

تشخیص موقعیت ۳ بعدی منابع صوتی یکی از قابلیت‌های بسیار مهم مغز انسان است که به صورت روزمره مورد استفاده قرار می‌گیرد. یکی از ویژگی‌های مهم برای برخورداری از چنین قابلیت‌ی وجود بیش از یک گیرنده (دو گوش) است. هنگامی که امواج صوتی از یک منبع ساطع می‌شوند، بر اساس فاصله و زاویه قرارگیری منبع نسبت به شنونده، با کمی تفاوت در دو گوش فرد دریافت می‌شوند. مغز در طول زمان آموزش دیده است تا این تفاوت‌ها را به صورت موقعیت‌های مختلف مکانی منبع تعبیر کند. در نتیجه، برای تولید یک صدای ۳ بعدی لازم است تا به روشی، تفاوت‌های لازم بین سیگنال‌های دریافتی دو گوش ایجاد شود. برای این منظور، علاوه بر پردازش‌هایی که در زیر شرح داده می‌شوند، باید شنونده از headphone برای شنیدن استفاده کند تا قابلیت دریافت سیگنال‌های مستقل برای دو گوش فراهم گردد.

برای مدل سازی نسبتاً دقیق سیگنال دریافتی هر گوش از یک منبع در موقعیت خاص، از روش اندازه گیری استفاده می‌شود. برای این منظور، دو میکروفون ریز در گوش یک فرد قرار داده می‌شود و با قرار دادن بلندگو در موقعیت‌های مختلف نسبت به سر فرد و ارسال سیگنال از بلندگو، پاسخ ضبط می‌شود. در صورتی که سر فرد و موقعیت بلندگو ثابت باشد، رابطه بین سیگنال ارسالی از بلندگو و سیگنال دریافتی در میکروفن داخل گوش خطی و تغییر ناپذیر با زمان است. در نتیجه، این رابطه را می‌توان با یک پاسخ ضربه بیان کرد. در این پروژه، یکی از مجموعه داده‌های ضبط شده توسط دانشگاه Aalto در اختیار شما قرار داده شده است (Impulse_Responses.mat) و بر حسب این داده‌ها، قصد داریم یک فایل صوتی ۳ بعدی تولید کنیم. دقت داشته باشید که این داده‌ها مربوط به یک فرد خاص است و به دلیل تفاوت شکل سر افراد (مثلاً فاصله دو گوش، فرم جمجمه و ...)، لزوماً فایل تولید شده با این مجموعه داده برای هر فرد ایده‌آل نخواهد بود، اما به هر حال تقریب قابل قبولی است.

در داده‌های ضبط شده، دو حالت فاصله بین بلندگو تا فرد لحاظ شده است. در حالت اول (far)، فاصله 135cm و در حالت دوم (near) فاصله 68cm است. برای هر یک از دو فاصله، 240 زاویه مختلف برای منبع ایجاد شده است (زوایای مربوط به far و near لزوماً بر هم منطبق نیستند) و پاسخ ضربه برای هر حالت قرارگیری منبع و هر یک از دو گوش چپ و راست محاسبه شده است. نحوه نمایش زوایا به صورت azimuth و elevation است. زاویه azimuth دور تا دور فرد را نشان می‌دهد؛ زاویه 0 یعنی رو به روی فرد، زاویه 90 یعنی سمت چپ فرد، زاویه 180 (یا -180) یعنی پشت فرد و زاویه 90- یعنی سمت راست فرد. زاویه elevation بالا و پایین بودن منبع نسبت به سطح افق دو گوش را نشان می‌دهد؛ مثلاً زاویه elevation صفر یعنی هم ارتفاع با دو گوش، مقدار زاویه منفی یعنی کم ارتفاع تر از دو گوش و زاویه مثبت یعنی مرتفع تر از سطح دو گوش.

برای تولید یک فایل صوتی ۳ بعدی، ابتدا منبعی را در یک موقعیت خاص فضا نسبت به سر در نظر می‌گیریم. برای این موقعیت منبع، باید پاسخ ضربه‌های لازم را به دست آوریم. در صورتی که محل منبع منطبق با نقاط اندازه‌گیری شده نباشد، باید پاسخ ضربه‌ها را به کمک درونیانی تقریب بزنیم. با توجه به داده موجود، فرض بر این است که فاصله منبع تا سر بین 68cm تا 135cm است. برای درونیایی، یکبار فرض می‌کنیم فاصله منبع تا سر 68cm است و یکبار 135cm. در هر یک از حالات، به صورت وزندار پاسخ ضربه‌های نقاط اندازه‌گیری شده مجاور را میانگیری می‌کنیم. در نهایت، بر حسب فاصله منبع تا سر، پاسخ ضربه‌های مربوط به فاصله‌های 68cm و 135cm را نیز به صورت وزندار میانگیری می‌کنیم تا پاسخ ضربه‌های تخمینی حاصل شوند. اکنون کافی است که سیگنال ارسالی منبع را به پاسخ ضربه‌های حاصل شده convolve کنیم.

در صورتی که منبع در حال حرکت باشد (حالتی که در این پروژه لحاظ شده است)، پاسخ ضربه‌ها در طول زمان عوض می‌شوند و رابطه به صورت یک colvolution ساده نخواهد بود. برای این منظور، داده‌های منبع را به تکه‌های کوچک تقسیم می‌کنیم (length_jump=100 نمونه در کد) و برای هر تکه، پاسخ ضربه جداگانه محاسبه می‌کنیم و خروجی هر تکه را به کمک colvolution به دست می‌آوریم. در نهایت خروجی تکه‌ها را صورت superposition و با لحاظ کردن تاخیر زمانی لازم ترکیب می‌کنیم. طبیعی است که هر چه طول تکه‌ها کوچکتر باشد، کیفیت ۳ بعدی خروجی بهتر خواهد بود، اما از لحاظ محاسباتی کندتر خواهد بود.

