پایان‌نامه برای دریافت درجه کارشناسیِ ارشد

در رشته مهندسی صدا

عنوان:

تولیدِ صدای سه بعدی با استفاده از روشِ رندرینگِ پاسخ ضربه‌ی مکانی(سه بعدی)

نگارش:

هدی ناصرالدین

استاد راهنما:

دکتر محمد عسگری

استاد مشاور:

دکتر ایوب بنوشی

تابستان ۱۳۹۱

بسم الله الرّحمن الرّحيم

تقدیم به:

پدربزرگوارم.

پدرم، هر چه هستم و هر چه دارم از توست.

مادر عزیزم که اسطوره‌ی صبر و تأمل من است.

همسر، خواهر و برادر عزیزم که همواره مرا همراهی و پشتیبانی می‌نمایند.

تشکر و قدردانی:

جا دارد از استادِ راهنما جنابِ آقای دکتر عسگری، استادِ مشاور جنابِ آقای دکتر بنوشی واستادِ گرامی سرکار خانمِ مهندس پورصادق و سایرِ دوستان و همکاران که من‌را در طولِ دوران انجام پروژه یاری رساندند تشکر و قدردانی نمایم. همچنین از جنابِ آقای دکتر زارعیان مدیرِ گروهِ مهندسیِ رسانه که با بردباریِ خود امکانِ انجامِ این پروژه را برای این‌جانب فراهم آوردند تشکر می‌نمایم. بی‌شک انجامِ این پایان‌نامه بدونِ راهنمایی و هم‌یاریِ این افراد امکان‌پذیر نبود.

چکیده

تولیدِ صدای سه بعدی به روشِ رندرینگِ پاسخ‌ضربه‌ی مکانی یکی از روش‌های جدید در بازتولیدِ اکوستیکِ فضاهای بسته به صورتِ چندکانالی است. در این روش میدانِ صوتیِ اصلی به صورتِ دقیق بازتولید نمی‌شود. بلکه با استفاده از مفاهیمِ اکوستیکِ روانی و فراسنج‌های مکان‌یابی، از تجزیه و تحلیلِ میدانِ صوتی به صورتِ فیزیکی استفاده می‌شود و صدایی بازتولید می‌شود که از لحاظِ ادراکی شبیه به صدای فضای اصلی است. بدین منظور باید پاسخ ضربه‌ی سالنِ مورد نظر، اندازه‌گیری شود و سپس پاسخ‌ضربه‌ی مطلوب، بر اساس تجزیه و تحلیلِ میدانِ صوتی ساخته شود. در صورتی که صدای هر منبعِ صوتی که در استودیو یا اتاقِ صامت(به صورتِ خام) ضبط شده است، با پاسخ‌ضربه‌ی ساخته شده به این روش، کانوالو شود، صدایی که شنیده می‌شود همان حسِ مکانی را دارد که گویی شنونده در سالنِ مرجع قرار گرفته است. پاسخ ضربه‌ی ساخته شده را در ریوربریتورها می‌توان موردِ استفاده قرار داد. برتری این روش به روش‌های پردازشی دیگر این است که امکان استفاده از هر نوع سامانه‌ی بلندگویی وجود دارد. یعنی سامانه‌ی بلندگوی موردِ استفاده، در بازتولیدِ صدا به این روش، محدود به ساختار و یا تعدادِ معینی بلندگو نیست و می‌توان با هر سامانه‌ی بلندگویی، بازتولیدِ صدا را انجام داد. بازتولیدِ صدا در این پایان نامه بر اساس سه بلندگو انجام شده است. برای اندازه‌گیریِ پاسخ ضربه‌ی سالن نیز باید از سامانه‌ی میکروفن‌گذاریِ چندکانالی استفاده کرد. زیرا یک میکروفن نمی‌تواند تمامی ویژگی‌های میدانِ صوتی را دریافت کند. بر این اساس در این پایان‌نامه برای اولین بار در ایران پاسخ ضربه یا به عبارتی پاسخِ سالن به روشی نوین اندازه‌گیری شده است. نوآوری انجام شده در اندازه‌گیریِ پاسخِ سالن به این روش، استفاده از سامانه‌ی میکروفن گذاریِ B-format و نیز نوعِ سیگنالِ تحریک است. سیگنالِ تحریک در این اندازه‌گیری بر خلافِ اندازه‌گیری‌های رایج به صورت جاروبِ سینوسیِ نمایی است که محدوده‌ی بسامدهای صوتی را پوشش می‌دهد. در مرحله‌ی پردازشِ سیگنال نیز، جهتِ به‌دست آوردنِ پاسخِ مطلوب، برای اولین بار در ایران، از پنینگِ برداریِ دامنه و دیفیوژن کانولوشن استفاده شده است.

