

TDK dolgozat

Varga Péter Zoltán

ALACSONY SZÁMÍTÁSI KAPACITÁSÚ ESZKÖZÖK FELHASZNÁLÁSA MOTOROS FŰRÉSZ HANGJÁNAK FELISMERÉSÉRE

Varga Péter Zoltán

Programtervező informatikus (Bsc) - 2010

Témavezető: Dr. Czúni László

Pannon Egyetem

Műszaki Informatika Kar, Villamosmérnöki és Információs Rendszerek Tanszék



2014

Tartalomjegyzék

Kivonat	4
1 Az erdővédelem korszerű eszközei	5
2 Szenzorhálózatok a kültéri erdei megfigyelésben	7
2.1 A fejlesztett rendszer topológiája.....	8
2.2 Feladat-specifikáció	9
2.3 A hangalapú detekció sajátosságai.....	12
2.4 Felhasználási lehetőségek	13
3 Akusztikus jelek feldolgozása	14
3.1 Előfeldolgozás.....	14
3.2 Tulajdonságok generálása	17
3.3 Osztályozó módszerek	19
3.4 Spektrális tér, Fourier transzformáció (FT)	20
3.5 Autokorrelációs szekvencia	21
3.6 TESPARE.....	23
4 Implementáció és tesztelés	26
4.1 Terepi mérések.....	26
4.2 Adatbázisok összeállítása.....	27
4.3 Autokorrelációs eljárás	28
4.4 TESPARE.....	36
5 Fúziós lehetőségek.....	49
7 Fejlesztési lehetőségek	57
8. Összegzés.....	61
Köszönetnyilvánítás.....	63
Hivatkozások	63

Kivonat

Az erdőterületek védelme kiemelt fontosságú feladat hazánk számára, hiszen országunk mintegy 20 százaléka erdő. Ennek ellenére napjainkban a természeti ártalmak és katasztrófák mellett az illegális fakitermelés ellen nem védekezünk elég hatékonyan. 2009-es adatok szerint a feketegazdaság az erdei munkafolyamatokban nagyrészt az engedéllyel végzett erdőtisztítás eredményeként keletkező fa tüzelőanyag illegális értékesítéséből, valamint engedély nélküli favágásból, erdőirtásból áll, melyek együttes eredménye éves szinten 3-3,5 millió köbméter illegális faanyag. Az engedély nélkül kitermelt fa árusításának ellenőrzésére rengeteg módszer létezik, az illegális kitermelés megelőzésére valamint a legálisan termelt faanyag továbbárusításának ellenőrzésére viszont igen csekély számú és korlátozott eredményű megoldásról olvashatunk.

A dolgozat célja egy komplex, vezeték nélküli szenzorhálózatot alkalmazó erdővédelmi rendszer akusztikus egységének kifejlesztésének bemutatása. A dolgozatban röviden ismertetem a már létező megoldásokat, azok korlátait, az általunk felhasznált technológia műszaki hátterét, a rendszer topológiai felépítését, illetve az általunk fejlesztett és vizsgált hangfeldolgozási és osztályozási módszerek működését, eredményeit. Több lehetséges algoritmikus megközelítést dolgoztunk ki, elemeztünk és hasonlítottunk össze, külön vizsgáltuk a jelfűzióból adódó előnyöket.

A kidolgozott módszerek képesek kis energiafogyasztású jelérzékelő és feldolgozó egységek alkalmazásával a belsőégésű motorok hangjának hatékony felismerésére a megfigyelt területen.

Dolgozatunkban akusztikus jelek feldolgozására mutatunk be szenzorhálózatot, mely a *GOP-1.1.1.-11-2011-0070 Szenzorhálózat alapú erdővédelmi felügyeleti rendszer* elnevezésű és a *TÁMOP-4.2.2.C-11/1/KONV-2012-0004* projekt keretén belül valósul meg.

1 Az erdővédelem korszerű eszközei

Az illegális fakitermelés napjainkban globális méretű probléma, az erdőterület csökkenése befolyásolja éghajlatunkat, az erdőgazdálkodási folyamatokat, és nagymértékben hozzájárul a természet biológiai sokféleségének csökkenéséhez. Ennek megelőzése érdekében az erdőterületek megfelelő felügyelete és gondos gazdálkodás szükséges.

A klasszikus vagyonvédelmi módszerek optimálisan nem alkalmazhatók optimálisan erdővédelemre, ám az elmúlt évek technológiai fejlődése az eddig létező eljárásokon felül új lehetőségek előtt nyit utat.

Röviden ismertetem az alapvető védelmi eljárások fajtáit:

Élőerő alkalmazása: A leggyakrabban alkalmazott eljárás az élőerő alkalmazása. Ebben az esetben a magánerdőket erdőkerülők, csőszök védik, míg az állami gazdaságokban az erdészeti munkatársai végzik a terület felügyeletét. Előnye, hogy látható védelmi egységként az öröknek elrettentő szerepe lehet, viszont nagy hátránya a módszernek, hogy a védeni kívánt terület nagy kiterjedésű, így a folyamatos védelem ellátásához szükséges létszámú őrgárda fenntartása nem gazdaságos.

RFID alapú azonosítás: Néhány éve több országban is megjelent az RFID alapú faárúkövetés. Többek közt Svédországban, Dániában és Malajziában a kivágott farönkhöz RFID azonosítókat erősítenek, a fafeldolgozó üzemek pedig csak ellenőrzött forrásból származó rakományt vesznek át, ezzel gátolva az illegális faértékesítést. A forrás ellenőrzését nem minden esetben RFID azonosító szolgálja, a mai napig nagyrészt égetett vagy festett vonalkódot használnak. Ezzel a módszerrel nagymértékben csökkenthető az illegális fakitermelés, viszont csak a már kivágott fa útvonalát lehet ellenőrizni. További hátrány, hogy a forgalom rejtett megfigyelésére nincsen mód, valamint az ellenőrző pontok könnyen kikerülhetők. Az „RFID Tags Applied for Tracing Timber in The Forest Products Chain” című cikk [Gjerdrum 2009] az RFID technológia alkalmazását mutatja be a kitermelt fa feldolgozási útvonalának követése szempontjából. A cikk elsődleges felhasználási módként a fa késztermékként történő árusítása kapcsán a faanyag minőségi, forrás, illetve feldolgozási útvonallal kapcsolatos adatainak ellenőrizhetővé tételét tekinti. Az RFID Journal 2009. november 2-i online cikkében [Friedlos 2009] egy malajziai fejlesztési projektről olvashatunk, melyben az RFID technológiát kifejezetten az illegális faértékesítés megelőzésére használják. A projekt rövid bemutatása szerint több ellenőrzési pontot állítottak fel minden helyen, ahol a faanyag őrizetét ellátó személyek, cégek változtak. Ilyenek a kitermelés helye, feldolgozó üzemek, rakodótelepek. A vizsgálatban egyedi azonosítóval 188 tönk közül mindössze 81 jutott el az utolsó ellenőrző pontig. A cikk ennek okát nem tárgyalja.

Műholdas felügyelet: Global Forest Watch [GFW 1997] egy a World Resource Institute által fémjelzett projekt, melynek célja a bolygó erdőterületének figyelemmel tartása. A GFW műholdak és emberi erő segítségével térképezi fel a nagy kiterjedésű zöldterület változásokat. Ahogy a neve is jelzi, globális megfigyelésre alkalmas rendszer, amelynek hátránya, hogy csak az erdőterület jelentős változását tudja detektálni, jelezve, hogy mely területeken kell nagyobb figyelmet fordítani az erdővédelemre. A rendszer egyik pozitívuma a pontos lokalizáció, és a változás mértékéről nyújtott információ. Ezzel szemben viszont, evidens gyengepontjai, hogy a műholdas felvételekről havas területeken nehéz a változást felismerni, és a kép frissítési frekvenciája korlátozott.

További műholdas megfigyelési megoldásokról olvashatunk a következő cikkekben:

- A [G. Guo 2004] írásban a MODIS felület hőmérsékleti adatai alapján történő erdőtűz kockázati elemzéssel foglalkoznak.
- A [Kyzirakos 2014] cikkben szintén az erdőtűzek felderítésére és monitorozására használnak műholdas felvételeket.

Videómegfigyelés: A képi információt használó módszerek jelenleg főként az erdőtűzek megelőzésére használatosak. Több országban alkalmaznak olyan rendszereket, ahol egy megfigyelőtornyon elhelyezett kamerák képét elemző automatika egy operátor felé továbbítja a jelzéseket. A gyanúsak ítélt képek alapján az őr tűzjelzést adhat, valamint manuális módba kapcsolhatja a rendszert és elemezheti az élőképet. Ezek a rendszerek nagyon eredményesek az erdőtűz elleni védelemben, viszont az illegális fakitermelés ellen sajnos nagyrészt használhatatlanok. A videó eszközök viszonylag drágák, nagy energiaigényűek, és a képi információ továbbításához nagy sáv szélességre van szükség és legfőképpen nehezen oldható meg nagy kiterjedésű területek felügyelete. A „Video-based Smoke Detection: Possibilities, Techniques, and Challenges” című cikk [Ziyu Xiong 2007] elsősorban a füst kép alapú detekciójára koncentrál, többek közt háttér kivonás, kontúrkeresés és kontúr osztályozás felhasználásával. A „The Development of a Real Time Forest Fire Monitoring and Management System” című írás [L. Trevis 2004] egy hőkamerákkal felszerelt rendszer felépítését és működését mutatja be.

A [Bo Lei 2013] cikkben szintén füst detektálására törekszenek videofelvételek segítségével.

2 Szenzorhálózatok a kültéri erdei megfigyelésben

A feladat áttekinthetőségének érdekében tisztáznunk kell mik is azok az ún. vezeték nélküli szenzorhálózatok. Vezeték nélküli szenzorhálózatnak nevezzük az olyan - térben elosztott - szenzoreszközök csoportját, amelyek környezetük fizikai állapotát figyelik, és a megfigyelési adatokat egy központi helyre továbbítják. Ilyen fizikai tulajdonság lehet például a hőmérséklet, páratartalom, hangnyomás, szélsébség vagy a levegő összetétele. A vezeték nélküli szenzorhálózatok fejlesztésének elsődleges célja katonai alkalmazások voltak, mint például harcér megfigyelés, célpontkövetés, orvlövész detektálás. Ennek ellenére napjainkban egyre több területen használnak ilyen hálózatokat, többek közt gyártási folyamatok felügyeletére, természeti katasztrófák előrejelzésére, vízminőségi, környezetvédelmi vagy meteorológiai adatok mérésére.

A szenzorhálózatok csomópontokból épülnek fel, melyek egy vagy több másik csomóponthoz kapcsolódnak. Az infrastruktúra legalsó szintjén található érzékelővel vagy kommunikációs eszközzel ellátott egységeit „mote”-nak szokás nevezni. Általánosságban egy mote felépítése szerint rendelkezik valamilyen energiaforrással, egy adattároló egységgel, rádiókommunikációs modullal, mikrovezérlővel, valamint egy - vagy több - hozzá csatolt mérő egységgel. Az ilyen csomópont eszközök mérete és ára is nagyon különböző lehet. Mérete pénzérmétől cipős doboz nagyságúig terjed, míg ára párezer forint és több tízezer forint közé eshet.

A szenzorhálózatok jellemző tulajdonságai:

- alacsony energiafogyasztás
- könnyű használat
- egyszerű skálázhatóság
- nagy ellenálló képesség környezeti hatásokkal szemben
- különböző típusú szenzoregységek
- gyakran előforduló hibák a távoli egységek kommunikációjában.

Néhány eddig publikált vezeték nélküli szenzorhálózatos alkalmazás erdővédelmi céllal:

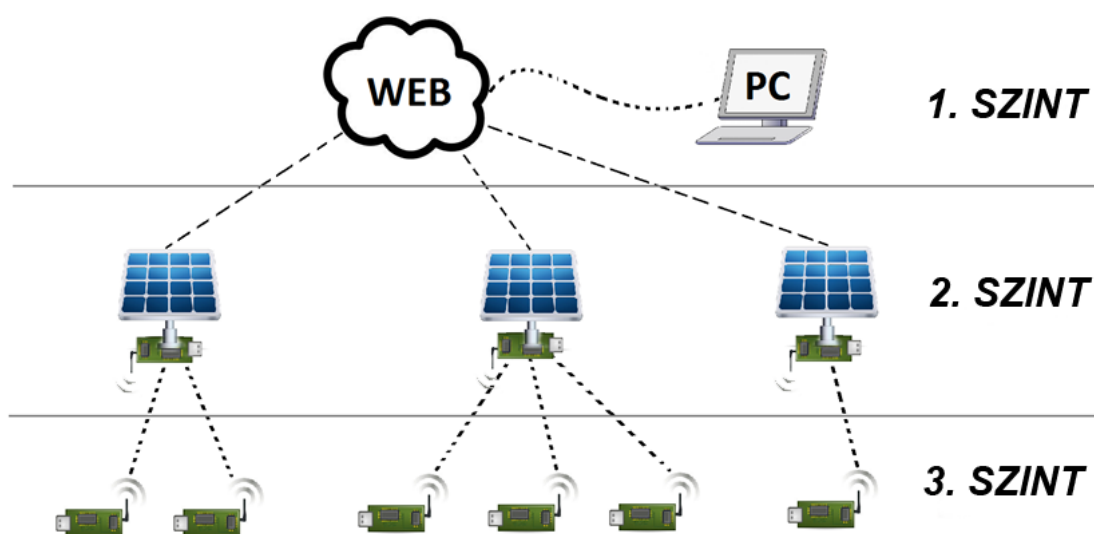
- A [Son 2006] valamint [Lloret 2009] cikkekben erdőtűz jelzésére használnak vezeték nélküli szenzorhálózatokat.
- A [Papán 2012] Illegális fakitermelés jelzésével foglalkozó írás az egyik általunk fejlesztett megoldáshoz hasonlóan a hangfelvételek autokorrelációja alapján próbálja motoros fűrészek hangját jelezni erdei környezetben. A cikk sajnos nem részletezi sem az osztályozás, sem a tesztelés módszerét, viszont megfelelő alapot képzett a fejlesztéshez.
- A [Babiš 2011] cikkben hangfelismerésen alapuló erdei tevékenységek detekcióról olvashatunk. Az írás elsősorban a szenzorhálózat architektúrájával foglalkozik, a lokális szenzoregységek detekciós algoritmusáról nem tesz említést.

Az általunk fejlesztett eljárások és a bemutatott eszközök egy komplex, multi-modális erdővédelmi rendszer részei. Az alkalmazott vezeték nélküli szenzorhálózat elsődleges célja az erdőterületek védelme gyorsulásmérők (accelerométerek), magnetométerek és akusztikus szenzorok felhasználásával. A földbe helyezett gyorsulásmérők és magnetométerek az elhaladó járművek detektálásáért felelősek, az akusztikus egységek pedig a belsőégésű

motorok hangjának felismeréséért. A rendszer logikai és kommunikációs topológia alapján mérési, szállítási és adatelemző rétegre bontható, a rétegeket a későbbiekben bővebben ismertetem. A bemutatott módszerek nagy előnye az eddig ismertetett eljárásokkal szemben a területek folyamatos felügyelete, valamint a védelmi rendszer várható alacsony költsége. Ideális esetben a rendszert adott területen elegendő egyszer kiépíteni, majd a továbbiakban az eszközök energiafogyasztásától függően tíz-tizenkét havonta akkumulátorokat cserélni, ill. feltölteni azokat. Az aktív felügyeletet az eszközök végzik, és elegendő egy emberi operátor időszakos beavatkozása a jelzések felülvizsgálatára, riasztás esetén a megfelelő lépések megtételére.

A dolgozat az akusztikus szenzorokon működő jelfeldolgozás lépéseit és módját mutatja be mint saját eredményeket, a többi szenzor fejlesztését más munkatársak végezték el.

2.1 A fejlesztett rendszer topológiája



1. ábra Erdővédelmi rendszer topológiája

Az 1. ábra a Pannon Egyetem Műszaki Informatikai Karán fejlesztett erdővédelmi rendszer logikai topológiáját jeleníti meg. Az egyes rétegekben található eszközök az akusztikus szenzorok szinkronizációs, valamint a felhasználó és az adatgyűjtő rendszer közti üzeneteken kívül, kizárólag a szomszédos rétegben lévő, hozzájuk kapcsolódó eszközökkel kommunikálnak.

A 3. rétegben az egy bázisállomáshoz kapcsolódó eszközcsoportokat szenzorszigeteknek nevezzük. Egyes szenzorszigetek gyorsulásmérői és magnetométerei egymástól független méréseket készítenek, míg az akusztikus eszközök akár egymással szinkronban is készítenek hangfelvételeket. Ezen mérések értékelését az eszközök a bázisállomások felé továbbítják.

A 2. rétegben található napelemes bázisállomások elsődleges célja az adatgyűjtés és továbbítás. Az ilyen központi állomások egy mikrokontroller egységből, egy GPRS vezérlőkártyából, 7Ah kapacitású akkumulátorból, valamint egy közel 0,15 m² méretű napelem panelből állnak. Ezen állomások a 3. rétegből érkező osztályozási információkat továbbítják a webes adatfeldolgozó rendszer felé, valamint a felsőbb réteg felől érkező konfigurációs utasításokat továbbítják a szenzoregységek irányába.

A legfelső (1-es) rétegben található az adatgyűjtő és feldolgozó rendszer, melynek feladata adatfúziós eljárások alapján végleges döntéshozatal és a felhasználó értesítése. A rendszer képes a feldolgozott adatok alapján utasításokat adni a különálló szenzoregységeknek, melynek előnye, hogy megfelelő utasítások küldésével javítható a detekció pontossága és csökkenthető az energiafogyasztás.

A bázisállomás és a szenzorok között látható 1:N kommunikációs vonal 868MHz-s rádió kapcsolaton keresztül valósul meg, az 1. és 2. réteg között pedig GPRS kapcsolat van.

A 3. rétegben található rezgés és mágneses szenzorok használatánál a jel alaptulajdonsága miatt lényegében csak zajszűrésre és egyszerű detekcióra van szükség. Ezzel szemben - a jelen eredményeink szerint - az akusztikus eszközök felvételeit alaposabban szükséges elemezni a jelentős környezeti zajok miatt, valamint a kommunikációs sáv szélesség és energiafogyasztás miatt a jelfeldolgozás jelentős részének az eszközön kell megvalósulnia.

2.2 Feladatspecifikáció

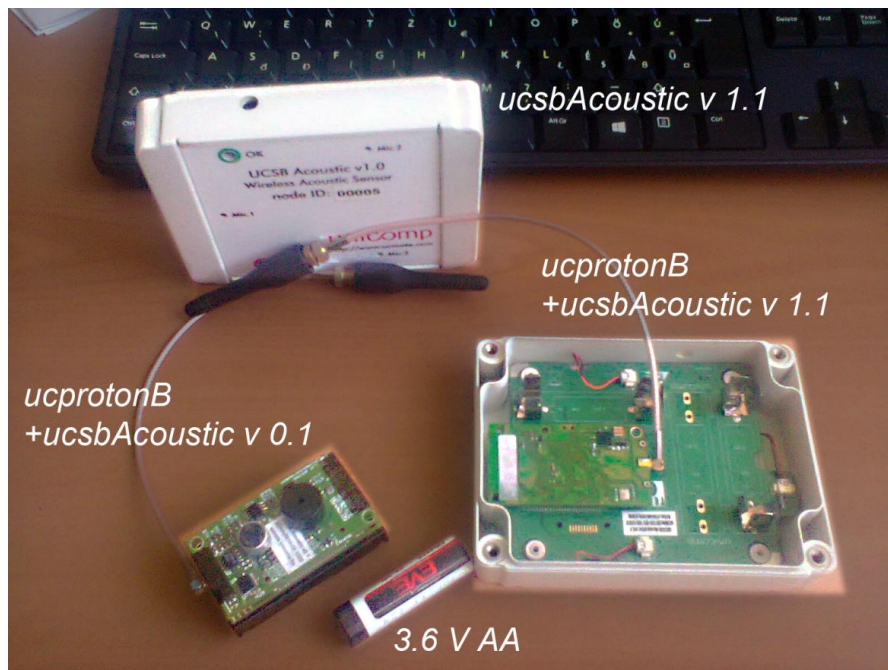
Az általam fejlesztett eljárások, a felvázolt topológia legalsó szintjén található akusztikus jelfeldolgozók felhasználásán alapulnak. Az egyes eszközök az Unicomp Kft. [Unicomp] által fejlesztett ucprotonB elnevezésű mikrovezérlős és ucsbAcoustic elnevezésű szenzormodulok egységéből állnak.

A következő táblázatban a mikrovezérlővel ellátott egység technikai adatai olvashatók, és a könnyebb értelmezés érdekében egy, a méretet reprezentáló kép is bemutatunk.

1. táblázat A céleszköz hardverspecifikációja

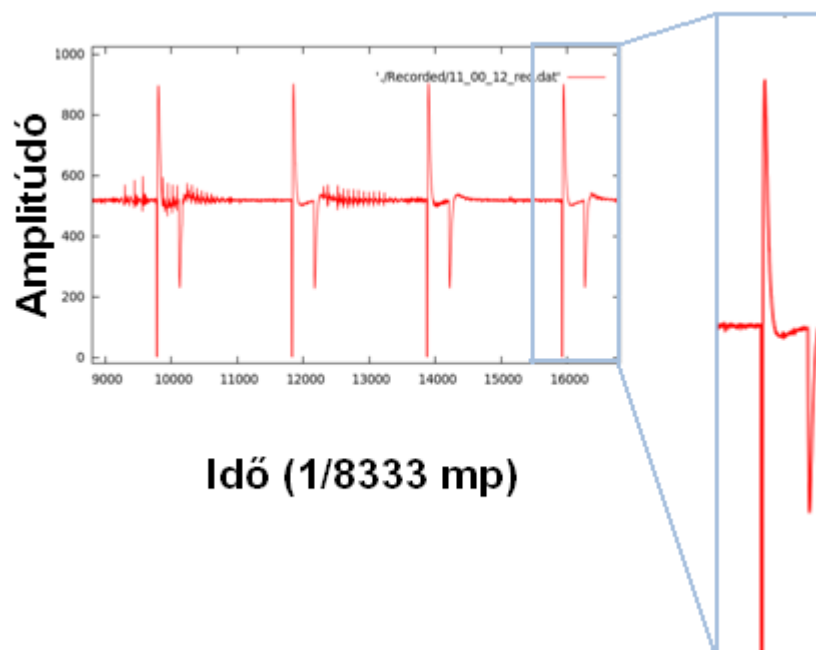
BOARD	ucprotonB
Mikrokontroller:	ATMega128RFA1, 8 bit, 16MIPS
Flash memória:	2 MB, SPI
RAM:	16 kB SRAM, mikrokontrollerbe integrálva
Felhasználói felület:	2+4 db LED (4 programozható), reset gomb
Csatlakozási lehetőségek:	Mikro USB, JTAG csatlakozó, 51 pin
Programozhatóság:	Mikro USB, JTAG, ISP
Szoftver:	TinyOS 2.x kompatibilitás
1. RÁDIÓ	
Chip:	ATMega128RFA1, mikrokontrollerbe integrálva
Antenna:	Omnipoláris chip antenna vagy MMCX csatlakozás
Standard:	IEEE 802.15.4
Frekvencia:	2400 MHz

Bitráta:	250/500/1000/2000 kb/s
RF teljesítmény:	Max. 3.5 dBm
Kommunikációs távolság:	>300 m
2. RÁDIÓ	
Chip:	Atmel AT86RF212
Antenna:	MMCX csatlakozás
Standard:	IEEE 802.15.4
Frekvencia:	784/868/915 MHz
Bitráta:	20/40/100/200/400/250/500/1000/ kb/s
RF teljesítmény:	Max. 10 dBm
Kommunikációs távolság:	>300 m
ELEKTROMECHANIKAI ADATOK	
Fogyasztás (CPU):	4 mA
alvó állapot:	7 uA
RX:	13 mA
TX:	15 mA
Áramforrás:	2 db AA 1,2-1,5 V cella
Tömeg:	14 g (elemek nélkül)
Méretetek:	57 mm x 31 mm x 8 mm



2. ábra Az algoritmusok célhardverei: ucprotnB és ucsbAcoustic eszközök

Fontos megemlíteni egy a specifikációban nem szereplő, viszont a mérések során tapasztalt hardveres korlátot. A hangfelvételek elején minden esetben megjelenik a mikrofon bekapcsolása által keletkező tranziens jelsorozat ezt ábrázolja a 3. ábra nagyított része. A tranziens hossza és megjelenésének ideje minden esetben más és más, viszont soha nem tart tovább $1/8$ másodpercnél, emiatt minden egyes mérésnél ennyivel hosszabb felvétel készítése szükséges, melynek első jeleit nem használhatjuk.



3. ábra A felvételekben megjelenő tranziens hiba

2.3 A hangalapú detekció sajátosságai

Jelen esetben az akusztikus szenzorok használatának fő előnye a másik két különböző típusú szenzorral szemben a nagyobb hatótávolság. Míg a gyorsulásmérők és magnetométerek hatótávolsága öt-tíz méter, az akusztikus eszközök képesek több mint negyven méteres távolságból (körülményektől függően) detektálni munkagépek zaját. Emellett fontos megjegyezni, hogy a felismerést nagymértékben megkönnyíti az erdei környezet, mivel az ilyen területeken alapvetően nem hallhatók munkagépek hangjai, csak bizonyos speciális esetekben (pl. közeli üzemek, építkezések, utak).

A hang, mint fizikai jelenség rengeteg informatikai alkalmazásban előfordul, ezek közül is a legérdekesebbek a hang változásával, jelentésével, vagy forrásával foglalkozó esetek. A hang változásával előjelezhetők mechanikai eszközök meghibásodásai, ultrahangos vizsgálattal feltárhatóak a fémtesteken szemmel nem látható repedések, visszaverődésének tulajdonságaiból radarképhez hasonló információt kaphatunk. A hang jelentésére vonatkozó alkalmazások állnak talán a legközelebb az emberek mindennapi életéhez. Léteznek különböző beszédfelismerő alkalmazások, használatosak a biometrikus biztonsági rendszereknél, automatikus kottázó programoknál, illetve népszerű zenekereső alkalmazásoknál, ahol egy zeneszám rövid részlete alapján az adatbázisból visszakapjuk a hallott hanganyagra vonatkozó adatokat. Minket jelen esetben a hang forrásával kapcsolatos eljárások foglalkoztatnak. Léteznek módszerek, amelyek képesek adott hang forrásának helyét, vagy személyét meghatározni, egy beszélgetésben szereplő partnerek hangjait megfelelően szegmentálni illetve kategorizálni, és sok a gépi zajok detektálásával foglalkozó megoldás is létezik.

