ارتقاء اسپکتروگرام پویا با استفاده از روش نمایش در باند پایه و نمایش رنگ معکوس

عباس نايبي
دانشگاه صنعتي شريف
info@nayebi.com

چکیده: در مصورسازي پوياي صدا انتقال هر چه بيشتر جزئيات اساسي و تشكيل دهنده صدا در زمان واقعي اهميت ويژه‌اي دارد. در كاربردهايي نظير كمك به افراد ناشنوا اهميت اين موضوع بسيار بيشتر است. اسپكتروگرام پويا و كورلوگرام دو روش عمده مصورسازي پويا هستند، اسپكتروگرام پويا در نمايش فاز و جزئيات ساختاري صدا بر كورلوگرام برتري دارد. در اين مقاله دو روش براي ارتقاء اسپكتروگرام پويا مطرح مي‌شود. روش نمايش در باند پايه كه كار دنبال كردن فرمنتها را راحت‌تر مي‌كند و روش نمايش رنگ معكوس كه كار تشخيص جزئيات لحظه‌اي را راحت‌تر مي‌كند. يك سري تست تشخيصي و يك تست جامع براي بررسي اثر اين روشها نيز انجام شده است كه نتايج آن آورده شده است.

واژه هاي كليدي: اسپكتروگرام پويا[1]، پردازش صحبت، مدل شنوايي انسان،مصورسازي صدا ، نمايش باند پايه

۱- مقدمه

مصورسازي صوت از ديرباز مورد توجه بوده است. شايد بتوان خط و نت نگاري موسيقي را از اولين انواع مصورسازي صدا دانست. گونة ديگري از مصورسازي صوت استفاده از نمايشهاي زمان-فركانسي است. يك نوع قديمي اين نوع نمايش نمودار اسپكتروگرام است [13]. در نمايشهاي زمان-فركانسي سعي مي‌شود تا تغييرات طيف سيگنال در بستر زمان نمايش داده شود. روشهاي زمان-فركانسي را بطور كلي ميتوان به دو دسته **پويا** و **ايستا** تقسيم كرد. روشهاي پويا روشهاي جديدتري هستند و زمان را بصورت واقعي (و نه انتزاعي) نمايش مي‌دهند همچنين سعي در بالا بردن سرعت درك توسط چشم انسان دارند. به همين جهت روشهاي پويا از انيميشن براي نمايش تغييرات زمان-فركانسي استفاده مي‌كنند. در روشهاي ايستا نمايش بصورت يك نمودار با دو محور افقي و عمودي است.

۲- سيستم شنوايي انسان و مدل آن

صدا چيزي است كه ما **مي‌شنويم**. بين مفهوم صدا و مفهوم شنيدن رابطه‌اي تنگاتنگ وجود دارد. بنابراين بهتر است چيزهايي را مصور كنيم كه در شنيدن صدا اهميت دارد. يك راه براي تحقق اين ايده آن است كه خود سيستم شنوايي انسان را مدل كنيم.

1 Dynamic spectrogram

۲-۱- مدل گوش خارجي و مياني

گوش خارجي شبيه يك فيلتر ميانگذر عمل مي‌كند كه فركانس تشديد آن حدود ۳ كيلو هرتز است. لذا ميتوان رفتار گوش خارجي و مياني را بصورت يك فيلتر بالا گذر با تابع سيستم زير در نظر گرفت :

$$H(z) = 1 + \alpha z^{-1} \qquad (۱)$$

مقدار α=-0.6 بصورت تجربي براي يك حالت متعادل انتخاب شده است.

۲-۲-مدل گوش داخلي

مي‌توان رفتار مكانيكي غشاء قاعده‌اي را با يك بانك فيلتر مدل كرد. پهناي باند هر فيلتر و فركانس مركزي آن براساس موقعيت فيلتر در غشاء قاعده‌اي مشخص مي‌شود [2,10] :

$$f_c = A(10^{ax} - 1) \qquad (۲)$$

در اين رابطه f_c فركانس مشخصه برحسب هرتز است و x فاصلة نرمال شده نقطة مورد نظر روي غشاء قاعده‌اي از اپكس است (۰‹ x ‹۱ ، x=۰ معادل اپكس و x=۱ معادل پايه حلزوني). A وa دو ثابت هستند كه در [10] براي حلزوني گوش انسان به ترتيب ۱٦۵/٤ و ۱/۲ پيشنهاد شده‌اند. پهناي باند مربعي معادل[1] هر فيلتر گوش براي فركانس مركزي f_c برابر است با

$$ERB = 6.23 f_c^2 + 93.39 f_c + 28.52 \qquad (۳)$$

در اين رابطه f_c فركانس مركزي فيلتر برحسب كيلو هرتز است و ERB برحسب هرتز بدست مي‌آيد. ميتوان هر فيلتر ميانگذر را با يك پنجره هنينگ[2] كه با يك سيگنال كسينوسي مدوله شده است مدل كرد [32].