واژه‌های کلیدی: صدای مکانی، رندرینگِ پاسخ‌ضربه، روشِ میکروفن‌گذاریِB-format، مکان‌یابی، تجزیه و تحلیلِ انرژیِ میدانِ صوتی، پنینگِ برداریِ دامنه، دیفیوژن کانولوشن.

فهرست مطالب

فهرست شکل‌ها ..... د
فهرست جدول‌ها ..... ز
فهرست علائم اختصاری ..... ح

فصل ۱: کلیاتِ تحقیق ۱
۱-۱ طرحِ مسأله ۴
۱-۲ ضرورت و اهمیتِ تحقیق ۵
۱-۳ اهدافِ تحقیق ۶
۱-۴ سؤالاتِ تحقیق ۶
۱-۵ فرضیاتِ تحقیق ۷
۱-۶ تعریفِ مفاهیم ۷

فصل ۲: مبانیِ نظریِ تحقیق ۹
۲-۱ بررسیِ تحقیقاتِ پیشین ۱۱
۲-۲ بررسیِ فیزیکیِ صدای سه بعدی ۱۱
۲-۲-۱ انتشارِ صدا ۱۲
۲-۲-۱-۱ سیگنالِ تحریک ۱۶
۲-۲-۱-۲ میکروفن‌های جهتی و آرایه‌ی میکروفن ۱۸
۲-۲-۲ تجزیه و تحلیلِ میدانِ صوتی بر اساسِ انرژی ۲۱
۲-۲-۲-۱ انرژیِ صدا ۲۱
۲-۲-۲-۲شدتِ صدا ۲۲
۲-۲-۲-۳ رابطه‌ی بینِ فشارِ صدا و سرعتِ ذره ۲۳
۲-۲-۲-۴ شدتِ فعال و غیرِفعال ۲۵
۲-۲-۲-۵ پخشایی ۲۶
۲-۲-۲-۶ تجزیه و تحلیلِ انرژی در حوزه‌ی بسامد ۲۸
۲-۲-۲-۷ روشِ اندازه‌گیریِ شدتِ صدا با استفاده از سیگنال‌های B-format [۲] ۲۹
۲-۳ شنواییِ سه‌بعدی ۳۰
۲-۳-۱ اعضاءِگوش ۳۱
۲-۳-۲ مکان‌یابی و فراسنج‌های مهم در آن ۳۵
۲-۳-۲-۱ اختلافِ ترازِ درون‌گوشی(ILD) ۳۶
۲-۳-۲-۲ اختلافِ زمانِ درون‌گوشی(ITD) ۳۸
۲-۳-۲-۳ توابعِ تبدیلِ مرتبط با سر(HRTF) ۴۱
۲-۳-۲-۴ همدوسیِ درون‌گوشی ۴۴
۲-۳-۲-۵ چرخشِ سر ۴۴
۲-۳-۲-۶ تأثیرِ سرنخ‌ها بر مکان‌یابی ۴۴
۲-۳-۲-۷ تصادمِ سرنخ‌ها با هم و منابعِ صوتیِ هم‌زمان ۴۶
۲-۳-۲-۸ اثرِ تقدم و مکان‌یابی در اتاق (فضای بسته) ۴۷
۲-۳-۳ وضوحِ زمانی و بسامدی در شنواییِ دوگوشی ۴۹

