibajavító szimbólumot rendelnek (P hibajavító kód). A hibajavító szimbólumok egy sávon belül két egymás melletti blokkba kerülnek (1 blokk 32 db. 8 bites adatszót tartalmaz).

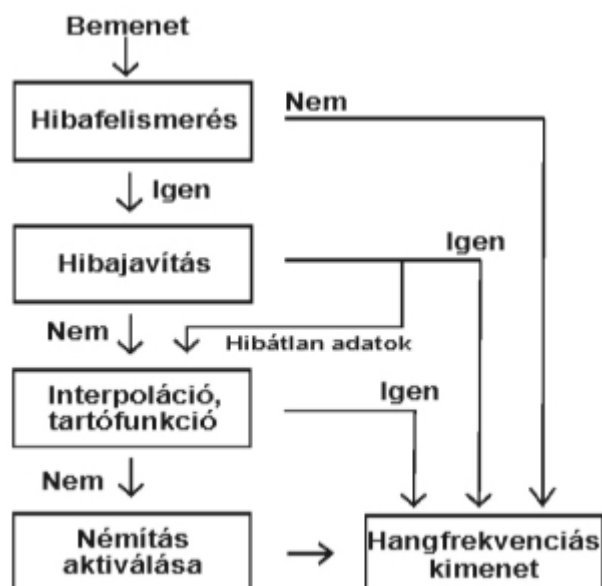
A C1 dekóder hibajavító képessége max. 4 szó, ha ismert a hiba helye, ellenkező esetben 2 szó.

A C2 kódolóban minden 26 szimbólumhoz 6 hibajavító szimbólumot rendelnek (Q hibajavító kód). A hibajavító szimbólumok egy sávon belül egymástól 4 blokknyira kerülnek elhelyezésre.

A C2 dekóder hibajavító képessége max. 6 szó, ha ismert a hiba helye, ellenkező esetben 3 szó.

Az ECC-interleaving elv a C1 és C2 hibajavító kódok egy sávon belüli elosztását teszi lehetővé.

Lejátszáskor először a PCM blokkok keresztbe összefésült szavai kerülnek helyreállításra, majd megtörténik a C1 és C2 dekóderek segítségével a hibafelismerés és hibajavítás. Azok a hibák, amelyek így módon nem javíthatók a hibaelfedés segítségével kerülnek "kijavításra". A teljes hibajavító rendszer működése a 28. ábrán látható. A hibajavító kapacitás összességében olyan nagy, hogy a lejátszás során egy fej teljes kiesése még nem hallható. Az adatok összekeverésének módja a választott üzemmód függvénye. A plusz információk mértéke a 4. táblázatban



28. ábra: a hibajavítás folyamata

látható. A hibajavítás miatt a hibajavító áramkör kimenetén az adatátviteli sebesség már 2,46 Mbps.

A következő lépés az alkód adatok beiktatása az adatfolyamba egy második multiplexer segítségével. A csatornakódolásra csak ezután kerülhet sor. Olyan csatornakódra volt szükség, amely a fej és a mágnesszalag által alkotott csatorna sávszűrő jellegéhez illeszthető. Ez az MNRZ-I $m=8$ és $n=10$ paraméterekkel (Eight-to-Ten Modulation). A bináris egy értéket a bitcella közepén bekövetkező élváltás jelenti. Az ATF frekvenciákat is a csatornakódolás során adják a jelhez.

A lehetséges 1024 db 10 bites kódszóból csak 512 kerül felhasználásra. Ezek egy táblázatban kapnak helyet (ROM), amely a 10 bites kódszón kívül tartalmazza a kódszó DC szintjét is. Ehhez a ROM-hoz kerülnek a 8 bites PCM- és szubkód adatok és ezek, illetve a DC szint alapján áll elő a 10 bites kódszó. A 8/10 moduláció nem okoz DC szint eltolódást. A 8/10 moduláció után a szavak NRZ-I átalakításon mennek át. $T_{\min} : T_{\max} = 1 : 4$ (két egyes között max. három 0 lehet). A kódolt csatornabitek impulzushossza:

min. $0,8T$ -től max. $3,2T$ -ig ($T = 1$ adatbit időbeli hossza).

$0,8T = 1 T_{ch}$ ($T_{ch} = 1$ csatornabit időbeli hossza).

Megállapodás szerint a kódolás során nem alakulhat ki olyan kód, amely egymás után kétszer a maximális impulzushosszat eredményezné, ez ugyanis a szinkron karakter számára van fenntartva.

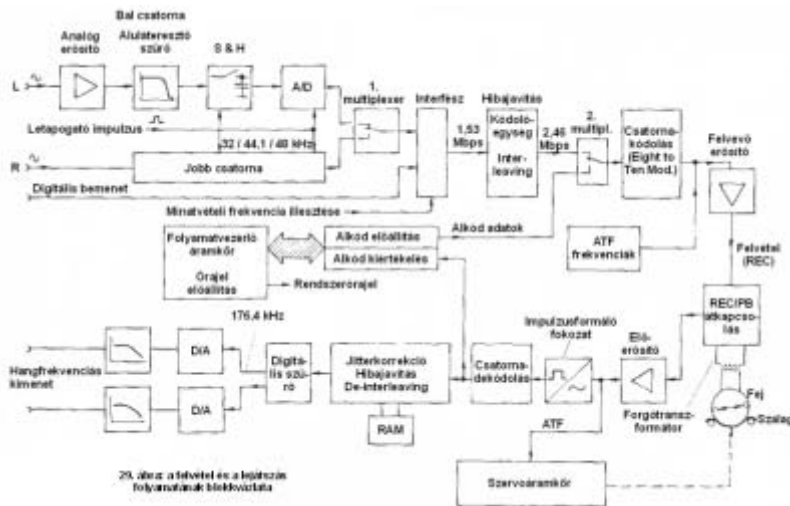
Az időbeli komprimálás miatt szükség van egy közbenső tárolóra, ahová a jelet felvételkor beírják és a komprimált jelnek megfelelő időközönként olvassák ki.

A végső adatátviteli sebesség 7,5 Mbps.

2.4.2. Lejátszás

Általában ugyanazokat a fejeket használják mint amit a felvételhez (kombinált fejek). Egyes High end, illetve profi készülékek külön felvevő és lejátszó fejekkel rendelkeznek, így lehetségessé válik a felvett anyagba történő belehallgatás már a felvétel alatt (monitoring). A fejdobon ebben az esetben 4 fej található, és minden csatornához külön kapcsolás szükséges a felvételhez és a lejátszáshoz is. Így lehetővé válik a hanganyag szerkesztése. A fejek által kiolvasott jeleket felerősítik és a fejátkapcsoló impulzus segítségével egyetlen hangfrekvenciás jellé alakítják. Az átviteli csatorna sávszűrő jellege miatt a jelelak szinuszos, amit impulzusformáló áramkörökkel négyszög jellé alakítanak. Következik az órajel regenerálás és az eredeti 8 bites blokkstruktúra helyreállítása. A fázisjitteres (szinkronizálási hibás) jelet tárolókba olvassák és a helyreállított órajel ütemében újra kiolvassák, majd elvégzik a hibajavítást. A hibajavító áramkör kimenetén már az eredeti 16, vagy 12 bites minták jelennek meg. Ezután jön a D/A átalakítás, deglitche áramkör, aluláteresztő szűrő. A rendszerdinamika javítása érdekében négyszeres oversampling eljárást alkalmaznak.

A felvétel és a lejátszás folyamatának blokkvázlata a 29. ábrán látható.



2.5. Szervorendszerek

A bonyolult mechanika és a különleges funkciók miatt szükséges pontos fejdobfordulatszám és szalagfutás vezérléséről négy szervo áramkör gondoskodik. Nevezetesen:

1. Capstan szervo.
2. ATF szervo.
3. Fejdobszervo.
4. Csévéelőtárca szervo.

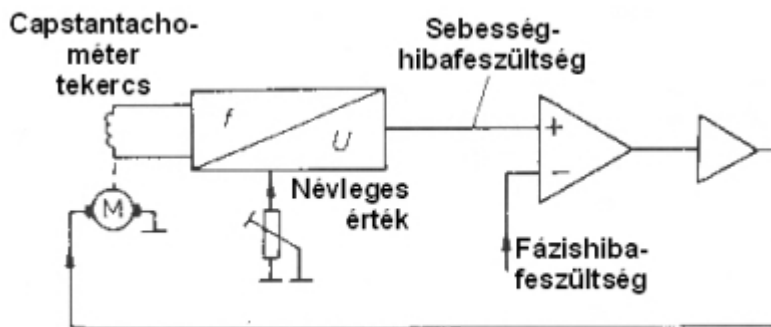
2.5.1. A capstan szervo

Feladata az állandó - 8,15 mm/s - sebességű szalagadagolás biztosítása normál üzemmódban, illetve az ATF szervoval közösen a ferde sávok korrekt sávközépi letapogatását végzi (tracking).

Egy sebesség- és egy fázisszabályozó áramkörből áll (30. ábra). A sebességszabályozó áramkörben egy tachométergenerátor állítja elő a szabályozójelet, amelynek ferkvenciájával lesz arányos a hibafeszültség. A szabályozófeszültség a referencia érték és a tachométergenerátor által adott érték összehasonlítása alapján képződik.

A fázisszabályozó áramkör a pontos sávletapogatásért felelős. Hibafeszültségét az ATF szervo adja. A sebesség - és fázisszabályozó áramkörök hibafeszültségei együttesen adják a capstan motor vezérlőfeszültségét.

A capstan szervonak még a normál szalagsebesség 16 -szorosánál is hibátlanul kell működnie.

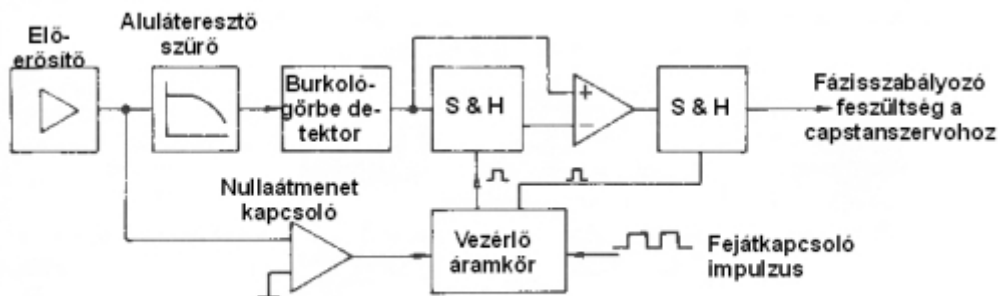


30. ábra: sebesség- és fázisszabályzó

2.5.2. Az ATF szervo

A pontos sávletapogatás alapjele az ATF tartomány 130,67 kHz -es jele (pilotjel). Ezt a frekvenciát úgy választották meg, hogy a fejek azimuthelyzete ne befolyásolhassa jelentős mértékben az amplitúdóját. Lejártszáskor ezért a fej a két szomszédos sáv áthallásából származó mérőjel komponenseket is letapogatja. Pontos sávletapogatás esetén a két áthallás értéke megegyezik. Az áramkör a 31. ábrán látható.

A nagyfrekvenciás jelből aluláteresztő szűrővel választják ki a mérőjelet, amit egyenfeszültséggé alakítanak. Mivel a mérőjelet mindkét szomszédos sávban kiértékelik, azonban az ATF tartományok nem egymás mellett helyezkednek el a sávokon, a kapott egyenfeszültségeket átmenetileg mintavevő és -tartó áramkörben tárolják, amelyeket a vezérlő áramkör kezel. Ha a letapogatás nem sávközépen történik, akkor a differenciálerősítő kimenetén megjelenő egyenfeszültség megváltozik, ennek hatására pedig a capstan motor visszaállítja a helyes sávkövetést.



31. ábra: az ATF szervo elvi kapcsolása

2.5.3. A fejdobszervo

Feladata a fejdob fordulatszámának pontosan 2000 1/min -es fordulaton tartása normál üzemmód esetén, illetve a forgás fázisának szabályozása. A fordulatszám és a fázis szabályozására külön áramköröket alkalmaznak. A különleges funkciók miatt a szalagról a normál sebesség 200 -szorosánál is ki kell tudni olvasni az adatokat. Ehhez a fejdob forgása és a szalagadagolás közötti relatív sebességet állandó értéken kell tartani. A fejdob sebessége a szalagadagolás sebességének megváltozásával a következőképpen változik:

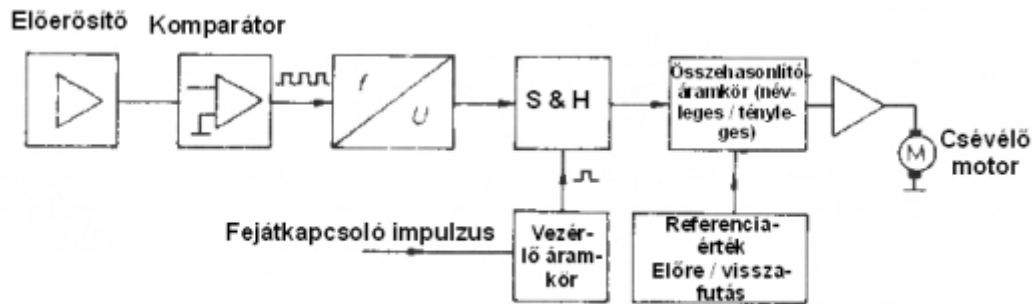
$$\text{szalag} = +200 \Rightarrow 3025 \text{ 1/min}$$

$$\text{szalag} = -200 \Rightarrow 964 \text{ 1/min.}$$

A sebességszabályozó áramkörben a tachométergenerátor jelét használják, míg a fázisszabályozóban helyzetjelző jeladó szolgáltatja a szabályozó és a fejátkapcsoló jelet.

2.5.4. A csévéltárcsa szervo

Feladata a szalaghúzás egyes üzemmódoktól függően változó, de különben állandó értéken tartása. A csévéltárcsa szabályozó jele a fejdob forgásából származik. A csévéltárcsa fordulatszámát a fejdob fordulatszámával szinkronban változtatják, így a gyors keresési funkció és a szalaghúzás közötti illesztés optimális. Egy lehetséges megvalósítás látható a 32. ábrán.