$$h(n) = h_b(n)\cos\left(2\pi \frac{f_c}{f_s}(n - \frac{N-1}{2})\right) \qquad (۴)$$

در اين رابطه $h_b(n)$ پاسخ ضربه فيلتر پايين‌گذر است. f_c فركانس مركزي فيلتر است و f_s نرخ نمونه‌برداري است.

1 Equivalent Rectangular Bandwidth
2 Hanning window

۲-۳- مدل تبديلات مكانيكي به الكتريكي

رفتار سلولهاي مويي داخلي با يكسوسازي خروجي فيلترها شبيه سازي مي‌شود. پديدة ديگري در تبديل ارتعاشات مكانيكي به الكتريكي وجود دارد. در فركانسهاي پايين (زير ۱ كيلو هرتز) دقت فركانسي بالا است و هارمونيك‌هاي فركانس پايه صداها كاملاً از يكديگر جدا مي‌شوند ولي در مورد فيلترهاي فركانسهاي بالاتر پهناي باند فيلترها زياد مي‌شود لذا ممكن است چند هارمونيك فركانس اصلي در خروجي يك فيلتر وجود داشته باشند لذا هيستوگرامي كه از بررسي فعاليتهاي عصبي در زمان حاصل مي‌شود، معادل پوش خروجي اين فيلترها است.براي مدل كردن اين خاصيت پس از انجام عمل يكسوسازي يك فيلتر پايين‌گذر با فركانس قطع ۸۰۰ هرتز را روي سيگنال اعمال مي‌كنيم [32].

۳- اسپكتروگرام پويا

اسپكتروگرام پويا [32] يك نمايش زمان-فركانسي است كه فاز خروجي فيلترها را حفظ مي‌كند و شكل موج آنها را به هم نمي‌زند. تحقيقات مختلف نشان داده است كه حفظ اين ساختار در شنوايي صداهاي مختلف مؤثر و مهم است [25,24,21] . براي توليد اسپكتروگرام پويا ابتدا يك اسپكتروگرام يا كوكلياگرام (اسپكتروگرام با مدل حلزوني گوش) ايجاد مي‌شود. سپس يک فريم ۲۰ ميلي ثانيه اي از آن به مدت ۲۰ ميلي ثانيه نمايش داده مي شود. سپس يک فريم ۲۰ ميلي ثانيه‌اي ديگر از آن به نمايش در مي‌آيد. فريم دوم با لغزاندن يک پنجره ۲۰ ميلي‌ثانيه‌اي روي يک بازه ۴۰ ميلي‌ثانيه‌اي از كوكلياگرام طوري انتخاب مي‌شود كه حداكثر تشابه را با فريم قبلی داشته باشد.اين كار باعث ثبات بيشتر انيميشن خروجي برای چشم می‌شود. در [32] جزئيات بيشتر آمده است.

۴- ارتقاء اسپكتروگرام پويا

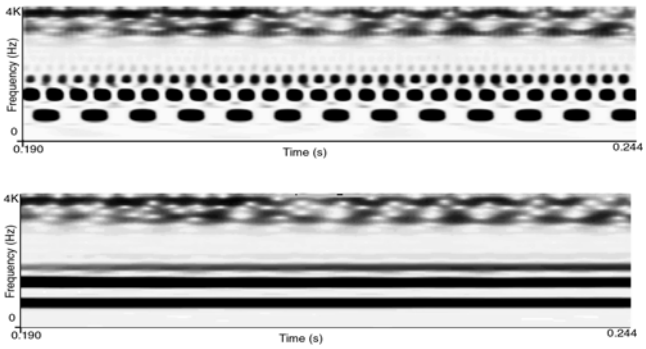
در اينجا پيشنهادهاي جديدي در مورد اسپكتروگرام پويا ارائه مي‌شود. اولين ايده استفاده از *نمايش رنگ معكوس* است. . در نمايشهاي زمان-فركانسي موجود هر جا كه مقدار نمودار در يک نقطه بيشتر شود رنگ مربوط به آن نقطه تيره‌تر مي‌شود. در