۲-۳-۳-۱ وضوحِ بسامدی ۴۹
۲-۳-۳-۲ وضوحِ زمانی ۴۹
۲-۳-۴ سایرِ ویژگی‌های ادراکِ سه‌بعدی ۵۰
۲-۴ روش‌های بازتولیدِ صدای چندکانالی ۵۵
۲-۴-۱ روش‌های میکروفن‌گذاریِ چندکانالی [۱] ۵۶
۲-۴-۱-۱ روش‌های میکروفن‌گذاری اصلیِ ۵ کانالی ۵۶
۲-۴-۱-۲ روش‌های میکروفن‌گذاری دو بخشی شامل بخش‌های جلویی و عقبی ۵۸
۲-۴-۱-۳ روش‌های منتخب میکروفن‌گذاری ۵۸
۲-۴-۲ پنینگِ دامنه ۶۹

فصل ۳: روشِ تحقیق ۷۱
۳-۱ فرضیاتِ موردِ استفاده در روشِ SIRR ۷۳
۳-۲ اندازه‌گیریِ فراسنج‌های اکوستیکی ۷۶
۳-۲-۱- اندازه‌گیری زمانِ واخنش ۷۶
۳-۲-۲ نوفه‌ی زمینه ۷۹
۳-۳ اندازه‌گیریِ پاسخِ سالنِ آمفی‌تئاترِ دانشکده‌ی صدا و سیما ۸۰
۳-۳-۱ روشِ اندازه‌گیری پاسخِ سالنِ آمفی‌تئاترِ دانشکده‌ی صدا و سیما ۸۰
۳-۴ رندرینگِ پاسخ‌ضربه‌ی مکانی برای بازتولیدِ صدای سه‌بعدی ۸۶
۳-۴-۱ آنالیزِ پاسخِ سالن ۸۸
۳-۴-۲ سنتزِ SIRR ۹۰
۳-۴-۲-۱ سنتزِ بخشِ غیرِ پخشا ۹۱
۳-۴-۲-۲ سنتزِ بخشِ پخشا ۹۳

فصل ۴: یافته‌های تحقیق ۹۵
۴-۱ یافته‌های حاصل از اندازه‌گیریِ زمانِ واخنش ۹۶
۴-۲یافته‌های حاصل از اندازه‌گیریِ نوفه‌ی زمینه ۹۹
۴-۳ پاسخِ اندازه‌گیری شده با استفاده از سیگنالِ جاروبِ سینوسی با بسامدِ متغیرِ نمایی و سامانه‌ی میکروفنیِ B-format ۱۰۰
۴-۴ رندرِ پاسخِ حاصل از سالنِ آمفی‌تئاترِ دانشکده‌ی صدا و سیما با روشِ SIRR ۱۰۵

فصل ۵: نتیجه‌گیری ۱۱۱
۵-۱- جمع‌بندی پایان‌نامه ۱۱۲
۵-۲-مقاله‌های تهیه شده ۱۱۴
۵-۳- پیشنهادها برای کارهای آینده ۱۱۴