32. ábra: a csévéelőtárcsa szervo egy lehetséges megvalósítása

Az alapelv az, hogy gyorskeresésnél a fejek több sávot kereszteznek és a nagyfrekvenciás feszültség szintje minden sávugrásnál csökken. A keresztezett sávok száma a keresés sebességével és irányával arányos.

A szintcsökkenéseket a komparátor alakítja négyszögjelekké, amiket a frekvencia-feszültség átalakító egyenfeszültséggé alakít. A fejek nem folyamatosan érintkeznek a szalaggal, így ebben az esetben is mintavevő és -tartó áramkört alkalmaznak, aminek a kimenetén jelenik meg a szabályozójel. Ezt összehasonlítják a referencia jellel és a különbségi jel vezérli a csévéelőmotort.

A következő részben különböző digitális audio tömörítési eljárásokkal foglalkozunk és megvizsgálunk egy komplett rendszert - a MiniDiscet - amely egy speciális audio tömörítési eljárásra épül.

3. Digitális audio tömörítés

3.1. Bevezetés

A legtöbb digitális adattípushoz hasonlóan - a digitális video kivételével - a tömörítetlen digitális audiohoz tartozó adatsebesség jelentős (a csatornán kialakuló végső adatátviteli sebesség DAT esetén 7,5 Mbps, CD esetén 4,3218 Mbps). A digitális audio tömörítés révén lehetővé válik a digitális audio hatékonyabb tárolása és átvitele. A különböző audio tömörítési eljárások különböző bonyolultságú kódolókat és dekódolókat, különböző minőségű tömörített audio jelet és különböző mértékű tömörítést biztosítanak. Az ADPCM kódoló egyszerű megoldás, kis tömörítési fokú, közepes minőségű hangot eredményez. Az MPEG/audio szabvány és a MiniDisc rendszer által használt ATRAC tömörítési eljárás bonyolult, nagy tömörítési fokú, kiváló hangminőséget eredményező algoritmusok. Ezeket az eljárásokat általános audio jelekre fejlesztették ki és nem speciálisan beszéd tömörítésére vannak beállítva. A továbbiakban az audio tömörítéshez szükséges pszichoakusztikai jellemzőkkel foglalkozunk, majd az ADPCM kódolás kerül terítékre, végül pedig az MPEG/audio szabványt és a MiniDisc rendszert vizsgáljuk meg részletesen.

3.2. Pszichoakusztika

3.2.1. Fletcher-Manson görbék

Az emberi fül által hallható hang tulajdonképp nem más, mint nyomásváltozás a közvetítő közegben. Erőssége kifejezhető a nyomásváltozás nagyságával. Általában ezt egy referenciértékhez viszonyítva, dB-ben adják meg. Ez a fizikai jellemző a hangnyomás szint:

$$L_p = 20 \lg \frac{p}{p_0}$$

, ahol p a hangnyomás időbeli változása, p_0 az $f = 1$ kHz frekvenciájú, emberi fül által még éppen hallható hangnyomás ($p_0 = 20$ mPa). Ezen kívül alkalmazzák még az intenzitás szintet, amelyet szintén dB-ben adnak meg:

$$L_i = 10 \lg \frac{I}{I_0}$$

, ahol $I_0 = 1$ pw/m^2 , szintén 1 kHz-es hangra vonatkozik. A fizikai jellemzők azonban közvetlenül nem hordoznak információt az emberi hallás szubjektív jellemzőiről. A pszichofizikai jellemzők a fizikai jellemzők áttételei az érzeti síkra. Ezek a jellemzők a hangerősség és a hangosság. A hangerősség (L_n) mértékegysége a phon. Egy hang hangerőssége annyi phon, ahány dB a vele azonos hangosságérzetet keltő 1 kHz-es szinuszos hang hangnyomásszintje. Így az éppen hallható hangerősség 0 phon. A hangerősség hátránya, hogy több hang együttes megszólalása esetén nem alkalmazható.

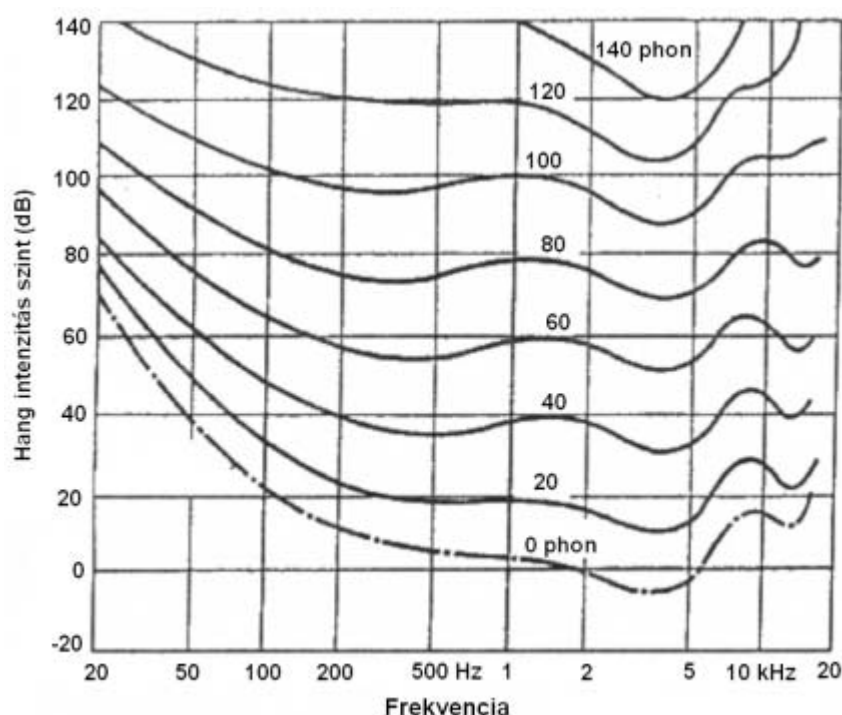
A hangosság az egyszerre megszólaló hangokra vonatkozik, jele: N ; mértékegysége a son:

$$\text{Ha } L_n > 40 \text{ phon, akkor } N = 2^{\frac{L_n - 40}{10}}$$

Ha a megszólaló hangok hangerőssége mind nagyobb, mint 40 phon, akkor egyszerű összegezéssel számítható a hangosság.

A fül érzékenysége a frekvencia függvénye. A legérzékenyebb az 1 kHz körüli frekvenciákra. Az a hangnyomás szint, amit 1 kHz-en épp meghallunk nem biztos, hogy meghallható más frekvencián. Általában két egyforma teljesítményű, de különböző frekvenciájú hangot nem azonos hangerejűnek hallunk. Az 33. ábrán láthatók a Fletcher-Manson görbék különböző hangerő értékekre. A "0 phon" feliratú görbe ábrázolja a minimális értékeket, amit a fül érzékelní képes különböző frekvenciákon. A görbék azt ábrázolják, hogy a fül bizonyos frekvenciákon érzékenyebb, mint másokon.

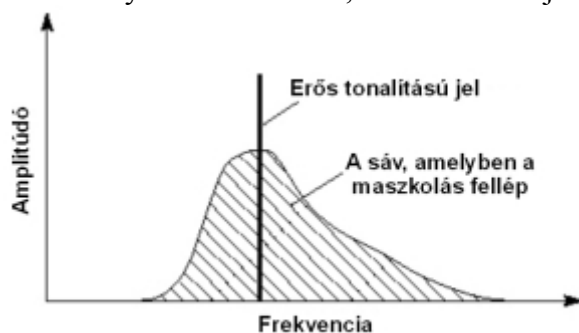
A Fletcher-Manson görbék hátránya, hogy átlagos érzetjellemzőket írnak le, azaz egy átlagos hallgatót definiálnak.



33. ábra Fletcher-Manson görbék

3.2.2. Maszkolási effektus

Maszkolásról akkor beszélünk, ha egy hangot egy másik hang hallhatatlanná tesz. Egyidejű maszkolás akkor lép fel, amikor két hang egyszerre szól, mint pl. amikor egy beszélgetést (a maszkolt jel) elnyom egy vonat dübörgése (a maszkoló). Visszafelé irányuló maszkolásról akkor beszélünk, ha a maszkolt jel előbb befejeződik, mint ahogy a maszkoló elkezdődik; előre irányuló a maszkolás, ha a maszkolt jel azután kezdődik, hogy a maszkoló véget ért.



34. ábra egyidejű maszkolás

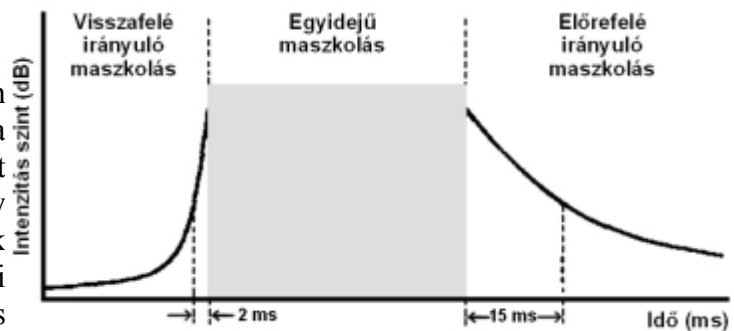
A maszkolás annál erősebb, minél közelebb áll egymáshoz - idő- és frekvencia tartományban egyaránt - a két jel. Pl. az egyidejű maszkolás erősebb a visszafelé- és az előre irányuló maszkolásnál is, mert a hangok egy időben szólalnak meg. A maszkolási kísérletekben általában maszkolóként keskeny sávzélességű

fehér zajt használnak és tiszta szinuszos hang hallásküszöbét vizsgálják különböző időben és frekvenciákon. Az egyidejű- és az időleges maszkolásokra a 34. és a 35. ábrán láthatunk példát.

Az ábrákból fontos következtetéseket vonhatunk le. Először is, az egyidejű maszkolás akkor a leghatásosabb, ha a maszkolt jel frekvenciája megegyezik, vagy nagyobb, mint a maszkolóé. Másodsor, míg az előre irányuló maszkolás jelentős ideig hatásos azután, hogy a maszkoló elhallgatott, a vissza irányuló kevesebb, mint 2-3 ms -ig él a maszkoló megszólalása előtt.

3.2.3. Kritikus sávok

A kritikus sávok ötlete a fül azon tulajdonságából adódott, hogy a hallható frekvencia tartományt alsávokra osztva analizálja. Az egy kritikus sávon belüli frekvenciák hasonlóak a fül érzékelési tulajdonságaihoz és más kritikus sávoktól elkülönítve dolgozzák fel őket. A kritikus sávokat hallási



35. ábra: a különböző maszkolások időtartama

kísérletekkel határozták meg, és levezethetők a belső fülben található érzékelő cellák elhelyezkedése alapján is. A kritikus sávok felfoghatók úgy, mint a fül által használt frekvencia sáv. Alacsonyabb frekvenciák esetén sokkal keskenyebbek ezek a sávok - tulajdonképp a kritikus sávok 3/4 része 5 kHz alatt található. Ez azt jelenti, hogy az alacsonyabb frekvenciákból több információt kap a fül, mint a magasabból.

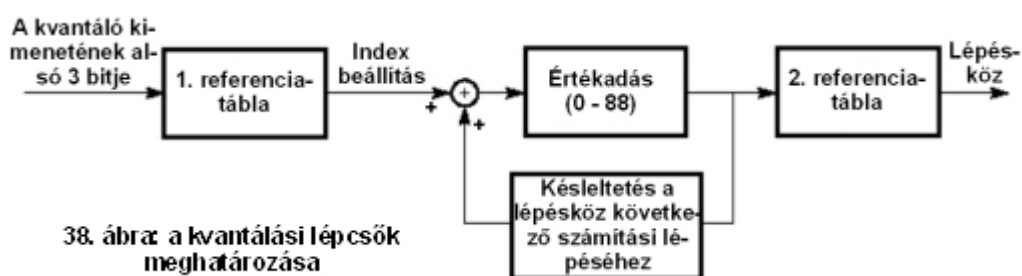
A digitális hangtömörítő algoritmusok a tömörítéshez jelentős mértékben hasznosítják a pszichofizikai jellemzőket az idő-frekvencia tartomány átalakítás és a bit kiosztás során.

4.2. Az IMA ADPCM algoritmus

Az IMA számítógépes hardver- és szoftver kereskedőkből álló konzorcium, akik együttműködnek egy tényleges multimédiás szabvány kifejlesztésén. Az IMA azt tűzte ki céljául, hogy egy nyilvános audio tömörítési eljárást hozzon létre, amely jó hangminőséget és tömörítési arányt eredményez. Plusz az algoritmusnak elég egyszerűnek kellett lennie ahhoz, hogy el tudja végezni a valós idejű, tisztán szoftveres dekompreszióját egy 44,1 kHz -es mintavételezésű jelnek egy 20 MHz -es 386 -os számítógépen. A kiválasztott ADPCM kódolás nemcsak a fentieknek megfelelő, hanem még a szoftveres, valós idejű tömörítésre is képes.

Az IMA ADPCM egyszerűsége a prediktorának nyersségében rejlik. A minta előre kiszámított értéke egyszerűen a közvetlenül előtte levő minta dekódolt értéke alapján képződik. Így a 36. ábrán látható prediktor blokk csupán egy késleltető, amelynek a kimenete és bemenete között egy minta időnyi késleltetés van. Ez a prediktor nem alkalmazkodó, így nem szükséges kísérő információ a prediktor rekonstruálásához.