واقع نمودار به صورت سياه در سفيد نمايش داده مي‌شود.در حاليكه چشم ما نقاط پر انرژي را روشن و نقاط كم انرژي را به صورت تاريک مي‌بيند لذا اگر نمودار بصورت سفيد در سياه نمايش داده شود به نظر مي‌آيد سرعت درك بالاتر باشد. آزمايشهاي انجام شده نيز اين موضوع را تأييد مي‌كنند.

ايده ديگر روش *نمايش در باند پايه* است. همانطوركه ذكر شد براي مدلسازي گوش، سيگنال ورودي به يک بانک فيلتر داده مي‌شود و خروجي فيلترها يكسو مي‌شود و از يك فيلتر پايين‌گذر عبور داده مي‌شود. پيشنهادي كه در اينجا مطرح مي‌شود آن است كه اين فيلتر پايين گذر حذف شود در عوض يک پروسس ديگر روي سيگنال انجام شود به اين صورت كه خروجي فيلتري كه فركانس مركزي آن f_c است را به اندازة f_c- در محور فركانس شيفت مي‌دهيم تا به باند پايه منتقل شود.

همانطور كه گفته شد پهناي باند فيلتر هاي فركانس بالا زياد است و چندين هارمونيک در خروجي يک فيلتر مي‌آيد و همين امر باعث مي‌شود كه خروجي اين فيلترها شبيه يک سيگنال مدوله شده با فركانس f_c (فركانس مركزي فيلتر) باشد. اعمال يک فيلتر پايين گذر، پوش سيگنال خروجي فيلتر را به ما ميدهد و شيفت فركانسي به اندازه f_c- نيز فركانس پايه را حذف مي‌كند و تقريبا همان اثر را دارد. اما در فركانس‌هاي پايين وضعيت متفاوت است. پهناي باند اين فيلترها كم است و يک هارمونيک بيشتر در آن قرار نمي‌گيرد لذا خروجي اين فيلترها سينوسي با فركانس پايين است و عبور دادن آنها از فيلتر پايين گذر تاثير خاصي روي شكل موج خروجي ندارد ولي شيفت فركانسي اثري كاملا متفاوت دارد و يک موج سينوسي با فركانس f_c را به يک مقدار ثابت (يک خط ممتد) تبديل مي‌كند(شكل ۱). مزيت استفاده از اين روش آن است كه تصوير بسيار نرمتري حاصل مي‌شود كه كار دنبال كردن فرمنت‌هاي فركانس پايين را راحتتر كند.

۵- آزمايشها



شكل ۱: مقايسه روش عادي و باند پايه. بالا) نمايش عادي اسپكتروگرام پايين) همان نمايش به روش باند پايه.

به طور كلي آزمايشات به دو روش انجام شده‌اند. در روش اول سعي شده است تا چند نوع آزمايش مختلف صورت گيرد كه هر يك ويژگي خاص خود را دارد. در اين روش تعداد نمونه‌ها بين ۲ تا ۵ نمونه انتخاب شده است و دو روش مصورسازي صوت در اين روش مقايسه شده‌اند. در روش دوم يك آزمايش جامعتر روي ۱۰ نمونه انجام شده است در اين روش چندين شيوة مصورسازي صوت مقايسه شده‌اند.

۵-۱-نتايج آزمايشهاي تشخيصي

دو روش مصور سازي مورد آزمايش قرار گرفته‌اند. روش اول: اسپكتروگرام پوياي عادي و روش دوم: اسپكتروگرام پويا با نمايش رنگ معكوس و نمايش در باند پايه است. اين آزمايش‌ها روي ٦ نفر انجام شده‌اند. هر شخص ابتدا يك تست از روش اول انجام داده است سپس يك تست از روش دوم. اين كار براي هر سري از نمونه ها سه بار تكرار شده است. امتياز ها با لحاظ كردن نمره منفي براي پاسخ هاي غلط در نظر گرفته شده است به نحوي كه دادن جوابهاي كاملا تصادفي نمره صفر را خواهد داشت. نتايج حاصله در جداول ۱ تا ۴ آمده است.براي انجام اين آزمايش‌ها يک نرم‌افزار خاص توليد شده است كه آزمايش را در سه فاز انجام مي‌دهد:در فاز اول (فاز آموزش) ۱۰ تكرار داريم و در هر تكرار هركلمه ۲ بار پخش مي‌شود و انيميشن مربوط به آن نيز همزمان به اجرا در مي‌آيد. در فاز دوم ۱۰ بار به طور تصادفي يک كلمه انتخاب مي‌شود و انيميشن آن اجرا مي‌شود و كاربر بايد تشخيص دهد كه مربوط به كدام كلمه است. پس از هر پاسخ، جواب درست به كاربر

