

دستگاه شنوایی انسان

پردازش صوت محدوده‌های گوناگونی را در بر می‌گیرد که همه به منظور ارائه صدا به شنوندگان انسانی ابداع شده‌اند. سه محدوده تکثیر موسیقی با کیفیتی به خوبی اصل همانند آنچه در سی‌دی‌های صوتی وجود دارد، ارتباط صوتی از راه دور که نام دیگر شبکه تلفنی است و، ترکیب صحبت که در آن کامپیوترها الگوهای صوتی انسان را تولید کرده یا تشخیص می‌دهند از دیگر قلمروهای دانش پردازش صوت مهم‌ترند. با وجود این که اهداف و مسائل این کاربردها متفاوتند همگی در یک نقطه مشترک به هم می‌رسند و آن گوش انسان است.

گوش انسان یک عضو به گونه‌ای فزاینده پیچیده است. قضیه وقتی پیچیده‌تر می‌شود که اطلاعات ارسالی از دو گوش در یک شبکه پیچیده گنج‌کننده که همانا مغز انسان باشد با هم ترکیب می‌شوند. به یاد داشته باشیم که بیان فوق یک گذر کلی بر قضیه است و تعداد زیادی از پدیده‌ها و آثار دقیق مرتبط با گوش انسان هنوز به درستی درک نشده‌اند.

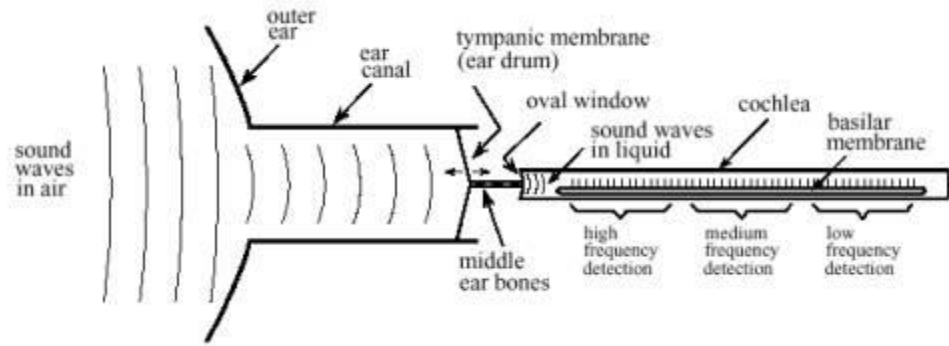
شکل ۱ قسمت اعظم ساختارها و پردازشهایی را که گوش انسان را در بر دارند به تصویر می‌کشد. گوش خارجی از دو بخش تشکیل شده است: نرمی پوست قابل مشاهده و غضروف متصل به کنار سر و کانال گوش که لوله‌ایست به قطر تقریبی ۰,۵ سانتیمتر و تا حدود ۳ سانتیمتر در داخل سر فرو می‌رود. این ساختارها صداهای محیط را به بخشهای حساس گوش میانی و گوش داخلی که در درون استخوانهای مجمله محافظت می‌شود راهبری می‌کنند. در انتهای کانال گوش یک ورقه نازک از نسوج که پرده صماخ یا طبل گوش نامیده می‌شود کشیده شده است. امواج صدا با برخورد به پرده صماخ باعث لرزش آن می‌شوند. گوش میانی مجموعه‌ای از استخوانهای کوچک است که لرزش مزبور را به حلزون گوش (گوش داخلی) انتقال می‌دهند و در آنجا این لرزشها تبدیل به ضربه‌های عصبی می‌گردند. حلزون گوش یک لوله پر از مایع است که به زحمت قطر آن به ۲ میلیمتر و طول آن به ۳ سانتیمتر می‌رسد. اگر چه حلزون گوش در شکل شماره ۱ به صورت یک لوله مستقیم نشان داده شده اما در واقع به دور خودش همانند صدف حلزون پیچ خورده است و وجه تسمیه آن که ریشه در کلمه‌ای یونانی به معنای حلزون دارد نیز این واقعیت است. وقتی یک موج صوتی سعی دارد از هوا وارد مایع شود تنها کسر کوچکی از آن از بین دو محیط عبور می‌کند و باقیمانده انرژی آن بازتابیده می‌شود. دلیل این امر مقاومت مکانیکی پایین هوا (ناشی از پایین بودن میزان فشار صوتی و سرعت بالایی ذرات هوا که به نوبه خود از چگالی پایین و تراکم‌پذیری بالایی آنها نشأت می‌گیرد) در برابر مقاومت مکانیکی بالایی مایع است. به عبارت ساده‌تر دلیل این امر مشابه دلیل این