A 37. ábrán látható az IMA algoritmus által használt kvantálási folyamat blokk diagramja. A kvantáló kiad egy 4 bitből álló mintázatot, amely az előjeles nagyságát mutatja a kvantálási szinteknek az egyes bemeneti mintákra. Az audio jelhez való alkalmazkodás csak a kvantáló blokkban valósul meg. A kvantáló a felbontást változtatja az éppen aktuális felbontás és a kvantáló előző bemenetre adott kimenete alapján. Ez a folyamat két, egymásra következő referenciatábla segítségével valósítható meg. Az első három bit - amely a kvantálási szintet adja meg - indexként szolgál az első referenciatáblában, aminek a kimenete index beállítást végez a második referenciatáblában. Ezt a beállító értéket egy tárolt, fix indexhez adják hozzá és a korlátozott tartományú eredményt használják indexként a második referenciatáblában. Az összegzett indexet letárolják és a következő iterációs lépésben kerül felhasználásra. A második referenciatábla kimenete az új kvantálási felbontás. Megjegyzendő, hogy egy, a második referenciatáblában található adott értékű index esetén az adaptációhoz használt adatok a kvantáló kimenetéből levezethetők; kísérő információkra nincs szükség. A 38. ábrán látható a kvantálási lépcsők meghatározásának blokkdiagramja, és a 7. és 8. táblázatban található a két referenciatábla adatai.



38. ábra: a kvantálási lépcsők meghatározása

7. táblázat: az első referenciatábla

Három bites felbontás	Index állítás
0	-1
1	-1
10	-1
11	-1
100	2
101	4
110	6
111	8

4.3. IMA ADPCM hibajavítás.

Egy szerencsés mellékhatása az ADPCM ezen típusának, hogy a dekóderben keletkező hibáknak - amelyeket okozhatnak az egyes kódszavakban fellépő hibák, szerkesztések, összeillesztések, vagy véletlenszerű hozzáférések a tömörített bit folyamathoz - általában nincs jelentős hatása a dekóder kimenetére. Ez általában nem igaz az olyan tömörítési eljárásokra, amelyek becslést (prediction) használnak. A becslés az előző minta helyes dekódolásától függ, így a hiba egyre inkább terjed. A továbbiakban megtudhatjuk, hogy ez a terjedés miért korlátos és miért nem végzetes az IMA ADPCM algoritmusra nézve.

A dekóder az $X_p(n)$ mintát az előző minta ($X_p(n-1)$), a kódszó végeredményének előjeles nagyságának ($C(n)$), a kvantáló lépésközének és a fél lépésköznyi offset összeadásával állítja helyre:

$$X_p(n) = X_p(n-1) + \text{lépésköz}(n) * C'(n)$$

ahol $C'(n) = C(n) \pm$ egy fél lépésköznyi offset.

Egy, a második referenciatáblán végzett vizsgálat kimutatja, hogy az egymást sorban követő bemenetek kb. 1,1 -szer hosszabak az őket megelőzőeknél. Amíg nem korlátozzuk az index nagyságát, a lépésköz értéke az előző lépésköz és a kódszó ($F(C(n-1))$) függvénye:

$$\text{lépésköz}(n) = \text{lépésköz}(n-1) * F(C(n-1))$$

A két fenti egyenlet átalakítható úgy, hogy a dekódolt mintát ($X_p(n)$) a lépésköz és a dekódolt minta értékének függvényeként írja le az egyenlet m időpillanatban, és kódszavak csoportjával m és n időpillanatok között:

$$X_p(n) = X_p(m) + \text{lépésköz}(m) * \sum_{i=m+1}^n \left\{ \prod_{j=m+1}^i F[C(j)] \right\} * C'(i)$$

8. táblázat: a második referenciatábla

Index	Lépésköz	Index	Lépésköz	Index	Lépésköz	Index	Lépésköz
0	7	22	60	44	494	66	4,03
1	8	23	66	45	544	67	4,43
2	9	24	73	46	598	68	4,87
3	10	25	80	47	658	69	5,36
4	11	26	88	48	724	70	5,89
5	12	27	97	49	796	71	6,48
6	13	28	107	50	876	72	7,13
7	14	29	118	51	963	73	7,85
8	16	30	130	52	1,06	74	8,63
9	17	31	143	53	1,17	75	9,49
10	19	32	157	54	1,28	76	10,44
11	21	33	173	55	1,41	77	11,49
12	23	34	190	56	1,55	78	12,64
13	25	35	209	57	1,71	79	13,9
14	28	36	230	58	1,88	80	15,29
15	31	37	253	59	2,07	81	16,82
16	34	38	279	60	2,27	82	18,5
17	37	39	307	61	2,5	83	20,35
18	41	40	337	62	2,75	84	22,36
19	45	41	371	63	3,02	85	24,62
20	50	42	408	64	3,33	86	27,09
21	55	43	449	65	3,66	87	29,79
						88	32,77

Figyeljük meg, hogy csak a kódszavakat összegezzük az $m+1$ -ik időpillanattól kezdődően. Egy hiba a kódszóban, vagy egy véletlen hozzáférés a tömörített bit folyamhoz q időpillanatban hibát okozhat a dekódolt kimenetben és a lépésközben a $q+1$ időpillanatban (lépésköz($q+1$)). A fenti egyenletből látható, hogy a hiba mértéke az $X_p(m)$ -ben egy konstans értéket ér el $X_p(n)$ jövőbeli értékeire. Ez az érték hallhatatlan mindaddig, míg a dekódolt kimenet nem lépi túl a megengedhető tartományt és így nem kell megvágni. A vágás egy pillanatnyi hallható torzításban jelentkezik, de ugyanakkor a konstans hiba részleges, vagy teljes javítását is szolgálja. Továbbá a dekóder kimenetére egy digitális felüláteresztő szűrőt

kötve ez a konstans hiba kiszűrhető. A fenti egyenletből az is kitűnik, hogy a lépésközben keletkező hiba a dekódolt kimenet nem kívánt erősítését, vagy csillapítását is okozhatja. A kimeneti jelalak mindaddig nem változik, amíg a második referenciatáblának szóló index nem skála határolt. A skála határolás a lépésköz értékének részleges, vagy teljes módosítását eredményezi.

A lépésköz adaptáció természeténél fogva behatárolja a hiba fellépés lehetőségét a lépésközben. Egy hiba a lépésköz(m+1) -ben, amit egy hibás kódszó okozott $1,1^9$ -szeres, vagy 7,45 dB -es változást okozhat a lépésköz értékében. Minden 88 kódszóból álló sorozat, amelyben a kódszavak hosszúsága 3 vagy kevesebb (az 7. táblázatnak megfelelően), teljes mértékben korrigálja a lépésközt a konstans hiba minimális értékére. Még a legkisebb mintavételi frekvencia esetén is - ami tipikusan 8 kHz - a 88 minta 11 ms -nyi audio jelnek felel meg. Így a véletlen hozzáféréshez és szerkesztéshez 11 ms-ként nyílik lehetőség.

5. MPEG/audio tömörítés

5.1. Az MPEG/audio tömörítési eljárás alapjai

Az MPEG/audio tömörítési algoritmus egy ISO, Hi-Fi minőségű audio tömörítési szabvány. Ez egy része egy három részből álló tömörítési szabványcsoportnak. A másik két résszel együtt a szinkronizált, digitális video és audio jel tömörítéséhez szükséges kompozit szabványt adja. Az így kialakuló adatsebesség durván 1,5 Mbps.

Bár az Mpeg algoritmus transzparens, azaz érzékelhető minőségromlás nélküli tömörítést valósíthat meg, azért mégiscsak veszteséges. Az MPEG/audio bizottság széleskörű, szubjektív hallástereszteteket végzett a szabvány fejlesztése során. A vizsgálatok azt eredményezték, hogy még 6:1 arányú tömörítés esetén (sztereo, 16bit, 48 kHz, 256 kbps -re tömörítve) és optimális hallgatási körülmények között sem tudták statisztikailag számottevő esetben megkülönböztetni egymástól a tömörített és az eredeti jelet még a "profí hallgatók" sem. Továbbá olyan jeleket választottak, amiket tipikusan nehéz tömöríteni.

9. táblázat: az MPEG algoritmus által használt kritikus sávok

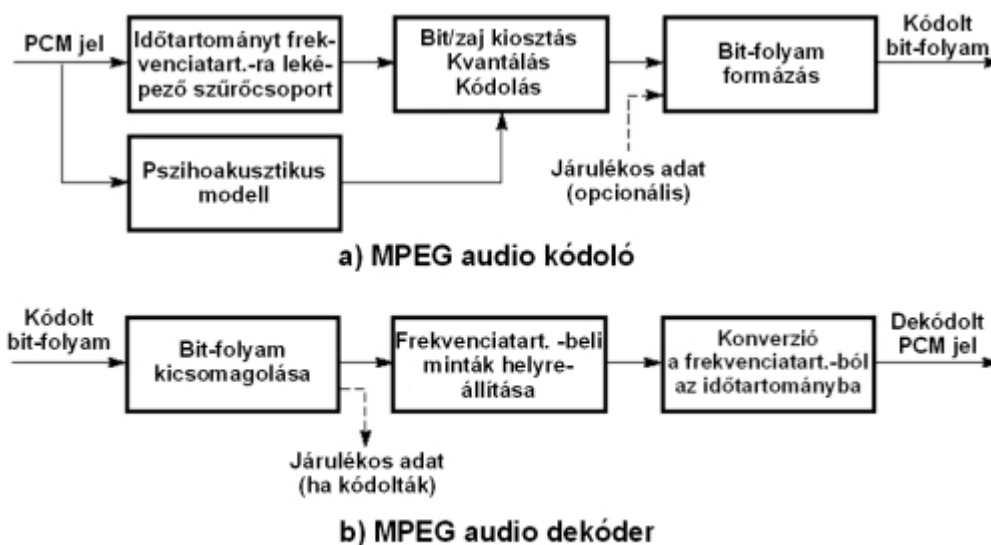
A sáv száma	Frekvencia (Hz)*	A sáv száma	Frekvencia (Hz)*
0	50	14	1,97
1	95	15	2,34
2	140	16	2,72
3	235	17	3,28
4	330	18	3,84
5	420	19	4,69
6	560	20	5,44
7	660	21	6,38
8	800	22	7,69
9	940	23	9,38
10	1,13	24	11,63
11	1,27	25	15,38
12	1,5	26	20,25
13	1,74		

* A frekvenciák a sáv felső határai.

A tömörítési algoritmus jó hatásfoka a maszkolási effektus kihasználásán alapszik. A fülnek van még egy, ide kapcsolódó hiányossága: korlátolt frekvencia szelektivitással rendelkezik, amely 100 Hz -től (a legmélyebb hallható hangokra) 4 kHz -ig (a legmagasabb hangokra) terjed. Emiatt lehetséges a hallható frekvenciatartomány felosztása kritikus sávokra, amelyek a fül frekvencia függő hangkezelését tükrözik. A 9. táblázat tartalmazza az algoritmus által használt kritikus sávok listáját.

A fül fenti tulajdonsága miatt a zaj maszkolás kizárólag egy frekvencia sávon belüli jelaktivitás függvénye. Ez a tulajdonság a tömörítési eljárás esetén úgy hasznosítható, hogy az audio jelet a frekvencia tartományba transzformálják, az így kapott spektrumot felosztják a kritikus sávoknak megfelelő alsávokra, és végül az egyes alsávokat a kvantálási zaj hallhatóságának függvényében kvantálják. Optimális tömörítés esetén minden egyes sávhoz éppen akkora kvantálási szintet kell biztosítani, amely feltétlenül szükséges ahhoz, hogy a kvantálási zaj hallhatatlan legyen. A továbbiakban részletesen foglalkozunk az MPEG/Audio algoritmussal.

5.2. MPEG/audio kódolás és dekódolás



39. ábra: az MPEG/audio kódoló és dekódoló

A 39. ábrán látható az MPEG kóder és dekóder blokkdiagramja. Ebben a megjelenítésben a kódolás nagyon hasonlít a fent leírt folyamatra. A bemeneti jel egy szűrőcsoporton halad keresztül, amely alsávokra osztja fel azt. A jel egyidejűleg áthalad egy pszichoakusztikai modellen is, amely meghatározza a jel-maszk viszonyt (signal-to-mask ratio = SMR) az egyes alsávokra. A bit- vagy zaj allokációs blokk használja az SMR-t a kvantáláshoz rendelkezésre álló összes bit alsávok közötti kiosztásakor, a kvantálási zaj minimalizálását célként kitűzve. A legutolsó blokk fogja a kvantált mintákat egy dekódolható bit folyamammá alakítani. A dekóder ezt az átalakítást egyszerűen megfordítja, felépíti a kvantált alsávok értékeit és végül visszatranszformálja ezeket egy időtartománybeli jellé. Az MPEG elvárásainak megfelelően az audio jelhez tartozó kísérő információkra nincs szükség, de alkalmazható a kódolt bit folyamban.

Az MPEG/audio szabványnak három, elkülönített rétege van. A Layer I tartalmazza az alap algoritmust, a Layer II és III pedig továbbfejlesztések, amelyek használják az I -es bizonyos

részeit. Minden egyes plusz réteg javítja a kódolás hatékonyságát, de növeli a kóder és dekóder bonyolultságát is.

5.2.1. Layer I

Az alap szűrőcsoportot használja, amit minden réteg tartalmaz. Ez a szűrőcsoport a jelet 32 állandó szélességű frekvencia sávra osztja. A szűrők viszonylag egyszerűek és a fül képességeihez mérten jó idő- és frekvencia tartománybeli felbontást biztosítanak. Ez egy kompromisszumos megoldás három fontos engedménnyel. Először, a 32 állandó szélességű frekvencia sáv nem tükrözi a fül kritikus sávjait (40. ábra). A sáv szélesség túl nagy az alacsony frekvenciák esetén, így a kvantálási bitek számát nem lehet speciálisan egy-egy kritikus sávra meghatározni. Ehelyett a sávba eső legnagyobb zajérzékenységű kritikus sáv határozza meg a kvantáló bitek számát. Másodszor a szűrőcsoport és inverze nem veszteségmentes átalakítások. Még a kvantálást elhagyva sem kapnánk vissza tökéletesen az eredeti bemeneti jelet. Szerencsére a szűrők által bevitt hiba kicsi és nem hallható. Harmadszor a szomszédos szűrők jelentős frekvenciaátlapolással rendelkeznek. Egy adott frekvenciájú jel így két szomszédos szűrőt befolyásolhat.