اعلام مي‌شود. فاز سوم دقيقا به همين صورت است با اين تفاوت كه جواب درست به كاربر اعلام نمي‌شود.

جدول ١: نتايج تست تشخيص كلمات «يك» ، «دو» ، «سه» ، «چهار» و «پنج»

	تست روش اول				تست روش دوم			
شماره شخص	امتياز ١ (درصد)	امتياز ٢ (درصد)	امتياز ٣ (درصد)	**ميانگين (درصد)**	امتياز ١ (درصد)	امتياز ٢ (درصد)	امتياز ٣ (درصد)	**ميانگين (درصد)**
١	٨٧	١٠٠	١٠٠	**٩٦**	٨٧	١٠٠	١٠٠	**٩٦**
٢	٦٢	٧٥	١٠٠	**٧٩**	٧٥	٨٧	١٠٠	**٨٧**
٣	٥٠	٦٢	٧٥	**٦٢**	٧٥	١٠٠	١٠٠	**٩٢**
٤	٥٠	٥٠	٧٥	**٨٥**	٣٧	٥٠	٨٧	**٥٨**
٥	٣٧	٣٧	٧٥	**٥٠**	٦٢	٨٧	٨٧	**٧٩**
٦	٦٢	٧٥	٨٧	**٧٥**	٦٢	٧٥	٨٧	**٧٥**

جدول ٢: نتايج تست تشخيص كلمات «Bed» و «Bad»

	تست روش اول				تست روش دوم			
شماره شخص	امتياز ١ (درصد)	امتياز ٢ (درصد)	امتياز ٣ (درصد)	**ميانگين (درصد)**	امتياز ١ (درصد)	امتياز ٢ (درصد)	امتياز ٣ (درصد)	**ميانگين (درصد)**
١	٨٠	٨٠	١٠٠	**٨٧**	١٠٠	١٠٠	١٠٠	**١٠٠**
٢	٢٠	٢٠	٤٠	**٢٧**	٨٠	١٠٠	١٠٠	**٩٣**
٣	٦٠	٦٠	٨٠	**٦٧**	٨٠	٨٠	١٠٠	**٨٧**
٤	٢٠	٤٠	٤٠	**٣٣**	٤٠	٦٠	٨٠	**٦٠**
٥	٢٠	٢٠	٤٠	**٢٧**	٨٠	١٠٠	١٠٠	**٩٣**
٦	٨٠	٨٠	٦٠	**٧٣**	١٠٠	١٠٠	١٠٠	**١٠٠**

جدول ٣: نتايج تست تشخيص كلمات «باد»، «پاد» و «داد»

	تست روش اول				تست روش دوم			
شماره شخص	امتياز ١ (درصد)	امتياز ٢ (درصد)	امتياز ٣ (درصد)	**ميانگين (درصد)**	امتياز ١ (درصد)	امتياز ٢ (درصد)	امتياز ٣ (درصد)	**ميانگين (درصد)**
١	٧٠	٨٥	٨٥	**٨٠**	٧٠	٨٥	٨٥	**٨٠**
٢	٥٥	٧٠	٨٥	**٧٠**	٥٥	٧٠	٧٠	**٦٥**
٣	٤٠	٧٠	٧٠	**٦٠**	٤٠	٧٠	٨٥	**٦٥**
٤	٤٠	٥٥	٧٠	**٥٥**	٥٥	٥٥	٧٠	**٦٠**
٥	٢٥	٥٥	٥٥	**٤٥**	٥٥	٢٥	٥٥	**٤٥**
٦	٤٠	٧٠	٧٠	**٦٠**	٥٥	٨٥	١٠٠	**٨٠**