Néhány példa hangfeldolgozásra:

- A [Redel-Macias 2011] cikkben mesterséges neuronhálózat használatával törekszenek hangforrás felismerésre gépjárműveknél.
- A [Taichi Nakashima 2014] írásban többszereplős dialógus hangforrás lokalizálására és szegmentálására törekszenek.
- A [Chih-Yi Chiu 2010] cikkben zene felismerésről olvashatunk hash illesztési és tartalom szűrési eljárások használatával.

A mi esetünkben a céleszköz az akusztikus szenzorkártya használatával 10 bit felbontású jeleket képes rögzíteni különböző mintavételi sűrűségekkel, miközben két digitális potméter segítségével állítható az eszközök érzékenysége. A két mintavétel között eltelt idő beállítási lehetőségei $1,75\ \mu\text{s}$ -tól $120\ \mu\text{s}$ -ig terjednek. Ezek alapján és a hardveres specifikációból jól látható, hogy a hangfeldolgozás ezen eszközökkel nehéz feladat, többek közt már a hang tárolására is korlátozott lehetőségekkel rendelkezünk. A szigorú memória és számítási korlátok egy relatív alacsony komplexitású hangfeldolgozási módszer fejlesztését teszik szükségessé.

A hangalapú detekció során eszközeink a belső égésű motorok által keltett zaj felismerésére törekszenek, ilyen lehet egy-egy szállítást, kitermelést végző munkagép vagy motoros fűrész. A gépi hang egy fontos tulajdonsága, hogy a jel monoton. Monoton, mert a hangban nem történik jelentős dinamikus változás, mint például az emberi beszédben. Például egy személyautó motorja állandó fordulatszámon haladva minden alkalommal ugyanazt a hangot fogja kelteni, vagy egy mosógép centrifugája folyamatosan gyorsul, amíg el nem éri a beállított fordulatszámot, eközben a hang alaphékvenciája fokozatosan nő, majd eléri a maximumot. Ezen tulajdonságokat felhasználva a bemutatott módszerekkel jól detektálható a motoros fűrészek által keltett hang, vagy egy gépjármű motorhang.

2.4 Felhasználási lehetőségek

Két példaszerű felhasználási eseten keresztül talán jól érzékeltethető az akusztikus szenzorok szerepe, illetve a használatukból adódó nehézségek.

Fűrészhang detektálás: Az erdei fakitermelést az erdészet megbízásából külső vállalkozó végzi. A vállalkozó előre meghatározott helyen, szigorú munkaidő betartása mellett dolgozhat. A vállalkozó által nem hasznosított faanyagot, ágszedők saját használatra összegyűjthetik. Az ágszedők parcellákat kapnak a területből, mely parcellán kötelesek a fahulladék összegyűjtésére. A vállalkozó jelenlétében azonban a területen munkaidő után semmilyen munkavégzés nem lehet. A vállalkozónak tarvágásra kiadott, ágszedésre kijelölt területekről pontos nyilvántartás áll rendelkezésre.

A fűrészhang detektálása egyszerűbb feladat a gépjárművek által keltett zajok felismerésénél az akusztikus szenzorok használatával, mert a felvételek időzítése nem annyira kritikus, viszont esetleg nehezebben lokalizálható. Amíg egy autó pár másodperc alatt elhalad a mérési pont mellett, addig az erdei munka több órát vesz igénybe, viszont a gépjárművek adott helyen levő utakon haladnak, emiatt könnyen lokalizálhatók. A szenzorok elhelyezésétől függően több módszer is lehetséges a hangok figyelésére, és ezeknél különböző elemzési paraméterek, módok használhatók. Egy nagyobb terület védelméhez akár néhány (fél tucat) szenzor elegendő a bázisállomásokkal, így néhány hektár felügyelete megoldható.

Rakodóterület védelme: A tarvágásra kijelölt terület egy jól megközelíthető, út menti részén a vállalkozó ideiglenes rakodóterületet alakít ki. A rakodótéren a kidöntött fák méretre vágása, osztályozása, illetve szállítás előtti tárolása történik. A rakodótérrel a faanyag közbeiktatott raktár nélkül a megrendelőkhöz kerül. Néhány esetben, megrendelés híján a faanyag a vágási munkálatok befejezése után is a rakodótéren maradhat. Lakott területtől távol eső hely esetén a vállalkozó a munkagépeket is otthagynak a munkaidő lejártakor. A rakodótéren elhelyezett, szállításra előkészített faanyag, illetve az ott hagyott gépek könnyen lopás célpontjaivá válhatnak.

A fűrészdetektáláshoz hasonlóan itt is érdemes az akusztikus szenzort használni. A rakodóterületről történő szállításhoz gépjárműveket használnak, így a gépjárművek hangjának detekciójával könnyen jelezhető az illetéktelen behatolás. A rakodóterületek alapterületének méretéből és helyének ideiglenességéből adódóan minimális mennyiségű eszköz megfelelően elláthatja a terület védelmét.

A fejlesztés célja olyan akusztikus mérőegységek szoftverének előállítása, melyek illeszkednek az ismertett kommunikációs topológiába, valamint elősegítik az illegális erdei munkafolyamatok detektálását. Esetünkben fő cél:

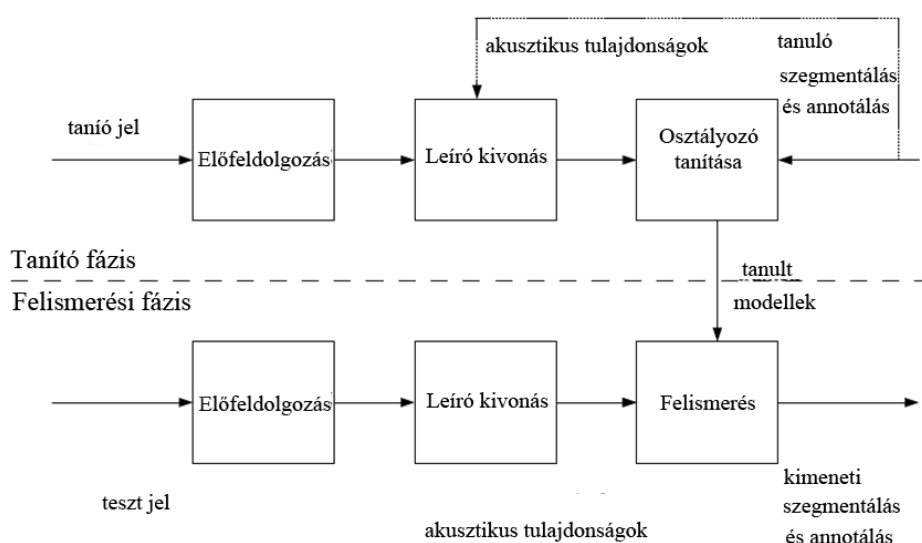
- közel valós idejű jelzések,
- magas találati arány alacsony fals pozitív jelzés mellett,
- minimális energiafogyasztás (azaz alacsony számítási igény),
- nagy hatótávolság

megvalósítása. Azonban a hardveres korlátok miatt kompromisszumot kell kötnünk néhány tekintetben. A megbízható, közel valós idejű riasztás, alacsony számítási igény elsődleges cél, az alacsony energiafogyasztás és a nagy hatótávolság kisebb prioritást élvez.

3 Akusztikus jelek feldolgozása

Az automatikus hangfelismerés egy olyan folyamat, amely során a digitálisan tárolt hangminták szűk osztályokba sorolására törekszünk. Ennek eléréséhez a hangmintából alacsony és magas szintű tulajdonságokat kinyerve tulajdonság vektorokat (deszkriptorokat) állítunk elő, és így alak-felismerési („pattern recognition”) feladattá transzformáljuk az eredeti problémát. A hangfelismerési feladatokat a hangforrás alapján általánosan három osztályba sorolhatjuk: beszéd, zene, és egyéb hangok, zajok. Felhasználási területtől függően más módszerek lehetnek eredményesek, a következő részben e módszerek generikus tulajdonságairól írok.

Az automatikus hangfelismerés általánosan három lépésből áll, melyek rendre: előfeldolgozás, leíró kinyerés, osztályozás. Ennek folyamatát mutatja be a 4. ábra.



4. ábra A hangfeldolgozási folyamat áttekintése

3.1 Előfeldolgozás

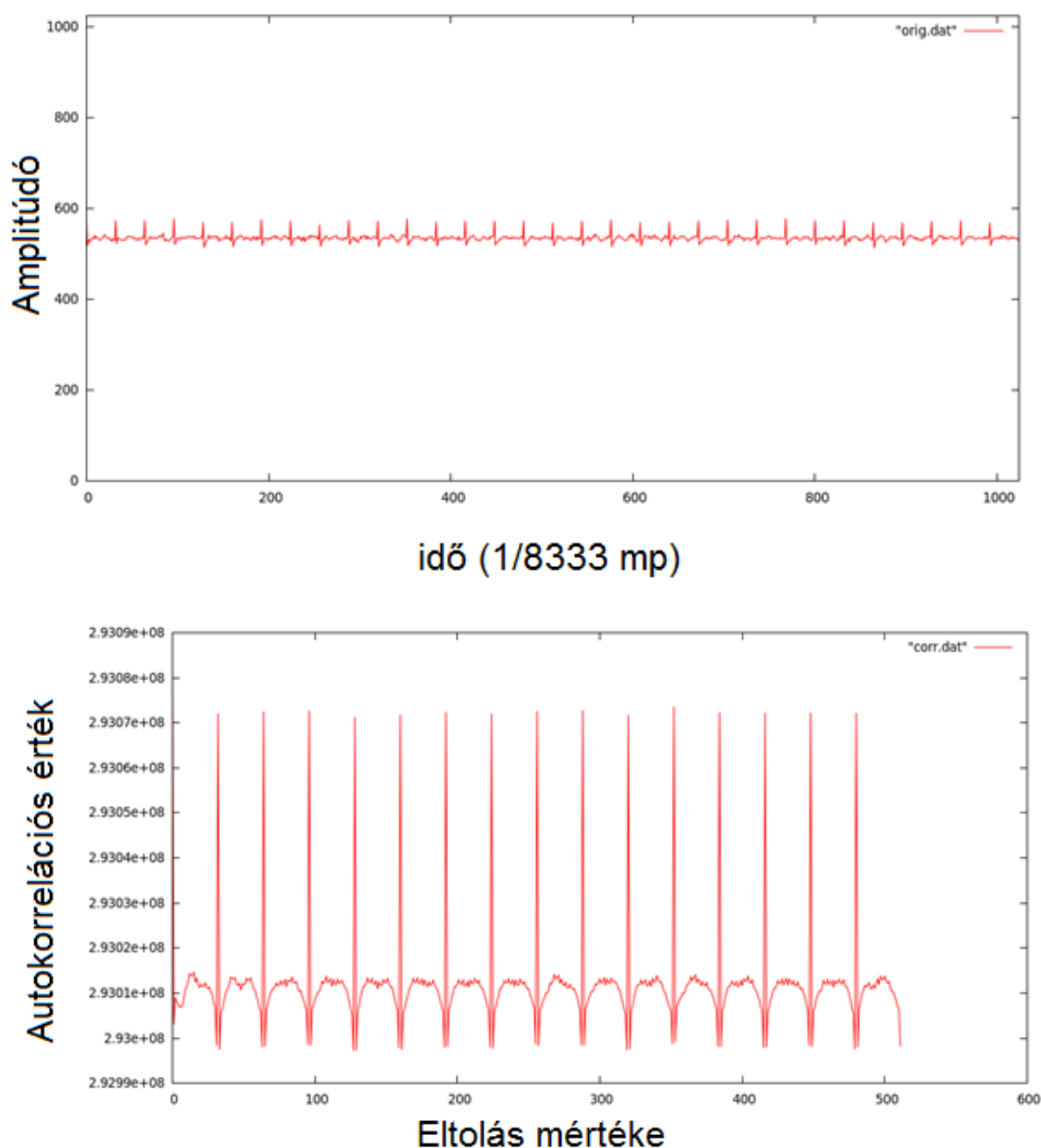
Az előfeldolgozási fázisban arra törekszünk, hogy a mintát (releváns) információvesztés nélkül olyan alakra hozzuk, amely nagyban megkönnyíti a felhasználni kívánt tulajdonságok leíróinak kinyerését. pl.: normalizálás, zajszűrés, újra-mintavételezés.

A leírók kinyerése tulajdonságoktól függően más- és más módszerekkel érhető el. Alapvetően három különböző típusú tulajdonságot különböztetünk meg: intenzitással, hangerővel kapcsolatos tulajdonságok, hangmagassággal, a hang alaphérfkvenciájával kapcsolatos tulajdonságok, és a hangszín. A hangszín egy komplex, pszicho-akusztikai tulajdonság, amely a hang minden olyan tulajdonságát magában foglalja, amely nem a hangerővel vagy alaphérfkvenciával kapcsolatos.

Egy példa – a fejlesztés során előforduló - előszűrésre:

A fejlesztés során rendelkezésre álló hardver egyik sajátossága, hogy a hangfelvételekben megjelenik egy periodikus zaj a rögzítéshez használt pufferek méretétől függően. A méréseink során 64 byte méretű puffereket használtunk az adattovábbításhoz, az egyszerűbb feldolgozás érdekében ezért a hangfelvételeket 32 jel hosszúságú pufferekben rögzítettük (a 10 bites jeleket 2 Byte-on tároljuk). A felvételekben ennek köszönhetően 32

jelenként jelent meg egy, a mikrofon bekapcsolási tranzienséhez hasonló, ám jóval kisebb intenzitású és rövidebb ideig tartó periodikus zaj. A zaj nagyban befolyásolja az osztályozás kimenetét, ugyanis a periodikusan megjelenő jel az autokorrelációs szekvenciát nagymértékbe torzítja, ezt az 5. ábra reprezentálja.



5. ábra Egy távoli zajos fűrészhang felvétel és autokorrelációja

A zaj szűrésére több lehetőség is van, viszont a lehető leggyorsabban elvégezhető, legkevesebb információvesztéssel járó eljárást kell választani, ami jelen esetben a periodikusan előforduló hibás jelek helyén történő interpoláció. Előnye, hogy időtartományban megvalósítható, így mindössze néhány számítási művelet szükséges a hiba javításához.

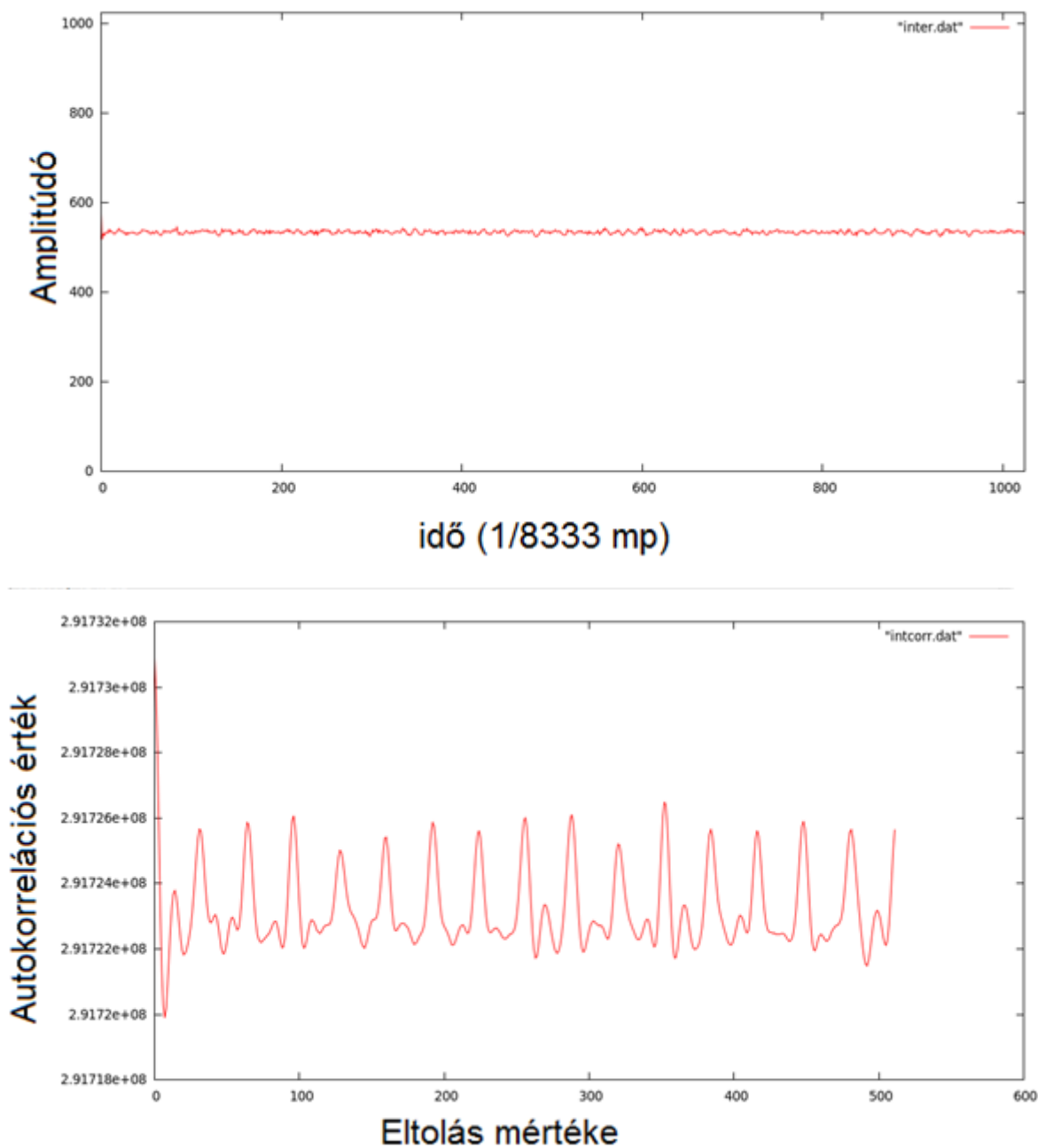
A zaj periodikusságát kihasználva több megoldás is lehetséges a jelek interpolálására, az általam választott módszer fokozatosan közelíti a zaj két szélén található jelet.

$$s(n) = \frac{s(n-1) + s(n+3)}{2}, \quad s(n+1) = \frac{s(n) + s(n+3)}{2},$$

$$s(n+2) = \frac{s(n+1) + s(n+3)}{2},$$

ahol $s(n)$ jelöli a felvétel n -edik jelét

(3.1)



6. ábra A jelsorozat interpolációs szűrés után és autokorrelációja

A 6. ábra az interpolációs szűrés eredményét mutatja be.

3.2 Tulajdonságok generálása

A tulajdonság leírók felhasználásának tekintetében különböző logikai szinteket különböztethetünk meg. Nyers adatok felhasználásánál alacsony szintről beszélünk, ilyen lehet például képek esetén egy-egy pixel érték vagy a hangfelvétel jelsorozatának használata. Egy jellemző példa az alacsonyszintű feldolgozásra a mintaillesztés. Közepes szintről beszélünk, amennyiben a nyers adat valamilyen tulajdonságát nyerjük ki, és ezt használjuk fel a felismerésben. A kivont leírókat felhasználva, megfelelő osztályozási eljárásokkal megállapítható a hang forrása, felismerhető a beszéd, vagy akár a zene dallam is.

Az előző részben tárgyalt három különböző típusú tulajdonsághoz számtalan leíró létezik, ezek közül csak néhányat sorolunk fel. Az egyenletekben s (a mintavétel után kapott) számsorozatot, míg s_i a számsorozat i -edik elemét jelöli.

A) **intenzitás:**

- STE (Short Time Energy) A jelsorozat rövid idejű energiája.

$$\sum_{n=0}^{|S|} s_n^2, n \in \mathbb{N} \quad (3.1)$$

- LGSTE (Logarithmic Short Time Energy)

$$\sum_{n=0}^{|S|} \log(s_n^2), n \in \mathbb{N} \quad (3.2)$$

- LOUD (Loudness) Adott érzékelési modell alapján számított hangerő

$$L = \log\left(\frac{10I}{I_0}\right), I = \left(\frac{P}{4\pi r^2}\right)$$

$$\text{ahol } I_0 = 10^{-12} \text{ Wm}^{-2}, P \text{ pedig az akusztikus jel energiája} \quad (3.3)$$

B) **Alapfrekvencia:**

- ZCR (Zero Crossing Rate) Nulla-átmeneti sűrűség

$$\frac{1}{1 - |S|} \sum_{n=1}^{|S|} I(\text{sign}(s_{n-1}) \neq \text{sign}(s_n)), n \in \mathbb{N} \quad (3.4)$$

- GIN (Gradient Index) Gradiens index

$$GIN = \frac{\sum_{n=1}^{|S|} \Delta \Psi(n) |s_n - s_{n-1}|}{\sqrt{\sum s_n^2}}, \Delta \Psi = \frac{|\Psi(n) - \Psi(n-1)|}{2}, \Psi(n) = \frac{s_n - s_{n-1}}{|s_n - s_{n-1}|}, n \in \mathbb{N} \quad (3.5)$$

- AC (Autocorrelation) Autokorrelációs szekvencia

$$R(m) = \sum_{n=-\infty}^{\infty} s_n \cdot s_{n+m},$$

$$\text{ahol } n, m \in \mathbb{N}, R(m) \text{ az autokorrelációs szekvencia } m\text{-edik eleme} \quad (3.6)$$

- SCENT (Spectral Centroid) A magnitúdó spektrum tömegközéppontja

$$C = \frac{\sum_{n=0}^{N-1} f(n)x(n)}{\sum_{n=0}^{N-1} x(n)},$$

ahol $x(n)$ a n -edik elem magnitúdója, és $f(n)$ n -edik érték frekvenciája (3.7)

- SHARP (Sharpness) Adott modell használatával számított érték, a hang magas frekvenciás komponenseinek arányára utal.

Ilyen például a Zwicker-féle érték:

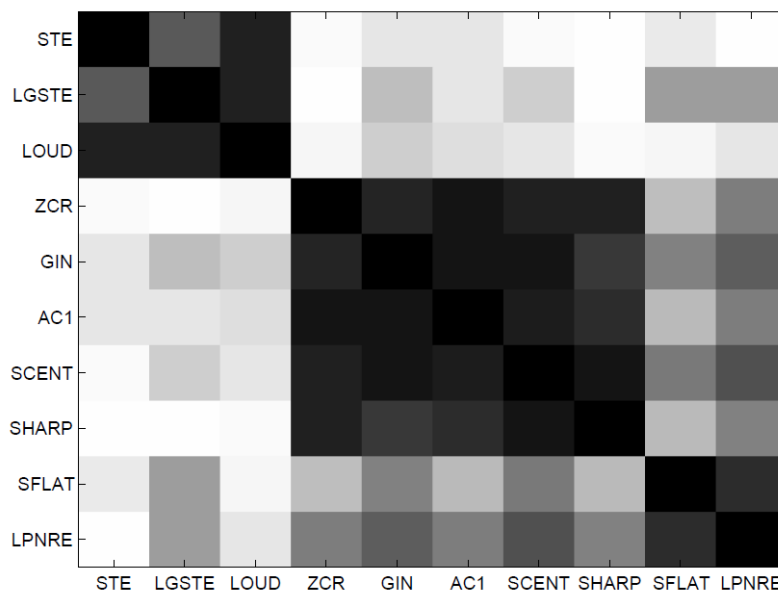
$$S = 0.11 \frac{\int_0^{24 \text{ Bark}} N' g(z) dz}{\int_0^{24 \text{ Bark}} N' dz},$$

ahol N' minden z kritikus sáv ereje a Bark skálán és $g(z)$ súlyozási faktor (3.8)

C) frekvencia és hangszín:

- SFLAT (Spectral Flatness) A magnitúdó spektrum geometriai és számtani átlagának aránya
- LPNRE (LP Normalized Residual Energy) LP normalizált reziduális energia

A fenti tulajdonságok sok esetben valamilyen mértékben a jeleknek hasonló tulajdonságát ragadják meg, ezen jelek korrelációjának mátrixos reprezentációja a 7. ábrán látható.



7. ábra A különböző hang tulajdonságok korrelációja: a sötétebb mező erősebb kapcsolódást jelez

A 7. ábra jól mutatja be, hogy egyes hangtulajdonságokhoz tartozó leírók egymással jól korrelálnak, azaz feltehető, hogy adott hangtulajdonság jellemzéséhez bármelyik az annak a tulajdonságnak megfelelő- leíró használható. Természetesen lehet különbség az egyes leírók zaj érzékenységében, invariáns tulajdonságaiban.

3.3 Osztályozó módszerek

A megmért, számszerűsített tulajdonságokat a későbbiekben elemezni fogjuk, osztályozásuk vezet a jelek felismeréshez. Tekintsünk át néhány –a hangfeldolgozásban is – gyakran használt osztályozási módszert.

Bayes: Bayes-féle osztályozók minden esetben valamilyen feltételes valószínűségek számításával rendelnek osztálycímket egy-egy mintához. Egyik legismertebb Bayes osztályozó eljárás a Naív Bayes, mely egymástól függetlennek tekinti az adott mintához tartozó leírókat.

kNN [D. Aha 1991] (k Nearest Neighbour): Az osztályozáshoz egy minta adatbázis felépítése szükséges, amelyben a leíróvektorokhoz osztálycímkeket rendelünk, majd a klasszifikáció során kiszámítjuk a vizsgált minta leíróvektorának távolságát a tanító adatbázisban levő k legközelebbi vektortól, majd azt a címkét rendeljük a vizsgált mintához, amely vektoraiból a legtöbb közel állót találunk.