(۱) ابتدا فایل ضمیمه را دریافت و unzip کنید.

(۲) قسمت زیادی از کدهای MATLAB به صورت آماده در اختیار شما قرار گرفته اند. فایل اصلی برای اجرا Main.m است که در پوشه S&S MATLAB Project قرار دارد. لطفاً این کد را تغییر ندهید.

۳) قسمت‌های مربوط به پروژه عبارتند از تکمیل فایل PathGeneration.m و تابع ImpResp_interpolator.m واقع در پوشه Functions_and_Subroutines، و تحلیل روش در قالب یک گزارش.

۴) در قسمت اول قصد داریم حرکت منبع روی یک خم حول شنونده را به فرم $azimuth$ ، $elevation$ و فاصله $range$ تبدیل کنیم. بر این منظور، جدولی در ابتدای فایل PathGeneration.m مشخص شده است. فرض بر این است که منبع در ابتدا رو به روی شنونده، با زاویه $elevation=20$ و فاصله 120cm قرار دارد و تا ثانیه 14 در همین وضعیت باقی می‌ماند. از ثانیه 14 منبع شروع به حرکت با سرعت ثابت می‌کند به نحوی که در ثانیه 22 به سمت چپ شنونده ($azimuth = 90$) با $elevation = 0$ در فاصله 70cm می‌رسد و به همین نحو تا ثانیه 45 مطابق با اعداد جدول حرکت می‌کند و از آن پس ثابت می‌ماند. در قسمت مشخص شده در فایل PathGeneration.m که لازم برای ایجاد بردار z و بردار فاصله را بنویسید.

۵) در مرحله بعد، هدف درونیانی پاسخ ضربه موقعیت دلخواه منبع بر اساس اطلاعات ضبط شده است. برای یافتن موقعیت‌های نزدیک به مکان منبع از لحاظ زاویه‌ای (فاصله را در این مرحله فراموش می‌کنیم) از ضرب داخلی استفاده می‌کنیم. به عبارت بهتر، راستای قرارگیری منبع نسبت به شنونده را با تمام 240 راستای اندازه‌گیری شده ضرب داخلی می‌کنیم. در محل مشخص شده در تابع ImpResp_interpolator.m که لازم برای محاسبه این ضرب داخلی‌ها را ایجاد کنید. در ادامه، ضرب داخلی‌های بالاتر از 0.9 به عنوان زوایای نزدیک به راستای منبع استخراج می‌شوند. پاسخ ضربه این راستاها به صورت وزندار میانگیری می‌شود (ضرب داخلی بیشتر معادل با وزن بیشتر). این عملیات یکبار برای اندازه‌گیری‌های far و یکبار برای اندازه‌گیری‌های $near$ انجام می‌شود و پاسخ ضربه‌های میانگیری شده هر کدام مجدداً بر اساس فاصله میانگیری می‌شوند تا پاسخ ضربه‌های نهایی حاصل شوند. قابل ذکر است که میانگیری بر اساس ضرب داخلی ایده‌آل نیست زیرا ممکن است زوایا دور راستای مورد نظر پخش نباشند و تنها در یک سمت آن قرار گیرند. اما از نظر محاسباتی این روش به سادگی قابل پیاده‌سازی است.

۶) پس از تکمیل دو قسمت قبلی، می‌توانید کد Main.m را اجرا کنید. با اجرای این کد، اطلاعات زاویه / فاصله تولید شده و فایل صوتی ۳ بعدی در پوشه Results ذخیره خواهند شد. در داخل MATLAB نیز هم فایل اصلی تک کاناله (غیر ۳ بعدی) و هم فایل ۳ بعدی تولید شده پخش می‌شوند.

۷) پس از پیاده‌سازی و اجرای صحیح فایل‌ها در MATLAB، علاوه بر ارسال گل پوشه، **ضمیمه کردن یک گزارش متنی تایپ شده به فرمت pdf ضروری است.** در فایل گزارش باید نحوه پیاده‌سازی‌ها، استدلال‌ها و توجیه نتایج شرح داده شوند. در نهایت تمرین ارسالی باید به صورت یک فایل zip شده شامل این گزارش باشد. لطفاً نام فایل zip را ترکیب نام و شماره دانشجویی خود قرار دهید.

۸) امتیازی: یک فایل صوتی مکالمه بین خود و فردی دیگر ضبط کنید و دو موقعیت منبع مکانی مختلف برای خود و فرد دیگر در نظر بگیرید. فایل صوتی ۳ بعدی تولید کنید به نحوی که در زمان صحبت کردن شما منبع در یک مکان باشد و در زمان صحبت کردن فرد دیگر، منبع در مکان دوم قرار گیرد.

۹) امتیازی: به جای استفاده از ضرب داخلی، روشی دقیقتر برای درونیایی پاسخ ضربه پیاده‌سازی کنید که در زمان معقول اجرا شود.

۱۰) این تمرین تنها از درگاه CW قابل ارسال است (از ارسال ایمیل خودداری کنید).

۱۱) مهلت تحویل پروژه چهارشنبه ۲۳ تیر (ساعت ۱۲:۰۰) مقرر شده است و تمدید نخواهد شد.

۱۲) این تمرین ۲ نمره از نمره نهایی درس را در بر می‌گیرد. در صورت مشاهده کپی بین چند تمرین مختلف، نمره ۲- برای تمامی افراد خاطی لحاظ خواهد شد.