40. ábra: az állandó szélességű frekvenciasávok

A szűrőcsoport 32 mintát készít - frekvenciasávonként egyet minden egyes audio mintára a 32-ből. A Layer I-es algoritmus csoportba foglal 12 mintát minden egyes sávból. Minden egyes 12 mintából álló csoport megkapja a bit allokációt és ha ez nem nulla, akkor a felbontási tényezőt is. Sztereo többletkódolás esetén - amit később tárgyalunk - ez másképp működik. A bit allokáció határozza meg, hogy egy mintára hány bit jut. A felbontási tényező egy szorzó, ami úgy méretezi a mintákat, hogy maximális legyen a kvantáló felbontása. A Layer I-es kóder 32 db 12 mintát tartalmazó csoportot foglal egy keretbe ($32 * 12 = 384$ minta). Az audio adat mellett minden keret tartalmaz egy fejrészt, egy opcionális CRC szekvenciát és esetleg egyéb többlet információkat.

5.2.2. Layer II

A Layer I egyszerű továbbfejlesztése. Úgy éri el a jobb hatásfokot, hogy az adatokat nagyobb csoportokban kódolja. A kereteket $3 * 12 * 32$ mintából készíti egy audio csatornára (csinál egy csoportot $3 * 12$ mintából, majd az így kapott csoportból 32-t egybefog és ez lesz a keret, ami 1152 mintát jelent). A Layer I 12 mintát fog csoportba alsávonként, míg a Layer II $3 * 12$ mintát. A kóder csak akkor használ egyéni felbontási tényezőt a 12 mintát tartalmazó

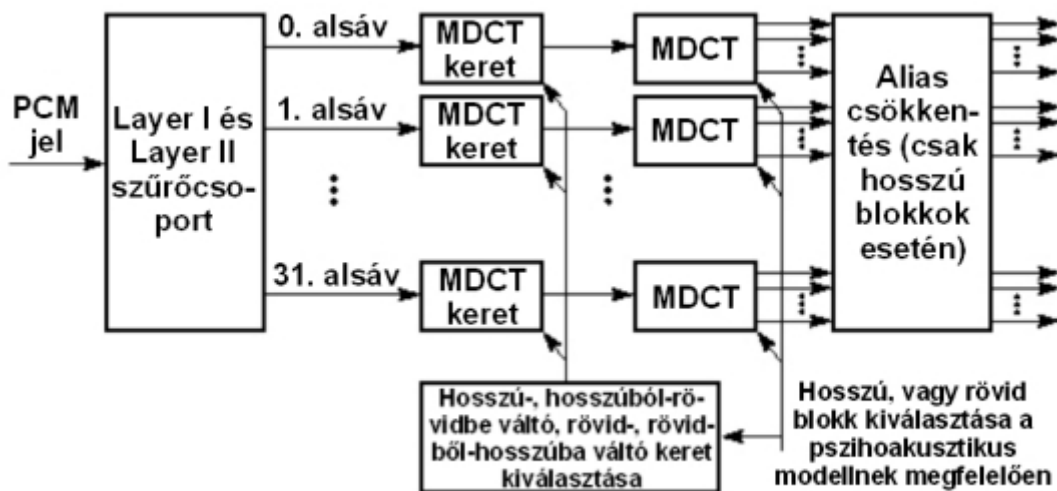
csoportok mindegyikére, ha ez szükséges a hallható tozítás elkerülése miatt. A kóder még két esetben osztja meg a felbontási tényező értékeit két vagy három csoport között:

1. amikor a felbontási tényezők értékei elég közel esnek egymáshoz.
2. amikor előre látható, hogy a fül időleges maszkolása elfedi a következményként fellépő torzítást.

A Layer II algoritmus azért is jobb hatásfokú a Layer I -nél, mert a bit allokációt, a felbontási tényező értékeit és a kvantált mintákat egy sokkal hatékonyabb kóddal rögzíti.

5.2.3. Layer III

Ez már egy sokkal kifinomultabb megoldás. Bár ez is a Layer I és II -ben alkalmazott szűrőcsoportra alapszik, annak néhány hiányosságát kompenzálja a szűrő kimenetének MDCT feldolgozásával (41. ábra).



41. ábra: a Layer III kódolás folyamata

Az MDCT -k tovább osztják a szűrők kimeneteit a frekvencia tartományban, hogy jobb spektrális felbontást biztosítsanak. Az elkerülhetetlen idő- és frekvencia tartománybeli felbontási különbség miatt a Layer III két különböző hosszúságú MDCT blokkot definiál: hosszú blokk, amely 36 mintából áll és rövid blokk 12 mintából. A rövid blokk javítja az idő tartománybeli felbontást és így megbírkózik a tranziensekkel. Egy rövid blokk hossza pont harmad része egy hosszú bloknak; így alkalmazásakor három rövid blokk helyettesít egyetlen hosszút. A hosszú- és rövid blokkok közti átváltás nem azonnali. Egy speciális hosszúból rövidbe, vagy rövidből hosszúba ablakkal rendelkező hosszú blokk végzi el az átalakítást. A Layer III -nak háromféle blokkolási módja van: két mód, amikor a szűrők kimenetei egyforma blokkhosszúsággal haladhatnak át az MDCT -n, és egy kevert mód, amikor a 2 alsó frekvenciasáv hosszú blokkot, a többi 30 pedig rövidet használ.

További jelentős fejlődések a Layer I -hez és II -höz képest:

- Alias csökkentés: a Layer III definiál egy eljárást az MDCT értékeire, amivel kiszűri a fölösleges részeket a jelből, amiket a szűrő egymást átlapoló sávjai okoznak.
- Nemlineáris kvantálás: a Layer III kvantáló bemeneti teljesítményét a teljes 3/4 részére növeli kvantálás előtt, hogy állandóbb J/Z viszonyt biztosítson a kvantálás teljes tartományán. A dekóderben lévő rekvantáló kimeneti teljesítményét 4/3 részre növeli, hogy helyreállítsa a linearitást.

- Az értékek entrópia kódolása: a Layer III Huffmann kódolást használ a kvantált minták kódolására a jobb adat tömörítés érdekében.
- Bit tároló használata: a Layer III bit folyamának kialakítása jobban alkalmazkodik a tömörített adat változó hosszúságához. Mint a Layer II a Layer III is 1152 mintából álló keretekbe foglalva dolgozza fel az adatokat. A Layer II -től eltérően azonban itt a mintákat reprezentáló kódolt adatnak nem szükséges egy fix hosszúságú keretbe illeszkednie a kódolt bit folyamában. A kóder adhat biteket a bit tárolóba, vagy elvehet onnan, amikor szükség van rá.
- Zaj allokáció bit allokáció helyett: a Layer I és II által használt bit allokációs eljárás csak a kvantálás által, egy adott számú bitre bevitt zajt becsli fel. A Layer III zaj allokáció iterációs hurkot használ. Ebben a hurokban a kvantálók szabályos sorrendben váltakoznak és az így adódó kvantálási zaj értékét specifikusan osztják szét az egyes alsávok között.

5.3. A pszihoakusztikai modell

Az MPEG kódoló kulcsfontosságú eleme a jó hatásfok elérésében. Feladata a bemenő audio jel analízálása és annak meghatározása, hogy a spektrumban hol és milyen mértékben lép fel a kvantálási zaj maszkolása. A kódoló ezt az információt arra használja, hogy eldöntse mennyire pontosan ábrázolja a bemenő audio jelet a korlátosan rendelkezésre álló bitjeivel. Az MPEG/audio szabvány két példát tartalmaz a pszihoakusztikai modell megvalósítására. Az alábbiakban a két modell által végzett számítások alapvető lépéseit írjuk le általánosan.

Időrendi audio adat: a pszihoakusztikus modellnek foglalkoznia kell a jel szűrőn való áthaladásakor keletkező késleltetésével és az offsettel is, így a lényeges adat az analízáló ablakban centralizálódik. Pl. a kettes számú pszihoakusztikai modell használatakor Layer I esetén a szűrők miatti késleltetés 256 minta és az offset ami a 384 minta centralizálásához szükséges az 512 pontos pszihoakusztikai analízáló ablakban: $(512 - 384)/2 = 64$ pont.

Az audio jel spektrális összetevőkre bontása: a pszihoakusztikai modell az idő tartományt a frekvencia tartományra egy 512, vagy 1024 pontos Fourier transzformáció segítségével képezi le. Az adatok egy hagyományos Hann súlyozással a Fourier transzformáció előtt beállíthatók úgy, hogy csökkenjenek a transzformációs ablak élhatai. A modell ezt az elkülönített és független leképezést használja a szűrőcsoportok kimenetei helyett, mert finomabb frekvencia felbontásra van szüksége, hogy kiszámítsa a maszkolási küszöbértékeket.

- A spektrális összetevők felosztása kritikus sávokra: a pszihoakusztikai számítások csökkentése érdekében a modell a spektrális összetevőket hallható mennyiségű csoportokba fogja.
- A küszöbérték egységesítése csendben: a modell tartalmaz egy empirikusan meghatározott abszolút maszkolási küszöbértéket. Ez a küszöbérték a legalsó határa a zaj maszkolásnak és a maszkoló jelek hiányában kerül kiszámításra.
- Tonális és nontonális részekre osztás: a modellnek fel kell ismernie a tonális és a zaj jellegű összetevőit az audio jelnek, mert a zaj maszkolási karakterisztikái a két összetevőnek eltérőek.
- Szórási funkció alkalmazása: a modell a zaj maszkolási küszöbértékeket egy empirikusan levezetett maszkolási, vagy szóródási függvény alkalmazásával számítja ki a jel összetevőire.
- A minimális maszkolási küszöbérték meghatározása minden egyes alsávra: a modell a maszkolási küszöbértékeket nagyobb frekvencia felbontással számítja, mint amit a szűrőcsoportok produkálnak. Ahol a szűrési sáv relatíve széles a kritikus sávhoz képest (a spektrum alsó végénél) a modell a szűrési sáv által lefedett minimális

maszkolási küszöbértéket választja. Ahol a szűrési sáv relatíve kicsi a kritikus sávhoz képest, a modell a szűrési sáv által lefedett maszkolási küszöbérték átlagát használja.

- Jel/Maszk viszony számítása: a pszihoakusztikus modell veszi a minimális maszkolási küszöbértéket és kiszámítja az SMR -t (Signal-to-Mask Ratio). A kapott értéket átadja a bit (vagy zaj) allokációs résznek a kódolóban.

5.4. Sztereo többlet kódolás (Stereo redundancy coding)

Az MPEG/audio algoritmus két típusát támogatja a sztereo többlet kódolásnak: az intenzitás sztereo kódolást és a közép/oldal (Middle/Side = MS) sztereo kódolást. Mindkét többlet kódolás a fül egy újabb érzékelési gyengéjét használja ki. A pszihoakusztikai eredményekből kitűnik, hogy a kb. 2 kHz -nél magasabb frekvenciákat lefedő kritikus sávokon belül a fül a sztereo hangtér leképezését már inkább a hangjel burkológörbéje, és nem a finom szerkezete alapján végzi.

Mindhárom réteg támogatja az intenzitás sztereo kódolást, de csak a III támogatja az MS-t.

Intenzitás sztereo kódolás esetén a kóder néhány magasabb frekvenciájú szűrő kimenetet egyetlen, összegzett jellel kódol, mintsem különálló kódokkal a jobb és a bal csatorna számára a 32 szűrő kimenet mindegyikére. A dekóder a jobb és bal csatornát egyedül a független jobb és bal csatorna felbontási tényezőire alapozva állítja helyre. Intenzitás sztereo kódolás esetén a spektrális alakja a jobb és bal csatornának ugyanolyan egy intenzitás kódolt szűrőcsoport jelen belül, de az abszolút értékek különböznek.

MS sztereo módban a kódoló a jel jobb és bal csatornáját egy közép- (a bal és jobb összege) és egy oldal csatornában (a bal és jobb különbsége) kódolja. Ezen kívül a kódoló speciálisan beállított technikákat alkalmaz az oldal csatorna további tömörítésére.

6. Az újraírható MiniDisc rendszer

6.1 Bevezetés

A MiniDisc rendszert azzal a nyilvánvaló céllal tervezték, hogy felváltsa a hagyományos kompakt kazettás rendszert. A MiniDisc formátum két különböző típusú hordozórétet definiál. Az egyik egy írható magneto-optikai lemez otthoni felvétel készítésre, a másik egy hagyományos csak olvasható lemez zene és szoftver publikáció céljára.

1992 -ben mutatták be a MiniDisc rendszert a fogyasztói piacon, mint egy új digitális hang lejátszó- és rögzítő rendszert. Ez csupán tíz évvel követte a Compact Disc (CD) bemutatását. Köztudott, hogy a CD a hagyományos bakelit lemezeket váltotta fel a hanglemez piacon. A CD technológia 16 bites kvantálásra és 44,1 kHz -en mintavételezett digitális hangrögzítésre épül. A CD hangminősége nagy előrelépést jelentett bármely más hétköznapi felvevő rendszerhez hasonlítva.

A CD elterjedése előtt sok mérnök a CD kifejlődését csupán a jobb hangminőségnek tulajdonította, de a CD lejátszó piacra dobása után kiderült, hogy a fogyasztónak nagy értékűvé vált az optikai lemezes rendszer gyors, véletlenszerű elérés módja. Ezen kívül a 12 cm -es lemezt könnyebb volt kezelni az LP -hez képest. Mind a lejátszó és a hordozó hosszabb élettartama erősen hozzájárult a CD formátum elfogadásához. A fejlesztés következő célpontja nyilvánvalóan az újraírható CD volt. A Sony és a Philips együtt dolgoztak ezen a rendszeren és 1989 -ben technikai valósággá tették. Két különböző CD rendszer alakult ki: az egyszer írható CD-R és az újraírható CD-MO.

Az audio kazetták eladási aránya csökkent 1989 óta és a Sony érezte, hogy az audio kazetta, mint hangrögzítő formátum életének végéhez közelít. Ha el is fogadta volna a fogyasztói társadalom az írható CD -t, még akkor is bonyolult lett volna betörni a hordozható készülékek piacára. Itt ugyanis a kompakt kazettás rendszerek uralkodtak az erős rázkódás állóságuknak és kis méretüknek köszönhetően. Egy új, lemezes rendszer kialakításának egyértelmű célja ezen gyengeségek kiküszöbölése volt. A Sony-nak sikerült ezt megvalósítania a MiniDisc-el (MD).