موضوع است که برای ایجاد موج با دست در درون آب به تلاش بیشتری به نسبت انجام این کار در هوا نیازمندیم. تفاوت موجود باعث بازتابش قسمت اعظم صوت در مرز هوا/مایع می‌گردد.

گوش میانی یک شبکه تطبیق مقاومت است که کسر انرژی صوتی وارد شده به مایع گوش داخلی را زیاد می‌کند. برای نمونه ماهی پرده صماخ یا گوش میانی ندارد چرا که نیازی به شنیدن در هوا ندارد. تغییر شدت، بیشتر ناشی از تفاوت مساحت پرده صماخ (که صدا را از هوا دریافت می‌کند) و دریچه بیضوی) که مطابق شکل ۱ صدا را به داخل مایع انتقال می‌دهد) می‌باشد. مساحت پرده صماخ حدوداً ۶۰ میلی‌متر مربع است حال آن که دریچه بیضوی حدوداً ۴ میلی‌متر مربع مساحت دارد. از آنجا که فشار برابر است با نسبت نیرو به مساحت، این تفاوت مساحت فشار موج صدا را حدوداً ۱۵ برابر افزایش می‌دهد.

در داخل حلزون گوش پرده اصلی قرار دارد که ساختاری را برای ۱۲۰۰۰ سلول حسی که شکل‌دهنده عصب حلزونی است ایجاد می‌کند. پرده اصلی در نزدیکی دریچه بیضوی بسیار سفت است و در انتهای دیگر انعطاف‌پذیرتر است که این امر به این عضو کمک می‌کند تا به عنوان تحلیگر طیف فرکانسی عمل کند. وقتی پرده اصلی در معرض یک سیگنال با فرکانس بالا قرار می‌گیرد در قسمت سفت‌تر طنین می‌اندازد که سبب تحریک سلولهای عصبی نزدیک به دریچه بیضوی می‌گردد. به همین ترتیب فرکانسهای پایین موجب تحریک انتهای دورتر پرده اصلی می‌شوند. این امر موجب پاسخگویی رشته‌های خاص عصب حلزونی در برابر فرکانسهای خاص می‌گردد. این سازوکار **اصل مکان** نامیده می‌شود و در سراسر مسیر به سمت مغز حفظ می‌شود.

طرح کدگذاری اطلاعات دیگری نیز در شنوایی انسان به کار می‌رود که **اصل رگبار** نامیده می‌شود. سلولهای عصبی اطلاعات را با تولید پالسهای الکتریکی کوچکی که پتانسیل کنش نامیده می‌شوند انتقال می‌دهد. یک سلول عصبی واقع بر پرده پایینی می‌تواند اطلاعات صوتی را با تولید یک پتانسیل کنش در پاسخ هر سیکل لرزش کدگذاری کند. برای نمونه یک موج صدای ۲۰۰ هرتزی می‌تواند توسط یک نورون ایجاد کننده ۲۰۰ پتانسیل کنش در ثانیه نشان داده شود. در هر صورت این روش تنها در فرکانسهای زیر حدوداً ۵۰۰ هرتز – بالاترین سرعت ممکن تولید پتانسیل کنش در نورونها – به کار می‌آید. گوش انسان برای غلبه بر این مشکل به نورونها اجازه می‌دهد که برای انجام این کار دسته‌جمعی عمل کنند. برای نمونه یک صدای ۳۰۰۰ هرتزی می‌تواند توسط ده سلول عصبی که هر کدام ۳۰۰ ضربه در ثانیه علامت می‌دهند نشان داده شود. این پدیده بازه کارایی اصل رگبار را تا ۴ کیلوهرتز گسترش می‌دهد که بالاتر از بازه عملیاتی اصل مکان می‌باشد.