مراجع ۱۱۶

فهرست شکل‌ها

شکل۱-۱) نمایشِ آزیموت، ارتفاع و فاصله [۴]..... ۷
شکلِ ۲-۱) نمایشِ الف- بازتابِ آینه‌ای پرتوهای صدا از سطحِ هموار، ب- بازتابِ پخشای پرتوهای صدا از سطحِ ناهموار، ج) شکستِ پرتوهای صدا از لبه‌های نوک تیز[۲]..... ۱۴
شکلِ ۲-۲) برخوردِ امواجِ تخت با زاویه‌ی θ به یک جفت میکروفنِ جهتی که در فاصله‌ی d نسبت به هم قرار گرفته‌اند [۲]..... ۱۹
شکل ۲-۳) المانِ مربع شکلِ از شاره به حجمِ $dV=dxdydz$ در هنگامِ انتشارِ صوت در شاره[۲]..... ۲۴
شکلِ ۲-۴) مدلِ کلی از شنوایی سه بعدی[۱۷]..... ۳۱
شکل ۲-۵) سطحِ مقطعِ‌گوش [۲]..... ۳۲
شکلِ ۲-۶) گوشِ میانی [۱۸]..... ۳۳
شکلِ ۲-۷) طرحِ ساده شده‌ی دالانِ جسمِ حلزونی[۱۸]..... ۳۴
شکلِ ۲-۸) جسمِ حلزونیِ گوشِ داخلی[۲۰]..... ۳۵
شکلِ ۲-۹) به‌وجود آمدنِ سایه توسطِ سر در بسامدهای بالا [۴]..... ۳۶
شکلِ ۲-۱۰) وابستگیِ سرنخِ ILD به بسامدهای مختلف [۴]..... ۳۷
شکلِ ۲-۱۱) سیگنالِ رسیده از یک منبعِ صوتی به دو گوش با هم اختلاف زمانی دارند [۴]..... ۳۸
شکلِ۲-۱۲) وابستگیِ ITD به زاویه‌ی آزیموت [۲]..... ۳۹
شکل۲-۱۳) مخروطِ اغتشاش[۴]..... ۴۱
شکل۲-۱۴) نمایی از نحوه‌ی اندازه‌گیری توابعِ HRTF[۴]..... ۴۲
شکلِ ۲-۱۵) توابعِ HRTF اندازه‌گیری شده در جهاتِ مختلف[۲]..... ۴۲
شکلِ ۲-۱۶) نمایشِ گرافیکیِ مفهومِ عرضِ منبعِ شنوایی(ASW) [۲۵]..... ۵۳
شکل ۲-۱۷) میکروفن‌گذاریِ ۵ کانالیِ اصلی..... ۵۷
شکل ۲-۱۸) نحوه‌ی دادن سیگنال خروجیِ تکنیک میکروفن گذاری ۵ کانالی اصلی به بلندگوها..... ۵۷
شکل ۲-۱۹) درخت دکا در ساختار فراگیر..... ۶۰
شکل ۲-۲۰) ساختار درخت فوکادا..... ۶۰
شکل ۲-۲۱) ساختار درخت OCT..... ۶۲
شکل ۲-۲۲) ساختارِ روشِ INA-3..... ۶۳
شکل ۲-۲۳) ساختارِ روشِ Near Coincident..... ۶۴
شکل ۲-۲۴) روشِ IRT-Cross..... ۶۶
شکل ۲-۲۵) نمایی دیگر از ساختار چیدمانِ IRT-Cross..... ۶۶
شکل ۲-۲۶) روشِ مربعِ هاماساکی..... ۶۷
شکلِ ۳-۱) فلوچارتِ روندِ انجامِ پایان‌نامه..... ۷۴
شکل ۳-۲) الف: نمایی از سالنِ آمفی‌تئاترِ دانشده صدا و سیما. در این نما شیب‌دار بودن سطح واضح است. همچنین پایین‌ترین نقطه و بالاترین نقطه در فضا مشخص شده است، ب: نمای شبیه سازی شده از سالن توسط نرم افزار ادئون [۲۷]. در این نما نیز بالاترین سطح و پایین‌ترین سطح مشخص شده است. ۷۷
شکل ۳-۳) : محل نقاط اندازه‌گیری زمان واخنش در سالنِ آمفی‌تئاتر دانشکده صدا و سیما نشان داده شده است..... ۷۸
شکل ۳-۴) الف- نحوه‌ی نصبِ تجهیزات جهتِ اندازه‌گیریِ زمانِ واخنش توسطِ آنالایزرِ ۲۱۴۴ و ب-