ANN (Artificial Neural Networks): Mesterséges neuronhálózatok többrétegű hálózathoz felépülő rendszerek. Két kiemelt fontosságú réteg létezik az input és az output réteg. Ezen két réteg között bármennyi feldolgozó réteg lehet. Minden egyes rétegben tetszőleges számú mesterséges neuron helyezkedhet el, melyek szinapszisaikon keresztül kapcsolódnak egymáshoz. Egy-egy neuron a beérkező adaton műveleteket hajt végre, (tipikusan súlyozást, több bejövő jel összegzését és szigmoid leképezést), majd a hálózat jellegétől függően továbbítja a következő réteg felé, vagy vissza is adhatja a küldő neuronnak. Mesterséges neuronhálózatok használatakor az osztályozandó vektort adjuk át az input rétegnek, a közbülső rétegek neuronjai különböző adott hálózathoz megfelelő műveleteket hajtanak végre, majd a lehetséges osztályozási címkék az output rétegnél megjelennek. A tanítás folyamata során helyes megoldáshoz vezető neuronok közötti kapcsolatok (súlytényezők) erősödnek a többi pedig gyengül. A [Michael Büchler 2015] cikkben hallókészülék fejlesztés céljából tesztelnek különböző módszereket az akusztikus környezet osztályozására, többek közt Bayes féle és mesterséges neuronhálózat alapú osztályozókat is.

SVM (Support Vector Machine): Az SVM-k alapvetően bináris osztályozók, bár fontos megjegyezni, hogy az osztályozás SVM-k kombinálásával kiterjeszthető több osztályra is. SVM-k használata során egy vágási hipersíkot keresünk a paraméterterben, amely sík két oldalán a lehető legjobban elkülönülnek a két osztályba tartozó vektorok. Léteznek lineáris és nemlineáris változatai. A [H.-J. Z. Lie Lu 2003] cikkben öt különböző hangosztály megkülönböztetésére törekcszenek SVM-ek használatával.

Döntési fák: A döntési fák hierarchikus osztályozó sémák. Lényegében a fa minden nem levélszintű csúcsa egy feltétel a vizsgált vektorral szemben, és a legtöbb esetben a levélszinten pedig osztálycímkeket rendel a vektorhoz. A döntési fák általában diszjunkt klaszterekre osztják a paraméter teret, például ilyen a CAR, REP vagy a BF fa. Viszont léteznek olyan fák, amelyek ágain párhuzamos feltételvizsgálatok hajtódnak végre, ezzel a klaszterek átlapolását idézve elő, ilyen például az AD fa. A [Sotris A Pavlopoulos 2004] cikkben döntési fákat használnak szívbetegség differenciál diagnózisához szívhangok alapján.

Heurisztikus szabályok: Nem minden osztályozási módszer háttérében található egzakt matematikai modell, lehetséges manuálisan, tapasztalat és intuíció alapján szabályrendszert felépíteni az osztályozáshoz. Amennyiben egyszerű a feladat, nem szükséges bonyolult osztályozókat használni, például sokszögek megkülönböztetésénél az oldalak vagy csúcsok száma optimális leíró, amit felhasználva könnyen írhatunk fel szabályokat. Természetesen

bonyolultabb feladatoknál, kevésbé jó leírókat használva a mesterséges osztályozók jobb eredményekkel szolgálhatnak. A [Lie Lu 2001] cikkben egy kétlépéses hangfelismerési eljárásról olvashatunk, az első lépésben kNN-t használnak a beszédhang és egyéb hangok differenciálására, majd a nem beszédhangokat osztályozzák tovább heurisztikus szabályok segítségével zene, környezeti zaj és csend kategóriákba.

Az eddig ismertetettek szerint leírókat nyerhetünk ki a spektrális illetve kepsztrális térből, korrelációs vagy autokorrelációs szekvenciából, valamint időtartományból. A mi feladatunk a hardveres korlátokat figyelembe véve a lehetőségek közül kettőt vizsgáltunk részletesen, az autokorrelációt és az időtartományból kivont leírókat.

3.4 Spektrális tér, Fourier transzformáció (FT)

A spektrális térből kivonandó leírókhoz a jelsorozat Fourier transzformációval, a kepsztrális térből kivonandókat pedig a transzformált értékek koszinusz transzformációjával érhetők el.

A [Marco F. Duarte 2004] cikkben vezeték nélküli szenzorhálózatokat alkalmaznak gépjármű osztályozásra, többek közt akusztikus, szeizmikus és infravörös szenzorok használatával. Az akusztikus eszközök FT együtthatókat használtak lokális osztályozáshoz 60%-70%-os találati aránnyal.

A [L. Couvreur 2004] robogó és dudaszó elválasztására törekedtek kepsztrális együtthatók és mesterséges neuronhálózatok felhasználásával, a cikkben 85%-95%-os találati arányról számolnak be.

A [Baljeet Malhotra 2008] cikkben gépjármű osztályozásra használt szenzorhálózatról olvashatunk, tömegközlekedési és egyéb járművek között kellett különbséget tennie az FT együtthatókon és további spektrális leírókon alapuló osztályozók használatával.

A [Khoirul Anwar 2011] cikkben alacsony komplexitású Fourier transzformáció közelítő módszerről olvashatunk. Az eljárás lényege, hogy a szinusz hullámot un. dupla négyszögjellel közelítik, így gyorsítva az algoritmus futási idejét.

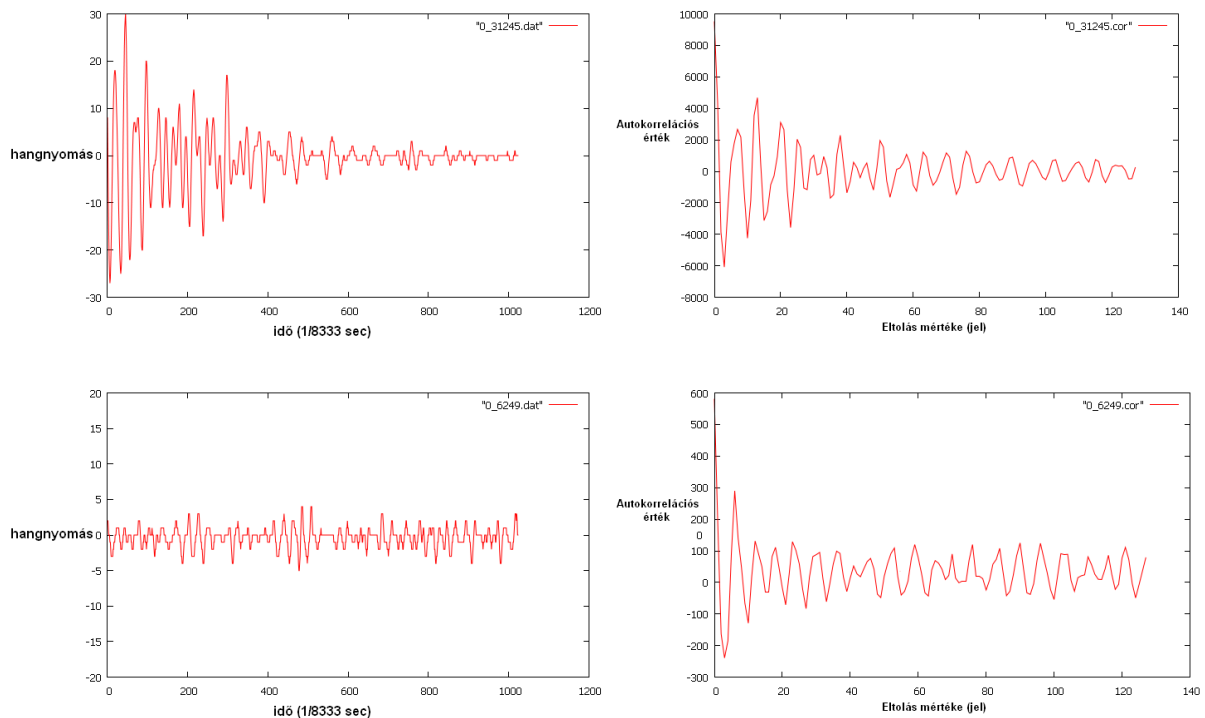
A spektrális, valamint kepsztrális térből kinyerhető deszkriptorokat pusztán a leírók kivonáshoz szükséges előfeldolgozás számítási és memória igénye miatt nem vizsgáltuk részletesen.

Szükséges feldolgozási lépés, a Fourier transzformált számítása, ezt a lehető legjobb leírók érdekében dupla pontosságú számokkal kell előállítani. A transzformált számítása szimpla diszkrét FT-t alkalmazva $O(n^2)$ időt vesz igénybe, FFT („Fast Fourier Transformation”) alkalmazásával ez az idő $O(n \log n)$ -re csökkenthető. A mikrovezérlő adatait áttekintve könnyen belátható, hogy a 8 bites processzor dupla pontosságú műveleteket elég lassan végez. Használhatnánk esetleg fix pontosságú műveleteket, vagy akár közelítő megoldásokat a transzformált előállításához, viszont a számítási igény csökkentéséhez a szinusz hullám tárolására lenne szükség. Arra, hogy elég hosszú felvételeket készítsünk, és nagy pontossággal és felbontással tároljunk szinusz hullámot a 16 KByte memória nem elegendő. Lehetséges lenne a nagyobb, 2 MByte méretű flash memóriában tárolni a hangfelvételeket, ez a tárhely azonban csak nagyon lassan elérhető, valamint az energiafogyasztás is növekedne a használatával. Mindent egybevéve elképzelhető olyan konstrukció, amelyben a spektrális tartományból vett leírók felhasználásával jó eredmények érhetők el, viszont jelen eszközön a feladat specifikációban felsorolt szempontok miatt nem alkalmazhatók.

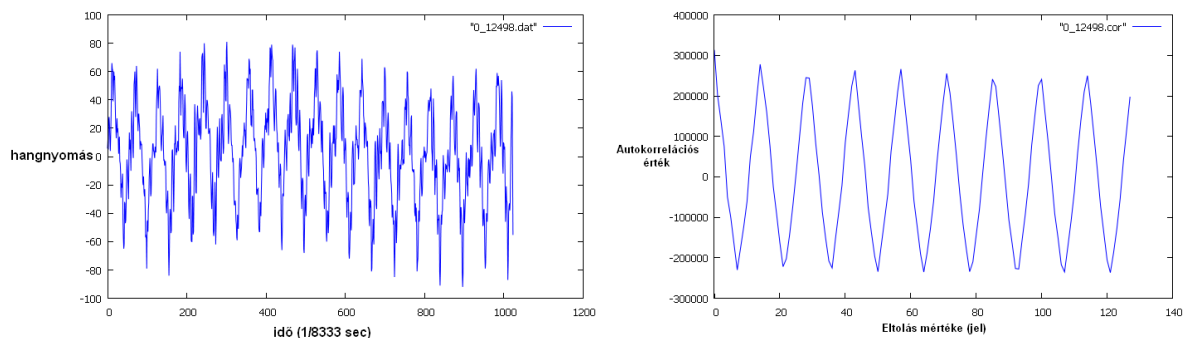
3.5 Autokorrelációs szekvencia

Az eddig ismertetteknek megfelelően tehát olyan tulajdonság leírókat kell választani, melyek számítása nem túl bonyolult, viszont kellőképpen leírja két hangosztály, jelen esetben a belsőégésű motor hangjának és az erdei környezeti hangok különbségét. A gyors működés érdekében minimalizálnunk kell a tört számok használatát, emiatt pedig az osztályozási pontosság érdekében az osztások, gyökvonások számát is.

A 2. fejezetben leírtakat figyelembe véve feltételezhetjük, hogy az erdei munkát végző gépek hangjai jól megkülönböztethetők egyéb környezeti hangoktól - és bizonyos esetekben egyéb gépi hangoktól is - az alapprojekcia és változásának használatával. A 8. ábrán beszédhang és autokorrelációja (fent), ill. szélhang és autokorrelációja (lent) látható. A 9. ábrán pedig motoros fűrészhangot és autokorrelációját ábrázoltuk. A leírók kinyerése után gyors osztályozás szükséges, mely szintén kis számítási igényű. A fent említettek közül a leggyorsabb a döntési fa illetve a heurisztikus szabályok. Ezeket egyszerű vezérlési szerkezetként lehet implementálni melyek gyors osztályozást eredményeznek.



8. ábra Beszéd- (fent) és szélhang (lent) felvételei, valamint autokorrelációs szekvenciáik



9. ábra Motoros fűrészhangja, valamint a felvétel autokorrelációs szekvenciája

Egy hang alaphfrekvenciáját megállapítani az említett korlátok mellett nehéz feladatnak tűnhet. A legegyszerűbb módszer a rögzített hangminta jeleit idő tartományból spektrális tartományba konvertálni Fourier-transzformációt használva, majd kiválasztani legnagyobb értékű lokális maximum helyét. A rendelkezésre álló céleszköz használatával ez a módszer nagyon lassú lenne. Egy másik megoldás, hogy a rögzített hang autokorrelációs szekvenciájában keressük a lokális maximumokat és azok távolsága alapján számítjuk az alaphfrekvenciát. Az autokorreláció számítása – a meglévő korlátokat szem előtt tartva – lényegesen gyorsabb, köszönhetően annak, hogy csak egész számokkal való szorzás és összeadás eredménye. A módszer másik nagy előnye, hogy az autokorrelációs szekvencia első eleme megegyezik az STE-vel, azaz számításával a hang intenzitásáról is kapunk információt (ez lényegileg igaz az FT-re is).

Osztályozáshoz használt leírók az autokorrelációs szekvenciából:

- lokális maximumok száma (mennyire jellemző a hangmintára az alaphfrekvencia)
- lokális maximumok átlagos távolsága (becslés az alaphfrekvenciára)

$$avgDist = \frac{1}{|max|} \sum_{i=1}^{|max|} max_i \quad (3.9)$$

- ezen távolságok szórásnégyzete (mennyire állandó az alaphfrekvencia)

$$variance = \frac{1}{|max|} \sum_{i=1}^{|max|} (dist_i - avgDist)^2 \quad (3.10)$$

- adott hang rövididejű intenzitása (milyen hangos a mért jel)

$$STE = \sum_{n=0}^{|S|} s_n^2, n \in N \quad (3.1)$$

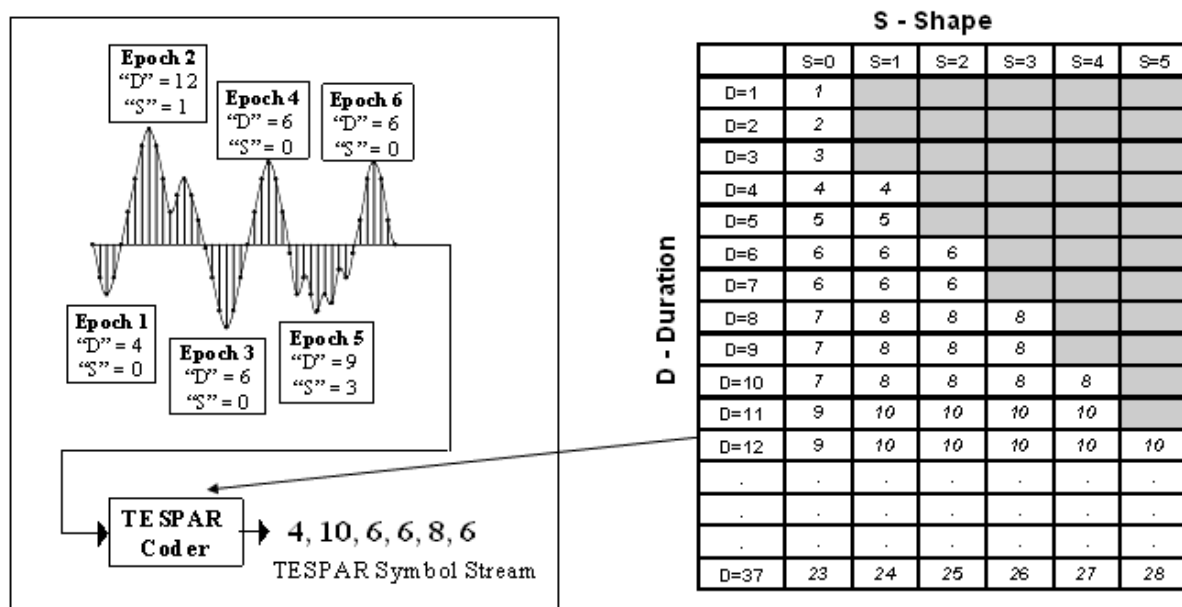
Egyszerűen belátható, hogy a hang intenzitása önmagában nem erős leíró, mivel nagymértékben függ az erősítési beállításoktól és a hangforrás távolságától. Annak érdekében, hogy mégis jól használható leírókat kapjunk, az erősítési beállításnak megfelelő értékkel normálható. A vizsgálatok során más leírókat is teszteltem, többek közt az autokorrelációs függvény csillapodását, a lokális maximumok értékének átlagát és szórását. Ezek azonban az osztályozást nem javították így ezek számítását nem részletezem a dolgozatban.

Az alaphfrekvencia megbecsülhető a hangminta autokorrelációs szekvenciájában történő lokális maximumok keresésével. (*Mintavételi sebesség / maximumok távolsága = alaphfrekvencia*)

3.6 TESPAP

A TESPAP [R.A. King 1999] elnevezés a Time Encoded Signal Processing And Recognition kifejezés rövidítése. A módszer lényegében időtartományban végzi a leírókivonást, a nulla-átmenetek közt eltelt idő (D érték), és adott időtartamban beérkezett jelek hullámalakjának (S érték) megállapításával. Ezen értékek alapján egy kódkönyv felhasználásával kiszámítható egy szimbólum folyam, amely reprezentálja a felvételt. A szimbólum sorozat alapján előállítható az ún. S mátrix (adott D-S párokhoz tartozó szimbólumok hisztogramja). Valamint számítható az ún. A mátrix, amely bármely két szimbólum együttes előfordulásainak számát tartalmazza „L” szimbólum távolságot alkalmazva (a képfeldolgozásban használatos co-occurrence mátrixhoz hasonlóan). Az említett kódkönyvek vektor kvantálással állíthatók elő.

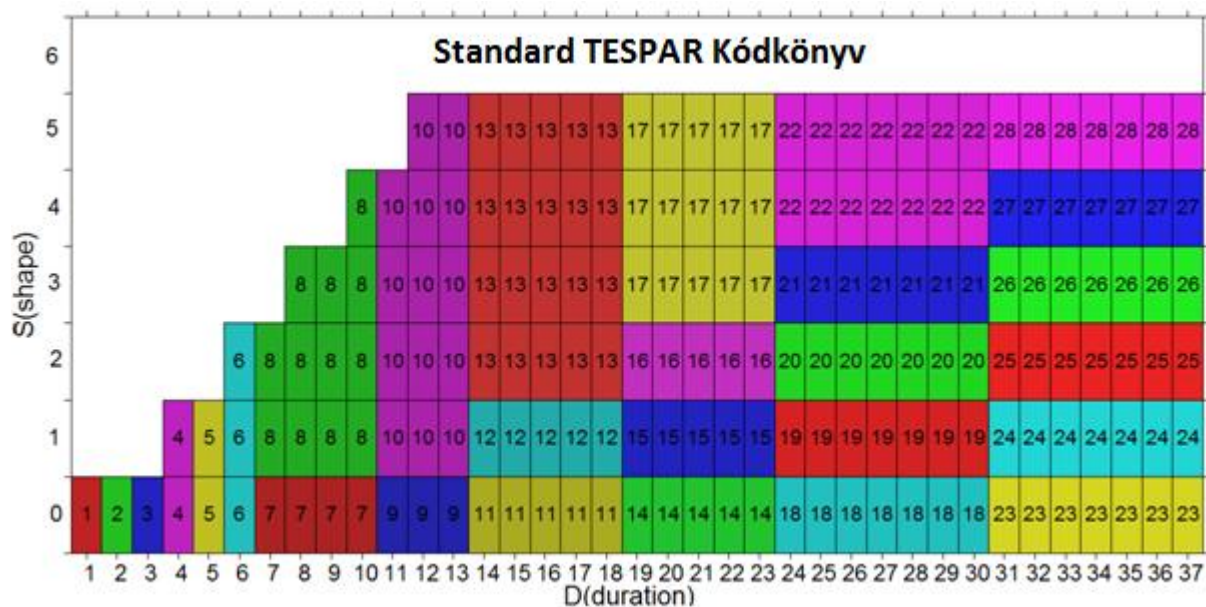
Érdekes megemlíteni, hogy a TESPAP módszer kísértetiesen hasonlít a szöveg és képfeldolgozásban ismert és igen népszerű ún. Bag of Words modellekre. A szövegfeldolgozásban a 90-es években jelentek meg ezek a módszerek [Joachims 1998], míg a képfeldolgozásban a 2000-es évek elején [L. Zhu 2002] [G. Csurka 2004]. A hangfeldolgozás irodalma nem tesz említést ezekről a kapcsolódásokról.



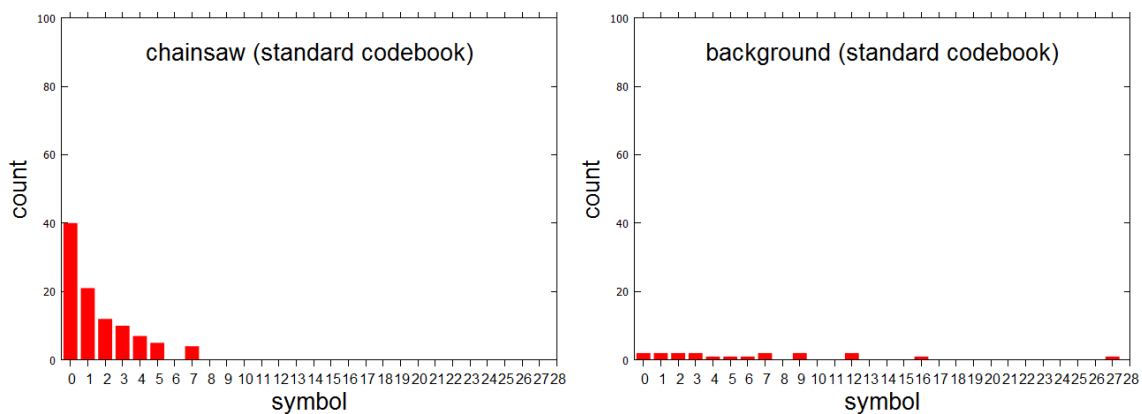
10. ábra TESPAP leíró komponensek [<http://www.tespar.co.uk>]

A 10. ábra jól mutatja be, hogy az S érték pozitív tartományban a lokális minimumok, negatív tartományban pedig a lokális maximumok száma, valamint, hogy a D értékek a nulla-átmeneti sűrűség számításához szükséges számok, ezáltal a hang alapfrekvenciájára utaló jelek. Azt is megfigyelhetjük, hogy az S értékek a hang komplexitására vonatkozó számok, minél több magas frekvenciájú komponensből áll egy hang, az S érték annál nagyobb lesz.

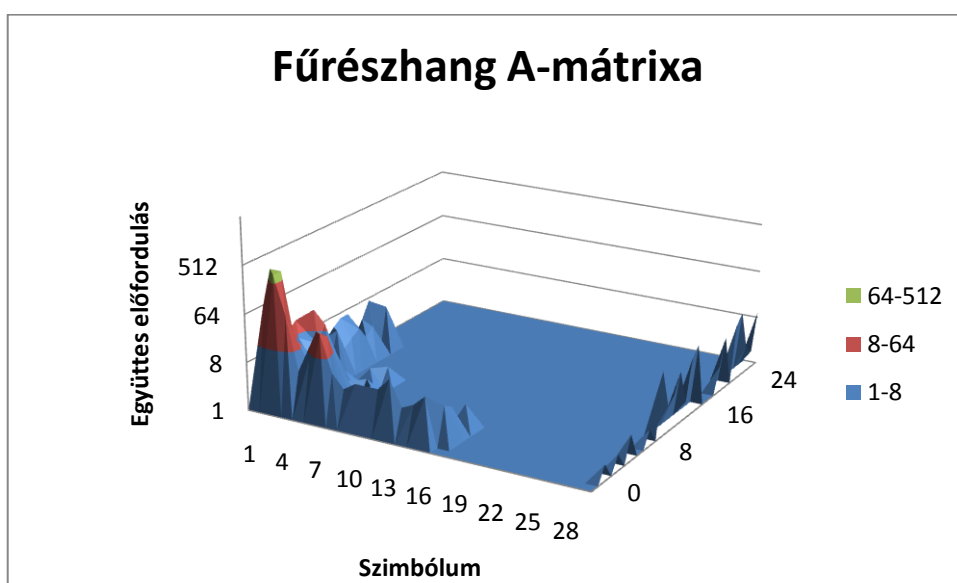
A hangfelvételek kiértékelésénél számított D-S párok és egy szimbólumszótár használatával előállíthatók a TESPAP leírók.



11. ábra A standard TESPAR kódkönyv



12. ábra Motoros fűrész (bal) valamint háttérzaj (jobb) felvételéből képzett S mátrixok



13. ábra Fűrészhang A-mátrixa

A keletkezett leírókkal való osztályozásnak több módja létezik, a leggyakrabban mesterséges neuronhálózatok használatosak, viszont bizonyos esetekben őstípusok képzését és az adott mintával való összehasonlítását alkalmazzák. Őstípusnak nevezzük a leíróvektorok tekintetében azt a prototípus vektort, amely az adott osztályban szereplő valamennyi objektumra jellemző tulajdonságokat tartalmaz. A [Marius Vasile Ghiurcaua 2012] cikkben TESPARE leírókkal való és Mel-frekvencia kepsztrális együtthatók felhasználásával működő osztályozási módokat hasonlítanak össze erdei környezetben. A cikkben írt módszer minden hangfelvételt besorol egy-egy speciális hangosztályba, a saját felhasználási eseteknél ez nem alkalmazható, ugyanis az erdei háttérzajok sokfélék lehetnek, így nem alkotnak jól definiálható hangosztályt.