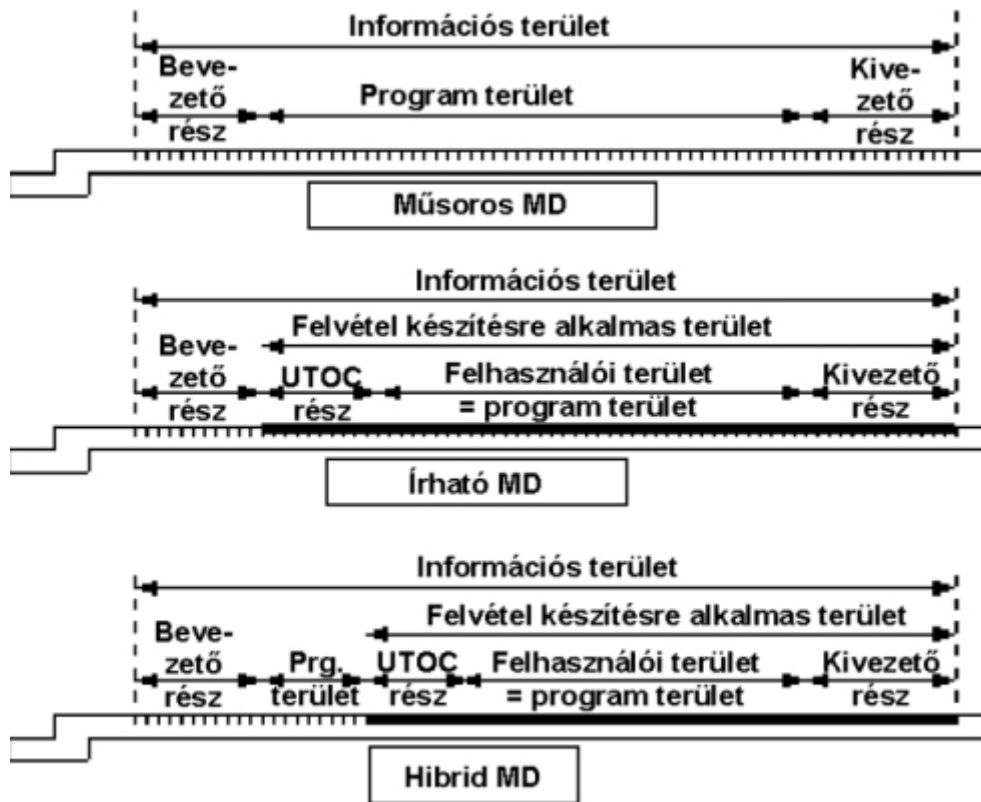
Magneto-optikai lemezes adattároló rendszereket pár éve már használtak számítógépes rendszerekben. Erre a technológiára alapozva kifejlesztettek egy közvetlenül felülírható rendszert ugyanakkora kapacitással, mint egy CD. Ezen kívül rázkódás álló memóriát alkalmaznak a hordozhatóság érdekében és egy digitális hang tömörítő eljárást használnak - amit ATRAC -nak hívnak (Adaptive TRansform Acoustic Coding) - aminek köszönhetően 64 mm -re csökkenthető a lemez átmérője. A félvezető elemek gyors technológiai fejlődése segített ennek megvalósításában.

A CD-ROM megvalósításához hasonlóan az MD Adat rendszer is az audio MD -re épül. A 140 Mbyte -os kapacitás és a célorientált file struktúra új lehetőségeket nyújt az MD adat lemezek különböző számítógépes konfigurációkba illesztésére. A MiniDisc rendszer szabvány leírását az ún. "Szivárvány könyv" tartalmazza (Rainbow Book).

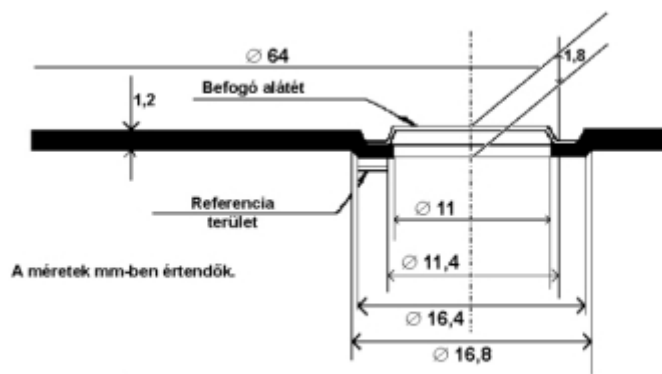
6.2. Rendszer konfiguráció

A kis házméretet a könnyű kezelhetőség és a lemez védelme érdekében alkalmazták. A ház fizikai méretei: 72 mm (sz), 68 mm (d), 5 mm (m), mely részletes tanulmányok eredménye. Két különböző típusú hordozó réteg és három féle lemez létezik. Az egyik a CD -hez hasonló szokványos lemez, amely alumínium fényvisszaverő réteggel és beleégetett jel mintázattal (pitek) rendelkezik. A második típus az írható lemez, amely széles barázdákkal bevont magneto-optikai hordozó réteggel rendelkezik. A harmadik típusú lemezen van előre égetett

és írható réteg is. A barázdákat egy 22,05 kHz -es vivő modulálja, amit pedig a cím adat modulál. Ez az előmodulációs eljárás elősegíti a helyes sávkövetést, a kerületi sebesség állandó értéken tartását és a hozzáférés vezérlését felvételkor. A 42. ábrán láthatók az MD lemezek típusai. Minden MD rendszernek képesnek kell lennie ezt a három típusú lemezt lejátszani. Ez azt jelenti, hogy az MD felvevő/lejátszó két funkciós optikai pick-up-al van felszerelve az előre rögzített bit mintázat és a magneto-optikai lemez olvasásához.



42. ábra: a három különböző MD lemez szerkezete



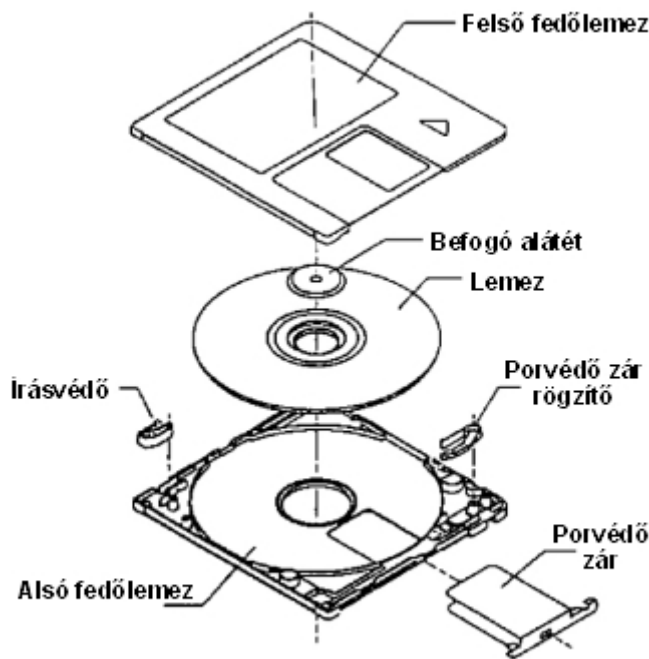
43. ábra: az MD lemez keresztmetszeti rajza

Az MD lemez hordozója fröccsöntési eljárásokkal készül. Így viszonylag egyszerűen kialakítható egy optimális illesztő terület, ami a lemez precíz betöltését, illetve rögzítését biztosítja. A 43. ábrán látható a lemez házba helyezett hordozófelületének keresztmetszeti rajza. Az optikai lemezt centralizálni kell miután a meghajtóba kerül. Mint a CD -nél, itt is a polycarbonát hordozó belső kerületét

használják erre, mint referenciát. Az egyszerű acélból készült alátét durván pozicionálódik a beállító terület szemközti oldalára, hogy pontos beállítást és centralizációt biztosítson. Ez az egyszerű eljárás csökkentheti a polycarbonát lemez középponti részeinek mechanikai igénybevételét is. Ily módon segít elkerülni a kettős optikai törést is a polycarbonát hordozón. A házban található egy porvédő zár. A gyári (műsoros) MD-n ez a zár csak az alsó részen van, míg az írható MD -nek mindkét oldalán található (44. ábra). Ez az egyszerű ház és lemez

felépítés olcsóvá teszi a gyártást. Ez pedig elengedhetetlen, ha zene rögzítésre, vagy szoftver kiadásokra akarjuk használni.

Az optikai paramétereit az MD lemeznek (laser hullámhossz, track pitch, az optikai pick up numerikus aperturája (NA), stb.) majdnem ugyanazok, mint a CD -nél (10. táblázat).

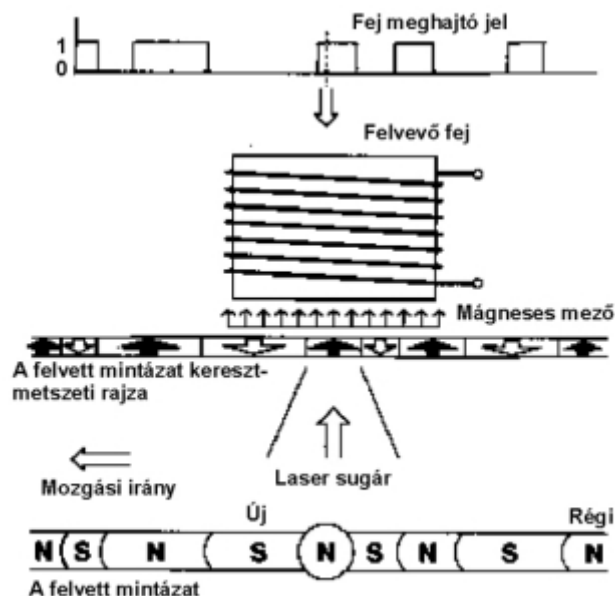


44. ábra: a MD lemez felépítése

mintázat hossza változó lesz.

6.3. Mágneses-mező modulációs felülírás (Magnetic-Field Modulation Overwriting)

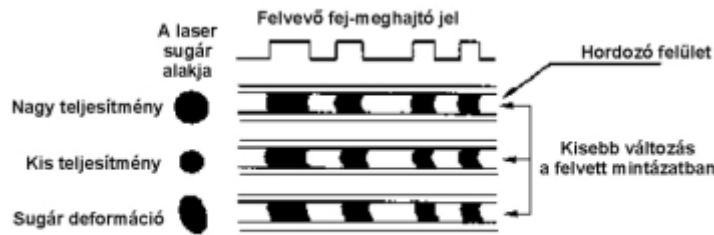
Egy másik rendszerben (a laser modulációs rendszerben) - ami a mágneses-mező moduláció mellett használt - a jelet a félvezető laser teljesítmény szintjének változtatásával rögzítik. A kétféle domén közül csak az egyiket használják. A terület amit a laser megvilágít 1 -es lesz, ahol nem rögzítünk semmit 0 lesz. A felvett mintázat asszimmetrikus lesz. Ezen kívül még egy probléma merül fel a laser modulációval kapcsolatban: érzékeny a felvételi teljesítmény ingadozásaira. Amikor a felvételi teljesítmény ingadozik a bitek kezdő- és vég pozíciója könnyen eltorzulhat, így a



45. ábra: a mágneses- és az optikai fej elhelyezkedése

A felvő/lejátszó MiniDisc -ek a digitális jeleket magneto-optikai úton rögzítik a mágneses-mező moduláció rendszer segítségével. Ezt a rendszert az újraírható CD technológia alapján alkalmazták. Az optikai fej és a lemezhez érő mágneses fej egymással szemközt helyezkednek el a lemez két oldalán, közrefogva azt (45. ábra). Mágneses-mező moduláció esetén egy

félvezető laser folyamatosan megvilágítja a felületet, mintegy 4,5 mW teljesítménnyel. Amikor a fénypont a lemezre esik, a felszíni réteg hőmérséklete kb. a Curie pontig emelkedik (kb. 180 °C). Amint a fénypont továbbhalad a réteg hőmérséklete csökkenni kezd. Ez a folyamat állandóan ismétlődik. Amikor N (É) vagy S (D) mágneses mezőbe kerül a pont amit a laser besugárzott egy-egy 1-t vagy 0 -t rögzít a rendszer a Curie hőmérséklet izotermáját határvonalként használva. Így a rögzített 1-es vagy 0 alakja és hossza meghatározott (46. ábra). Ha a mágneses mező elég gyorsan változik, lehetségessé válik, hogy 0,3 μm -es osztással írjunk egy pitbe, még akár egy 780 nm -es hullámhosszúságú fénysugár és 0,45 NA -s lencse kombinációjával is. Továbbá az "1" és a "0" területek szimmetrikusak lesznek, ami a mágneses-mező moduláció jellemzője.