شکل شماره ۱- توضیحات مربوط به شکل: نمودار کارکردی گوش انسان. گوش خارجی امواج صوتی را از محیط می‌گیرد و آنها را به سوی پرده صماخ (طبل گوش) که ورقه نازکی از بافت است و هماهنگ با شکل موج هوا می‌لرزد راهبری می‌کند. استخوانهای گوش میانی (استخوانهای چکشی، سندان و رکابی) این لرزشها را به درجه بیضوی که پرده‌ای منعطف واقع در حلزون گوش پر از مایع است انتقال می‌دهند. در داخل حلزون گوش پرده اصلی قرار دارد که ایجاد کننده ساختاری برای ۱۲۰۰۰ سلول عصبی شکل‌دهنده عصب حلزون گوش است. بسته به سفتی متغیر پرده پایینی، هر سلول فقط به بازه کوچکی از فرکانسهای صدا پاسخ می‌دهد که این پدیده گوش را تبدیل به یک تحلیلگر طیف فرکانسی می‌نماید.

شکل شماره ۲ رابطه بین شدت صدا و بلندی مشاهده شده را نشان می‌دهد. غالباً شدت صدا را با یک اندازه لگاریتمی که دسی‌بل اس.پی.ال. (سطح توان صدا) نامیده می‌شود نشان می‌دهند. در این معیار ۰ دسی‌بل اس.پی.ال. موج صدایی با قدرت ده به توان منفی شانزده وات بر سانتیمتر مربع است که حدوداً ضعیف‌ترین صدای قابل تشخیص توسط گوش انسان است. صحبت معمولی حدوداً ۶۰ دسی‌بل اس.پی.ال. است و صدایی با شدت ۱۴۰ دسی‌بل اس.پی.ال. برای گوش دردناک و زیان‌آور است.

	Watts/cm ²	Decibels SPL	Example sound
	10 ⁻²	140 dB	Pain
	10 ⁻³	130 dB	
	10 ⁻⁴	120 dB	Discomfort
	10 ⁻⁵	110 dB	Jack hammers and rock concerts
	10 ⁻⁶	100 dB	
	10 ⁻⁷	90 dB	OSHA limit for industrial noise
	10 ⁻⁸	80 dB	
	10 ⁻⁹	70 dB	
	10 ⁻¹⁰	60 dB	Normal conversation
	10 ⁻¹¹	50 dB	
	10 ⁻¹²	40 dB	Weakest audible at 100 hertz
	10 ⁻¹³	30 dB	
	10 ⁻¹⁴	20 dB	Weakest audible at 10kHz
	10 ⁻¹⁵	10 dB	
	10 ⁻¹⁶	0 dB	Weakest audible at 3 kHz
	10 ⁻¹⁷	-10 dB	
	10 ⁻¹⁸	-20 dB	

شکل شماره ۲ - واحدهای شدت صدا. شدت صدا به صورت توان بر واحد مساحت تعریف می‌شود (مثلاً وات بر سانتیمتر مربع) یا به صورت معمول‌تر با استفاده از یک اندازه‌گیری که دسی‌بل اس‌پی‌ال خوانده می‌شود. همچنان که این جدول نشان می‌دهد قوه شنوایی انسان بیشتر به صداهای بین ۱ کیلوهرتز تا ۴ کیلوهرتز حساس است.