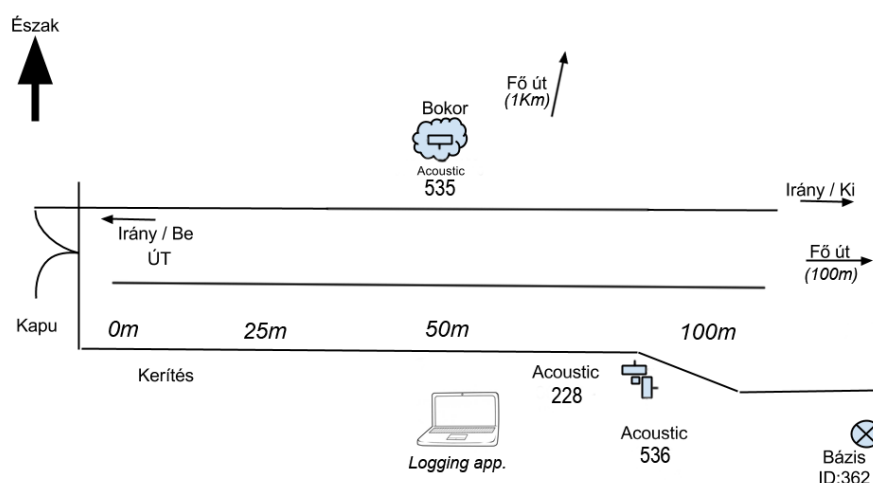
4 Implementáció és tesztelés

Az előző fejezetben tárgyalt korlátok miatt a spektrális, illetve kepsztrális térből kinyerhető leírók felhasználásával történő hangfelismerést nem vizsgáltuk, viszont az autokorrelációs és TESPARE módszerek nyújtotta lehetőségeket nagy mélységben teszteltük.

4.1 Terepi mérések

A tanító és tesztadatbázisok felépítéséhez több mérést is végeztünk. Többek között 2013 októberében, a Pannon Egyetem területén, a Műszaki informatikai Kar épülete mögött, valamint egy külvárosi környezetben családi háznál végeztünk egymástól független méréseket. A mérések során különböző távolságokból, különböző erősítési beállítások alkalmazásával rögzítettük a motoros fűrészek hangját, valamint mértük a gépek által keltett hang erejét. Kiderült, hogy a fűrészhang még 60 méteres távolságból is nagyjából 53 dB hangerővel hallható, míg a környezeti zajok alig haladták meg a 40 dB-t. A mérések eredményeként készült el az (AGDB)-vel jelzett adatbázis.

Mint már említettem, prototípus eszközökről lévén szó, egyszerű eszközhiány miatt, csak három, autokorrelációs eljárást használó szenzort telepítettünk egy erdőszéli területre 2014 áprilisában. A területről a fontosabb tudnivalók, hogy 1 kilométerre található egy nagy forgalmú főút, valamint pár száz méteres körzetében található egy kisebb főút, egy fűrészüzem, valamint rengeteg szántóföld. A szántóföldek közelsége alapvetően nem befolyásolná az akusztikai környezetet, viszont a mezőgazdasági munkagépek, illetve a permetezést végző repülőgépek hangja jól hallható. A 14. ábra a mérés helyrajzát mutatja be. A leírás alapján könnyen érthető a kritikus zóna elnevezés. A teszt célja adatgyűjtés volt az eszközök hatékonyságáról, valamint az adatfúziós lehetőségekről.



14. ábra Kritikus zóna mérés helyrajza

Hatótávolsággal kapcsolatos tesztjeink alapján az autokorrelációs eljárás 50-60 méteres távolságból még megfelelő találati aránnyal bír, ezért az eszközöket egymáshoz képest ilyen távolságban helyeztük el. Arra számítottunk, hogy a bekötőúton haladó autókat néhány esetben jelezni fogják az eszközök, viszont ezen kívül pozitív jelzéseket nem reméltünk. Ennek ellenére az eszközök szinte véletlenszerűen, folyamatosan jeleztek, 24 óra leforgása alatt több ezer jelzés érkezett.

Ezen problémán túl, kommunikációs nehézségeink is akadtak. Az eszközökön található kommunikációs interfész teljesítménye jóval elmarad a dokumentációban találhatóaktól, emiatt a szinkronvezérlés nem működött, az eszközök egyszerűen túl távol voltak egymástól ahhoz, hogy a szinkronizációs üzeneteket megkaphassák.

Az első jelenséget nem okozhatta hardveres hiba, hiszen mindhárom eszköz „fals” pozitív jelzést adott. A problémát a szoftveres oldalon kellett keresnünk. Azt gyanítottuk, hogy eszközeink túl érzékenyek, ennek megfelelően egy új mérést végeztünk, hasonló környezetben, nagy forgalmú főútvonalhoz közel, szintén több eszközzel, fix, alacsony erősítési beállítások használatával. Az újabb mérés során sikerült viharban készítenünk hangfelvételeket, ezen anyagok szél, eső, és mennydörgések hangját is tartalmazzák. Ennek eredményeként keletkezett az (FGDB)-vel jelzett adatbázis.

A laboratórium munkatársai korábban diktafonnal készítettek hangfelvételeket erdei munkafolyamatokról, ezen hangfelvételeket, valamint szabadon felhasználható internetes hangmintákat használva képeztük az (SWDB)-vel jelölt adatbázist.

A mérések során rögzítettük az időjárási körülményeket, az eszközök, fontosabb tereptárgyak elhelyezkedését, ill. a háttérzaj mértékét. A hangnyomás számszerűsítéséhez CEL 430/1 típusú integráló zajszintmérő készüléket használtunk.

A mérések során felmerülő problémákat az eszközökön futó algoritmusok fejlesztésénél figyelembe vettük, a lehetőségekhez mérten megfelelően kezeltük.

4.2 Adatbázisok összeállítása

A szenzorok egyéni döntési folyamatának kidolgozása során a paramétertér komplexitása miatt gépi tanulási módszer alkalmazása indokolt. A gépi tanuláshoz tanító és tesztadatbázis összeállítása szükséges. Különböző mérések alapján, valamint internetes forrásból származó hangfelvételeket felhasználva több tanító és tesztadatbázist hoztunk létre.

- Céleszközzel készített felvételek különböző előerősítési beállításokkal (AGDB)
 - Tanítás: 490 darab 1/8 másodperc hosszúságú fűrészhang, valamint 330 háttérzaj felvétel
 - Teszt: Független mérés alapján 340 fűrészhang, 230 háttérzaj felvétel
- Céleszközzel készített felvételek rögzített erősítési beállításokkal, alul-mintavételezéssel (FGDB)
 - Tanítás és tesztelés: 228 darab 6/8 másodperc hosszúságú fűrészhang, 174 erdei és út menti háttérzaj, valamint 60 vihar hangminta 50-50 százalékban véletlenszerűen kettéosztva tanításhoz és teszteléshez.
- Diktafonnal készített felvételek, valamint internetes forrásból származó hanganyagok (SWDB)
 - Tanítás és tesztelés: 140 darab 6/8 másodperc hosszúságú fűrészhang, 370 háttérzaj, valamint 360 darab vihar hangminta 50-50 százalékban véletlenszerűen kettéosztva tanításhoz és teszteléshez.

A céleszközzel készített felvételekből összeállított adatbázisok mintái különböző távolságokról (5-75 méter) készített hangfelvételek motoros fűrészhangjáról, valamint környezeti zajokról. A 15. ábra egy 2013. októberi mérés során készült képeket mutat be, baloldalon a méréshez használt eszközök: zajszint mérő, és laphoz csatlakoztatott akusztikus szenzor, jobboldalon közel a mote a hozzácsatolt akusztikus érzékelővel, távolabb motoros fűrészhang.



15. ábra Egy 2013. októberi mérés és a felhasznált eszközök

4.3 Autokorrelációs eljárás

Az autokorrelációs eljárás használatakor célunk a hangfelvételek alapfrekvenciájának becslése, és ezen alapuló osztályozás. Az implementáció során figyelembe kell vennünk számítási és memória igényeket is: a leírókinyeréshez szükséges a hangfelvétel valamint az autokorrelációs szekvencia tárolása is. A hangfelvételek 10 bites felbontásban készülnek, azonban e jelsorozatot csak 16 bites memóriablokkokban tudjuk tárolni. Az jelsorozat hosszától függően az autokorreláció számításánál bármilyen nagy számok előfordulhatnak, amelyeket már nem tudnánk az adott architektúrán reprezentálni és tárolni, emiatt korlátokat kell bevezetnünk a felvételek hosszára vonatkozóan. A tesztek során arra jutottunk, hogy 1024 jel hosszúságú sorozat alapján ($1/8$ másodperc $8333,3\text{Hz}$ mintavételi sebességnél) már megfelelően detektálható a számunkra érdekes frekvencia tartomány ($100\text{-}300\text{Hz}$). Amennyiben e korlátot alkalmazzuk, az autokorrelációs sorozatban lehetséges legnagyobb szám $2^{10} \cdot 2^{10} \cdot 2^{10} = 2^{30}$ elfér 4 Byte-on, ennek megfelelően az autokorrelációs szekvencia tárolása 2 KByte memóriát vehet igénybe, így lehetőségünk van akár hat darab $1/8$ mp hosszúságú felvétel készítésére amennyiben mindig csak egy 1024 jel hosszúságú szelet autokorrelációját tároljuk.

4.3.1 Előfeldolgozás

A felhasznált eszköz két digitális potméterrel rendelkezik, melyeket használva különböző erősítési beállításokat alkalmazhatunk a felvételeknél. Amennyiben megfelelő beállításokat alkalmazunk, az eszköz kevésbé érzékeny a távolság tekintetében.

AutoGain algoritmus

Mindkét potméter szabadon programozható, $0\text{-}128$ erősítési értékig. A továbbiakban az erősítési beállítást $G(g_1, g_2)$, $g_1, g_2 \in \{0, 1, 2, \dots, 128\}$ jelöli.

Elkészítettem egy adaptív autogain algoritmust, amely minden felvételhez a maximális erősítést állítja be, ami még nem okozza a jelek túlcsoportulását illetve a dinamikatartomány

telítettségét. Ezt egy előzetes felvétel első negyedének vizsgálatával szűri. Ennek pszeudokódja a 16. ábrán látható. Amennyiben az erősítés megfelelő, a felvételt elemzése megkezdődik, ellenkező esetben újabb felvétel készül az új erősítési beállításokkal.

```

AUTOGAIN (g1,g2, changed)
1   peak=0
2   overflow = 0
3   CIKLUS i=1-től 1024-ig
4       HA minta [ i ] >= 900 AKKOR
5           HA minta[i]>minta[i-1] ÉS minta[i]>minta[i+1] AKKOR
6               peak=peak+1
7           ELÁGAZÁS VÉGE
8       HA minta[i]>=1000 AKKOR
9           overflow=overflow+1
10          ELÁGAZÁS VÉGE
11      ELÁGAZÁS VÉGE
12  CIKLUS VÉGE
13  HA overflow>(peak/8) AKKOR
14      HA g1>0 AKKOR
15          g1=g1-1
16          changed=1
17      KÜLÖNBEN
18          HA g2>0 AKKOR
19              g2=g2-1
20          ELÁGAZÁS VÉGE
21      changed=0
22  ELÁGAZÁS VÉGE
23  KÜLÖNBEN
24      HA changed=1 AKKOR
25          g2=g2+1
26          changed=1
27      ELÁGAZÁS VÉGE
28  ELÁGAZÁS VÉGE
29  AUTOGAIN VÉGE

```

16. ábra Az autogain algoritmus pszeudokódja

Ezzel a módszerrel a lehető legnagyobb erősítés érhető el, viszont viszonylag lassú. Amennyiben növeljük a léptéket és nem egyesével váltunk, a művelet gyorsítható. Intervallumfelezéses algoritmussal gyorsan elérhető lenne az optimális erősítés, viszont nem szükséges mindig a lehető legnagyobb amplifikáció elérése, így előre meghatározott erősítési beállításokkal a művelet még gyorsabbá tehető.

Az eszközön implementált algoritmus a lehető leggyorsabb beállítás érdekében G csak a $\{(96,96),(64,64),(32,32),(16,16)\}$ értékeket veheti fel.

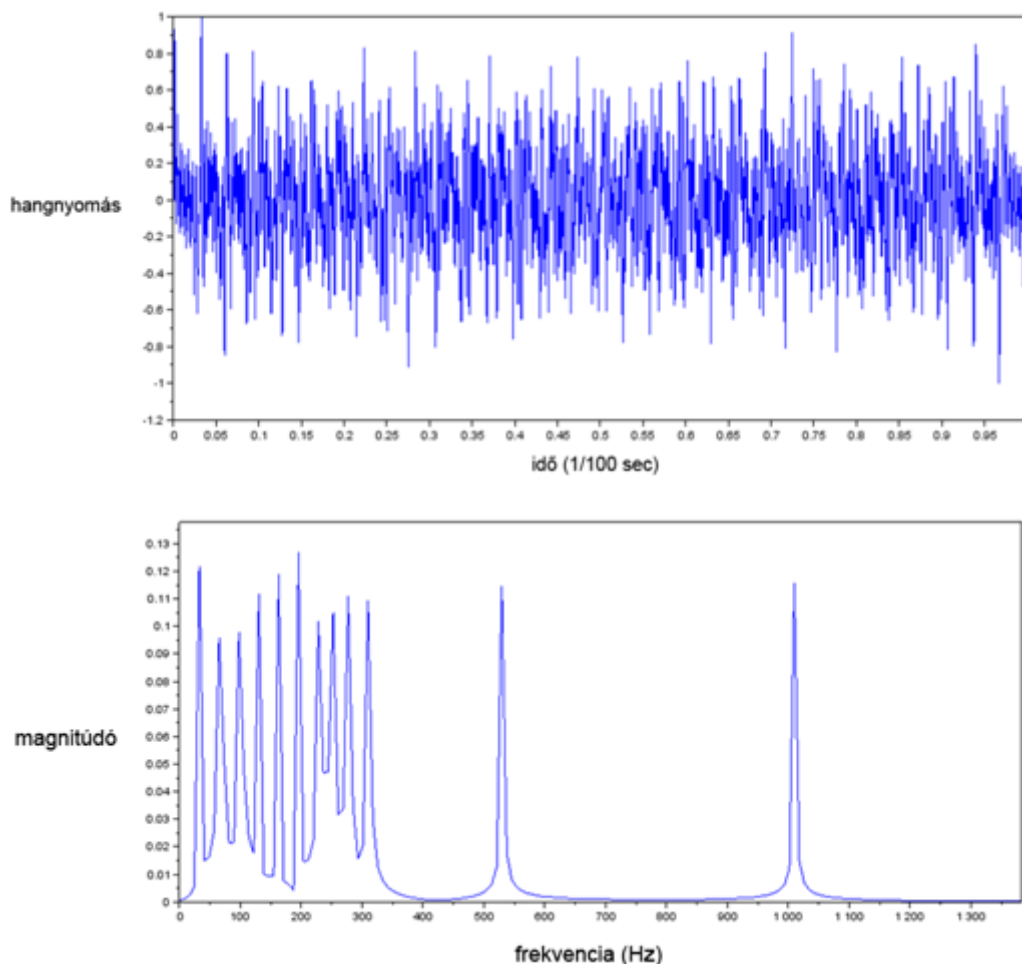
Feltételezhetjük, hogy az erdőben alapvetően csend van, és a motoros fűrész hangja távolról hallatszik. Ennek megfelelően a legnagyobb erősítési beállítással szükséges inicializálni minden mérést. Amennyiben a dinamikataromány túltelített, az eszköz a leírtaknak megfelelően csökkenti az erősítést, mindaddig, amíg a telítettség megszűnik, vagy eléri a legkisebb beállítási lehetőséget.

Alul-mintavételezés

Az akusztikus szenzor által rögzített hangok osztályozásának gyorsítása érdekében a rögzített adatok alul-mintavételezése lehetséges. Az újra-mintavételezéssel jelentősen csökkenthető a számítási és memória igény, pozitívan befolyásolható az eszköz működése. A rögzített adatok 8333KHz mintavételi sebességgel készülnek (ez a sebesség adott, az eszköz által fizikailag meghatározott legalacsonyabb érték). Viszont a feladat keretein belül felismerendő hangok az 50-300Hz tartományba esnek, pontosabban a motoros fűrész hangja alapjáraton 60-80Hz, vágás közben pedig 100-300Hz jelentős összetevőt tartalmaz. Ahhoz, hogy ilyen alaphérfkvenciát meg tudjunk állapítani, nem szükséges a 8KHz-es mintavétel. A

Nyquist-Shannon mintavételezési tételből adódik, hogy a 300Hz felismeréséhez elegendő 600Hz-es mintavételi sebesség. Ennek megfelelően a felvételeket akár 1/14 arányú mintavételezése is lehetséges, viszont az ilyen mértékű információcsökkenés nagymértékben befolyásolja a fent leírt módszer működését. Annak érdekében, hogy a számítási igény jelentősen csökkenjen, és az autokorreláció alapú módszer is megfelelően működjön, a felvételek negyedére történő újra-mintavételezése szükséges. A megfelelő alul-mintavételezéshez egy olyan low-pass filter keresése indokolt, amelyet felhasználva az adott frekvencia tartományban levő hangok detektálhatósága nem változik, és nem keletkeznek fals minták az újra-mintavételezés során.

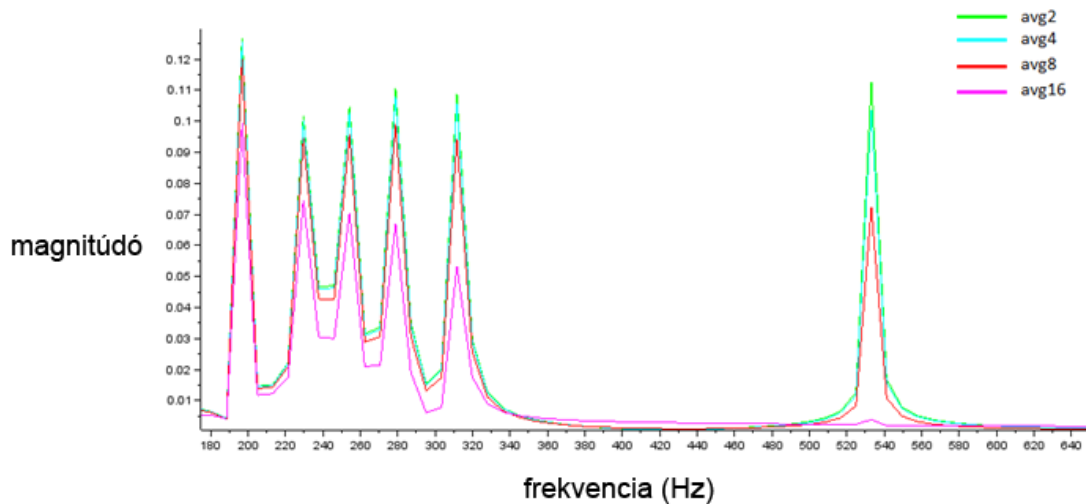
Az első lépés a vizsgálatok során egy jelsorozat generálása, amely különböző frekvenciájú szinusz hullámok összessége. A frekvenciák kiválasztásánál elsődleges szempont, hogy elég távol álljanak egymástól ahhoz, hogy az alul-mintavételezett hang Fourier-transzformációs elemzésénél könnyedén megkülönböztethető legyen, hogy melyik maximum, melyik frekvenciához tartozik, valamint, hogy a kiválasztott frekvenciák ne legyenek egymás közös többszörösei. A generált jelet (fent) és spektrumát (lent) a 17. ábra mutatja be.



17. ábra A generált jelsorozat, valamint Fourier transzformáltja

Második lépésként szintén az eszköznek megfelelően egy 1024 hosszúságú jelsorozatot szükséges választani a további vizsgálatokhoz, véletlenszerű eltolással. A szűrő kiválasztásánál elsődleges szempont a minimális számítási igény. Emiatt a szűrő keresése a különböző méretű boksz filterekre korlátozható, ennek nagy előnye, hogy nem igényli törtek használatát. Amennyiben 2, 4, 8, 16, stb. jel átlagát vizsgáljuk a számítási igény tovább

csökkenthető, mivel az osztás numerikus módszerek használata nélkül, egyszerű shift művelettel elvégezhető. A kettes és négyes átlagolás nem vált ki jelentős változást a vizsgált frekvenciáknál, viszont a nyolcas már kiszűri a mintából az 1000Hz feletti jeleket és csökkenti valamelyest a 300-500Hz körüli maximumok értékét. A tizenhatos átlagolás nagyjából 200Hz-ig szinte változatlanul hagyja az értékeket, viszont a 200-500Hz tartományban nagymértékben változtatja a maximumok értékét, az 500Hz feletti frekvenciákat szűri. Ezen adatok jól láthatók a 18. ábra grafikonján, ahol a különböző görbék a különböző kernelmérettel végzett simítás utáni transzformáltat jelölik.



18. ábra A Fourier transzformált az alul-áteresztő szűrők használata után

Konklúzió: az eszközzel való munka során biztonságosan használható négyes szélességű simító kernel és a jel negyedelése, mivel az újra-mintavételezés során a jel alakja lényegileg nem változik.

Ennek és az autokorrelációs sorozat szimmetriájának köszönhetően az autokorrelációs képletünk módosul:

$$R(m) = \sum_{n=0}^{|s|} (s_n - \bar{s}) \cdot (s_{(n+m) \bmod |s|} - \bar{s})$$

ahol s_n az n -edik jel, $\bar{s}$ a jelsorozat átlaga, $|s| = 256$, és $m \in \{0, 4, 8, 16, \dots, 254\}$ (4.1)

4.3.2 Leírókivonás

Jelen módszernél az autokorrelációs szekvenciából nyerjük ki az előző fejezetben már említett deskriptorokat.

A leírókivonásnak három paramétere van:

- a maximumkeresési ablakméret (maxWindow) megadja, hány szomszédos jel közül kell, hogy maximális legyen az adott jel
- a simítási kernelméret (kernelSize), simítást alkalmazva lehetséges az autokorrelációs szekvenciában levő, zajból adódó kis különbségeket eltüntetnünk
- egy vertikális alsó korlát (bound), amely meghatározza, hogy minimum mekkora értékkel rendelkező maximumokat vegyünk figyelembe

A maximum keresés folyamatának pszeudokódját a 19. ábra mutatja be.

```
FINDMAX(maxWindow, kernelSize, bound)
1   HA kernelSize<0 AKKOR
3       simítás a paraméternek megfelelő kernellel
4   ELÁGAZÁS VÉGE
5   maxCounter=0
6   maxIdx[0]=0
7   direction=0
8   CIKLUS i=1-től 128-ig
9       HA corr[i]>corr[i-1] AKKOR
10           direction=1
11       KÜLÖNBEN
12           HA corr[i]<corr[i-1] AKKOR
13               HA direction=1 AKKOR
14                   direction=0
15               HA corr[i-1]>bound ÉS corr[i-1] > MAX(i-1,maxWindow) AKKOR
16                   maxCounter=maxCounter+1
17                   maxIdx[maxCounter]=i -1
18               ELÁGAZÁS VÉGE
19           ELÁGAZÁS VÉGE
20       ELÁGAZÁS VÉGE
21   ELÁGAZÁS VÉGE
22   CIKLUS VÉGE
FINDMAX VÉGE
```

19. ábra A maximumkeresés pszeudokódja

A talált maximumok listáján végighaladva kiszámoljuk a maximumok átlagos távolságát és a távolságok szórásnégyzetét, ismerjük a talált maximumok számát, valamint az első maximum értéke megegyezik a rövid idejű energiával. A számított jellemzőknek elegendők az alapfrekvencia becsléséhez, a becslési pontosság megállapításához, valamint a megfelelő osztályozáshoz.

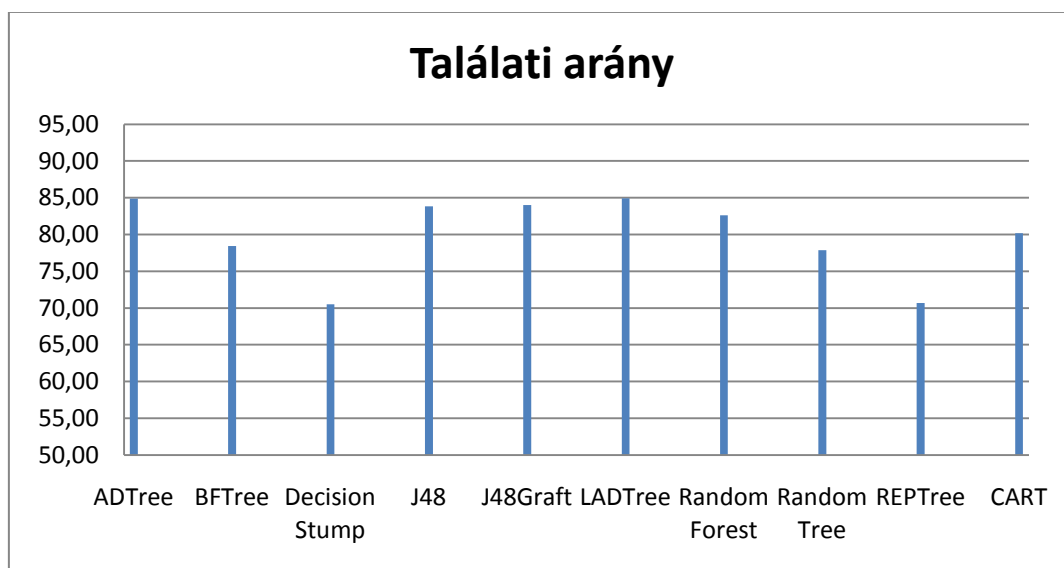
4.3.3 Osztályozás

Az eszközök egyéni döntésének meghozatalához gépi tanulási módszereket alkalmazva különböző osztályozók készíthetők. A szenzoron való implementálhatóság érdekében alacsony memória és számításigényű osztályozó keresése szükséges, a leggyakrabban használtakat a 3. fejezetben már áttekintettük. Az egyik legkönnyebben implementálható osztályozó típus a döntési fa. A gyors munkavégzés érdekében az osztályozók összehasonlításához a Waikato Egyetem által fejlesztett [WEKA] elnevezésű szoftvert használtam.