جدول ٤: نتايج تست تشخيص كلمات «مدال» و «مثال»

	تست روش اول				تست روش دوم			
شماره شخص	امتياز ١ (درصد)	امتياز ٢ (درصد)	امتياز ٣ (درصد)	**ميانگين (درصد)**	امتياز ١ (درصد)	امتياز ٢ (درصد)	امتياز ٣ (درصد)	**ميانگين (درصد)**
١	٧٠	٩٠	٩٠	**٨٣**	٩٠	١٠٠	١٠٠	**٩٧**
٢	١٠	٤٠	٦٠	**٣٧**	٤٠	٧٠	٩٠	**٦٧**
٣	٤٠	٤٠	٦٠	**٤٧**	٩٠	٩٠	١٠٠	**٩٣**
٤	٤٠	٣٠	٤٠	**٣٧**	٧٠	٥٠	٨٠	**٦٧**
٥	٣٠	٣٠	٥٠	**٣٧**	٥٠	٧٠	٨٠	**٦٧**
٦	٧٠	٧٠	٩٠	**٧٧**	٩٠	١٠٠	١٠٠	**٩٧**

جدول ٥: متدهاي مختلف مصورسازي آزمايش شده در آزمايش جامع

شماره	استفاده از روش باند پايه	نمايش رنگ معكوس	فاكتور كاهش سرعت اجراي انيميشن	انتخاب مجموعه تست از مجموعه تمرين
١	X		١	X
٢	X		١	
٣			١	X
٤			١	
٥	X	X	١	
٦	X		٦	
٧	X	X	٦	
٨	X		١	

٥-٢- نتايج آزمايش جامع

در اينجا تعداد نمونه‌هاي انتخاب شده ١٠ نمونه است كه به ترتيب عبارتند از: «ميز»، «تيز»، «تيغ»، «نيم»، «بيم»، «خيز»، «خَز»، «نَم»، «بم»و «كم». در اين آزمايش سعي شده است تا متدهاي مختلفي مورد آزمايش قرار گيرند (جدول ٥). اين متدها به همان ترتيبي كه در جدول ٥ آمده‌اند تست شده‌اند.

در [19] نتايج آزمايشات انجام شده به صورت ماتريس ابهام آمده است. در جدول ٦ ميانگين نتايج حاصل از آزمايشات آمده است. با مطالعه اين جداول و اشكال موارد زير مشخص مي‌شوند:

تفاوت چنداني بين حالت اجراي عادي و حالت اجراي آهسته انيميشن وجود ندارد و اين نشان مي‌دهد كه سرعت اجراي

انيميشن در حالت عادي آن‌قدر بالا نيست كه نتوان اطلاعات لازم را دريافت كرد.

بين حالتي كه مجموعه تست از مجموعه تمرين جداست و حالتي كه مجموعه تست از مجموعه تمرين انتخاب مي‌شود تفاوت چنداني وجود ندارد و اين نشان مي‌دهد تست قابل اعتماد است.

تفاوت بين ميانگين آزمايشهاي ٢ و ٣ همچنين تفاوت بين آزمايشهاي ٥ و ٦ كم يا منفي است و اين به خاطر آن است كه شخص تست شونده با يك روش جديد آشنا مي‌شود.

تفاوت بين ميانگين آزمايشهاي ٤ و ٥ مقدار قابل ملاحظه‌اي است در حاليكه در اينجا نيز شخص تست شونده با يك روش جديد آشنا شده است و اين نشان مي‌دهد كه متد ٥ يعني بكارگيري توأم روش نمايش رنگ معكوس و نمايش در باند پايه متد موفقي بوده است.

با ملاحظه جداول و نمودارهاي ماتريس ابهام مشخص مي شود كه بيشترين پراكندگي ها مربوط به تشخيص كلمات «خيز» و «تيز» است. اين مورد در تمام نمودارها كم و بيش ديده مي‌شود و همانطور كه ملاحظه مي‌شود به طور تدريجي با افزايش مهارت شخص تست شونده درصد اين اشتباه كاهش مي‌يابد.

با ملاحظه ماتريس ابهام مشخص مي‌شود كه متدهاي ٣ و ٤ كه اسپكتروگرام عادي هستند در تشخيص كلمات «بيم» ، «نيم» و «تيغ» ضعيفند.