محل قرارگیری بلندگوی همه جهتی و آنالایزر در سالنِ آمفی تئاتر دانشکده صدا و سیما. ۷۹
شکلِ ۳-۵) طیفِ جاروب سینوسی نمایی (۴۰۰Hz-۲۱kHz) سیگنالِ تحریک در نرم افزارِ Adobe Audition. محورِ افقی بسامد و محورِ عمودی، دامنه‌ی طیف را در بسامدهای مختلف نشان می‌دهد. ۸۱
شکل ۳-۶) نحوه‌ی چیدمانِ سامانه‌ی میکروفنِ B-format ۸۲
شکلِ ۳-۷) نحوه‌ی ارتباطِ تجهیزاتِ ضبط(الف) و (ب) پخشِ سیگنالِ جاروبِ سینوسیِ نمایی در سالن جهتِ اندازه‌گیریِ پاسخِ سالن ۸۳
شکل ۳-۸) محلِ نقاطِ اندازه‌گیریِ پاسخِ سالن و محلِ قرارگیریِ بلندگوها. ۸۴
شکل ۳-۹) نمایی از محلِ قرارگیریِ تجهیزاتِ ضبط و پخش در یکی از اندازه‌گیری‌های پاسخِ سالنِ آمفی‌تئاتر. در این اندازه‌گیری بلندگو در نقطه‌ی s1 و میکروفن‌ها در نقطه‌ی p3 قرار گرفته است. ۸۴
شکل ۳-۱۰) نمونه‌ای از سیگنال‌های حاصل از اندازه‌گیریِ پاسخِ سالنِ آمفی‌تئاترِ دانشکده‌ی صدا و سیما در نرم‌افزار Adobe Audition در حوزه‌ی زمان. سیگنالِ سمتِ راست محدوده‌ی بسامدیِ ۴۰۰Hz-۲۱KHz و تقریباً معادلِ ۴ ثانیه و سیگنالِ سمتِ چپ محدوده‌ی بسامدیِ ۳۰Hz-۴۵۰Hz و تقریباً معادلِ ۲ ثانیه را نشان می‌دهد. ۸۰
شکلِ ۳-۱۱) بلوک دیاگرامِ پردازشِ سیگنال‌ها در مرحله‌ی آنالیز. ۸۷
شکلِ ۳-۱۲) بلوک دیاگرامِ پردازشِ سیگنال‌ها در مرحله‌ی سنتز و به‌دست آمدنِ سیگنال‌های نهایی. ۸۸
شکلِ ۳-۱۳) ساختارِ چینشِ سه بلندگوی فرضی بر اساس پنینگِ برداریِ دامنه. ۹۲
شکلِ ۳- ۱۴) بخشی از برنامه‌ی نوشته شده در نرم‌افزارِ Matlab جهتِ تولیدِ سیگنالِ پخشا در مرحله‌ی سنتز ۹۴
شکلِ ۴-۱) نمودارِ میانگینِ زمان واخنشِ اندازه‌گیری شده در ۲۱ نقطه. ۹۹
شکل ۴-۲) مقایسه‌ی مقادیر اندازه‌گیری شده برای نوفه‌ی زمینه در سالن با مقدار استانداردِ NC30. ۱۰۰
شکلِ ۴-۳) سیگنالِ اندازه گیری شده در نقطه۱ با بلندگوی ۱. سیگنال‌ها به ترتیب از بالا (ردیفِ دوم) به پایین عبارتند از سیگنالِ خروجی از میکروفنِ همه‌جهتی، سیگنالِ خروجی از میکروفنِ جهتی در راستای x، سیگنالِ خروجی از میکروفنِ جهتی در راستای y و سیگنالِ خروجی از میکروفنِ جهتی در راستای z – الف) اولین اندازه‌گیری، ب) دومین اندازه‌گیری، ج) سومین اندازه‌گیری. سیگنالِ نشان داده شده در ردیفِ اول، سیگنالِ منبع است ۱۰۲
شکلِ ۴-۴) سیگنالِ اندازه‌گیری شده در نقطه۱ با بلندگوی ۲. سیگنال‌ها به ترتیب از بالا (ردیفِ دوم) به پایین عبارتند از سیگنالِ خروجی از میکروفنِ همه‌جهتی، سیگنالِ خروجی از میکروفنِ جهتی در راستای x، سیگنالِ خروجی از میکروفنِ جهتی در راستای y و سیگنالِ خروجی از میکروفنِ جهتی در راستای z – الف) اولین اندازه‌گیری، ب) دومین اندازه گیری. سیگنالِ نشان داده شده در ردیفِ اول، سیگنالِ منبع است ۱۰۲
شکلِ ۴-۵) سیگنالِ اندازه‌گیری شده در نقطه۲ با بلندگوی ۱. سیگنال‌ها به ترتیب از بالا (ردیفِ دوم) به پایین عبارتند از سیگنالِ خروجی از میکروفنِ همه‌جهتی، سیگنالِ خروجی از میکروفنِ جهتی در راستای x، سیگنالِ خروجی از میکروفنِ جهتی در راستای y و سیگنالِ خروجی از میکروفنِ جهتی در راستای z – الف) اولین اندازه‌گیری، ب) دومین اندازه‌گیری، ج) سومین اندازه‌گیری. سیگنالِ نشان داده شده در ردیفِ اول، سیگنالِ منبع است. ۱۰۳
شکلِ ۴-۶) سیگنالِ اندازه‌گیری شده در نقطه۲ با بلندگوی۲. سیگنال‌ها به ترتیب از بالا (ردیفِ دوم) به پایین عبارتند از سیگنالِ خروجی از میکروفنِ همه‌جهتی، سیگنالِ خروجی از میکروفنِ جهتی در راستای x، سیگنالِ خروجی از میکروفنِ جهتی در راستای y و سیگنالِ خروجی از میکروفنِ جهتی در راستای z – الف) اولین اندازه‌گیری، ب) دومین اندازه‌گیری. سیگنالِ نشان داده شده در ردیفِ اول، سیگنالِ منبع است ۱۰۴
شکلِ ۴-۷) سیگنالِ اندازه‌گیری شده در نقطه۳ با بلندگوی ۱. سیگنال‌ها به ترتیب از بالا (ردیفِ دوم)