A vizsgált, WEKA-ban implementált döntési fák a következők:

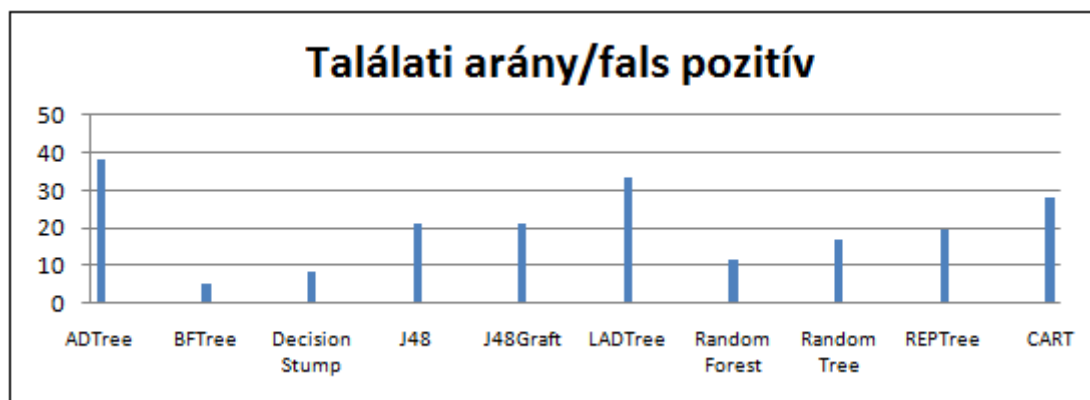
Alternating Decision Tree [Freund 1999], Best-first decision tree [Shi 2007], Decision Stump, J48 (C4.5) [Quinlan 1993], J48Graft [Webb 1999], LogitBoost Alternating Decision Tree [Freund 1999], Random Forest [Breiman 2001], Reduced-Error Pruning Tree [Mark Hall 2009], Simple Classification and Regression Tree [Leo Breiman 1984]

A kivont leírókat felhasználva különböző osztályozók különböző eredményeket szolgáltatnak. A feladat tekintetében kijelenthető, hogy az újra-mintavételezett hangfelvételek autokorrelációs szekvenciájának leíró vektorai alapján a különböző döntési fák 70-85% pontossággal képesek megkülönböztetni a fűrészhangot a környezeti zajoktól.



20. ábra A tesztelt döntési fák találati arányai

A 20. ábra grafikonján jól látható, hogy több döntési fa is 85% körüli találati aránnyal rendelkezik. A megfelelő fa kiválasztásához más szempont figyelembevétele is szükséges. A feladat tekintetében az osztályozó fontos jellemzője a fals pozitív osztályozások aránya. Erdei munka detektálásánál a fals pozitívok száma jóval kritikusabb tulajdonság, mint a fals negatívok száma. Az erdei munkafolyamatok során a különböző gépek hangjai több órán keresztül is hallhatók, ezalatt akár több száz felvételt értékelhet ki az eszköz, tehát nem annyira kritikus, ha néhány felvétel esetében a szenzor nem jelzi a munkavégzést. Fordított esetben, viszont fontos, hogy a szenzor ne küldjön fals riasztást az erdőben történő munkavégzésről, mivel a riasztás ellenőrzése általában egy járőr kiküldését igényli, ami jelentős plusz költséget jelent. Ezek alapján a következő vizsgálati szempont a fals pozitívok száma.



21. ábra A vizsgált döntési fák találati aránya a fals pozitívok tekintetében

A 21. ábra grafikonján látható, hogy az ADTree (Alternating Decision Tree) teljesít a legjobban. ADTree egy „boosting” algoritmussal növesztett fa, amely önmagában gyenge osztályozók strukturált egyesítése egy jobb teljesítményű osztályozóként. A végeredmény egy

bináris osztályozó, melynek kimenete egy valós szám. A kimeneti érték előjele határozza meg, hogy a szenzor melyik osztályba sorol egy hangmintát. A tesztek szerint ezzel az algoritmussal előállított döntési fa összességében 84,9%-os találati aránnyal rendelkezik, míg az esetek 2%-ában fals pozitív, 13%-ában pedig fals negatív jelzést ad.

Az eszközön implementálható osztályozók közül ez bizonyult a leghatékonyabbnak, ezért ez került implementálásra. Az eszköz a százalékos kimeneti értéket az osztályozás eredményének nullától való relatív távolsága alapján határozza meg.

Előzetes fúzió:

Az akusztikus szenzorok egy fél másodperces felvételt négy részletben értékelnek ki. Feltehetjük, hogy a felvétel ideje alatt a fűrészhang alaphfrekvenciája jelentősen nem változik, így a négy autokorrelációs szekvenciából kinyert leírók egyszerű átlagolása jó eredménnyel használható egy egyszerű osztályozóval. Az átlagolás problémája, hogy amennyiben csak motoros fűrészhangokat tesztelünk jó eredménnyel szolgál, viszont egyszerű környezeti hangok átlaga eredményezhet fals pozitív jelzést. Ennek elkerülése érdekében plusz leírókat szükséges bevezetni, melyek segítségével jobban elkülöníthető a két hangosztály. Ilyen leírók lehetnek a hangszeletek osztályozási eredményeinek átlaga, valamint a számított leírók varianciája. Az új deskriptorokat használva ismételten egyszerű osztályozókat készíthetünk gépi tanulási módszerekkel, azonban a szenzorral készített valós felvételek tesztjei során kiderült, hogy a fals pozitívak nagymértékű szűréséhez elegendő az osztályozási eredmények, százalékos esélyének átlagát figyelembe venni. A többi leíró bevonásával kismértékben csökkenthető a fals negatívok száma, viszont a minimális javulás a számítási igény – a haszonhoz képest - nagy növekedésével jár.

A tanításhoz 240 darab négy szeletet tartalmazó hangfelvételt, míg a teszteléshez 120 független, szintén négy szeletből álló hangfelvételt használtam (AGDB). Egyszerű küszöb használatával a keveredési mátrix a következő:

2. táblázat Egyéni döntéshozatal, keveredési mátrix

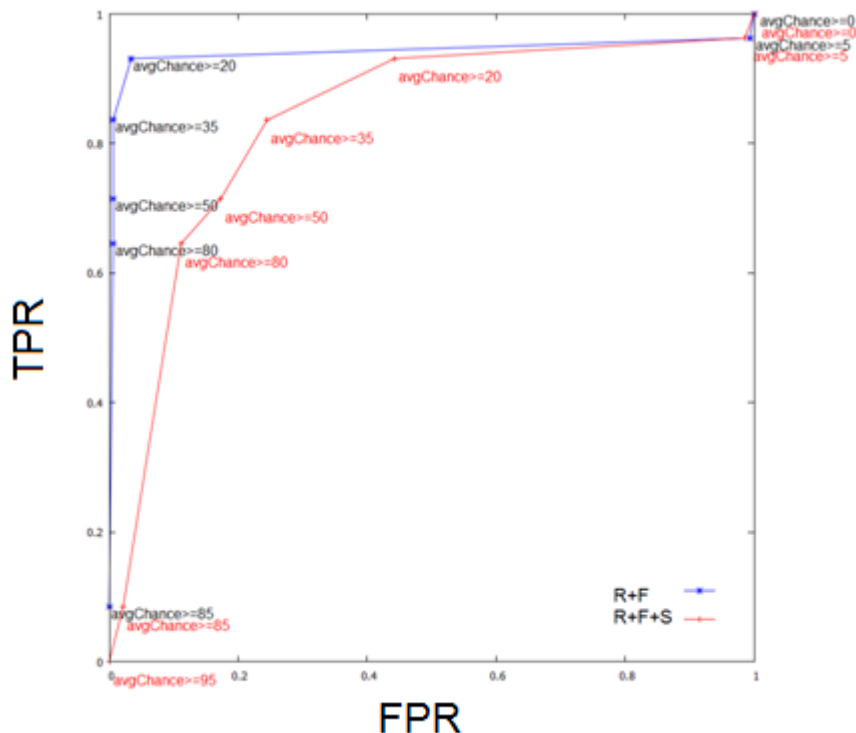
DecisionStump (valószínűségi átlag>=42)			
osztályozott	pozitív	negatív	megfigyelt
	53	18	pozitív
	1	48	negatív

A 2. táblázatból jól látható, hogy a találati arány 84,16%, a fals negatív jelzés 15%, a fals pozitív pedig 0,83%. Ezek szerint a tesztadatok tekintetében a fals negatívok aránya ugyan nőtt valamelyest, a találati arány jelentősen nem változott, viszont a fals pozitív jelzéseket sikerült 1% alá szorítani, ezzel is segítve a hatékony energiagazdálkodást és a pontosabb riasztást.

A mérések során felmerült érzékenységi problémát az előerősítési beállítások rögzítésével orvosoltuk, az új (FGDB) adatbázison tesztelve az algoritmust újabb problémába ütköztünk. Az autokorrelációs módszer nem tud megfelelően különbséget tenni a mennydörgések és a belsőégésű motorok hangja között. A viharban hallható természeti hangok és a távoli motorhangok intenzitása közel azonos nagyságrendű. Ennek köszönhetően kizárólag az alaphfrekvencián, illetve a frekvencia változásán alapulhat a döntés, viszont a mennydörgések hangjának alaphfrekvenciája gyakran azonos egy belsőégésű motor hangjáéval, valamint a vizsgált rövid időintervallumban szinte alig változik. Az eredményeket

ROC (Receiver Operating Curve) görbékén ábrázoltam, ezen típusú grafikonokon a tesztek során tapasztalt valós pozitív arány (TPR) és fals pozitív arány (FPR) kapcsolatát figyelhetjük meg.

Az előzetes fúzióval az FGDB adatbázison elért eredményeket a 22. ábra mutatja be.



22. ábra Az előzetes fúzióval elért eredmények, viharhang figyelembevételével (piros) és anélkül (kék)

Az ábrán jól látható, hogy amennyiben a tesztet viharhangok nélkül futtatjuk (R+F), az AGDB adatbázison futtatott tesztnek megfelelő eredményeket érünk el, viszont viharhangok hozzáadásával (R+F+S) a rendszer teljesítménye jelentősen romlik, a legjobb esetben is 12% fals pozitív tartozik a 64% valós pozitívhoz.

Megjegyezzük, hogy a tesztekben viszonylag rövid időintervallumok feldolgozását és elemzését végeztük, mutatjuk itt be. Ha folyamatos, hosszú percekig tartó favágás történik, akkor az esetleges jelzés gyakorisága érezhetően nagyobb, mint a néhány percenként előforduló mennydörgések esetében.

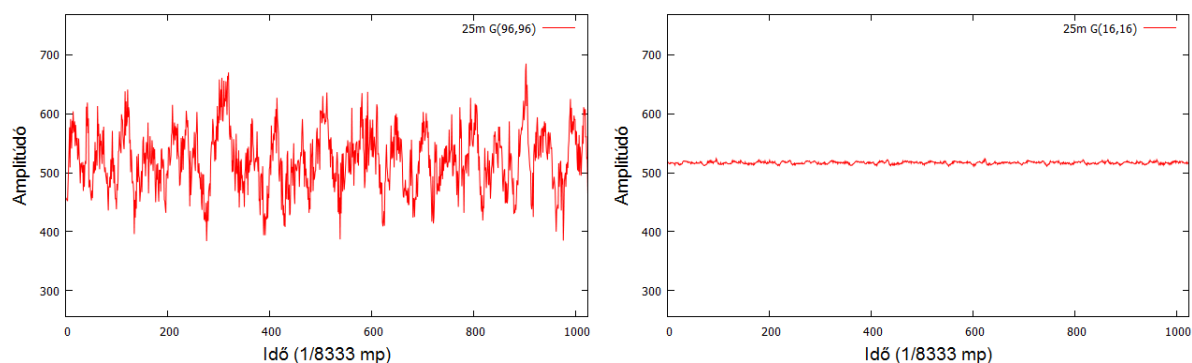
4.4 TESPAR

Az időtartományban történő feldolgozás egyik rendkívül nagy előnye, hogy elegendő a hangfelvételeket, valamint a kinyert leírókat tárolni, nem szükséges a mért jelek transzformációja a deskriptorok számításához. Ennek köszönhetően a TESPAR rendkívül gyors, közel valós idejű feldolgozást tesz lehetővé, valamint memória igénye kisebb az autokorrelációs módszernél. TESPAR használatánál a memóriakorlátok lazábbak, mivel elegendő a kizárólag a hangfelvételeket és a kinyert leírókat tárolni. Ennek köszönhetően az eszköz képes akár 6/8 másodpercnyi hangfelvétel tárolására is. Természetesen itt is adódik – a korábban ismertetett – hardveres probléma, emiatt ebből csak 5/8 hasznosítható.

Eleinte ennél a megoldásnál is teszteltük a rövidebb 1/8 másodperc hosszúságú felvételekkel elérhető eredményeket, viszont arra jutottunk, hogy a hosszabb hangfelvételek alapján pontosabb képet kaphatunk a hangforrásról.

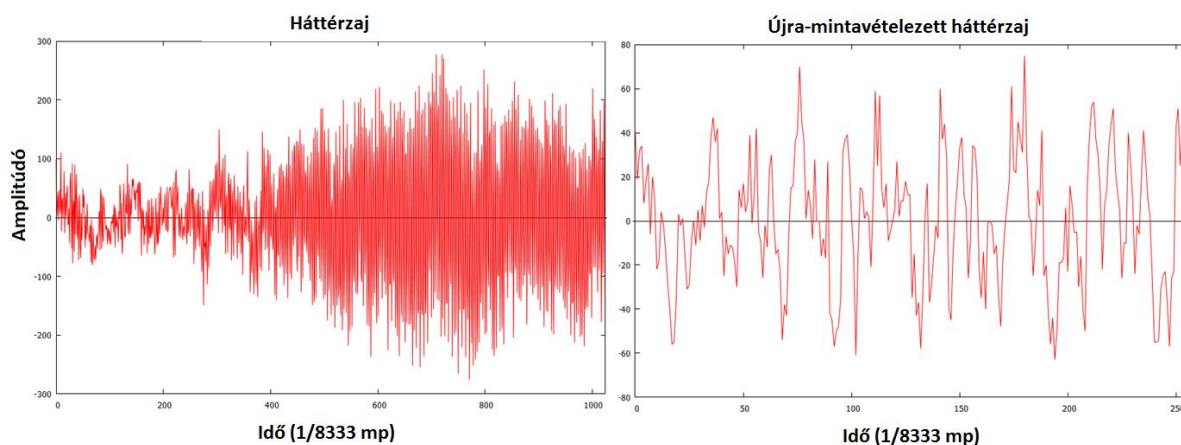
4.4.1 Előfeldolgozás

Első ránézésre az autokorrelációs módszernél alkalmazott autogain algoritmus itt jól használható, azonban különböző tesztekkel igazoltuk, hogy míg az erősítési beállítások módosításával a hang alaphfrekvenciája nem változik, az időtartományban található jelek nagymértékben torzulnak. Magasabb előerősítést alkalmazva megjelennek a háttérzajt jelentő magas frekvenciájú komponensek, amelyek minden nulla-átmenet párnál módosítják az S értéket. Ennek következtében az eljárás azonos hang (pl. természetes erdei háttérzaj) esetén, különböző erősítési beállításoknál különböző leírókat generál. Ezt mutatja be a 23. ábra.



23. ábra Fűrészhang felvétel 25m távolságból G(96,96) (bal) és G(16,16) (jobb) erősítési beállításokkal

Az első TESPAR tesztek végzése során abba a hibába estem, hogy evidensnek tekintettem az alul-mintavételezés lehetőségét, viszont sajnálatos módon a sebesség optimalizálásra használt alul-mintavételezés nem alkalmazható megfelelően a TESPAR-t használó eljárás esetén. Utólag átgondolva ez könnyen belátható. Amennyiben negyedeljük a hangfelvételt, a legkisebb észlelhető nulla-átmeneti pár hossza az eredeti hangfelvétel tekintetében négy egység, valamint az alul-áteresztő szűrő alkalmazásával nagymértékben befolyásoljuk az adott párokra vonatkozó S értékeket is. Ez jól látható a 24. ábra felvételein.



24. ábra Szélhang felvétele, valamint alul-mintavételezett mása

4.4.2 Leírókivonás

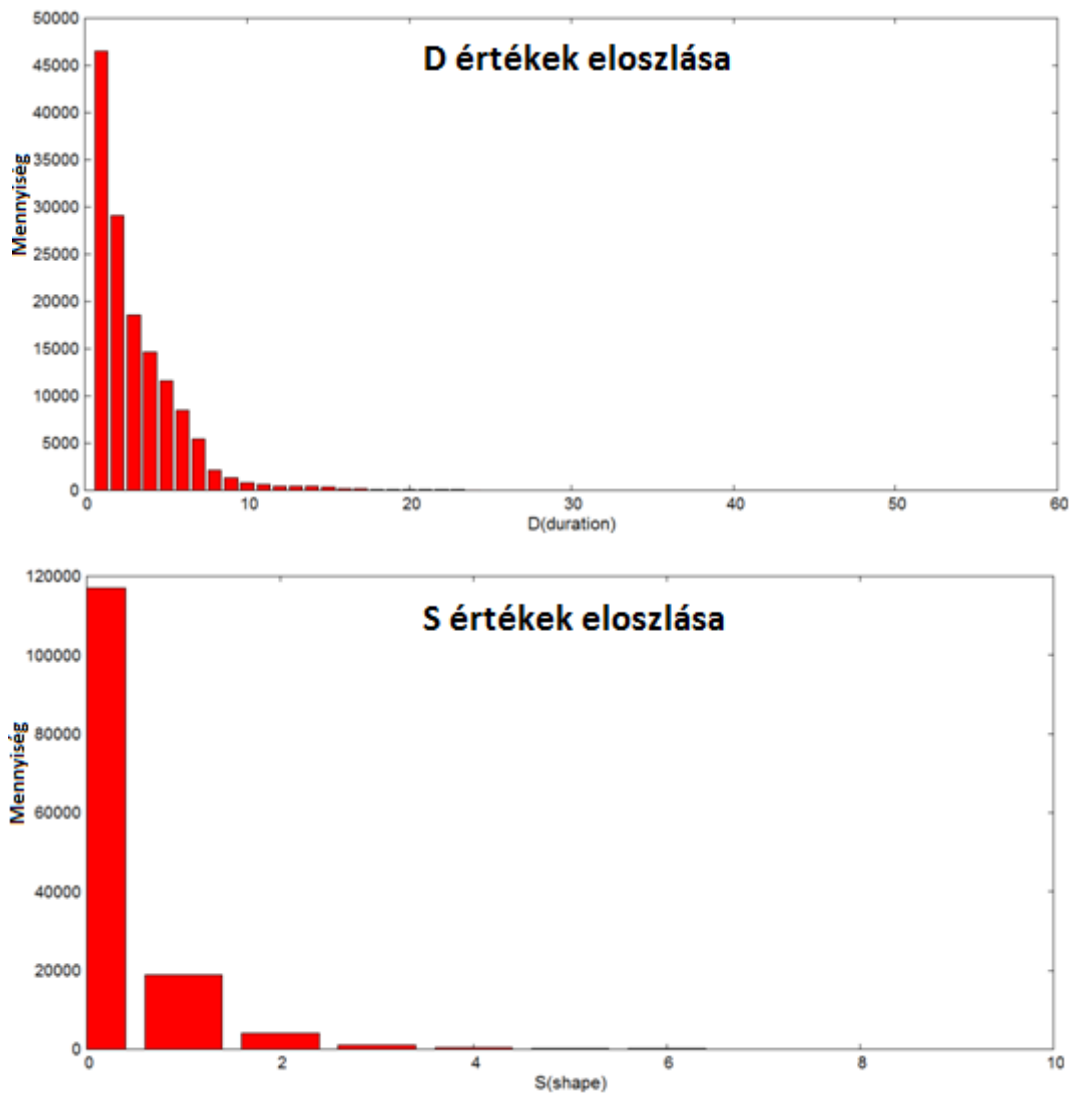
A TESPAP módszer alapvetően két absztrakt leírót használ, az S (hisztogram) és A mátrixot (együttes előfordulási mátrix). Könnyen belátható, hogy a rögzített jel időtartományban való pásztázásával akár valós időben is előállítható a jel által generált szimbólumsorozat, ennek következtében a szimbólum hisztogram számítása sem vesz plusz időt igénybe. Ez az A mátrix esetében is fenn áll, viszont az „L” paraméternek megfelelő ablakot kell használni a szimbólumsorozaton, azaz az L-nek megfelelő számú szimbólumot mindig tárolnunk kell egy FIFO szerkezetben. Felmerül a kérdés, vajon mi lehet L megfelelő értéke jelen használati esetben. Erre a kérdésre a választ sajnálatos módon csak fordított irányban tudjuk meghatározni. Adott osztályozási módszerhez kereshetjük az optimális L értéket, azonban célunk egy új detektálási rendszer kidolgozása, így egyelőre az A mátrixszal nem foglalkozhatunk.

Az említett okokból adódóan csak az S mátrix felhasználására hagyatkozhatunk. Az osztályozáshoz egyszerű, primitív eljárást szeretnénk alkalmazni, ennek érdekében östípusokat kell képeznünk a hangfelvételek által generált hisztogramokból.

Bár a hisztogramképzés valós időben is megoldható, a hangeszköz mintavételi sebességének pontossága érdekében először rögzítjük a feldolgozni kívánt hangfelvételt, majd ezek után állítjuk elő a hisztogramokat. (Az eszköz sajátosságai miatt két jel rögzítése között végzett művelet időigénye nem konstans.)

Az első tesztek során az internetes forrásból származó hangfelvételeket használtam (SWDB), az eljárás felhasználhatóságának vizsgálatához. Ennek oka egyszerűen az volt, hogy magas szintű nyelveket használhattam a gyorsabb implementációhoz.

Végeztünk néhány vizsgálatot a kódkönyvekre vonatkozóan is, és arra a megállapításra jutottunk, hogy a különböző vektor-kvantálási algoritmusokkal előállított szimbólum szótárak használata igen kis mértékben befolyásolja az osztályozás eredményét, sőt, esetünkben az S értékek megkülönböztetése elhanyagolható tényező. Ez jól látható a 25. ábra hisztogramjain.



25. ábra A hangfelvételek elemzésénél előforduló D és S értékek eloszlása

A hisztogramok készítéséhez az összes rögzített hangfelvételt használtam. Az ábrákon jól látható, hogy az S egyszer sem vesz fel 6-nál nagyobb értéket, míg a D közel 150.000-ból nagyjából 1000 esetben vesz fel 20-nál nagyobb értéket. Ezeket figyelembe véve az új szótárak előállításánál e tartományon alkalmaztam különböző vektor kvantálási algoritmusokat, viszont az új kódkönyvek nem javítottak az osztályozáson.

Az S értékek elhagyása lehetőséget ad ugyanakkora szimbólumszám mellett több D érték megkülönböztetésére, azonban az osztályozást ez sem befolyásolta. Emiatt kizárólag a standard ábécé használatával elért eredmények bemutatására hagyatkozom.

A hisztogram előállításának algoritmus a következő:

1. A felvétel első jele az első nulla-átmenet
2. Keressük meg a következő nulla-átmenetet vagy a jelsorozat utolsó elemét, majd számítsuk ki kettő távolságát (D)
 - a. amennyiben a nulla-átmeneti szakasz pozitív tartományba esik, számoljuk meg hány lokális minimum található a szakaszon (S)
 - b. amennyiben a nulla-átmeneti szakasz negatív tartományba esik, számoljuk meg hány lokális maximum található a szakaszon (S)

3. Inkrementáljuk a számított D-S értékeknek megfelelő bin értékét a hisztogramban
4. Ismételjük a 2,3 lépéseket mindaddig, amíg az utolsó jelhez nem érünk

Az őstípus képzése a tanító adatbázisban fellelhető motoros fűrészhangok által generált hisztogramok átlagolásával történik.

A környezeti hangok sokfélék lehetnek, nem hozható létre olyan őstípus, amely minden háttérzajt megfelelően lefedne, így az első tesztek során kizárólag egy pozitív hangminta őstípust képeztem. Azonban a későbbiekben fel kellett ismernünk, hogy az ilyen alacsony szintű leírók tekintetében egy-egy fűrészhang különböző S mátrixokat is generálhat, tehát nem elegendő a pozitív mintákat sem egy őstípussal lefedni.

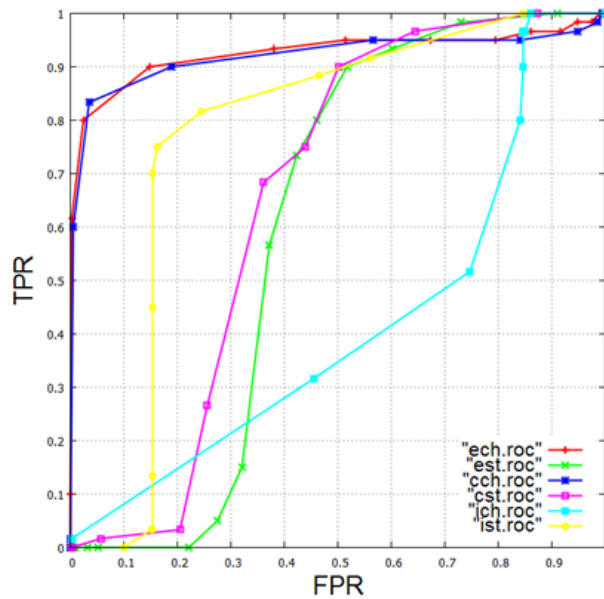
A különböző őstípusok képzéséhez a „farthest first” [Gonzales 1985] klaszterezési eljáráshoz hasonló algoritmust használtam.

Az eljárásnak két paramétere van:

- a minimális klaszterméret, amelyet még felbonthatunk
 - valamint két hisztogram közötti minimális távolság a klaszteren belül
1. Számítsuk ki a hisztogramok egymástól vett euklideszi távolságát
 2. Az így keletkezett mátrixban keressük meg a maximális értéket, majd az ehhez tartozó két hisztogramot tegyük külön klaszterekbe.
 3. Minden fennmaradó hisztogramot tegyük abba a klaszterbe, amelyik kezdő értékéhez közelebb áll.
 4. Amennyiben valamelyik klaszterben:
 - a. a hisztogramok száma kisebb, mint a paraméterként meghatározott minimum,
 - b. vagy a klaszteren belüli hisztogramok egymástól vett távolságának maximuma kisebb, mint a paraméterben meghatározott minimális távolság, a klasztert nem vizsgáljuk tovább, a tartalmazott hisztogramok átlaga új őstípus
 5. Ismételjük az 1.-4. lépést, amíg a klaszterek már nem bonthatók tovább.