اختلاف بلندترین و ضعیف‌ترین صداهایی که انسان می‌تواند بشنود ۱۲۰ دسی‌بل است که از لحاظ دامنه معادل بازه‌ای حدود یک میلیون است. شنونده تغییر بلندی صدا را وقتی صدا حدود ۱ دسی‌بل (۱۲٪ در دامنه) تغییر کند تشخیص می‌دهد به عبارت دیگر تنها ۱۲۰ سطح بلندی صدا از ملایم‌ترین نجوا تا بلندترین تندر قابل تشخیص است. حساسیت گوش آنقدر جالب توجه است که هنگام شنیدن به ضعیف‌ترین صداها پرده صماخ به اندازه‌ای کمتر از قطر یک ملکول به لرزش درمی‌آید!

احساس بلندی صدا با توان صدا رابطه‌ی توانی با نمایی ۳/۱ دارد. به عنوان نمونه اگر شما توان صدا را ده برابر کنید شنوندگان آن صدا دو برابر شدن بلندی صدا را احساس و گزارش می‌کنند.

این مسأله یک مشکل بزرگ برای حذف صداهای محیطی ناخواسته به وجود می‌آورد. برای نمونه فرض کنید که شما ۹۹٪ دیوار را با عایق صوتی پوشانده‌اید و تنها ۱٪ که مربوط به درها، گوشه‌ها، منافذ و... هستند باقی مانده‌اند. با وجود آن که توان صدا تا اندازه ۱٪ مقدار اولیه آن کاسته شده بلندی صدا تنها به اندازه ۲۰٪ کاهش پیدا کرده‌است.

بازة شنيداري انسان بين ۲۰ هرتز تا ۲۰ كيلوهرتز در نظر گرفته مي‌شود، حال آن كه بيشتر صداهاي قابل حس در بازة ۱ كيلوهرتز تا ۴ كيلوهرتز قرار دارند. براي نمونه شنوندگان مي‌توانند صدائي به ميزان صفر دسي‌بل را در فرکانس ۳ كيلوهرتز بشنوند حال آن كه براي شنيدن يك صدائي ۱۰۰ هرتزي حداقل مقدار آن بايد ۴۰ دسي‌بل باشد. شنوندگان مي‌توانند بگويند كه دو صدا متفاوتند اگر فرکانس آنها بيش از حدود ۰,۳٪ در ۳ كيلوهرتز متفاوت باشد. به عنوان نمونه كليدهاي كنار هم در بيانو به اندازه حدود ۶٪ تفاوت فرکانس دارند.

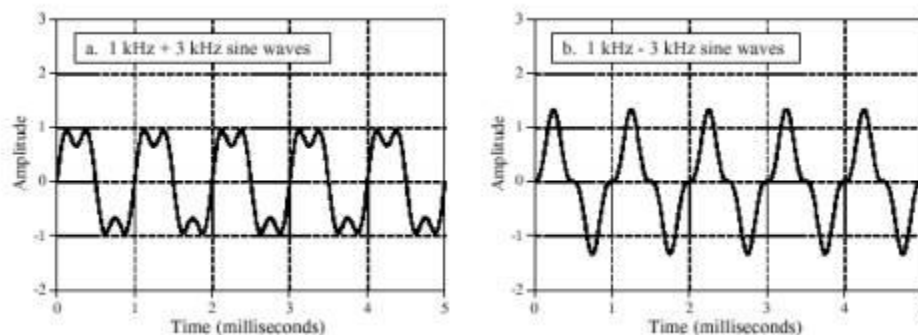
مهم‌ترين مزيت داشتن دو گوش تشخيص جهت صداست. شنوندگان انساني مي‌توانند تفاوت بين دو منبع صدا را كه فاصله‌اي به كمي ۳ درجه دارند (حدوداً برابر با عرض يك انسان در فاصله ده متری) تشخيص دهند. اين اطلاعات جهتي به دو روش جداگانه به دست مي‌آيند. اولاً فرکانسهاي حدوداً بالاي ۱ كيلوهرتز به شدت زير ساپة سر قرار مي‌گيرند. به بيان ديگر گوشي كه به منبع نزديك‌تر است سيگنال قوي تري را به نسبت گوشي كه در جهت مخالف دارد دريافت مي‌كند. روش ديگر تشخيص جهت آن است كه گوش دورتر به خاطر فاصله بيشترش از منبع صدا را كمي ديرتر از گوش نزديك‌تر دريافت مي‌كند. به واسطه اندازه معمول سر (حدوداً ۲۲ سانتيمتر) و سرعت صوت (حدود ۳۴۰ متر در ثانيه) تفاوت‌گذاري زاويه‌اي سه درجه دقت زماني حدود ۳۰ ميكروثانيه نياز دارد. چون اين فاصله زماني نيازمنند اصل رگبار است اين روش جهتيابي براي صداهاي داراي فرکانس كمتر از حدود ۱ كيلوهرتز به كار مي‌رود.