به پایین عبارتند از سیگنالِ خروجی از میکروفنِ همه‌جهتی، سیگنالِ خروجی از میکروفنِ جهتی در راستای x، سیگنالِ خروجی از میکروفنِ جهتی در راستای y و سیگنالِ خروجی از میکروفنِ جهتی در راستای z – الف) اولین اندازه‌گیری، ب) دومین اندازه‌گیری، ج) سومین اندازه‌گیری. سیگنالِ نشان داده شده در ردیفِ اول، سیگنالِ منبع است. ۱۰۴
شکلِ ۴-۸) سیگنالِ اندازه‌گیری شده در نقطه۳ با بلندگوی ۲. سیگنال‌ها به ترتیب از بالا (ردیفِ دوم) به پایین عبارتند از سیگنالِ خروجی از میکروفنِ همه‌جهتی، سیگنالِ خروجی از میکروفنِ جهتی در راستای x، سیگنالِ خروجی از میکروفنِ جهتی در راستای y و سیگنالِ خروجی از میکروفنِ جهتی در راستای z – الف) اولین اندازه‌گیری، ب) دومین اندازه‌گیری، ج) سومین اندازه‌گیری. سیگنالِ نشان داده شده در ردیفِ اول، سیگنالِ منبع است..... ۱۰۵
شکلِ ۴-۹) بخشی از طیفِ سیگنالِ همه‌جهتیِ ضبط شده در سالن که در نرم افزار Matlab نشان داده شده است. محورِ افقی، بسامدِ سیگنال و محورِ عمودی، دامنه را نشان می‌دهد..... ۱۰۶
شکلِ ۴-۱۰) سیگنالِ همه‌جهتیِ ضبط شده در سالن که در نرم افزار Matlab در حوزه‌ی زمان نشان داده شده است. محورِ افقی، زمان و محورِ عمودی، دامنه را نشان می‌دهد..... ۱۰۷
شکلِ ۴-۱۱) توزیعِ زمانی- بسامدیِ بردارِ شدتِ فعال در راستای x..... ۱۰۷
شکلِ ۴-۱۲) سیگنالِ جهتیِ به‌دست آمده از مرحله‌ی سنتز در راستای x..... ۱۰۸
شکلِ ۴-۱۳) سیگنالِ جهتیِ به‌دست آمده از مرحله‌ی سنتز در راستای y..... ۱۰۹
شکلِ ۴-۱۴) سیگنالِ جهتیِ به‌دست آمده از مرحله‌ی سنتز در راستای z..... ۱۰۹
شکلِ ۴-۱۵) سیگنالِ پخشای به‌دست آمده از مرحله‌ی سنتز. این سیگنال به صورت مساوی به هر سه بلندگو ارسال می‌گردد. ۱۱۰