كاهش سرعت در متدهاي ٦ و ٧ باعث شده است تا تشخيص كلمه «كم» راحتتر انجام شود و اين به‌خاطر لحظه‌اي بودن تلفظ «ك» است. البته به نظر مي‌رسد با افزايش مهارت افراد از متدهاي ديگر نيز بتوان همين نتايج را گرفت.

برخلاف كلمه «كم»، كاهش سرعت اجراي انيميشن در متدهاي ٦ و ٧ باعث شده است تا تشخيص كلمات «بم» و «نم» مشكلتر شود و اين به‌خاطر آن است كه در اين حالت تشخيص سرعت تغيير فرمنت‌ها و مقايسه آن با حالت واقعي تلفظ اين دو كلمه مشكلتر است.

با مطالعه اشتباهات مربوط به تشخيص كلمه «خَز» مشخص مي‌شود كه روش نمايش رنگ معكوس كه در متدهاي ٥ و ٧ استفاده شده است، موفقيت بيشتري داشته است و اين به‌خاطر آن است كه تشخيص اشكال پيچيده ايجاد شده در فركانسهاي بالا در اين روش راحتتر انجام مي‌شود و اين اشكال هم در حرف «خ» و هم در حرف «ز» وجود دارند.

متد ٥ در تشخيص كلمات «ميز» و «تيز» از يكديگر از همه متدها موفق‌تر بوده است و اين به خاطر استفاده از روش نمايش رنگ معكوس و روش باند پايه است. علت آنكه متد ٧ به اين اندازه موفق نبوده است كاهش سرعت اجرا در اين روش و غير واقعي بودن آن است كه همانطور كه ذكر شد كار تشخيص سرعت تغيير فرمنت ها را مشكل مي‌كند.

٦- نتيجه‌گيري و كارآينده

نتايج آزمايشات انجام گرفته نشان داد كه روش نمايش رنگ معكوس در نمايش بعضي خصوصيات مؤثرتر است و نقاط كم انرژي را بهتر نمايش مي‌دهد. همچنين روش نمايش باند پايه كار دنبال كردن فورمنت‌ها در فركانس‌هاي پايين را ساده مي‌كند. حقيقت آن است كه مصورسازي صوت به شكل جديد آن يعني نمايشهاي زمان-فركانسي پويا موضوعي بسيار جوان است و تا رسيدن به اهداف نهايي آن راه درازي باقي است. يكي از اهداف مصورسازي صوت آن است كه چشم دقيقاً جاي گوش را در يك فرد ناشنوا پر كند. براي ادامة كار در اين مرحله ايده‌هاي زير مطرح مي‌شود:

-مدل كامل گوش

-استفاده از رنگ براي ايجاد ثبات بيشتر در اثرات لحظه‌اي

-آموزش و آزمايش مصورسازي صدا با يك برنامه ٦ ماهه

جدول ٦: ميانگين امتيازهاي آزمايشهاي انجام شده

شماره متد	١	٢	٣	٤	٥	٦	٧	٨
ميانگين امتياز	٣٤/٧	٤٨/٤	٤٨/٩	٥٠/٧	٥٧/٤	٥٦/٨	٦١/٢	٦٣/٣

روي يک گروه ۵۰ تا ۱۰۰ نفري

-استفاده از نمونه‌هاي موسيقي به جاي گفتار با هدف انتقال ريتم وملودي قابل درك براي چشم

۷- مراجع

[1] J. B. Allen. Cochlear modelling. IEEE-ASSP Mag., pp. 3-29, January 1985.

[2] G. von Békésy. Experiments in hearing. McGraw-Hill, New York, 1960.

[3] Martine Cooke, Steeve Beet, Malcom Crawford. Visual Representation of Speech Signal. Wiley Inc., 1993.

[4] J. R. Deller, J. G. Proakis, J. H. L. Hansen. Discrete-Time Processing of Speech Signals. Maccmillan, 1993.

[5] P.B. Dense and E.N. Pinson. The speech chain, the physiology and biology of spoken language. W. H. Freeman and Company, New York, 1993.

[6] R. M. Fano. Short-Time Autocorrelation Fuctions and Power Spectra. J. Acoustical Soc. Amer., vol. 22, pp. 546-550, Sept. 1950.

[7] O. Ghitza. Auditory models and human performance in tasks related to speech coding and speech regcognition. IEEE Trans. on Speech and Auditory processing, 2(1, part II):115-132, Jan 1994.