در حالي كه قوه شنوائي انسان مي‌تواند جهت صدا را تشخيص دهد در تشخيص فاصله منبع صدا مشكل دارد. اين امر بدان علت است كه چيزهاي كمي در موج صدا وجود دارد كه اطلاعات اين گونه را در اختيار بگذارد. شنوائي انسان به صورت ضعيفي در مي‌يابد كه منابع صداهاي با فرکانس بالا نزديكند و صداهاي با فرکانس پايين از فاصله دورتر پخش مي‌شوند. اين به آن دليل است كه صداها در فاصله‌هاي دور از ميزان فرکانسشان كاسته مي‌شود. پژواك روش ضعيف ديگري براي تشخيص فاصله است و با استفاده از آن مثلاً مي‌توان ابعاد يك اتاق را حدس زد. براي نمونه صداهاي موجود در يك تالار بزرگ پژواكهايي با وقفة ۱۰۰ ميلي ثانيه دارند، حال آن كه براي يك دفتر كار كوچك اين مقدار ۱۰ ميلي ثانيه است. بعضي از موجودات با استفاده از دستگاه طبيعي تشخيص فاصله صوتي مسأله فاصله‌يابي را حل کرده‌اند. مثلاً خفاشها و دلفينها صداهايي مثل تيك و جيغ توليد مي‌كنند كه از سوي اشياء نزديك بازتابيده مي‌شوند. با اندازه‌گيري ميزان وقفة بازتاب اين صداها اين جانوران مي‌توانند با دقت ۱ سانتيمتر اشياء را

مکانیابی کنند. تجربیات نشان داده‌اند که بعضی انسانها به خصوص نابینایان تا حد کمی از روش مکانیابی با استفاده از پژواک استفاده می‌کنند.