فهرست جدول‌ها

جدولِ(۴-۱) مقادیرِ به‌دست آمده از اندازه‌گیریِ زمانِ واخنشِ سالنِ آمفی‌تئاترِ دانشکده‌ی صدا و سیما در ۲۱ نقطه. ۹۸
جدولِ (۴-۲) میانگینِ زمان واخنشِ اندازه‌گیری شده در ۲۱ نقطه. ۹۸
جدولِ (۴-۳) میانگینِ نوفه‌ی اندازه‌گیری شده در سالنِ آمفی‌تئاتر دانشکده‌ی صدا و سیما. ۱۰۰

فهرست علائم اختصاری

Pa	پاسکال
Hz	هرتز
kHz	کیلوهرتز
SIRR	رندرینگِ پاسخ ضربه‌ی مکانی
dB	دسی‌بل
dB_{SPL}	ترازِ فشارِ صوتی برحسبِ دسی‌بل

فصل ۱:

کلیاتِ تحقیق

امروزه تولیدِ صدای سه‌بعدی[1] ، مانند تصویرِ سه‌بعدی و نیز تصویرِ با وضوحِ بالا[2] (HD) از جمله فناوری‌های به‌روزی است که تحقیقاتِ فراوانی در خصوصِ آن در حالِ انجام است. توقعِ شنوندگان و بینندگان از صدا و تصویر، دیگر فقط شنیدنِ یک پیام و یک تصویرِ معمولی با حداقل کیفیت نیست و به عبارتی فقط محتوا برای افراد اهمیت ندارد. بلکه کیفیتِ صدا و تصویر نیز از اهمیتِ بالایی برخوردار است. در زمینه‌ی تصویر پیشرفت‌های وسیعی به‌وجود آمده است. اغلب بینندگان امروزه خواهانِ تصاویرِ با وضوحِ بالا و سه‌بعدی به جای تصاویرِ استاندارد[3] (SD) هستند. از این‌رو پخشِ تصاویر نیز با فرمتِ HD صورت می‌گیرد. گیرنده‌های تلویزیونی نیز از حالتِ سنتی خارج شده‌اند و فناوریِ ساختِ آن‌ها به منظورِ دریافتِ سیگنال‌های HD و نیز نمایشِ تصاویرِ سه‌بعدی بسیار پیشرفت کرده است. مشاهده‌ی این تصاویر با صدای تک‌کاناله[4] و یا دوکاناله[5] دیگر دلپذیر نیست و توقعِ شنوندگان نیز در سطحِ بالاتری قرار گرفته است. از این‌رو پخشِ صدا به صورتِ فراگیر[6] توسطِ تجهیزاتِ پخشِ خانگی در حالِ افزایش است. شنوندگان انتظار دارند صدایی بشنوند که نظیرِ صدای طبیعی موجود در محیط (فضای باز یا فضای بسته) باشد. این صدا، صدای سه‌بعدی است. یعنی صدایی که از جهاتِ مختلف و یا حداقل از سه جهت در راستای محورهای مختصاتِ دکارتی (x، y و z) به دو گوشِ شنونده می‌رسد. در واقع این صدای درک شده توسطِ دو گوش شنونده، برایندِ تمامیِ صداهای مستقیم، غیرِ مستقیم، بازتاب شده، تداخل پیدا کرده و اثرِ اتاق (در فضای بسته) یا فضای باز است.