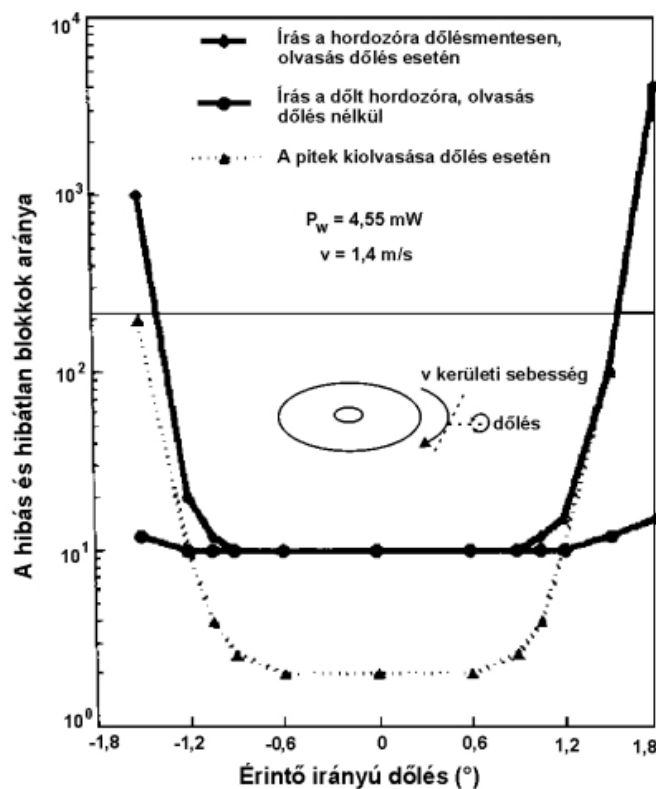
A MiniDiscben használt mágneses-mező modulációs eljárás



Laser modulációs eljárás

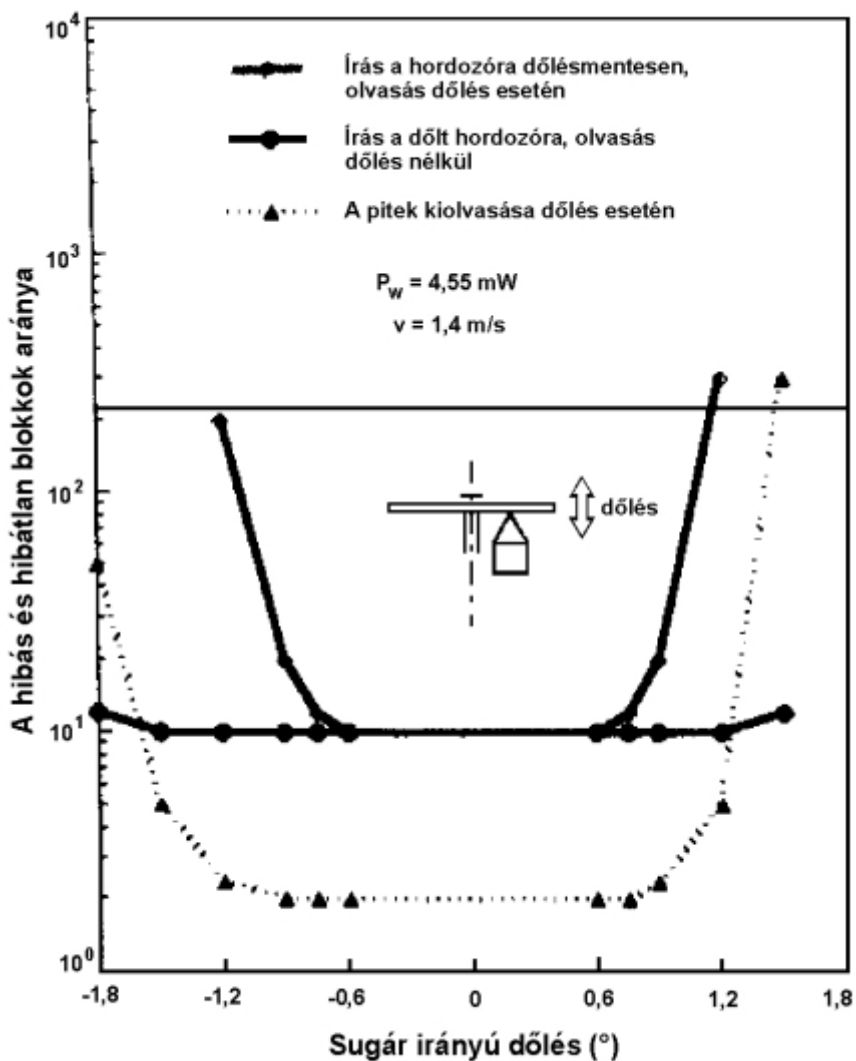
46. ábra: a mágneses-mező- és a laser moduláció összehasonlítása

A mágneses-mező modulációnak még egy nagy előnye, hogy rendkívül ellenálló a lemez dőléssel szemben. Amikor a lemez meghajlik, a fénypont eltorzul. A laser moduláció esetében a felvett bitmintának az alakját a fénypont szabja meg. A torzulás elég jelentős lehet. A mágneses-mező moduláció során a lasert a magneto-optikus réteg hőmérsékletének növelésére használják és nem befolyásolja a felvett bitmintázat alakját. Így ha felvétel közben a lemez megdől, vagy elhajlik kisebb lesz a torzulás.



47. ábra: érintő irányú dőlés türése

A 47. és a 48. ábrán láthatók egy kísérlet eredményei. A 47. ábra a hibás blokkok arányának ingadozását mutatja a lemez dőlésének függvényében felvételkor és lejátszáskor. A lemez akár +/- 1,5 fokkal is megdőlhét a barázdák irányában, a blokk hibaarány még akkor is majdnem stabil marad, mindaddig, míg olvasáskor is meg nem dől a lemez. A 48. ábrán látható a mérés eredménye sugár irányú dőlés esetén. Nagyon hasonlóak a párhuzamos dőlésnél mért értékekhez.



48. ábra: sugár irányú dőlés tűrése

Amint azt már elmondtuk számos előnye van a mágneses-mező modulációnak. Ezeknek a tulajdonságoknak köszönhető a tömegtermelésben készülő MD rendszerek nagy megbízhatósága felvételkor és lejátszáskor, habár bonyolultnak tűnik a mágneses fej elkészítése. Egy a lemezzel érintkező író fej alkalmazása viszonylag egyszerű megoldás. Az MD lemez felső részét be kell vonni egy vékony védőréteggel, ami megvédi a hordozó réteget a környezeti hatásoktól és ezzel a réteggel beállítható a súrlódási erő a megfelelő értékre. Így, ha az író fej csak durván érintkezik is ezzel a réteggel, akkor is megfelelő nagyságú mágneses mezőt hoz létre. Rendkívül megbízható ez a módszer, akár milliószor is körbefordulhat a lemez. Lényeges, hogy a fej meghajtó árama elég kicsi ahhoz, hogy telepről működjön.

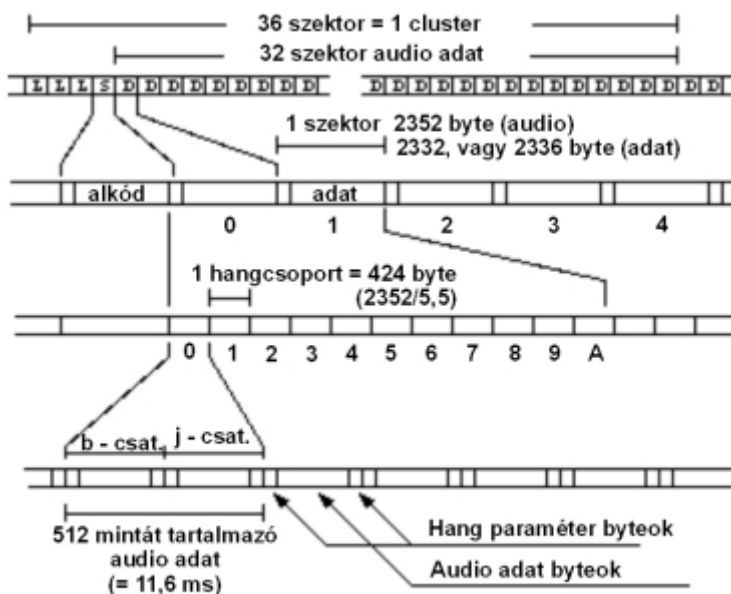
6.4. Adat struktúra

Az előre elkészített és az írható MD rögzített jele nagyon hasonlít a CD -nél alkalmazotthoz. EFM -et (Eight-to-Fourteen Modulation) és CIRC -t (Cross Interleaved Reed-Solomon Code) alkalmaznak. Mint ismeretes, az EFM és a CIRC kombinációja potenciális hibajavítási lehetőségeket és nagy adatsűrűséget eredményez. A "csupasz" CD lemez is hosszú távon képes az adatok biztonságos tárolására. Így, ha a lemezt egy tokba helyezzük, elvárható a megbízhatóság növekedése. Ez erősen befolyásolta a tervezőket, hogy ugyanezt a rendszert használják.

Az audio CD -től eltérően kissé megváltoztatták a CIRC szekvenciákat egy olyan keret formátumra, ami alkalmas az egyes audio szekvenciák közti szünetek és a folyamatos adat szekvenciák kezelésére is. A neve: ACIRC (Advanced CIRC).

A MD rendszer csoportokba rendezve rögzített, tömörített audio adatot használ. A blokk formátuma nagyon hasonló a CD-ROM mode-2 -höz. A CD-ROM esetében a blokk cím alapja a CD subcode címe, így a fizikai cím amit használunk a játékidőn alapszik. A CD és a CD-ROM percek, másodpercek és kereteket használ. Az MD-nél felváltották az idő alapú címezést a bináris kód alapúval, ami keret (blokk) egységeken alapszik. Egy ilyen egységet (keret, blokk) szektornak nevezünk. A hosszú ACIRC hibajavító kód miatt három szektort "bekötő" szektorként kell használni. Ha a felhasználó megváltoztat, vagy új adatokat rögzít az MD lemezre, két vagy három szektort kell rögzíteni az új adat kezdő és vég pozíciójánál.

A rögzítendő adat hossza változó, ami problémát okoz a címezés szempontjából, ezért bevezették a cluster (csoport) fogalmát, ami meghatározza a rögzítendő, vagy felülírandó adat minimumát. Egy cluster 36 szektorból áll. Ebből 32 szektort használnak az információ tárolására, a további négyet pedig az adatok összekötésére (linking - feladatuk a CD-nél alkalmazott mergin bitekhez hasonló), illetve plusz információk rögzítésére (49. ábra). Feljebb elmondtuk, hogy a hordozó előkészítését (pregroove) a 22,05 kHz -es vivő végzi, amit a cím adat modulál. A cím a szektor és a cluster számmal egyezik meg. A meghajtó egység bármely írási és olvasási folyamat közben megtalálja a helyes címet a modulált hordozó segítségével.



49. ábra: az MD lemez adatstruktúrája

Az információ hordozó felület belső részén található a tartalomjegyzék (TOC = Table Of Contents), amit előre égetett felületen rögzítenek (gyárilag). Néhány szektor a lemez azonosítását, illetve az adott lemez típushoz tartozó plusz információkat tartalmazza. Az írható MD lemezen a TOC -ban található néhány olyan paraméter is, amit felvétel készítéskor használnak (optimális felvevő teljesítmény, a felhasználói terület kezdő és vég címe, stb.). A gyári MD TOC -ja tartalmazza a sáv felosztás táblázatát (track allocation table), a track -ek nevének táblázatát, a felvétel dátumát, stb. A track felosztás táblázatnak van egy

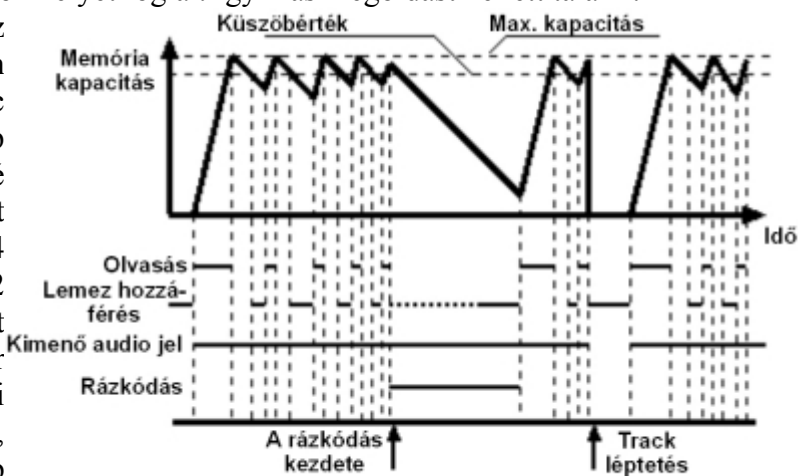
része, ami a zeneszámok kezdő és vég címére mutat a program területen. A track név, és dátum táblázat opció. A felhasználói tartalomjegyzéket (UTOC = User TOC) felvétel

készítés, vagy szerkesztés után rögzítik a lemezre. A UTOC track felosztás táblázata ugyanúgy kezeli a program területeket. Maximum 255 zeneszámot definiál a track felosztás táblázat, plusz lehetőség nyílik a feldarabolt track -ek összekötésére, folyamatosként feltüntetve azokat. A rázkódásálló memória alkalmazásával az MD képes automatikusan összefűzni a szét darabolt részeket. Így az MD rendszer könnyű és megbízható működést biztosít.

6.5. Rázkódás védelmi memória

Egy hagyományos optikai lemezes rendszer könnyen hibázhat, ha rázkódás, vagy ütés éri. Emiatt sokáig úgy gondolták, hogy a mágnesszalagos média a legjobb megoldás kültéri használatra. Bár az autókba szerelhető CD lejátszókat mechanikai felfüggesztéssel védik a rázkódástól, a hordozható berendezésekben - mint a Walkman - bonyolult mechanikai védelmet alkalmazni, mert túl sok helyet foglal. Így más megoldást kellett találni.