2- ویژگیهای امواج صوتی

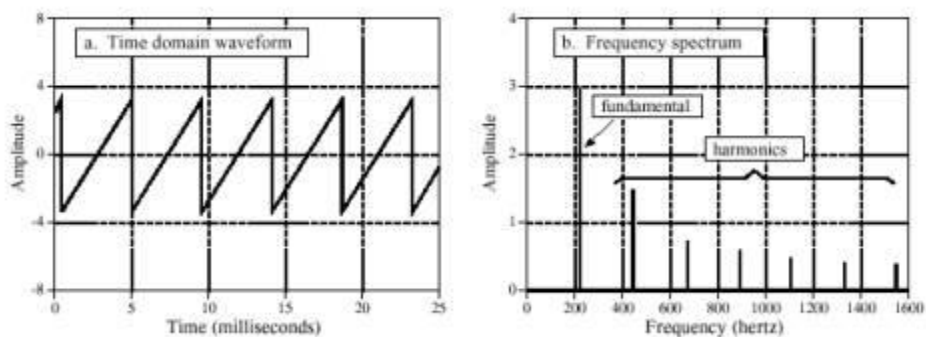
غالباً برای درک یک صوت پیوسته مثل نت یک ابزار موسیقیایی سه بخش مجزا را باید تشخیص داد: بلندی صدا، زیری یا بمی صدا (پیچ) و طنین صدا. بلندی همانگونه که قبلاً توضیح داده شد معیاری برای شدت موج صوتی است. پیچ، فرکانس جزء اصلی صدا – فرکانسی تکرار موج صوتی توسط خودش – می‌باشد. طنین صدا از دو جزء قبلی پیچیده‌تر است و با تعیین محتوای همساز صدا تعیین می‌گردد. شکل شماره ۳ دو موج را که هر دو از جمع یک موج سینوسی یک کیلوهرتزی با دامنه یک و یک موج سینوسی سه کیلوهرتزی با دامنه یک دوم به وجود آمده‌اند نشان می‌دهد. تفاوت آنها در آن است که در شکل b جزء با فرکانس بالاتر ابتدا معکوس شده و سپس با موج دوم جمع شده است. علی‌رغم موجهای در دامنه زمان بسیار متفاوت این دو صوت یکسان به نظر می‌رسند. این به خاطر آن است که شنوایی انسان بر اساس دامنه فرکانسهاست و نسبت به فاز آنها بسیار غیر حساس است. شکل موج صوتی در دامنه زمان فقط به صورت غیر مستقیم با شنوایی رابطه دارد و معمولاً در سیستمهای صوتی در نظر گرفته نمی‌شود. عدم حساسیت گوش به فاز صدا با توجه به روش پخش شدن آن در محیط قابل درک است. فرض کنید که شما در یک اتاق به صحبت‌های فردی گوش می‌دهید. بیشتر صداهایی که گوش شما دریافت می‌کند حاصل بازتاب صدای اصلی از دیوارها، سقف و کف اتاق است. از آنجا که انتشار صدا بستگی به فرکانس آن دارد و میرایی، بازتاب و مقاومت در برابر صدا بر روی آن تأثیرگذار است فرکانسهای متفاوتی از مسیرهای متفاوت به گوش می‌رسد. این به این معنی است که وقتی شما جایی خود را در اتاق عوض می‌کنید فاز هر یک از فرکانسها تغییر می‌کند. چون گوش این تغییر فازها را نادیده می‌انگارد با وجود تغییر مکان شما تغییری در صدای شخص صحبت کننده احساس نمی‌کنید. از دیدگاه فیزیکی فاز یک سیگنال صدا در هنگام پخش در یک محیط پیچیده به صورت تصادفی تغییر می‌کند. از طرف دیگر گوش به فاز صدا غیر حساس است زیرا این جزء دارای اطلاعات قابل استفاده بسیار کمی می‌باشد.



شکل شماره ۳ - تشخیص فاز توسط گوش انسان. گوش انسان نسبت به فاز نسبی سینوسیهای مرکب بسیار غیر حساس است. برای نمونه این دو موج یکسان به نظر خواهند رسید، زیرا دامنه اجزاء آنها یکسان است اگر چه فاز نسبی آنها متفاوت است.

در حالت کلی نمی‌توان گفت که گوش نسبت به فاز کاملاً ناشناخت است. چرا که تغییر فاز می‌تواند باعث تغییر آرایش زمانی یک سیگنال صوتی شود. اما چنین امری یک پدیده نادر است که در محیط‌های شنیداری طبیعی اتفاق نمی‌افتد.