4.4.3 Osztályozás

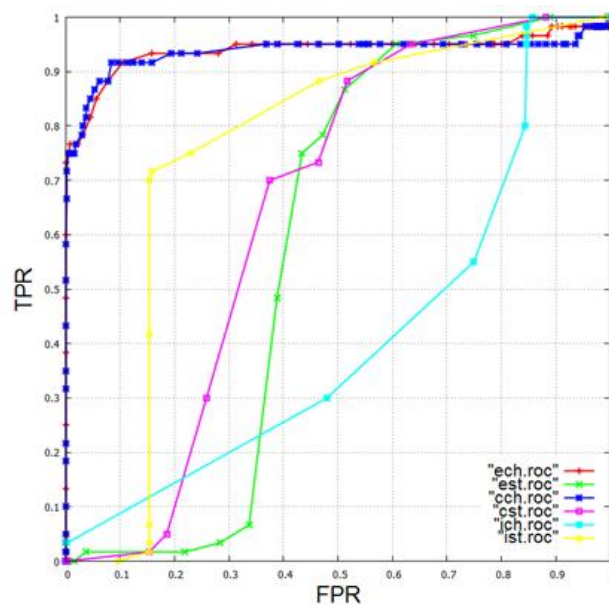
Eleinte több metrikát is teszteltünk, az őstípus és a beérkező hangminta hasonlóságának megállapításához. Ezek az euklideszi, Manhattan távolságok, illetve a hisztogram metszetek voltak. Néhány esetben a Mahalanobis távolságot is vizsgáltuk, az új metrika viszont nem javított az eredményeken, ám a számítási igényt jelentősen növelte. Teszteltük a motoros fűrészhang és viharhang őstípustól való távolság alapján történő osztályozást is. Az osztályozás kizárólag távolság alapján egyszerű küszöbértékkel befolyásolható, a következő ábrákon az adott távolságokat használva készített ROC görbék láthatók (e*= euklideszi távolság, cb*=Manhattan távolság, i*= metszet, m*= Mahalanobis távolság, *ch=fűrészhang, *st=viharhang)



26. ábra 1/8 hosszúságú minták hasonlítása egy őstípushoz

Jól látható, hogy az euklideszi és a Manhattan távolság nyújtja a legjobb TPR/FPR arányt, valamint hogy az eljárás eredményeit nem rontja a viharhangok jelenléte. Azt is érdemes megfigyelni, hogy a viharhang őstípustól vett távolság az autokorrelációs eljáráshoz hasonló eredményeket mutat.

Az eddigi tesztek eredményei alapján az sejthető, hogy hosszabb idejű megfigyelés alapján jobb döntés hozható a megfigyelt jelenséggel kapcsolatban. Az alapötlet, hogy a maximálisan rögzíthető ötszörös hosszúságú felvételeknél a motoros fűrész hangjában nem történik nagy változás, azaz a hangban előforduló szimbólumok száma szimbólumonként nagyjából megötszöröződik, a negatív hangok esetében viszont újabb szimbólumok kerülhetnek a hisztogramba.

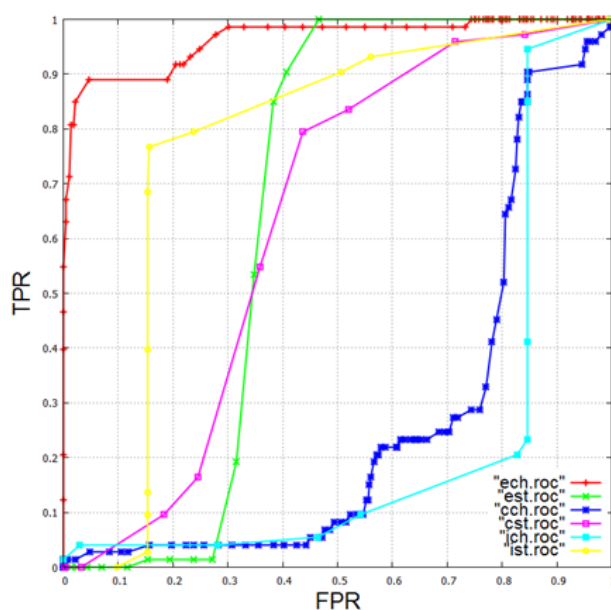


27. ábra 5/8 mp hosszúságú minták hasonlítása egy őstípushoz

Hosszabb mintákat felhasználva az ROC görbék a vártak megfelelően módosultak. A 90% körüli helyes felismerés mellett metrikától függően 6-10% fals pozitív érhető el. Szintén jelentős különbség, hogy fél százalék fals pozitív engedélyezése mellett euklideszi és Manhattan távolságot alkalmazva is 75%-os TPR valósítható meg a korábbi 60%-hoz képest.

Az előző részben ismertetett őstípus képzési algoritmus használatával állítottam elő több archetípust mivel a pozitív hangok is lehetnek egymástól nagyon eltérőek. A klasszifikáció az összes őstípushoz viszonyított távolságok minimumán alapszik, vagyis minden esetben az az őstípus a mérvadó, amelyikhez a legközelebb áll a minta.

Az eljárás eredményeként keletkezett nyolc fűrészhang és hat viharhang archetípust felhasználva az eredmények a következő ROC görbén láthatók.



28. ábra 8 fűrészhang és 6 viharhang őstípust használó osztályozók ROC görbéi

Az ábrán jól látható, hogy a vizsgált esetben egyedül az euklideszi távolság használható megfelelően az osztályozáshoz. Ezen metrika használatával a 2%-5% fals pozitív rátának megfelelő tartományban jelentős TPR növekedés látható az egy őstípust használó verzióhoz képest (az intervallum határain 76%-ról 81%-ra, valamint 82%-ról 89%-ra nőtt), viszont 0% körüli fals pozitív mellett jelentősen csökken a TPR (75%-ról 67%-ra).

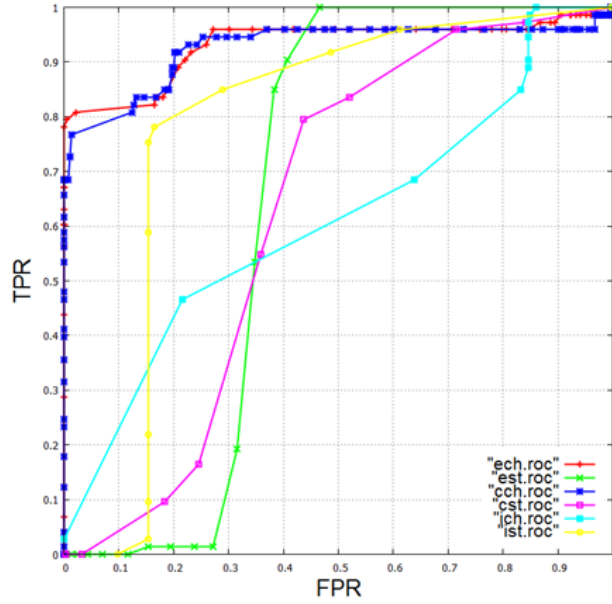
Amennyiben célunk a fals pozitívok számának csökkentése, ellenőriznünk kell, hogy adott őstípusok mennyire jellemeznek bizonyos negatív hangokat. Ennek érdekében külön háttérzaj őstípusokat is képeztem, majd kiszámítottam a pozitív archetípusok euklideszi távolságát a negatívoktól.

3. táblázat Az archetípusok egymástól számított távolsága

	CH1	CH2	CH3	CH4	CH5	CH6	CH7	CH8
CH1	-							
CH2	991,093	-						
CH3	1561,083	671,232	-					
CH4	1374,752	492,420	215,548	-				
CH5	1364,982	470,067	231,950	112,765	-			
CH6	1324,152	418,692	265,370	136,132	92,715	-		
CH7	1268,688	351,101	335,187	164,168	136,143	111,297	-	
CH8	1457,276	560,311	221,323	223,341	168,473	206,366	233,542	-
BG1	1108,650	592,237	1152,866	1002,641	980,377	940,712	859,221	1017,448
BG2	887,144	220,869	869,365	688,831	668,539	624,655	542,781	741,553
BG3	1292,947	468,042	527,199	461,595	412,306	381,232	355,237	341,791
BG4	1691,862	819,386	228,539	362,934	402,575	446,145	493,848	399,009
BG5	877,764	242,101	691,182	501,611	493,522	453,795	397,310	600,431
BG6	1184,098	813,773	1377,850	1229,923	1210,026	1169,636	1090,333	1247,629
BG7	1421,777	1077,971	1602,025	1467,597	1449,140	1411,021	1335,475	1479,487
BG8	292,961	738,249	1328,209	1134,986	1128,621	1087,351	1028,193	1233,921
ST1	747,791	292,289	871,477	692,826	668,198	622,781	556,916	738,133
ST2	605,319	429,524	1049,030	865,962	845,012	799,517	730,364	917,087
ST3	604,958	410,639	1005,529	822,001	801,525	757,287	692,601	879,143
ST4	319,517	699,616	1307,600	1118,406	1104,125	1061,217	996,819	1190,528
ST5	473,399	550,181	1164,471	978,029	960,423	916,403	849,254	1039,166
ST6	1678,432	805,238	212,521	347,364	387,174	430,857	478,768	385,169

Min.Táv.:	991,093	351,101	215,548	112,765	92,715	92,715	111,297	168,473
-----------	---------	---------	---------	---------	--------	--------	---------	---------

A 3. táblázatban az egyes őstípusok közti távolságok láthatók, valamint minden fűrészhang archetípus a többi fűrészhang típustól vett távolságainak minimuma (a CH kezdetű címkék motoros fűrés, a BG kezdetűek háttérzaj, míg az ST kezdetűek viharhang archetípust jelölnek). Ezen adatok jól mutatják, hogy például a CH1 nevű őstípus jobban hasonlít 8 negatív hangtípusra (pirossal jelölt értékek), mint bármely más fűrészhang típusra. Ezek alapján arra a következtetésre juthatunk, hogy adott negatív hangok elég közel lehetnek pozitív őstípusokhoz, ahhoz hogy pozitív hangmintaként értékelje azokat egy osztályozó. Ennek kiküszöbölésére az ilyen őstípusok elhagyhatók, viszont számítanunk kell a TPR csökkenésére, ugyanis így adott pozitív hangokat nem fednek le megfelelően a fennmaradó őstípusok.



29. ábra 5 fűrészhang és 6 viharhang őstípust használó osztályozók ROC görbéi

A fennmaradó öt fűrészhang és hat viharhang archetípus használatával az eredmények a vártak megfelelőek, a TPR szinte minden FPR értéknél csökkent, viszont a 0% körüli fals pozitív aránynál a helyes jelzések aránya 67%-ról közel 80%-ra nőtt. Érdekes különbség még az előző teszthez képest, hogy a Manhattan távolságot használó osztályozó nagymértékben javult, az euklideszi távolságot használó osztályozóhoz hasonló értékeket mutat. Ezt mutatja be a 29. ábra

Az eddigi eredményeinkkel bemutattuk, a különböző metrikák, őstípusok osztályozásra gyakorolt hatásait. Kiderült, hogy a különböző ábécék használata csak kis mértékben befolyásolja az osztályozást, sőt azt is igazoltuk, hogy jelen használati esetben a nulla-átmeneti szakaszokban található hullámalak kevés többlet információt tartalmaz a szakasz hosszához képest. Nyilvánvalóvá vált, hogy a vizsgált eseteknél, általánosan az euklideszi távolság a legjobban alkalmazható metrika. Valamint kiderült az is, hogy több őstípus alkalmazása javít az osztályozáson.

A következő tesztél azt vizsgáltam, hogy egy-egy 5/8 másodperc hosszúságú internetes hangminta alapján készült hisztogram hol helyezkedik el az őstípusokhoz képest a paraméterterben. Ennek vizsgálatához kNN eljárást alkalmaztam, a távolságok alapján történő súlyozással. A súlyozásnál a legtávolabb eső őstípus képezte a normalizálás alapját, mely után a súlyok $1/\text{távolság}$ értéket vettek fel. (A legtávolabb eső minta egy szavazatot ér, a közelebbi minták egynél nagyobb súlytal számítanak.)

$$\text{dist}(aT, q) = \sqrt{\sum_{i=1}^{|q|} (q_i - aT_i)^2},$$

ahol aT jelöli az adott őstípust, q_i a felvétel hisztogram i -edik vödrének értéke. (4.2)

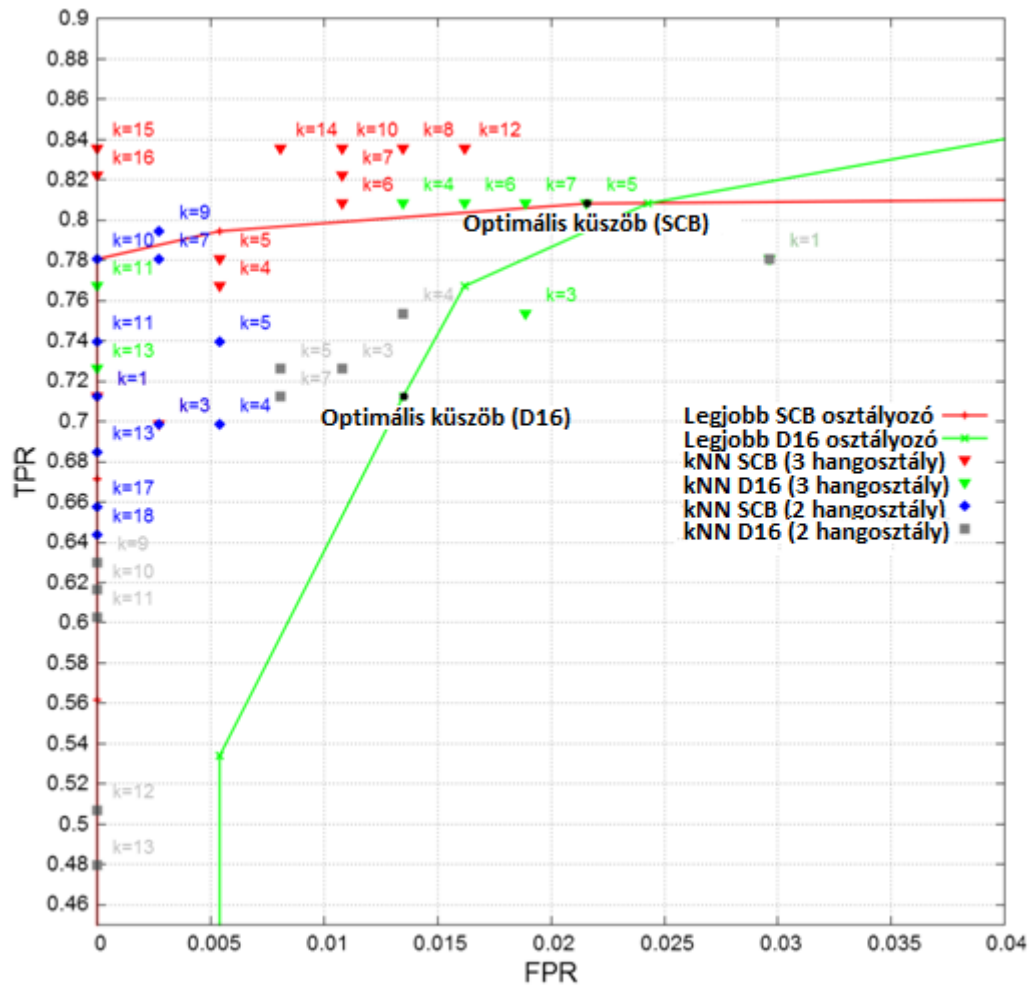
$$\text{weight}(aT_i) = \frac{1}{\text{dist}(aT_i, q) * \frac{1}{\max(\text{dist}(aT_1), \text{dist}(aT_2), \dots, \text{dist}(aT_n))}} \quad (4.3)$$

A tesztnél két különböző szótárat vizsgáltam, a standard TESPARET ábécét, valamint egy, csak a nulla-átmenethosszok alapján kvantáló, 16 szimbólumból álló szótárat. Az új szótár gyakorlatilag a nulla-átmenethosszokat tartalmazza, azzal a megkötéssel, hogy a 16 egységnél hosszabb nulla-átmeneti szakaszok a 16-os szimbólumot kapják. Mindkét esetben az adott szótárnak megfelelő őstípusokat alkalmaztam, háttérzaj hangmintákból képzett őstípusokkal kiegészítve. Ezen felül mindkét szótárnál külön vizsgáltam azt az esetet, amikor kettő (pozitív, negatív), illetve három (pozitív, háttér, vihar) hangosztályt különböztetünk meg.

A tesztek során az őstípusok adták a hangosztályok egyedeit:

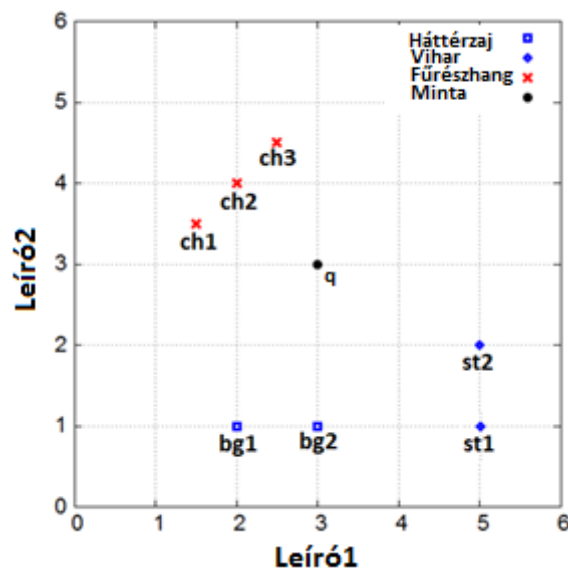
- Standard ábécével készített őstípusok száma:
 - o 5 fűrészhang egyed
 - o 8 háttérzaj egyed
 - o 6 viharhang egyed
- D értékek alapján képzett ábécével készített őstípusok száma:
 - o 3 fűrészhang egyed
 - o 7 háttérzaj egyed
 - o 3 viharhang egyed

Az említett paraméterek mellett végzett vizsgálatok eredményeit az ROC görbéknél megszokott diagramon ábrázoltam. Az ábrán feltüntettem az eddig legjobban teljesítő osztályozók eredményeit, az azoknál meghatározott optimális TPR/FPR arányt jelentő pontot, valamint az új teszt eredményeit különböző k értékeket használva. A vizsgálatoknál a k paraméter egytől az őstípusok számáig terjedő tartományban vesz fel értékeket. Amennyiben egy adott TPR-FPR pár több módszerrel is elérhető, a legkisebb k értékkel rendelkezőt tüntettem fel.



30. ábra A súlyozott kNN módszer eredményei

Az ábrán jól látható, hogy a kNN eljárás a legtöbb esetben nagymértékben csökkenti a fals pozitívok számát, valamint, hogy három hangosztály megkülönböztetése javít a fűrészhangok felismerésén. Ezt könnyű belátni a 31. ábra segítségével.



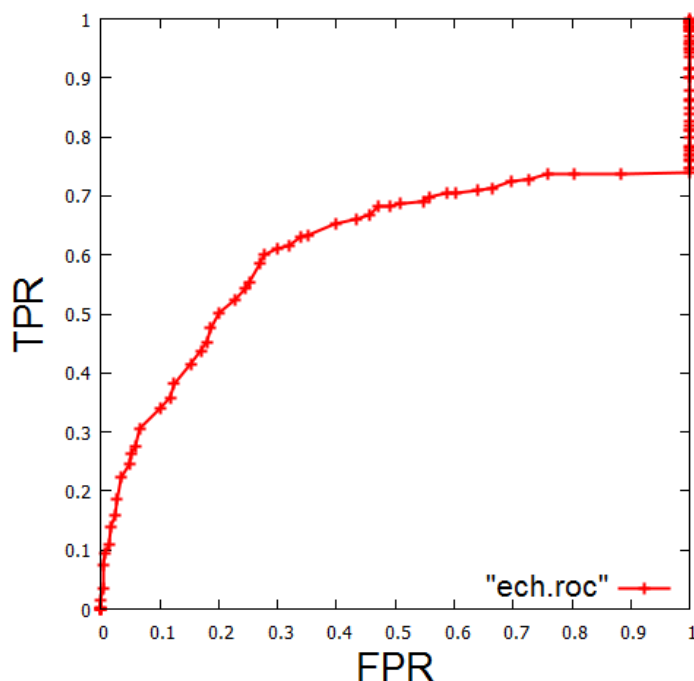
31. ábra Hangosztályok számának kNN-re gyakorolt hatása

Az ábrán a paraméterter két dimenziója látható, kizárólag a motoros fűrész pozitív hang, a háttérzaj és a viharhang is negatív hangok. Egyszerű esetet tekintve $k=7$, súlyozás nélkül, könnyen belátható, hogy a beérkező minta a megkülönböztetett osztályok számától függően más értékelést fog kapni. A körülötte levő hét egyedből három pozitív és négy negatív hang, tehát a mintát negatívként értékelnénk, viszont három hangosztály esetén a hét mintából három pozitív, kettő háttérzaj, kettő pedig viharhang, tehát a mintát pozitívként kell értékelnünk. Az ábra alapján az is könnyen belátható, hogy a hangosztályok egyedeinek száma nagymértékben befolyásolja a döntést. Amennyiben k elég nagy, az osztályozás során nagyobb valószínűséggel sorolunk egy hangmintát a nagyobb hangosztályokba.

Amennyiben az osztályozó preferenciáit minimalizálni szeretnénk, súlyozás nélküli esetben a k értéket a legkisebb hangosztály méretének kétszeresében kell maximalizálnunk. Súlyozott esetben ez a korlát lehet nagyobb is, hiszen nagyobb súllyal számítanak a közelebb eső hangok.

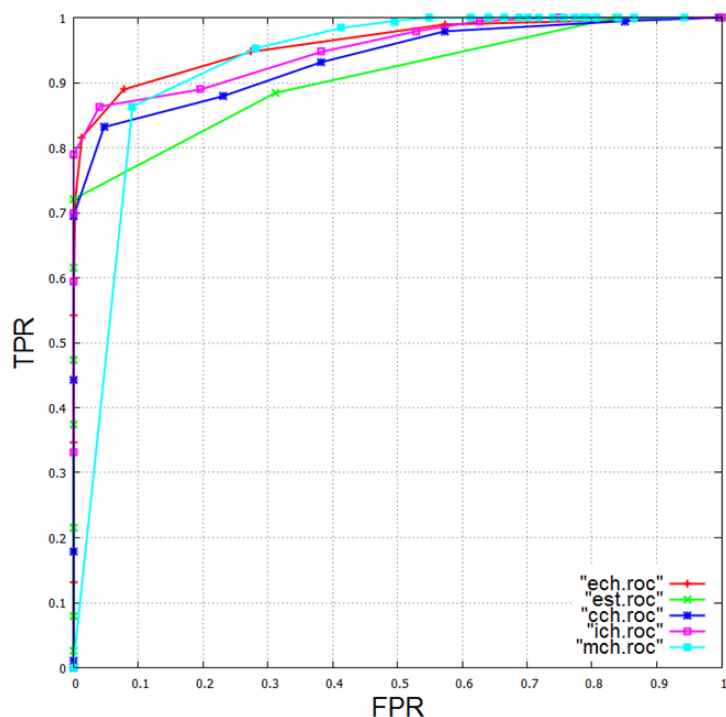
Fontos megemlíteni, hogy bár a kNN-nel kapcsolatos tesztek azt mutatják, hogy az osztályozási pontosság javulhat, a számítási igény nagymértékben nő, emiatt a céleszközön nem érdemes ezt az osztályozót alkalmazni.

Az SWDB adatbázison végzett vizsgálatok után szükséges volt a módszer valós adatokon történő tesztelése. A tesztek során kizárólag a legjobban teljesítő, hosszú felvételek, több őstípustól mért euklideszi távolság alapján való osztályozást értékeltük. A több távolságból, különböző erősítési beállításokkal rögzített felvételekből képzett AGDB-n futtatott teszteknel a módszer gyenge teljesítménye volt várható. A változó érzékenység és intenzitás nagymértékben variálja a D és S értékeket, így nagy erősítéssel rögzített háttérzaj hisztogramja hasonlíthat egy alacsony erősítéssel rögzített motorhang hisztogramjához.



32. ábra AGDB-n végzett teszt eredménye

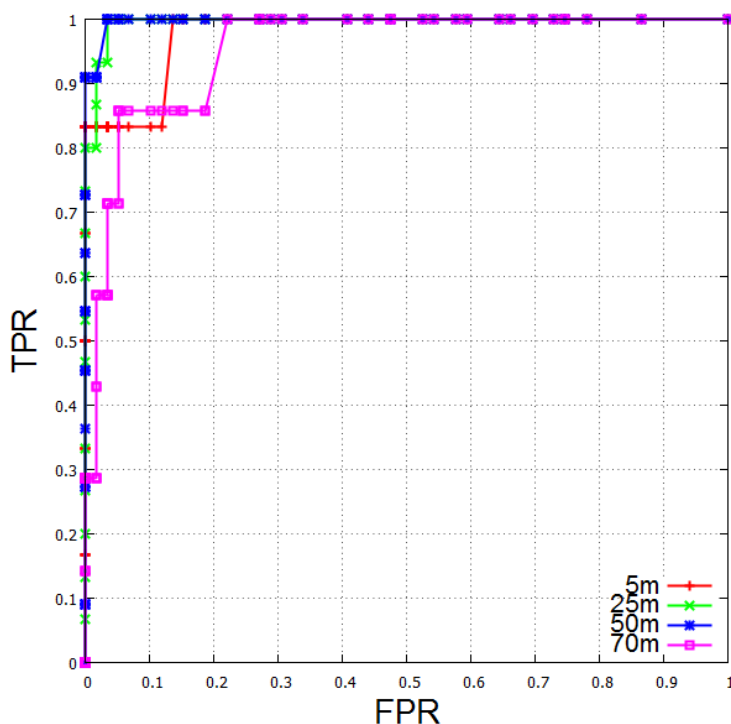
Az ábrán jól látható, hogy sokkal rosszabb eredményeket ér el a módszer a különböző erősítési beállítások használatánál, mint a normalizált felvételek esetében, viszont ez várható volt. Következő tesztként a fix erősítési beállításokkal rögzített felvételeken vizsgáltuk a módszert, mely eredményei rendkívül jónak tekinthetők.