[8] Ben Gold, Nelson Morgan. Speech and Audio Signal Processing , Processing and Perception of Speech and Music. John Wiley and Sons, Inc., 2000.

[9] M. H. Goldstein. Auditory periphery as speech signal processor. IEEE Engineering in Medicine and Biology, 13(2):186-196, 1994.

[10] D. Greenwood. A cochlear frequency-position function for several species-29 years later. J. Acoust. Soc. Amer., 87(6):2592-2605, 1990.

[11] T. Irino. ‘gammachrip’ function as an optimal auditory filter with the mellin transform. In ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing – Proceedings, vol. 2, pp. 981-984. NTT Basic Research Lab, Kanagawa, Jpn, IEEE, Piscataway, NJ, USA, 1996.

[12] S. M. Khanna and D. G. B. Leonard. Basilar membrane tuning in the cat cochlea. Science, 215:305-306, 1982.

[13] W. Koening, H. K. Dunn, L. Y. Lacey. The Sound Spectrograph. J. Acoustical Soc. Amer., vol. 18, pp. 19-49, 1946.

[14] R. F. Lyon. Automatic gain control in cochlear mechanics. in: The Mechanics and Biophysics of Hearing (ed. P. Dallos, C. D. Geisler, J. W. Matthews, M. A. Ruggero & C. R. Steele). Springer-Verlag, New York, 1990.

[15] B. C. J. Moore. Introduction to the Psychology of Hearing. Academic Press, 2nd edition, 1989.

[16] B. C. J. Moore and B. R. Glasberg. Auditory filter shapes derived in simultaneous and forward making. J. Acoust. Soc. Amer., 70:1003-1014, 1981.

[17] B. C. J. Moore and B. R. Glasberg. Suggested formula for calculating auditory filter bandwidth and excitation patterns. J. Acoust. Soc. Amer., 74:750-753, 1983.

[18] B. C. J. Moore and B. R. Glasberg. Formulae describing frequency selectivity as a function of frequency and level, and their use in calculating excitation patterns. Hearing Res., 28:209-225, 1987.

[19] A. Nayebi, http://www.Nayebi.com/sv/test1_result.pdf

[20] A. V. Oppenheim. Speech Spectrograms Using the Fast Fourier Transform. IEEE Spectrum, vol. 7, pp. 57-62, August 1970.

[21] R. D. Patterson. A pulse ribbon model of monaural phase perception. J. Acoust. Soc. Amer., 67:229-245, 1980.

[22] R. D. Patterson , I. Mimmo-Smith, D. L. Weber, and R. Milory. The deterioration of hearing with age: Frequency selectivity, critical ratio, the audiogram and speech threshold. J. Acoust. Soc. Amer., 72:1788-1803, 1982.

[23] J. O. Pickles. An introduction to the physiology of hearing. Academic Press, London, second edition, 1998.

[24] R. Plomp. Pitch of complex tones. J. Acoust. Soc. Amer., 41(6):1527-1533, 1967.

[25] R. Plomp and H. J. M. Steeneken. Effect of phase on the timbre of complex tones. J. Acoust. Soc. Amer., 46:409-421, 1969.

[26] Ralph K. Potter, George A. Kopp, and Harriet Green Kopp. Visible Speech., Dover Publications, Inc., New York, 1996.

[27] J. G. Proakis, D.G. Manolakis. Digital Signal Processing Principles, Algorithms and Applications. second edition, McMillan, 1992.

[28] W. S. Rhode. Observations of the vibration of the membrane in squirrel monkeys using the mössbauer technique. J. Acoust. Soc. Amer., 49:1218-1231, 1971.

[29] M. A. Ruggero. Response to sound of the basilar membrane of mammalian cochlea. Opinion Neurobiol, 2:449-456, 1992.

[30] M. R. Schroeder and B. S. Atal, Generalized Short-Time Power Spectra and Autocorrelation Functions. J. Acoustical Soc. Amer., vol. 34, pp. 1679-1683, Nov. 1962.

[31] A. A. Soltani Farani. A Performance Comparison of Four Parametric Representations of the Spectral Envelope of Speech. Mphil/Phd Transfer Report, university of Surrey, UK, 1995.

[32] A. A. Soltani Farani. Sound Visualisation As an Aid For The Deaf, A New Approach. Ph.d. thesis, university of Surrey, UK, 1998.