فرض کنید از یک نوازنده ویولون خواسته‌ایم نت‌ها را بنوازد. وقتی که موج صوتی ایجاد شده بر روی اسیلوسکوپ نشان داده شود یک موج دندانه‌اره‌ای مانند شکل شماره ۴ (a) مشاهده می‌شود. شکل شماره ۴ (b) نشان می‌دهد که این صوت چگونه توسط گوش دریافت می‌شود. گوش یک فرکانس اساسی (در مثال شکل ۲۲۰ هرتز) را و همسازهایی را در ۴۴۰، ۶۶۰، ۸۸۰ و... هرتز دریافت می‌کند. اگر این نت بر روی ابزار دیگری نواخته شود هنوز هم همان ۲۲۰ هرتز (همان فرکانس اساسی) را دریافت می‌کند. و از این لحاظ دو صوت مشابهند که گفته می‌شود این دو صوت پیچ یکسانی دارند ولی چون دامنه همسازها متفاوت است دو صوت یکسان نیستند و گفته می‌شود که طنین دو صوت متفاوت است.



شکل شماره ۴ - موج صوتی ویولن. ویولن موج دندانه‌اره‌ای ایجاد می‌کند (شکل a) ، صدای دریافت شده شامل فرکانس اساسی و همسازهای آن است (شکل b)

اغلب گفته می‌شود که طنین صدا از روی شکل موج صوتی تعیین می‌گردد. این مسأله درست است ولی کمی گمراه کننده است. احساس طنین صدا از روی میزان هارمونیک‌های تشخیص داده شده توسط گوش تعیین می‌گردد. در حالی که هارمونیک‌ها از روی شکل موج صوتی تعیین می‌گردد عدم حساسیت گوش به فاز رابطه را بسیار یک طرفه می‌کند. به همین دلیل هر موج صوتی فقط یک طنین دارد حال آن که یک زنگ خاص متعلق به تعداد بی‌نهایتی از موج‌های صوتی است.

گوش بیشتر برای شنیدن هارمونیک‌های اساسی تنظیم شده است. اگر یک شنونده به صدایی که حاصل ترکیب دو موج صوتی سینوسی ۱ کیلوهرتز و ۳ کیلوهرتز است گوش دهد آن را مطلوب و طبیعی توصیف خواهد کرد حال آن که اگر از موج‌های ۱ کیلوهرتز و ۳,۱ کیلوهرتز استفاده شود برای شنونده شکایت برانگیز خواهد بود. این مسأله اساسی برای اندازه‌ها و اختلاف‌های استاندارد ابزارهای موسیقیایی فراهم می‌آورد.

3- روش‌های دیجیتالی ذخیره صدا

در طراحی یک سیستم صوتی دیجیتال دو پرسش وجود دارند که باید پاسخ داده شوند: ۱- چقدر لازم است صوت خوب به نظر برسد؟ ۲- چه نرخ داده‌ای قابل تحمل است؟ جواب به این پرسش‌ها غالباً به یکی از این سه انتخاب منجر می‌شود: اول موسیقی با وفاداری بالا که در آن کیفیت صدا مهم‌ترین چیز است و تقریباً هر نرخ داده‌ای قابل قبول است. دوم ارتباط تلفنی که نیازمند طبیعی به نظر رسیدن صحبت و یک نرخ داده پایین برای کاهش هزینه سیستم است. سوم صحبت فشرده شده که در آن کاهش نرخ داده بسیار مهم است و مقداری غیر طبیعی به نظر رسیدن کیفیت صدا قابل تحمل است. این مورد در بر دارنده ارتباطات نظامی، تلفن‌های سلولی و صحبت ذخیره شده به صورت دیجیتال برای پست الکترونیکی صوتی یا کاربردهای چند رسانه‌ای است.

شکل شماره ۳ بده بستانه‌ای موجود در انتخاب هر یک از این سه روش را نشان می‌دهد.

در حالی که موسیقی نیازمند پهنای باند ۲۰ کیلوهرتز است صحبتی که طبیعی به نظر برسد فقط به پهنای باندهای ۳,۲ کیلوهرتز نیازمند است. در این حال هر چند پهنای باند به اندازه ۱۶٪ مقدار اولیه محدود می‌شود ولی فقط ۲۰٪ اطلاعات اولیه از دست می‌رود.