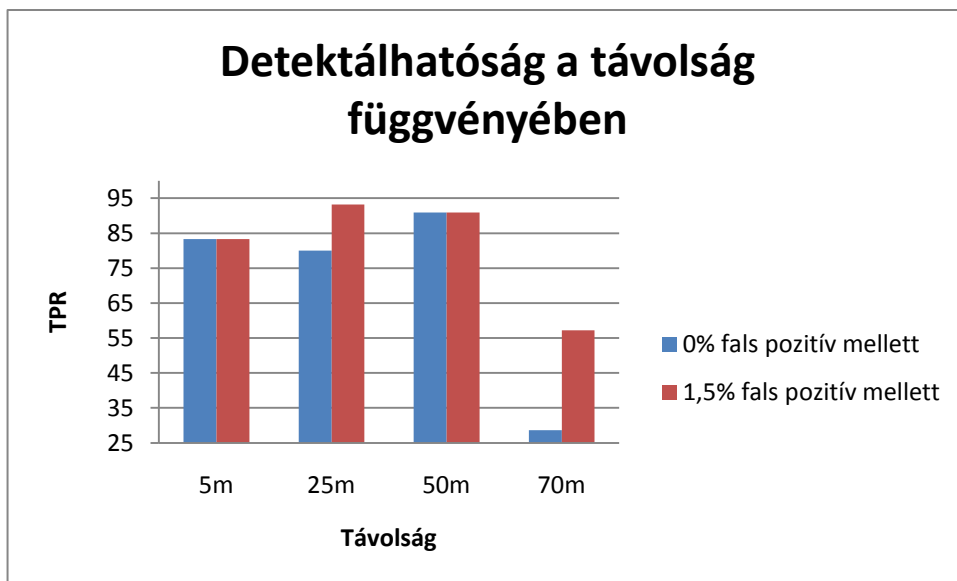
33. ábra Az FGDB-n végzett tesztek eredményei

Az ábrán jól látható, hogy rögzített erősítési beállításokkal alkalmazva a módszert, a TESPAP jelentősen meghaladja az autokorrelációs eljárás eredményeit. A tesztadatbázisban szereplő eseteknél 80%-ban jelezte belsőégésű motor hangját fals pozitívok nélkül.

Ezek után már csak a hatótávolsággal kapcsolatos tesztek maradtak hátra. A legjobb eredményt elérő osztályozót az FGDB adatbázis különböző távolságból rögzített hangjain vizsgáltam. Ennek eredményeit mutatja be a 34. és 35. ábra



34. ábra TESPAP eredmények a távolság függvényében (FGDB)



35. ábra Detektálhatóság a távolság függvényében

Az elvégzett TESPARE tesztek eredményei alapján jól látható, hogy kizárólag az S mátrixok felhasználásával határozottan jó eredmények érhetők el, viszont a távolság függvényében az osztályozás 50 méteres távolság felett nagymértékben romlik.

5 Fúziós lehetőségek

Az előző fejezetben áttekintettük, hogy a különböző módszereket használva az eszközök egyenként milyen pontossággal végzik az osztályzást egy felvétel alapján. Könnyen belátható, hogy egy mérés alapján nem jelenthetünk ki semmilyen állítást a vizsgált jelenségről, emiatt vizsgálnunk kell az információfúzió lehetőségeit.

Az információ fúzió egy olyan eljárás, mely során különböző műveletek sorozatával egy adathalmazból egy újabb adathalmazt képzünk csökkentve a bizonytalanságot. Az egyesíteni kívánt adatok jellege alapján megkülönböztetünk homogén és heterogén rendszereket. A különböző fúziós eljárásokat alacsony, közép, vagy magas szintűnek tekintjük az információ jellege, valamint centralizált, elosztott vagy hibrid fúziónak, a fúzió megvalósulásának helye alapján. Az alacsony szintű adatfúzió több mérés nyers adatát, közép szintű fúzió a mérési adatból kinyert leírókat, míg a magas szintű a leírók felhasználásával kinyert információt egyesíti. Centralizált adatfúziónál az összes eszköztől érkező információ egy helyen egyesül, míg elosztott esetben minden eszköz saját méréseit egyesíti, és az alapján hoz döntést. Hibrid esetben a szenzorok előzetes döntést hoznak a megfigyelt jelenségről, majd ezt továbbítják a fúziós egység felé, amely az összes szenzortól érkező információ alapján hoz döntést. Az akusztikus szenzorok tekintetében a technológiai korlátok miatt csak közép és magas szintű, homogén, hibrid adatfúzió valósulhat meg a rendszerben.

A szakirodalom rendkívül gazdag heterogén rendszerek tekintetében, viszont homogén leírófúzió kapcsán nem találhatók leírások.

Az „Extended Confidence-Weighted Averaging in Sensor Fusion” című írásban [Angela Schörgendorfer 2006] a leírók megbízhatóságán alapuló súlyozott fúzióról olvashatunk.

A „Feature Level Sensor Fusion” című cikk [Tamar Peli 1999] egy a fals riasztások szűrésére használt fúziós eljárást mutat be katonai célpontkövető rendszerekben.

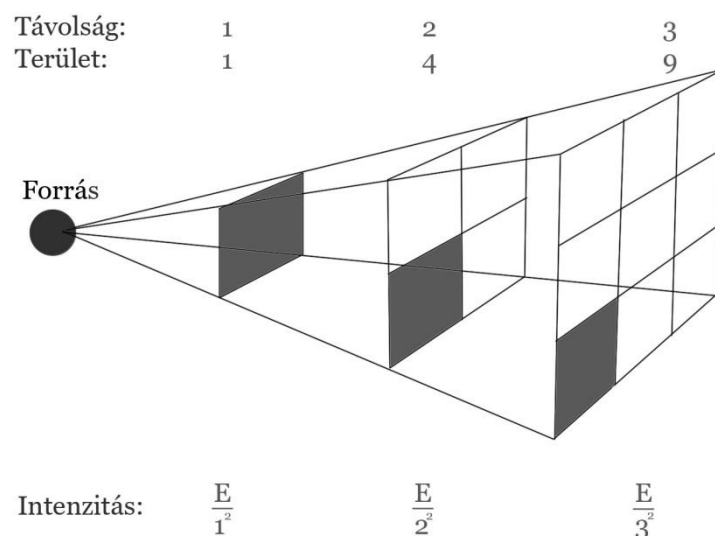
A nyers adatból kinyert numerikus leírók fúzionálása egy homogén rendszerben alapvetően két módon valósulhat meg: az összetartozó leírók konkatenációjával, vagy algebrai műveletsorozatok eredményeként új leírók készítésével. Konkatenációt használva az új osztályozó működését nagymértékben befolyásolja, hogy egy-egy információ melyik egységtől érkezett, azaz nagymértékben függ az akusztikai környezettől is. Emiatt a végleges helyén telepített szenzorok adataival készített modell optimális lehet lokalizált szenzorszigetekenél. Célunk azonban egy olyan rendszer fejlesztése, amely független a szenzoregységek elhelyezésétől. Hasonlóan, a magas szintű eljárások esetében, előzetes információkra alapozott fúziós módszerek, mint a központi határeloszlás tételt, Kálmán szűrőt, vagy lineáris meggyőződési függvényt használó eljárások sem alkalmazhatók.

Vizsgálataink alapján kijelenthető, hogy a fejlesztett eljárások közül kizárólag az autokorrelációs módszer esetében merül fel a leírófúzió lehetősége. TESPARE esetén alacsony szintű leírókat használunk, amelyek egyesítése kizárólag abban az esetben valósítható meg, amennyiben az eszközök tökéletes szinkronban vannak egymással, és közel azonos távolságról rögzítik a hangot. Ez az esetek többségében csak akkor következhet be, amennyiben a szenzorok egymás mellett találhatóak, viszont így nem növeljük a hatótávolságot, és nagy valószínűséggel az osztályozás pontosságát sem. Emiatt a leírófúziót kizárólag az autokorrelációs eljárás használatánál vizsgáltuk. A javasolt algoritmusokat megfelelő szinkronizáció hiányában sajnos csak szimulációs környezetben tesztelhetjük.

Szinkronizált felvételek létrehozása

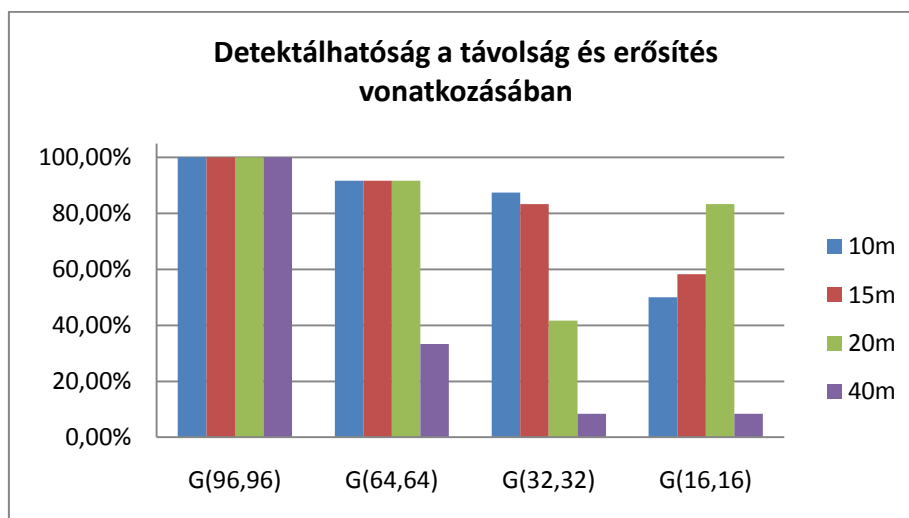
Prototípus eszközökről lévén szó az eszközök darabszáma miatt szinkronizált felvételek létrehozására azonos típusú eszközökkel eleinte nem volt lehetőség, a későbbiekben pedig a mérések során felmerülő szinkronizációs problémák miatt nem tudtunk megfelelő szinkronizált hangfelvétel adatbázist felépíteni. Annak érdekében, hogy a fúziós eljárások tesztelhetők legyenek, szükség volt egy szimulátor létrehozására. A szimulációs program adott forrásokra emulálja az akusztikus szenzorok működését, valamint a kiértékelések eredményeit továbbítja egy log-elemző modulnak. A log-elemző egység szerepe, hogy lehetőséget ad a fúziós algoritmus fejlesztéséhez és teszteléséhez.

A program wav formátumú hangfájlokat olvas be, melyeket paraméterként meghatározott számú, szinkronban működő szenzormodul vizsgál. A szimuláció inicializálása során a szenzorok hangforrástól való relatív távolságát egységnyinek, duplának, vagy négyszeresnek állítja be, valamint egy zajokat tartalmazó mappából kiválaszt egy az adott szenzorhoz tartozó zajforrást. A hangforrás távolságának szimulációja egyszerű feladat, amennyiben a hangfájl készítésekor fennálló távolságot egységnyinek tekintjük, majd ehhez képest relatív távolságokat szimulálunk az inverz négyzetes törvény segítségével. Inverz négyzetes törvény a hang vonatkozásában azt jelenti, hogy adott hang intenzitása fordítottan arányos a hangforrástól való távolság négyzetével (ennek vizuális reprezentációja a 36. ábrán látható). A log-elemző modulban implementált algoritmusok segítségével tesztelhetővé váltak a fúziós eljárások.



36. ábra Az intenzitás és a hangforrás forrás távolságának összefüggése

Intenzitás alapján súlyozott leíró fúzionálás

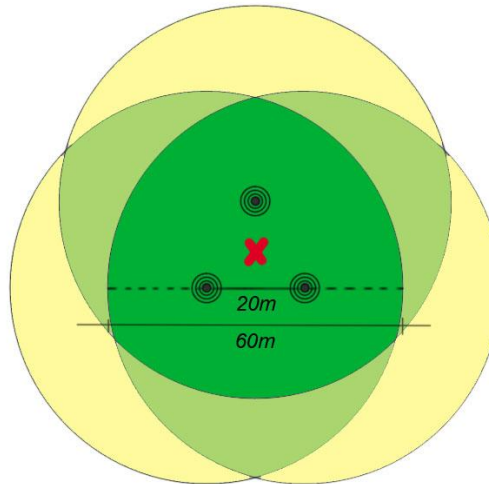


37. ábra Az autokorrelációs módszer hatékonysága a távolság és az erősítési beállítások függvényében

Az 37. ábra diagramján jól látható a hatótávolsággal kapcsolatos tesztek eredménye, itt kizárólag a valós pozitív jelzéseket vettem figyelembe (a cél a detektálhatóság mérése, nem a találati arány megállapítása). Ezen vizsgálatok során a különböző távolságokból az akusztikus eszközzel rögzített motoros fűrész hangokat a már említett előzetes fúziót alkalmazó autokorrelációs egyéni döntéshozatali eljárással értékeltem ki. A grafikonon jól látható, hogy a legnagyobb erősítéssel készített független hangminták még 40 méteres távolságból is nagy biztonsággal felismerhetők.

A távolsággal kapcsolatos tesztekkel igazolt, hogy az akusztikus eszközök felhasználásával, sík terepen, terepakadályok nélkül negyven méteres távolságból még jól detektálható a láncfűrész hangja G(96,96) beállítást használva. Ez alapján ideális esetben a 38. ábra rajzának megfelelően elhelyezett három szenzoregység erdei környezetben is képes 1500-2000 m² terület hatékony megfigyelésére. Az ábrának megfelelő elhelyezést a fúzió lehetőségének fenntartása indokolja, egyéni döntéshozatal esetén az eszközök távolabb is tehetők egymástól. Fontos megjegyezni, hogy ezek a területadatok nagy biztonsági határral számolt becslések, ennél valószínűleg sokkal nagyobb területet képesek lefedni az ábrának megfelelően elhelyezett eszközök.

-  Akusztikus szenzor
-  Bázis álmás
-  3 szenzor által fedett terület
-  2 szenzor által fedett terület
-  1 szenzor által fedett terület



38. ábra Három autokorrelációt alkalmazó eszköz lehetséges elhelyezése a leírófúzió optimalizálásához

A fals pozitív jelzések elkerülése érdekében egyszerű kritériumként megadható, hogy amennyiben csak egy szenzor jelzi a belső égésű motor jelenlétét, nem tekintjük érdemi jelzésnek, viszont későbbi heurisztikus módszerek használatára tároljuk az adatokat. A feltétel alkalmazása következtében a fúziós algoritmus csak abban az esetben fut le, ha legalább két szenzor egység pozitív jelzést adott. Amennyiben csak két eszköztől érkezik jelzés, az csak kölcsönös egyetértés esetén tekinthető pozitívnak. Három jelzés esetén a rendszer több információval rendelkezik a megfelelő döntés meghozatalára.

Ugyan az adatelemző rendszernek nincs semmilyen információja az egyes eszközök megbízhatóságáról, mégis képes sorrendet állítani annak tekintetében, hogy egy adott időben készített felvétel kiértékelése melyik szenzornál várhatóan megbízhatóbb. Jelen esetben a sorrendet a hangfelvételek rövid idejű intenzitása alapján állapítja meg az elemző modul.

A jobb jel/zaj viszony miatt, három szenzor által, azonos jelenségről készített szinkronfelvételek közül nagy valószínűséggel az a legjobban osztályozható, amelyiknek a legnagyobb a hangereje. Azonban előfordulhat olyan eset, amikor a legnagyobb intenzitású hangfelvételtől hozott döntés hibás, emiatt nem használható egyszerűen az intenzitási értékek egymáshoz viszonyított arányán alapuló súlyozás. Mindenképpen garantálni kell, hogy egy szenzor hibás döntését kettő másik javítani tudja, ez rögzített súlyozással érhető el. Három szinkronfelvételből nyert leírókat az intenzitás sorrendjétől függően rendre 0,40, 0,35 és 0,25 súllyal átlagolja az elemző, majd az egyéni döntéshez hasonlóan szórásokat is számítja.

Egyszerű többségi szavazat

Teoretikus esetben, ha az eszközök az általuk negatív mintaként értékelt felvételek osztályozási adatait is továbbítják az adatelemző rendszer felé, a három szenzor döntési folyamatát három független osztályozónak tekinthetjük, és így lehetőségünk nyílik felső szintű adatfúzió megvalósítására. Mivel az eset csak elméleti, és a valóságban a rendszer kommunikációs korlátai miatt nem valósítható meg, csak egyszerű többségi szavazat alkalmazását vizsgáltam. Egyszerű többségi szavazatnál értelem szerűen a többség által hozott döntést tekintjük helyesnek. A szimulációs programot nem kötik kommunikációs korlátok, így az elv tesztelése lehetséges.

A vizsgálatokhoz negatív mintaként internetes forrásból származó, különböző időjárási környezetben – derős, szeles, esős, viharos - készített természeti hangokat, míg pozitív mintaként szintén internetes forrásokat, illetve egy 2012-es mérés erdei munkáról készített hangfelvételeit használtam.



39. ábra A 2012 októberi mérésről készült fotók

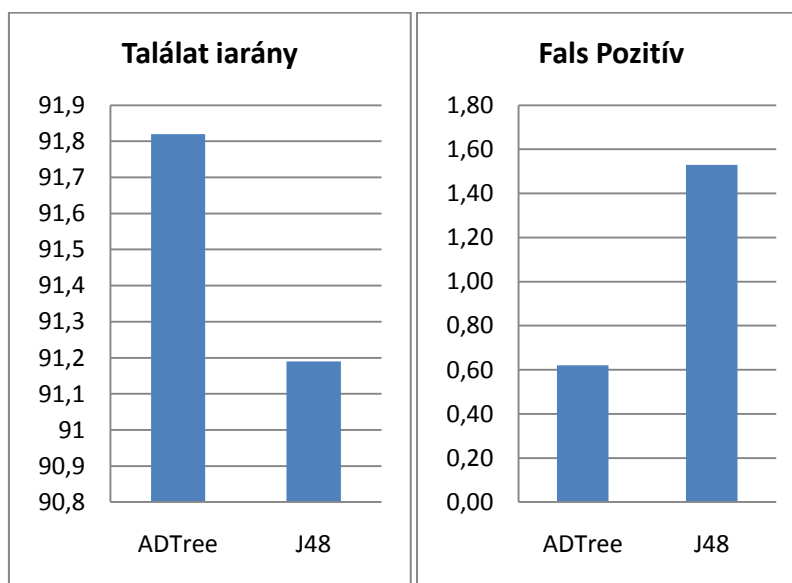
A 39. ábrán a 2012-es mérésről készült felvételek láthatók. Baloldalt fadóntó gép (forwarder), jobboldalt ideiglenes rakodótér.

Az eszköz többségi szavazás felhasználásával nagymértékben csökkentette a fals pozitív jelzések számát. Előfordult, hogy egy-egy madárhang esetén egyik szenzor pozitív jelzést adott, viszont a másik kettő negatívát. A vizsgálatoknál az eszközök egyenként 5-6%-ban küldtek fals pozitív jelzéseket, a többségi szavazást használva ezek száma 1%-ra csökkent.

Kiterjesztett döntési fa

Mivel valós esetben az eszközöktől nem érkezik negatív jelzés - így fals negatív sem -, a jelzések pontosítása csak a fals pozitívok csökkentésével valósulhat meg. Az egyéni döntéshozatalnál mutatott jó teljesítmény alapján evidens volt, hogy érdemes a döntési fákot tesztelni az új leírók felhasználásával. A már említett döntési fák szimulált adatokon, az új,

fúzionált leírók használatával javultak. A legjobb teljesítményt az ADTree és a J48 (C4.5) nyújtotta. Ez a 6.15 ábrán látható.



40. ábra A legjobb eredményeket nyújtó döntési fák találati aránya és a fals pozitívok száma fúziós alkalmazásnál (%)

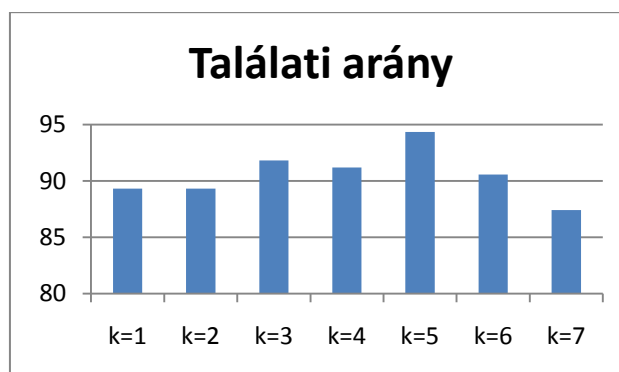
Fontos megjegyezni, a szimulált környezet korrelációját a való világgal nehéz megállapítani, ennek tekintetében az eredmények nem feltétlenül reprezentatívak, valamint manuálisan összeválogatott adatbázis mérete elég csekély pontos eredmények kijelentéséhez, viszont a működési elv jól tervezhető és ellenőrizhető. Amennyiben a szimulált környezet és hangok kis hibával megfelelnek a valóságnak, hasonló eredmények érhetők el a valódi eszközökkel.

Példa alapú osztályozók

Az egyéni döntéshozatalhoz hasonlóan a szinkronizált felvételek új leíróinak felhasználásával, a példa alapú osztályozók a döntési fákhoz hasonló eredményt mutattak. Különböző paraméterezésekkel vizsgáltam kNN, K* és SVM osztályozókat.

kNN:

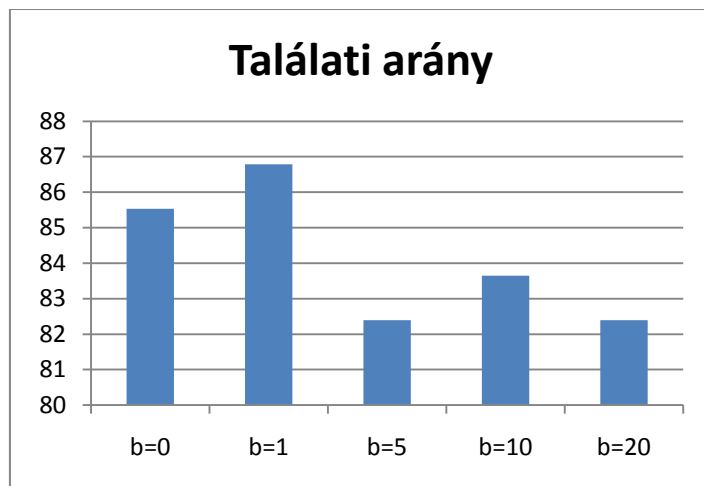
A vizsgálatok során k értéként a kis adathalmazra való tekintettel kis számokat alkalmaztam. A WEKA-ban implementált algoritmus engedélyezi az euklideszi távolság alapú súlyozást, így páros számú szomszédok vizsgálata is lehetséges. A legjobb eredmény 5 szomszéd vizsgálatával született, amely 94,34%-os találati arány mellett 0,56% fals pozitív és mindössze 3,7% fals negatív jelzéssel szolgált.



41. ábra kNN használatával elért találati arányok a vizsgált k paraméterek tekintetében

K* [John G. Cleary 1995]:

A K^* egy példaalapú osztályozó, amely nem hipergeometriai távolságokat használ a közeli szomszédok meghatározásához, hanem entrópia alapú távolságot. Lényegében a távolságokat az alapján számítja az algoritmus, hogy egy-egy leíróból milyen hosszú műveletsorozattal lehetséges egy másik leíró előállítani. Az algoritmusnak egy paramétere van. A b társító paraméter a tanító adatbázis méretének százalékos arányában határozza meg a paraméterter egy gömb alakú klaszterét, amelyben található példákat szomszédnak tekinti.

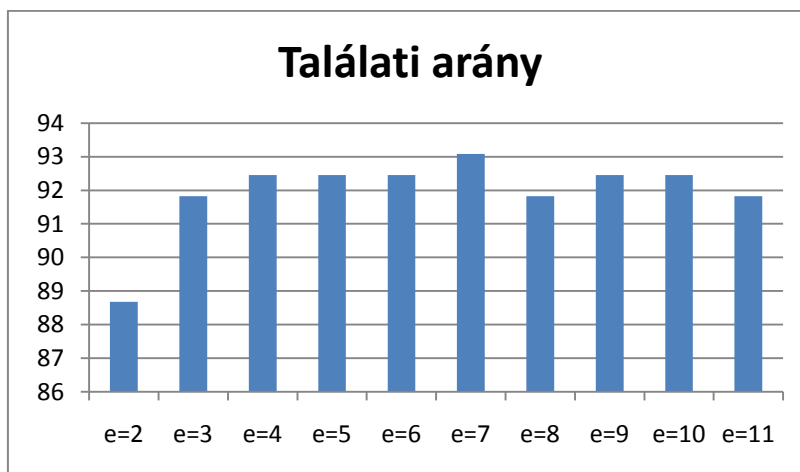


42. ábra A K^* használatával elért találati arányok a vizsgált b paraméterek tekintetében

42. ábrán jól látható, hogy a találati arány tekintetében a K^* teljesítménye rosszabb a kNN-nél még szűk környezet vizsgálatánál is, emiatt a fals pozitívok és fals negatívok tekintetében nem szükséges további vizsgálat.

SVM:

A WEKA-ban implementált SMO [Platt 1998] algoritmust különböző kezdőexponensű normalizált polinom kernelekkel használva vizsgáltam a szimulált adatokat. A kernel exponensének megválasztása nagymértékben befolyásolja az osztályozás teljesítményét. A mindössze hét dimenziós paraméterter miatt elegendő viszonylag kevés számú kernelt tesztelni.



43. ábra SVM használatával elért találati arányok a vizsgált exponens paraméterek tekintetében

A 43. ábrán láthatóan a legmagasabb találati arány 93,08% volt 0,6% fals pozitív és 7% fals negatív jelzéssel.