Tapasztalatból tudjuk, hogy az ütések és a rázkódások nem folyamatosak. Ezért a MiniDisc rendszer egy az optikai pick-up és az ATRAC kódoló közé helyezett félvezető memóriát használ adat pufferelésre. Egy 4 Mbit -es memóriában kb. 12 másodpercnyi tömörített hanganyag tárolható. Amikor az optikai pick-up elkezd olvasni az adatokat a lemezről, néhány másodpercnél kevesebb idő alatt feltöltődik a memória, mert az optikai olvasás és az



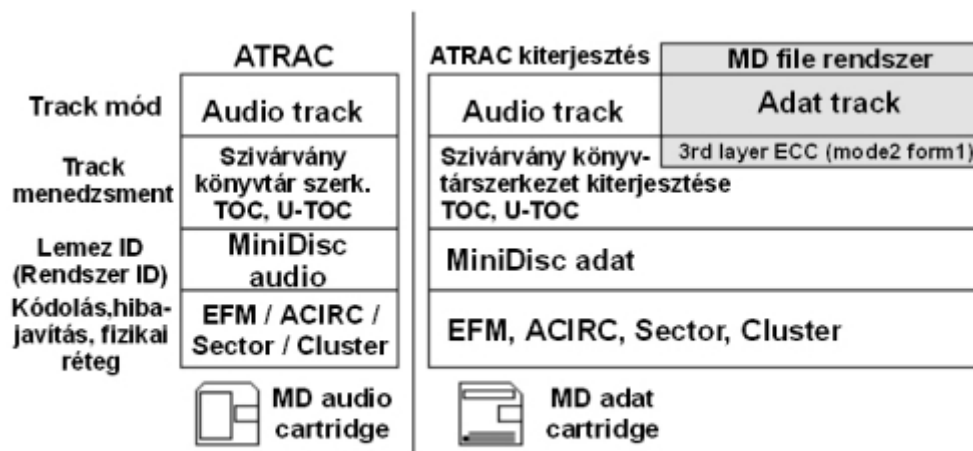
50. ábra: a rázkódásvédelem működése

ATRAC dekóder között ötszörös bitsebességbeli különbség van. Így normál lejátszáskor az optikai pick-up szakaszosan olvassa az adatot a lemezről, hogy folyamatosan feltöltve tartsa a memóriát. Ha a pick-up kiesik a pozíciójából ütés/rázkódás miatt és az adatfolyam megszakad, az adatokat a dekóder továbbra is olvashatja a memóriából így biztosítva a folyamatos lejátszást még 12 s-ig. Amint a pick-up visszanyeri a pozícióját, újra feltölti a memóriát. Mivel a lemez egész területén kb. 13,3 ms -ként található cím információk, a pick-up általában 1 s alatt visszanyeri a pozícióját (50. ábra). Mostanában a CD lejátszó gyártók is alkalmaznak egy majdnem ugyanilyen módszert. Így ez az optikai lemezes rendszer már megfelel kültéri használatra.

6.6. MD-adat szabvány

Az audio MiniDisc szabványra épülő MD-adat szabványt 1993 -ban fejlesztették ki. Az MD-adat lemez és a kulcsfontosságú alkatrészek (optikai pick-up és LSI chip) majdnem ugyanazok, mint az audio MD rendszerénél. Így általában ugyanaz az infrastruktúrája az MD-adat lemez gyártásának is, mint az audio MD lemezé. Tulajdonképp hasonlít a CD-ROM előállítás folyamatához. Az MD-adat lemeznek kb.140 Mbyte a kapacitása (51. ábra). A különböző lemeztípusoknak (csak olvasható, írható és hibrid lemez) és a kis méretnek köszönhetően az MD-adat rendszer a jövő adathordozója és publikációs médiája lehet. Annak ellenére, hogy az audio és az adat MD jellemzői nagyon hasonlóak, az audio MD fizikai jellemzőit továbbfejlesztették. Lényeges, hogy az audio vásárló ne keverje össze a két

lemeztípust. Emiatt az adat lemez házának alakja eltérő (51. ábra). Szükség volt a jobb adatminőség és a magasabb adatátviteli sebesség biztosítására, ezért egyes jellemzők megváltoztak, hogy megfeleljenek a kívánalmaknak. Plusz az MD célorientált file struktúrával rendelkezik, ami különböző platformokon futó alkalmazásokat támogat.



51. ábra: az audio- és az adat MD összehasonlítása

Az adat lemez file struktúráját és terjedelmét úgy tervezték, hogy megbírkózzon az MD fizikai korlátjaival. Az adat sávot - ugyanúgy, mint az audio lemeznél - a TOC -ban és a UTOC -ban definiálják. Az első cluster a boot cluster, a további 16 pedig a VMA (Volume Management Area) -t tartalmazza, amiben az összes file és könyvtár kezelési információ található. A VMA fizikailag középen helyezkedik el és általában egy félvezető memóriával cache -elik, a hozzáférések illetve az újraírások száma minimális. Továbbá meggondolva, hogy a cluster egy minimális egységet jelent, az adat lemez file struktúrája 2 - 64 kbyte -ig terjedő méretű blokkokat képes kezelni. A blokkméretet automatikusan választja ki a rendszer a felhasználó által használt alkalmazásoknak megfelelően. A file rendszer támogatja a hierarhikus könyvtárszerkezetet, rövid- és hosszú file neveket, plusz információkat, stb., így könnyen csatolható különféle gazdarendszerekhez, illetve lehetővé teszi az adatcserét a különböző rendszerek közt (10. táblázat).

10. táblázat: a MiniDisc rendszer paraméterei		
Fő paraméterek	Audio MD	Adat MD
Kapacitás	Max. 74 perc	140 Mb
Cartridge méret	68(d) x 72(sz) x 5(m) mm	
A diszk jellemzői		
Átmérő	64 mm	
Vastagság	1,2 mm	
A középső lyukátmérő	11 mm	
A programterület kezdetének átmérője	32 mm	

A bevezető rész kezdetének átmérője	Max. 29 mm	
Pit nagyság	1,6 mm	
Kerületi sebesség	1,2 - 1,4 m/s	1,2, vagy 2,4 m/s
Optikai paraméterek		
A laser hullámhossza	780 nm	
Lencse NA	0,45	
Laser teljesítmény rögzítéskor	2,5 - 5 mW	
Rögzítési eljárás	Mágneses-mező moduláció	
Általános audio jellemzők		
Csatornák száma	Sztereo/mono	
Frekvencia tartomány	5 Hz - 20 kHz	
Dinamika	105 dB	
Nyávgóság, lebegés	Quartz precíztség	
Jelformátum / adatstruktúra		
Mintavételezési frekvencia	44,1 kHz	
Tömörítés	ATRAC	
Moduláció	EFM	
Hibajavító eljárás	ACIRC	
Szektor méret		2048, vagy 2336 byte
Cluster méret		32 szektor
Maximális cluster szám		kb. 2200
Maximális szektor szám		kb. 70000
Blokkméret		2 kbyte - 64 kbyte

6.7. ATRAC: Adaptive Transform Acoustic Coding (Alkalmazkodó Leképezésű Akusztikus Kódolás)

A lecsökkent tárolókapacitás ellenére szükséges volt, hogy a MD jó hangminőséget adjon és 74 perces játékidővel rendelkezzen. Ezért az ATRAC -ot úgy tervezték, hogy megfeleljen a következő követelményeknek: a 16 bites, 44,1 kHz -es sztereo audio jelet kevesebb, mint 1/5 - ére tömörítse össze a hangminőség minimális romlása mellett. Hordozható készülékekben is alkalmazható egyszerű és olcsó hardverrel is megoldható legyen a kódolás és a dekódolás.

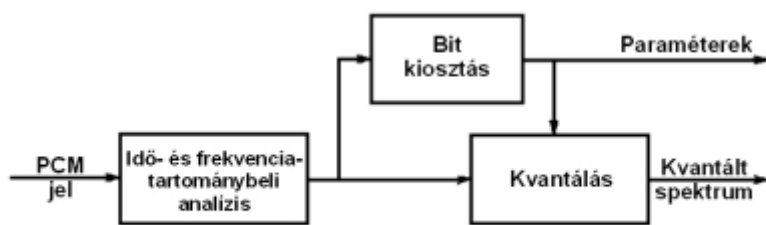
A digitális audio jel tömörítésekor bizonyos mennyiségű zaj adódik a jelhez. A legtöbb audio kódolási eljárásnak az az alapelve, hogy az ilyen zaj idő- és frekvencia tartománybeli eloszlását úgy kezeljék, hogy az hallhatatlan legyen az emberi fül számára. Ha ez sikeres, akkor az ilyen módon előállított jel megkülönböztethetetlen lesz az eredetitől.

Az audio kódoló rendszerek általában úgy működnek, hogy felosztják a jelet különböző idő- és frekvenciatartománybeli egységekre. Ezt a felosztást használva a jelet a pszihoakusztikai törvényszerűségeknek megfelelően elemzik. Az elemzés során kimutatható, hogy mely egységek a kritikusak és ezért nagy pontossággal kódolandók, és mely egységek kevésbé jelentősek és így elviselnek némi kvantálási zajt a kívánt hangminőség romlása nélkül. Erre az információra alapozva rendelik az elérhető biteket az egyes egységekhez. A spektrális összetevőket ezután kvantálják ezeknek a biteknek a felhasználásával. A dekóderben a kvantált spektrumot visszaalakítják a bit kiosztásnak megfelelően és ezután audio jellé szintetizálják.

Az ATRAC rendszer is a fentieknek megfelelően - csak számos kiegészítéssel - működik. A pszihoakusztikus jellemzőket ugyanis nem csak a bit kiosztás során használja, hanem az idő- és a frekvencia tartomány felosztásakor is. A bemeneti jelet nem azonos frekvencia osztásokban analizálják, és így kiemelik a fontos, alacsony frekvenciás összetevőket. Ezen kívül változó hosszúságú blokkokat használnak, ami a bemeneti jelhez igazodik. Így a statikus részek is hatékonyan kódolhatók és az átmeneti, gyorsan változó részekre nem pazarolják az időbeli felbontást.

Elsőként az ATRAC kódolót vizsgáljuk meg az idő- és frekvencia tartománybeli felosztás, a spektrális összetevők kvantálása és a bit kiosztás alapján, majd az ATRAC dekódert mutatjuk be.

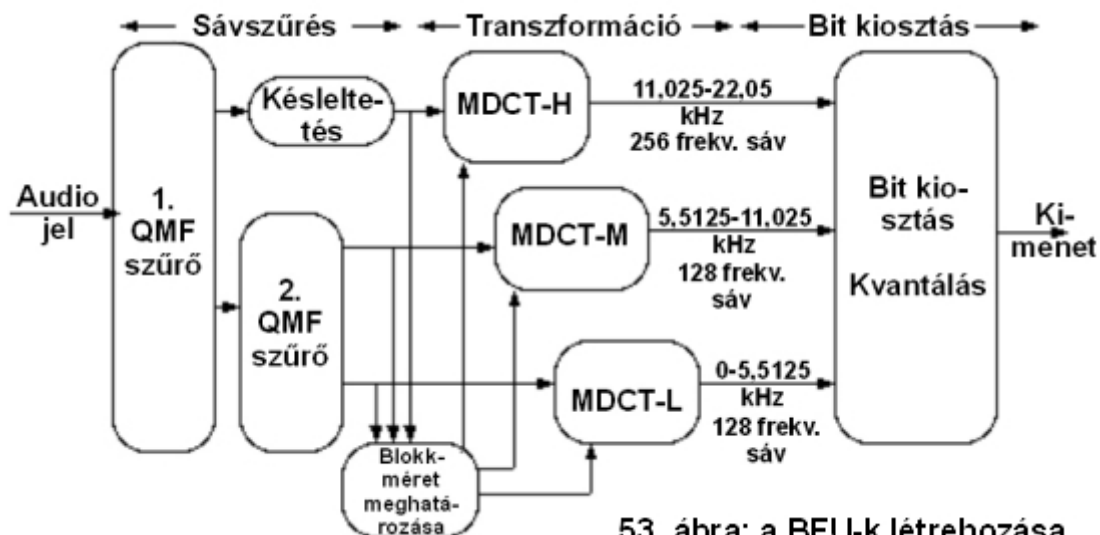
6.7.1 Az ATRAC kódoló



52. ábra az ATRAC kódoló blokkdiagramja

Az 52. ábrán látható a kódoló blokkdiagramja. A kódoló három részből áll. Az analízis rész spektrális összetevőire bontja a jelet, amelyeket csoportokba (BFU = Block Floating Unit) rendeznek. A bit allokáció rész elosztja a

rendelkezésre álló biteket az egyes BFU -k között, kevesebb bitet osztva a kisebb jelentőségű csoportoknak. A kvantálás rész kvantálja az egyes spektrális összetevőket a meghatározott szóhosszúságra.



53. ábra: a BFU-k létrehozása

Idő - frekvencia analízis

Ez a rész generálja a BFU -kat három lépésben, az alsáv- és az átviteli kódolás segítségével.(53. ábra). Először a jelet felosztják három alsávra: 0-5,5 kHz, 5,5-11 kHz és 11-22 kHz. A kritikus frekvenciasávok listáját a 11. táblázat tartalmazza. Ezután minden egyes alsávot átranszformálnak a frekvencia tartományba, így létrehozva a spektrális összetevőket. Végül ezeket a spektrális összetevőket nem azonos méretű csoportokba - a BFU -kba - fogják.

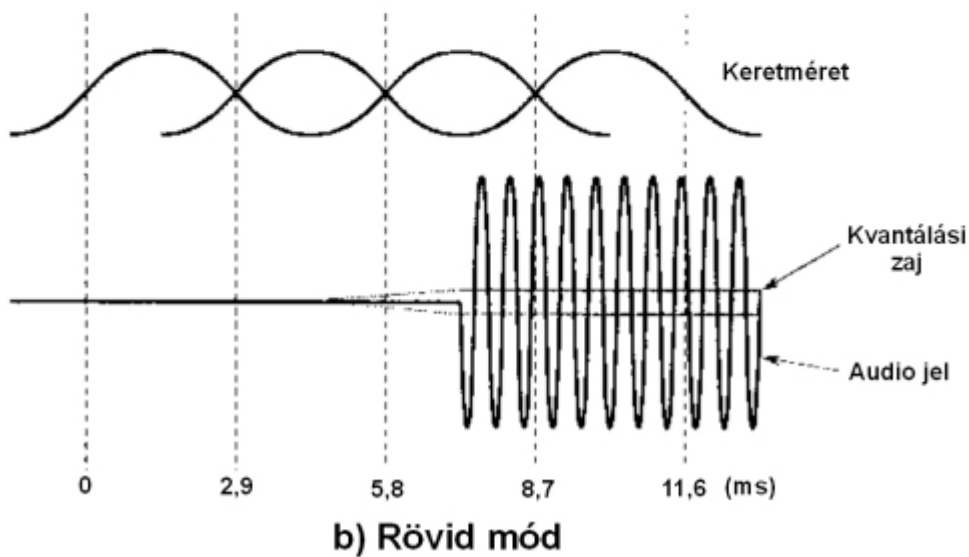
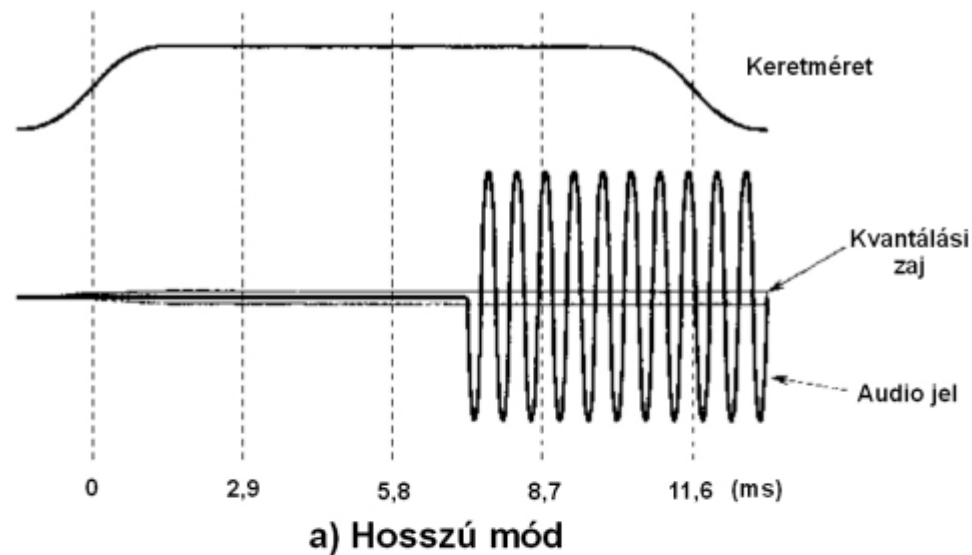
11. táblázat: az ATRAC által használt kritikus frekvenciasávok