سیستمهای ارتباطی راه دور اغلب از نرخ نمونه برداری در حدود ۸ کیلوهرتز استفاده می کنند که اجازه انتقال صحبت را با کیفیتی در حد طبیعی می دهد ولی اگر از آن برای انتقال موسیقی استفاده شود تا میزان بالایی از کیفیت آن از دست می رود. شما احتمالاً با تفاوت این دو میزان آشنایی دارید: ایستگاههای رادیویی اف.ام با پهنای باندهای در حدود ۲۰ کیلوهرتز اقدام به پخش می کنند حال آن که ایستگاههای ای.ام محدود به ۳,۲ کیلوهرتز هستند. صحبت و صداهای معمول روی ایستگاههای نوع دوم طبیعی به نظر می رسد حال آن که موسیقی این گونه نیست.

Sound Quality Required	Bandwidth	Sampling rate	Number of bits	Data rate (bits/sec)	Comments
High fidelity music (compact disc)	5 Hz to 20 kHz	44.1 kHz	16 bit	706k	Satisfies even the most picky audiophile. Better than human hearing.
Telephone quality speech	200 Hz to 3.2 kHz	8 kHz	12 bit	96k	Good speech quality, but very poor for music.
(with companding)	200 Hz to 3.2 kHz	8 kHz	8 bit	64k	Nonlinear ADC reduces the data rate by 50%. A very common technique.
Speech encoded by Linear Predictive Coding	200 Hz to 3.2 kHz	8 kHz	12 bit	4k	DSP speech compression technique. Very low data rates, poor voice quality.

شکل شماره ۳ - نرخ داده صوتی در برابر کیفیت صدا. کیفیت صدای یک سیگنال صوتی دیجیتال به نرخ داده آن که برابر با حاصل ضرب نرخ نمونه برداری آن در تعداد بیت های آن در هر نمونه بستگی دارد که به سه بخش تقسیم می شود: موسیقی با وفاداری بالا (۷۰۶ کیلوبیت بر ثانیه)، صحبت با کیفیت تلفن (۶۴ کیلوبیت بر ثانیه) و صحبت فشرده شده (۴ کیلوبیت بر ثانیه)

سیستمهایی که فقط با صدا (و نه موسیقی) سر و کار دارند می توانند مقدار دقت را از ۱۶ بیت به ۱۲ بیت بدون از دست رفتن دقتی قابل توجه کاهش دهند. این میزان می تواند با انتخاب اندازه نامتساوی برای گام مقدارگزینی می تواند به ۸ بیت در هر نمونه نیز کاهش یابد. یک نرخ نمونه برداری ۸ کیلوهرتز با دقت ای.دی.سی ۸ بیت در هر نمونه به نرخ داده ۶۴ کیلوبیت بر ثانیه می انجامد. این یک حد نهایی برای طبیعی به نظر رسیدن صحبت است. دقت کنید که صحبت نیازمند نرخ داده ای معادل ۱۰% نرخ داده موسیقی با وفاداری بالاست.

نرخ داده ۶۴ کیلو بیت بر ثانیه نمایانگر کاربرد نهایی نظریه نمونه برداری و مقدارگزینی برای سیگنالهای صوتی است. روشهای کاهش نرخ داده به اندازه‌های بیشتر از این مبتنی بر فشرده‌سازی جریان داده با حذف تکرارهای ذاتی سیگنال صحبت است. یکی از کاراترین روشهای موجود ال.پی.بی است که انواع و زیرگروه‌های متعدد دارد. بر اساس کیفیت سیگنال صحبت مورد نیاز این روش می‌تواند نرخ داده را تا اندازه‌ای بین ۲ تا ۶ کیلو بیت بر ثانیه کاهش دهد.

www.deconik.net