Végleges jelzés, heurisztikák

Az eddig ismertetett eljárások felhasználásával a fals pozitív jelzések esélye kisebb mint 1 %, ám még így is előfordulhat, hogy a rendszer hibásan erdei munkát jelez egy-egy alkalommal. Annak érdekében, hogy felesleges riasztás ne történjen, érdemes hosszabb időszakok, esetleg nagyobb terület, több szenzorsziget megfigyeléseire alapozni a riasztást. A szimulációs környezetben lehetőségem volt hosszabb hangfelvételeken tesztelni a bemutatott osztályozási eljárásokat, valamint egyszerű heurisztikákat.

A vizsgálatok során internetes forrásból származó környezeti zajokat tartalmazó hosszú hangmintákat, valamint a már említett mérések hangfelvételeit használtam.

A heurisztikus szabályok szenzorszigetenként a következők voltak:

- Egy eszköztől érkező jelzést nem tekintünk pozitívnak, viszont tároljuk.
- Legalább két eszköztől egyszerre érkező jelzést pozitívnak tekintünk, amennyiben az előző egy percben legalább tíz pozitív jelzés érkezett, különben tároljuk.
- Három eszköztől egyszerre érkező jelzést pozitívnak tekintjük, amennyiben az előző egy percben legalább öt jelzés érkezett.

Az eredmények a következők:

- Tizenöt perces esős időjárású hangmintán nem érkezett jelzés a szimulált eszközöktől. Az eszközök 1182 mérést végeztek a hangfájlt felhasználva, és nem jelezték motoros fűrész zajának jelenlétét.
- Egy perces derűs időben madárhangokkal készült felvételt vizsgálva 70 mérésből egyszer sem jelezték motorhang jelenlétét.
- Közel kétperces út menti erdei felvétel 130 mérése alapján nem érkezett pozitív jelzés az eszközöktől.
- Három perces erdei munkáról készített felvételt használva az eszközök 206 mérés alapján 26 alkalommal jelezték belsőégésű motorok hangját.

7 Fejlesztési lehetőségek

Amint láthattuk, az eszközök egyéni döntései arra elegendőek, hogy egy emberi operátor felé jelezzük, hogy hallható-e belső égésű motor a vizsgált területen. A végső cél azonban az, hogy automatikusan jelezzük az erdei munkafolyamatokat. Ennek érdekében hosszabb időszakokat kell vizsgálnunk, majd ezen időszakról kell valamilyen megállapítást tennünk.

Ilyen esetekben az első gondolat általában egy heurisztikus modell felépítése vagy valamilyen létező matematikai modell a rendszerre való ráerőltetése. Mi is a létező megoldások módosításán gondolkodtunk. Az egyik ötlet Markov, illetve rejtett Markov modellek használata volt. Nézzük meg, mik is ezek.

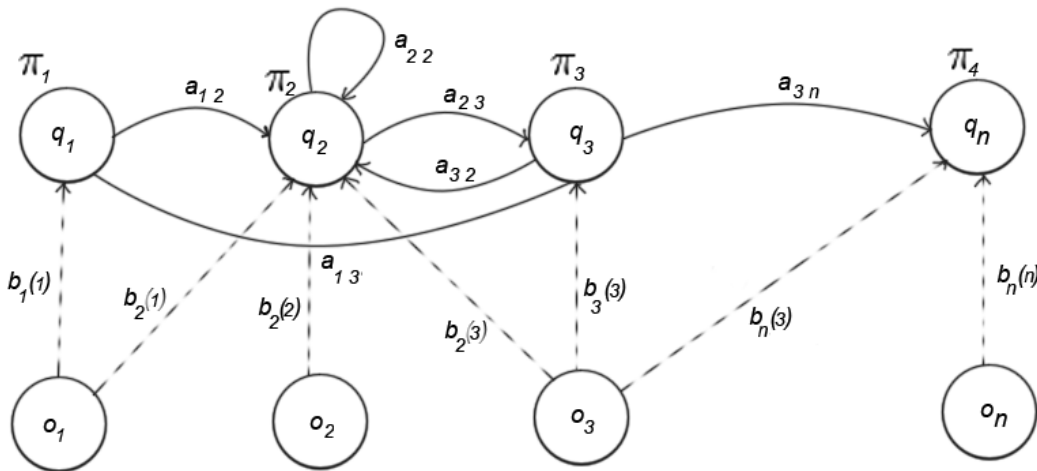
Markov modell

A diszkrét Markov folyamatok felfoghatók egy N állapotú rendszerként, amelyben minden egyes időpillanatban állapot váltás következik be. Ez az állapot váltás nem jelenti szükségszerűen, hogy a rendszer új állapotba kerül (adott állapotból a rendszer „válthat” az aktuális állapotba is).

Minden egyes állapotátmenethez tartozik egy-egy átmeneti valószínűség, amely meghatározza, milyen valószínűséggel vált a rendszer egyik állapotból a másikba. (A)

Minden egyes állapothoz tartozik egy kezdeti valószínűség, amely azt jelöli, mekkora valószínűséggel van a rendszer adott állapotban a t_0 időpillanatban. (π)

Amennyiben az állapotok közvetlenül nem megfigyelhetők, az állapotok valószínűségi függvényeit (B: kibocsátási valószínűség) felhasználva egy olyan sztochasztikus folyamatot kapunk, amely magába ágyazva tartalmazza az eredeti folyamatot, viszont közvetetten, valamilyen kibocsátásuk megfigyelhető (O).



44. ábra Egy rejtett Markov modell (Hidden Markov Model – HMM) sematikus ábrája

Az ábrán $\forall q \in S$ rejtett állapot, melyek közti átmeneti valószínűségeket $a_{ij} \in A$ jelöl. π_i jelöli a kezdeti valószínűséget, $b_i(j)$ jelöli az i -edik állapot j -edik megfigyeléshez tartozó kibocsátási valószínűségét, valamint $\forall o_i \in O$ jelöli az O megfigyelés-sorozat i -edik megfigyelését. (Az ábra nem teljes, az állapotok száma és a megfigyelés-sorozat hossza nem

feltétlenül azonos, viszont minden állapothoz tartozik minden megfigyelésre vonatkozó kibocsátási valószínűség.)

Jelen esetben több probléma is van a Markov modellek használatával.

Milyen állapotai vannak a rendszernek?

Az erdei munkafolyamatokról a módszereinkkel két féle információt kaphatunk, adott pillanatban hallható vagy nem hallható belső égésű motor hangja. Magától értetődő lenne, hogy ez a két állapot szerepel a rendszerünkben.

Milyen kezdeti valószínűségeket alkalmazzunk?

A kitelepített rendszernél helyszíntől, évszaktól, időjárási körülményektől függően különbözőek lehetnek a kezdeti valószínűségek. Könnyen belátható, hogy vannak olyan területek, ahol rendszeres erdei munka zajlik, illetve vannak olyan magánterületek, ahol alapvetően soha nem hallható belső égésű motor. Az erdei munkálatokat többnyire ősszel kezdik, majd a faáru szállítását az első fagy után kezdik meg.

Milyen átmeneti valószínűségeket határozzunk meg?

Amennyiben kizárólag egy erdei munkafolyamatot figyelünk meg, és hallható a motoros fűrész hangja, nagy valószínűséggel a következő időpillanatban szintén hallható lesz. Viszont amennyiben egy erdő általános akusztikai környezetét vizsgáljuk, amennyiben nem hallható belsőégésű motor hangja, rendkívül kicsi az esély arra, hogy a következő időpillanatban hallható lesz.

Szimpla Markov modellt nem alkalmazhatunk, hiszen a hangfelismerő rendszerünk nem tökéletes, emiatt nem tudjuk, hogy adott időpillanatban milyen állapotban van a rendszer, viszont valószínűségi értékeket tudunk adni. Tehát megfigyeléseink vannak, ebből következik, hogy rejtett Markov modellt kellene alkalmaznunk.

Ezek után felmerül a kérdés, hogy milyen megfigyeléseket generál a rendszer? Használjuk-e a módszerek által kinyert leírókat megfigyelésként, vagy a számított valószínűségi értékeket?

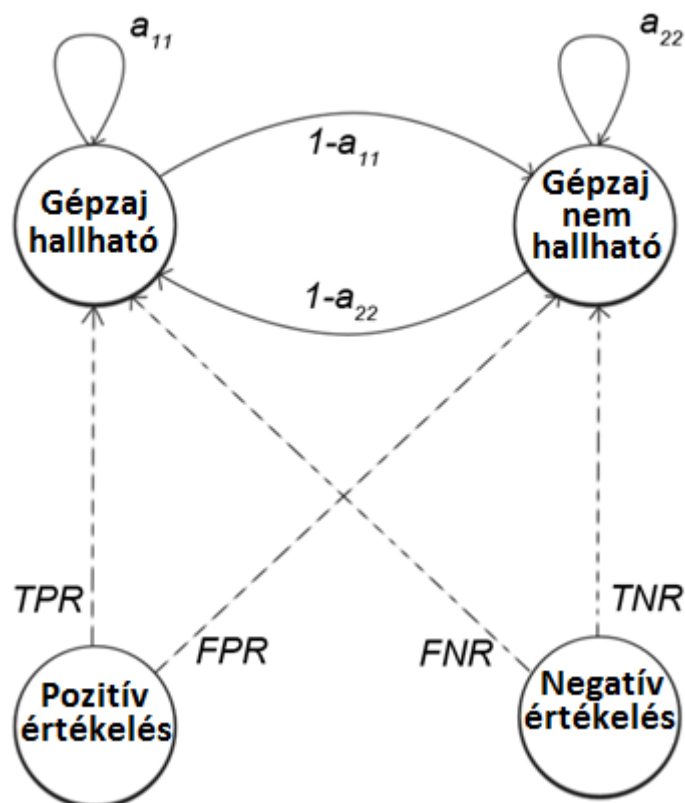
Amennyiben minden kérdésre tudunk megfelelően válaszolni még mindig akadnak problémák a modellek használatával.

Hányad rendű Markov láncról van szó?

Az elsőrendű Markov lánc esetén minden állapot pontosan annyira függ az előző állapottól, mint az azelőtti összes többitől. Ez esetünkben nem feltétlenül igaz, viszont amennyiben nem használunk szigorú feltételeket, és kijelentjük, hogy elsőrendű Markov láncról van szó, még mindig megmarad az alapvető problémánk, hogy az erdei munkafolyamatok nem strukturáltak, nem jól definiálhatók a módszereink által képzett leírókkal. Nem tudjuk meghatározni, hogy motoros fűrészszel milyen hosszú ideig vágnak, egy autó hangja meddig hallható, mennyi idő telik el egy újabb fűrészhang vagy gépjármű hallatáig, ezek mind jelentősen változó időtartamok.

Mindezek ellenére a rejtett Markov modellek bizonyos paramétereit felhasználhatjuk egy heurisztikus osztályozó rendszer felépítéséhez.

Tegyük fel, hogy létezik olyan rejtett Markov modell, amely a fent írt feltételek nagy részét teljesíti. Egy ilyen rendszer felépítése látható 45. ábra diagramján.



45. ábra Egy lehetséges HMM ábrája

Az ábrán jól látható, hogy a rendszerünk két állapotához kétféle megfigyelés tartozik, amikor pozitív, illetve negatív hangot jelez az eszköz. Az ezekhez tartozó kibocsátási valószínűségek könnyen megállapíthatók a valós és fals pozitívok illetve negatívok arányának meghatározásával.

A kibocsátási és átmeneti valószínűségeket felhasználva módosíthatjuk az eszközök egyéni döntésére vonatkozó valószínűségi értékeket. Míg a kibocsátási valószínűségek bizonytalansági mértéket rendelnek a megfigyeléseinkhez, az átmeneti valószínűségek az előző állapot ismeretében plusz támpontot adnak az aktuális mérésre vonatkozó valószínűség megállapításában.

Az említett megoldás elméletben javíthatja az eszközök egyéni döntését, mivel plusz információt használunk fel korábbi tapasztalatok alapján, azonban ez még mindig nem oldja meg a munkafolyamatok automatikus jelzését. Ennek érdekében több megoldás is lehetséges, ezek közül kettőt emelnék ki.

Feltételezzük, hogy a pontosított egyéni döntések alapján tökéletesen felismerhető a motorok hangja, így előállítható adott hosszúságú időintervallumra egy bináris sorozat. A képzett bitsorozatot használva leíróként vizsgálhatjuk a sorozatban található pozitív és negatív jelzések arányát, ez alapján átfogóbb képet alkotva a megfigyelt időszakról.

Egy másik lehetőség, hogy a módosított valószínűségi értékekből álló sorozat összegét tekintjük leírónak. Olyan időszakok vizsgálatával, amelyeknél nem fordulnak elő gépi zajok, mindkét esetben számítható egy küszöb érték, amely felett már jeleznünk kell az anomáliát. Egyszerűen megfogalmazva megbecsülhető a fals jelzésekben adódó maximális érték.

Mindent egybevéve, lehetséges, hogy léteznek olyan matematikai modellek, amelyeket felhasználva jó eredményeket érhetünk el, ezek felderítése azonban további, jelentős kutató munkát igényel. A Markov modellek jelen esetben nem alkalmazhatók, viszont a javasolt heurisztikus rendszer rövid időn belül implementálható.

8. Összegzés

Néhány mondatban kiemelném a projektben végzett saját munkám, illetve fejlesztéseim. A projektben végzett munkám során áttekintettem a hazai és nemzetközi szakirodalmat, ennek ellenére sajnos nem találtam a számunkra megfelelő módszert, ezért új eljárások kidolgozására volt szükség.

Megismertem a fejlesztés során alkalmazott hardver és szoftvereszközöket, ismert eljárásokat teszteltem, adatgyűjtést végeztem tesztkörnyezet kialakításához.

Tapasztalatot szereztem a digitális jelfeldolgozásban, matematikai és statisztikai programcsomagok használatában, hangkódolási eljárásokban.

Lehetőséget kaptam saját módszerek, algoritmusok fejlesztésére, szimulációs környezetet készítettem a leírófúzió teszteléséhez. Különböző előfeldolgozási, klaszterezési, leíró kivonási és osztályozási algoritmusokat fejlesztettem.

Az új módszereket valós és virtuális környezetben is teszteltem, ellenőriztem a fejlesztett eljárások hatékonyságát a távolság függvényében, különböző időjárási, illetve földrajzi körülmények között.

Nagy mélységben vizsgáltam az autokorreláció és a TESPARE leírók adta lehetőségeket, kísérletekkel igazoltam, hogy a hang komplexitási leíró használatával és anélkül közel azonos eredmények érhetők el saját felhasználási esetünkben.

Egy valós erdővédelmi rendszerhez készítettem gyakorlati alkalmazásra szánt szoftvert.

Célunk egy létező rendszerbe illeszkedő új modul fejlesztése volt. Tekintsük át, milyen lehetőségeink voltak, és mit sikerült elérnünk.

A 3. fejezetben három lehetséges módszert ismertettem, melyek korlátait a célhardver specifikációi alkották. Az említett módszereknek különböző előnyei és hátrányai léteznek, ezek láthatók a következő táblázatban. Bár a Fourier transzformáción alapuló módszereket nem implementáltunk, a fellelhető irodalmi adatok alapján általánosan kijelenthető néhány előnyük és hátrányuk.

4. táblázatA vizsgált eljárások összehasonlítása

	Spektrális tér, FFT	Autokorrelációs eljárás	TESPAR
Előnyök:	• Rengeteg leíró kinyerését teszi lehetővé • Kiváló eredmények érhetők el • Hatótávolsága jóval nagyobb is lehet, mint a két másik módszernek	• Néhány leíró használatával megfelelő becslés a hang alapfrekvenciájára • Viszonylag nagy hatótávolság • Szabályozható előerősítés • Gyors osztályozás	• Valós idejű leírókivonás • Pontos osztályozás időjárástól függetlenül • Könnyen bővíthető eljárás • Számítási komplexitás • $O(n)$ • Alacsony memóriaigény
Hátrányok:	• Számítási komplexitás • Bár az FFT $O(n \log n)$ futási idő elérését teszi lehetővé, a duplapontos tört számok használata miatt eszközünkön túl lassú • Jelentős memóriaigény	• Számítási komplexitás • $O(n^2)$, viszont alul-mintavételezés használatával kevés művelet kell elvégezni • Viszonylag lassú leírókivonás • Időjárási zajokra való érzékenység	• Rövid hatótávolság • Az osztályozáshoz használt távolságok megállapításához gyökvonás használata • Rögzített előerősítés

A fejlesztés során felhasznált szoftverek, programnyelvek:

- Weka
- Scilab
- Eclipse
- GnuPlot
- Notepad++
- ANSI C
- C++
- Java SE

Az implementáció TinyOS környezetben nesC programnyelv felhasználásával történt.

Köszönetnyilvánítás

Köszönöm témavezetőm Dr. Czúni László bizalmát és útmutatásait, Dr. Simon Gyula a hangfeldolgozással kapcsolatos tanácsait, Lipovits Ágnes statisztikai programcsomagokkal kapcsolatos útmutatásait, Vakulya Gergely, Bíró András és Salamon Gábor az eszközök programozásának megismerésében nyújtott segítségét, valamint Kiss Péter József a méréseknél nélkülözhetetlen segítségét.

A dolgozatban a bemutatott kutatás a *GOP-1.1.1.-11-2011-0070 Szenzorhálózat alapú erdővédelmi felügyeleti rendszer* elnevezésű és a *TÁMOP-4.2.2.C-11/1/KONV-2012-0004* számú projekt nélkül nem valósulhatott volna meg.

Hivatkozások

Angela Schörgendorfer, Wilfried Elmenreich. „Extended Confidence-Weighted Averaging in Sensor Fusion.” *Proceedings of the Junior Scientist Conference 2006*, page, 2006: 67-68.

Babiš, Ďuriček, Harvanová, Vojtko. „Forest Guardian - monitoring System for Detecting Logging Activities Based on Sound Recognition.” *Student Research Conference on Informatics and Information technologies*. Bratislava, 2011. 354-359.

Baljeet Malhotra, Ioanis Nikolaidis, and Janelle Harms. „A Simple Vehicle Classification Framework for Wireless Audio-Sensor Networks.” *Journal of Telecommunications and Information Technology on Wireless Ad-Hoc Networks*, 2008: 43-50.

Bo Lei, Zhijie Zhang Song Yue, Chensheng Wang. „Forest smoke detection in video images based on constant speed rotating platform.” *Proc. SPIE 8908, International Symposium on Photoelectronic Detection and Imaging 2013: Imaging Sensors and Applications 89082K (August 21, 2013)*, 2013: doi:10.1117/12.2035026.

Breiman, Leo. „Random Forests.” *Machine Learning, Volume 45 Issue 1*, 2001: 5-32.

Chih-Yi Chiu, Bountouridis, Ju-Chiang Wang, Hsin-Min Wang. „Background music identification through content filtering and min-hash matching.” *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, 2010: 2414-2417.

D. Aha, D. Kibler. „Instance-based learning algorithms.” *Machine Learning, Volume 6 Issue 1*, 1991: 37-66.

Freund, Mason. „The alternating decision tree learning algorithm.” *Proceeding of the Sixteenth International Conference on Machine Learning, Bled, Slovenia*, 1999: 124-133.

Friedlos. „RFID in the Forest.” *RFID Journal*. 2009. nov 2. www.rfidjournal.com/article/view/?5350 [hozzáférés dátuma: 2014. nov 8].

G. Csurka, C. Dance, L.X. Fan, J. Willamowski, C. Bray. „Visual categorization with bags of keypoints.” *Proc. of ECCV International Workshop on Statistical Learning in Computer Vision*, 2004: 1-22.

G. Guo, M. Zhou. „Using MODIS Land Surface Temperature to Evaluate Forest Fire Risk of Northeast China.” *IEEE Geoscience and Remote Sensing Letters*, vol. 1, no. 2, 2004: 98-100.

GFW. *Global Forest Watch*. 1997. <http://www.globalforestwatch.org/> [hozzáférés dátuma: 2014. nov 08].

Gjerdrum. „RFID tags applied for tracing timber in the forest products chain.” *Proceedings of the 5th meeting of the Nordic-Baltic Network in Wood Material Science and Engineering (WSE)*, 2009: 109-114.

Gonzales, T.F. „Clustering to minimize the maximum intercluster distance.” *Theoretical Computer Science*, 1985: 38(2-3): 293-306.

Joachims, T. „Text categorization with support vector machines: Learning with many relevant features.” *Lecture Notes in Computer Science Volume 1398 [ECML]*, 1998: 137-142.

John G. Cleary, Leonard E. Trigg. „K*: An Instance-based Learner Using an Entropic Distance Measure.” *12th International Conference on Machine Learning*, 1995: 108-114.

Khoirul Anwar, Minoru Okada. „Low Complexity Fourier Transforms using Multiple Square Waves, Fourier Transforms - Approach to Scientific Principles.” *Fourier Transforms - Approach to Scientific Principles, Prof. Goran Nikolic (Ed.)*, ISBN: 978-953-307-231-9, 2011: doi:10.5772/16089 .

Kyzirakos, Karpathiotakis, Garbis, Nikolaou, Bereta, Papoutsis, Kontoes. „Wildfire monitoring using satellite images, ontologies and linked geospatial data.” *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 24, 2014: 18–26.

L. Couvreur, M. Laniray. „Automatic Noise Recognition in Urban Environments Based on Artificial Neural Networks and Hidden Markov Models.” *The 33rd International Congress and Exposition on Noise Control Engineering*, 2004.

L. Trevis, Dr. N. El-Sheimy. „The Development of a Real Time Forest Fire Monitoring and Management System.” *ISPRS XXXV conference proceedings*, 2004: 65-70.

L. Zhu, A. Rao, A. Zhang. „Theory of Keyblock-based image retrieval.” *ACM Transactions on Information Systems*, 20, 2002: 224-257.

Leo Breiman, Jerome H. Friedman, Richard A. Olshen, Charles J. Stone Wadsworth International Group, Belmont, California. *Classification and Regression Trees*. London, UK: Chapman and Hall, 1984.

Lie Lu, Hao Jiang, and HongJiang Zhang. „A robust audio classification and segmentation method.” *Proceedings of the ninth ACM international conference on Multimedia*, 2001: 203-211.

Lie Lu, Hong-Jiang Zhang, Stan Li. „Content-based audio classification and segmentation by using support vector machines.” *Multimedia Systems, Volume 8, Issue 6*, 2003: 482-492.

Lloret, Garcia, Bri, Sendra. „A wireless sensor network deployment for rural and forest fire detection and verification.” *Sensors 9 (11)*, 2009: 8722-8747.

Marco F. Duarte, Yu Hen Hu. „Vehicle Classification in Distributed Sensor Networks.” *Journal of Parallel and Distributed Computing* , Volume 64 Issue 7, 2004: 826-838 .

Marius Vasile Ghiurcaua, Corneliu Rusua, Radu Ciprian Bilcub, Jaakko Astola. „Audio based solutions for detecting intruders in wild areas.” *Signal Processing, Volume 92, Issue 3*, 2012: 829–840.

Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, Ian H. Witten. „The WEKA Data Mining Software: An Update.” *SIGKDD Explorations, Volume 11, Issue 1*, 2009: 10-18 .

Michael Büchler, Silvia Allegro, Stefan Launer, Norbert Dillier. „Sound classification in hearing aids inspired by auditory scene analysis.” *EURASIP Journal on Applied Signal Processing*, 2015: 2991-3002.

Papán, Jozef, Matús Jurecka, and Jana Púchyová. „WSN for Forest Monitoring to Prevent Illegal Logging.” *Proceedings of the Federated Conference on Computer Science and Information Systems*, 2012: 809-812.

Platt, J. *Fast Training of Support Vector Machines using Sequential Minimal Optimization*. Cambridge, MA, USA: MIT Press, 1998.

Quinlan, Ross. *C4.5: Programs for Machine Learning*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1993.

R.A. King, Mr T.C. Phipps, Shannon. „TESPAR and approximation strategies.” *Computers & Security, Volume 18, Issue 5*, 1999: 445-453.

Redel-Macías, Gutiérrez, Cubero-Atienza, Hervás-Martínez. „Sound Source Identification in Vehicles a Combined Linear Evolutionary Product Unit Neural Network Model.” *Soft Computing Models in Industrial and Environmental Applications, 6th International Conference*, 2011: 379-386.

Shi, Haijian. *Thesis: Best-first decision tree learning, Master of Science*. Waikato, NZ: The University of Waikato, 2007.

Son, Byungrak, Yong-sork Her, and J. Kim. „A design and implementation of forest-fires surveillance system based on wireless sensor networks for South Korea mountains.” *International Journal of Computer Science and Network Security, Vol 6.*, 2006: 124-130.

Sotris A Pavlopoulos, Antonis CH Statis, Euripides N Loukis. „A decision tree – based method for the differential diagnosis of Aortic Stenosis from Mitral Regurgitation using heart sounds.” *BioMedical Engineering Online*, 2004: doi:10.1186/1475-925X-3-21.

Taichi Nakashima, Kazunori Komatani, Satoshi Sato. „Integration of Multiple Sound Source Localization Results for Speaker Identification in Multiparty Dialogue System.” *Natural Interaction with Robots, Knowbots and Smartphones*, 2014: 153-165.

Tamar Peli, Mon Young, Robert Knox, Ken Ellis, Fredrick Bennett. „Feature Level Sensor Fusion.” *DTIC*. 1999. www.dtic.mil [hozzáférés dátuma: 2014. 11 08].

Unicomp. *Unicomp*. 2014. www.unicomp.hu [hozzáférés dátuma: 2014. 11 08].

Webb, Geoff. „Decision Tree Grafting From the All-Tests-But-One Partition.” San Francisco, CA, 1999.

WEKA. <http://www.cs.waikato.ac.nz/ml/weka/> [hozzáférés dátuma: 2014. 11 08].

Ziyou Xiong, Rodrigo Caballero, Hongcheng Wang, Alan M. Finn, Muhidin A Ledic, Pei-Yuan Pend. „Video-based Smoke Detection: Possibilities, Techniques, and Challenges.” *IFPA, Fire Suppression and Detection Research and Applications - A Technical Working Conference(SUPDET)*, 2007.