Kritikus sáv	Frekvencia (Hz)			Kritikus sáv	Frekvencia (Hz)		
	Alsó	Felső	Szélesség		Alsó	Felső	Szélesség
0	0	100	100	13	2000	2320	320
1	100	200	100	14	2320	2700	380
2	200	300	100	15	2700	3150	450
3	300	400	100	16	3150	3700	550
4	400	510	110	17	3700	4400	700
5	510	630	120	18	4400	5300	900
6	630	770	140	19	5300	6400	1100
7	770	920	150	20	6400	7700	1300
8	920	1080	160	21	7700	9500	1800
9	1080	1270	190	22	9500	12000	2500
10	1270	1480	210	23	12000	15500	3500
11	1480	1720	240	24	15500	22050	6550
12	1720	2000	280				

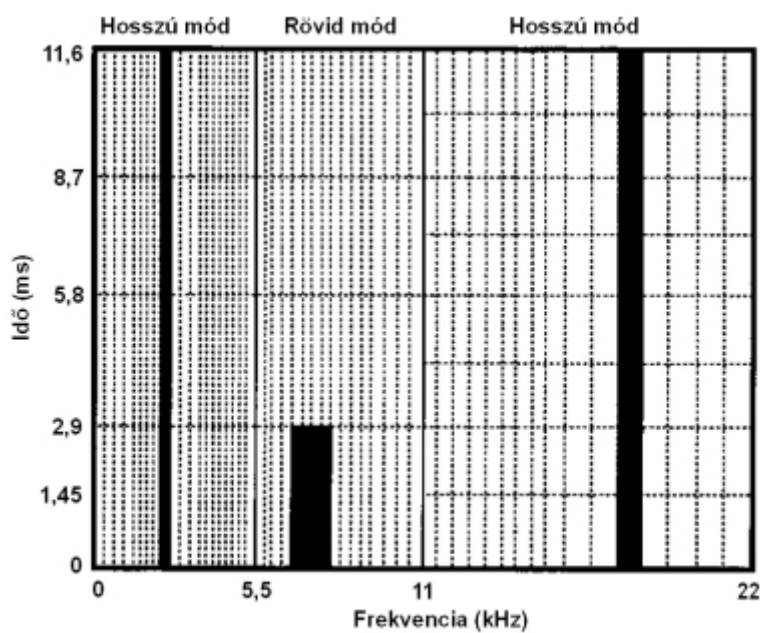
Az alsávokra történő felosztást QMF (Quadrature Mirror Filter) szűrők használatával oldják meg. A bemenő jelet felső- és alsó sávra osztja az első QMF szűrő, majd a másodikat ismét két részre osztja a második szűrő. A QMF -ek használata biztosítja, hogy az alsávokra osztás miatt keletkezett időtartománybeli aliasing megszűnjön a rekonstrukció során.

Ezután mindhárom alsávot átranzformálják a frekvencia tartományba az MDCT segítségével (Modified Discrete Cosine Transform). Az MDCT -t használva 50 % -os átlapolás engedhető meg az időtartománybeli ablakok között, ami jobb frekvencia felbontást eredményez a kritikus mintavételezés megtartása mellett. Fix keretméret helyett az ATRAC az egyes sávokhoz a jelek karakterisztikáitól függő kerethosszot választ. Két üzemmód van: hosszú (11,6 ms) és rövid (1,45 ms a magas frekvencia tartomány számára és 2,9 ms a többinek). Normál esetben a hosszúat használják, a jó frekvencia felbontás érdekében. Azonban a jel felfutásakor, illetve bizonyos részeinél hibák adódhatnak. Általában a kvantálási zaj terjed szét az egész blokkon és a kezdeti kvantálási zajt még semmi sem maszkolja; ezt hívják elővisszhangnak. Ennek kiküszöbölésére a kódoló átkapcsol rövid módba, a jel felfutásakor. Ilyenkor a jel előtt csak egy rövid részen található zaj, amit elnyom a visszafelé irányuló maszkolás. A visszafelé irányuló maszkolás nem hatásos a hosszú mód esetén, mert nagyon rövid ideig áll fenn. Így az ATRAC kódoló hatásos a statikus és az átmeneti jelek kódolásakor is (54. ábra).

Vegyük észre, hogy a rövid mód nem szükséges a jel gyengülésekor, mert ilyenkor a kvantálási zajt az előre irányuló maszkolás nyomja el, ami sokkal tovább tart, mint a visszafelé irányuló. A maximális rugalmasság biztosításához minden egyes sávhoz külön választható a blokk méret üzemmódja.



54. ábra: a kvantálási zaj hosszú- és rövid mód esetén



55. ábra: nemlineáris frekvencia felosztás

Ezután az MDCT spektrális összetevőket BFU -kba csoportosítják. Minden csoport fix számú összetevőt tartalmaz. Hosszú mód esetén egy ilyen egység 11,6 ms -ot jelent egy keskeny frekvencia sávra; rövid módban minden csoport időben kevesebb, de nagyobb frekvencia felbontást biztosít (55. ábra). A BFU -k sűrűsége alacsony frekvenciákon nagyobb, mint magas frekvenciákon; ez az emberi fül pszichoakusztikus tulajdonságait tükrözi. Minden egyes hang keret hossza fix 212 byte. A

lemezre 11 sztereo hang keretet rögzítenek minden két szektorba. Mono felvétel készítésekor kétszeres mennyiségű hang rögzíthető a lemezre.

Spektrális kvantálás

A spektrális értékeket két paraméter - a szó hossz és a skála felbontás alapján kvantálják. A skála felbontás a kvantálás teljes sávszélességét definiálja, míg a szó hosszúság a felbontás precízségét ezen a skálán belül. Az egy spektrumhoz tartozó BFU-knak ugyanakkora ez a két paraméterük, ami az egy csoportba foglalt frekvenciák pszihoakusztikai hasonlóságát tükrözi. A skála felbontási tényezőt - ami a spektrális összetevők számosságát tükrözi egy BFU -n belül - fix értékek közül választják. A szó hosszt a bit allokációs algoritmus határozza meg. Minden egyes hangkerettel (ami 512 bemeneti pontból áll) együtt a következő információkat is tárolják a lemezen:

- az MDCT blokk méret módja (hosszú, vagy rövid).
- adat a szó hosszról az egyes BFU -kra.
- skála felbontási paraméter az egyes BFU -kra.
- kvantált spektrális összetevők.

Bit allokáció

A bit allokációs algoritmus osztja el a rendelkezésre álló biteket az egyes BFU -k között. A sok bittel rendelkező egységek mentesek lesznek a kvantálási zajtól, míg a kevésse rendelkezőkben jelentős lesz a zaj. A jó hangminőség érdekében az algoritmusnak elegendő bitet kell biztosítania a kritikus egységek számára és ügyelnie kell arra, hogy a zaj a nem kritikus egységekben ne legyen hallható.

Az ATRAC nem definiál speciális bit allokációs algoritmust - bármely megfelelő algoritmus használható. Az egyes BFU -k szó hossza a MiniDiscen tárolódik a kvantált spektrummal együtt, így a dekóder teljesen független a bit allokációs algoritmustól. Ez biztosítja a kódoló evolúciós fejlődését a MiniDisc szabvány, vagy a dekóder megváltoztatása nélkül.

Nagyon sok a szóba jöhető algoritmus az egészen egyszerűektől a szélsőségesen bonyolultakig. A hordozható MD rekorderek lehetőségei ezen a téren kissé korlátozottak, mert olcsó és kis teljesítmény igényű, kompakt hardver elemekből kell felépíteni őket. Az ATRAC azonban egyszerű bit allokációs algoritmus esetén is jó hangminőséget ad a pszihoakusztikai jellemzői miatt. Már a kódoló változó idő-frekvencia struktúrája is pszihoakusztikai törvényszerűségeken alapul, tehermentesítve ezzel az allokációs algoritmust. Az egyik javasolt algoritmus fix és változó bitek kombinációját használja. A fix bitek hangsúlyozzák a fontos, kis frekvenciájú tartományokat, kevesebb bitet adva a magas frekvenciákat tartalmazó BFU-knak. A változó biteket az egyes BFU-kban található spektrális összetevők logaritmusának megfelelően osztják ki. A totális bit allokáció ($b_{tot}(k)$) a fix bitek ($b_{fix}(k)$) és a változó bitek ($b_{var}(k)$) súlyozott összege:

$$b_{tot}(k) = T b_{var} + (1-T) b_{fix}$$

A T súlyozó tényező a hang tonalitásának mértéke: tiszta szinuszos hang esetén 1, fehér zaj esetén 0 az értéke. Ez azt jelenti, hogy a fix és a változó bitek aránya változtatható. Ezért tiszta hangok esetén a rendelkezésre álló biteket kevesebb BFU között osztják el. Zajosabb jelek esetén az algoritmus a fix biteket fogja kihangsúlyozni, hogy csökkentse a kevésbé érzékeny, magas frekvenciákhoz kiosztott bitek számát.

A fenti egyenlet nem vonatkozik a teljes bitsebességre, és általában több bitet oszt ki, mint ami rendelkezésre áll. A fix adatsebesség biztosítása érdekében az összes BFU -ra egységesen vonatkozó offsetet számolnak (b_{off}). Ezt az értéket minden $b_{\text{tot}}(k)$ -ből levonják, így áll elő a végső bit kiosztás ($b(k)$).

$$b(k) = \text{integer}[b_{\text{tot}}(k) - b_{\text{off}}]$$

Ha a kivonás negatív számot eredményez, akkor a kérdéses BFU -nak 0 bitet oszt ki az algoritmus.

6.7.2. Az ATRAC dekóder

Az 56. ábrán látható a dekóder blokk diagramja. A dekóder először a spektrális összetevőket állítja elő a kvantált értékekből a

szó hosszát és a skála felbontási paramétert használva. Ezekből a spektrális összetevőkből állítja elő az eredeti audio jelet. Az összetevőket először visszatranszformálják az idő tartományba az inverz MDCT segítségével, hosszú- és rövid módot használva a paramétereknek megfelelően. Végül a három, időtartománybeli jelből szintetizálják a kimeneti jelet QMF szintézis szűrőkkel.



Összefoglalás

A bemutatott rendszerek különböző felépítésének és működésének köszönhetően más-más alkalmazási területeken mutatkoznak meg előnyeik, illetve hátrányaik.

A DAT rendszer széleskörűen elterjedt a stúdiotechnikában, mert jó minőségű és megbízható hangrögzítést tesz lehetővé. A CD rendszerrel szemben a hanganyag szerkesztése egyszerűbben megoldható, a mester szalag elkészítéséhez nincs szükség speciális körülményekre. Az otthoni felhasználásban a készülékek ára (egy DAT kazetta kb. 1300 Ft) és a hagyományos kompakt kazettás, illetve a CD rendszerek tömeges jelenléte miatt nem terjedt el még nagy mértékben. Számítógépes háttértárként való alkalmazása különösen nagy adatbázisok, vagy egyéb nagy tárolókapacitás igényű adatok (pl. digitális video) tárolását, archiválását segíti elő.

A MiniDisc rendszer már jobban megfelel az otthoni felhasználás által támasztott követelményeknek. A DAT rendszerhez képest olcsóbb üzemben tartási költségek, egyszerűbb kezelhetőség, gyorsabb hozzáférés a rögzített anyaghoz a hétköznapi felhasználó számára jó alternatívaként tünteti fel ezt a rendszert. Ma már kb. 50000 Ft-tól kapható a boltokban MiniDisc rekorder, a lemez ára kb. 700 Ft. Stúdió célokra azonban kevésbé alkalmazható az alkalmazott tömörítési eljárás miatt. Ugyanis bármennyire is hallhatatlan a bekövetkező minőségromlás, a jel esetleges további feldolgozása, szerkesztése során már problémát jelenthet. Háttértárként a mindennapi adattárolásban és -szállításban játszhat szerepet kapacitásának köszönhetően. Jellemzői alapján elterjedt adathordozóvá válhat, azonban ez a mai napig még nem következett be.

Az MPEG/audio formátum rétegeződésének köszönhetően alkalmazható digitális diktafonok audio formátumaként, közepes, vagy akár majdnem CD minőségű zenei anyagok tárolására. Segítségével megoldották az ISDN összeköttetésen alapuló internetes kapcsolaton keresztül történő, kiváló minőségű audio jel valós idejű átvitelét. Nem egy adott hardver rendszerre specializált tömörítési eljárás, így az MPEG/audio formátumú audio jelek gyakorlatilag bármilyen digitális adathordozóra rögzíthetők: merevlemezre, floppyra (persze csak korlátozott minőségben, vagy mennyiségben), CD-re, streamerekre, DAT-ra, MiniDisc-re, illetve a piacon elterjedt flash kártyákra, amelyeket a hordozható berendezésekben alkalmaznak előszeretettel.