[1] Three Dimentional
[2] High Definition
[3] Standard Definition
[4] Mono
[5] Stereo
[6] Surround

از طرفی برخی فضاهای بسته از جمله سالن‌های موسیقی وجود دارند که شنیدنِ ارکسترِ موسیقی در آن‌ها بسیار لذت‌بخش و خوشایند است. شنوندگان دوست دارند که برای شنیدنِ موسیقی به این سالن‌ها بروند. در واقع، فضا می‌تواند به صدا گرمیِ خاصی ببخشد. همچنین این سالن‌ها از نظرِ افرادی که موسیقیِ کلاسیک را اجرا می‌کنند بسیار مهم هستند. این سالن‌ها بیش از یک قرن قدمت دارند. فهمِ عمیقِ پدیده‌های اکوستیکی پایه و به ویژه مواردِ مرتبط با نحوه‌ی ادراکِ انسانی از جمله تحقیقاتِ بسیار به روز برای این سالن‌ها می‌باشند، که در حالِ انجام است. به علاوه امکانات تولید و تجربه‌ی محیط‌های صوتیِ فراگیر و متنوع، با توسعه‌ی فناوری ضبط و بازتولید صدا چندین برابر شده است.

ولی این سالن‌ها همه جا در دسترس نیستند و امکانِ حضور در این سالن‌ها برای تمامیِ افراد وجود ندارد. بنابراین به نظر می‌رسد که اگر صدای ارکسترِ موسیقی که در چنین سالنی می‌نوازد را به‌گونه‌ای بتوان ضبط کرد که در هنگامِ بازشنودِ آن، حسِ حضور در آن سالن، ایجاد شود، یکی از راه‌هایی است که همیشه بتوان از صدای این سالن‌ها استفاده کرد. بر این اساس مهندسِ صدا باید به طریقی صدا را ضبط کند که هم صدای ارکسترِ موسیقی و هم اکوستیک و حسِ سالن را دریافت کند که در هنگامِ بازتولیدِ آن، حسِ قرارگرفتن در سالنِ اصلی ایجاد شود. ولی با توجه به محدودیتِ نوع و ساختارِ میکروفن‌ها این کار بسیار مشکل و گاهی امکان‌ناپذیر است. از این‌رو این پایان‌نامه به دنبالِ روشِ جایگزینی است که به چنین هدفی بتوان دست‌یافت. به همین دلیل بر اساسِ بررسی‌های به عمل آمده در موردِ روش‌های تولیدِ صدای سه‌بعدی که عبارتند از :

۱- میکروفن گذاریِ چندکانالی [۱]،
۲- میکسِ چندکانالی [۱]و
۳- پردازش [۱].

در این پایان‌نامه برای اولین بار در ایران، روشِ پردازشیِ رندرینگِ پاسخ ضربه‌ی مکانی[1]، برای بازتولیدِ صدای سالن به صورت سه‌بعدی استفاده می‌شود. هدف این است که اکوستیکِ سالنِ موردِ نظر را در فضای دیگری بتوان بازتولید کرد. در صورتِ رسیدن به این هدف، صدای هر منبعِ صوتی (مثلِ ساز یا سخنران) را به صورتِ خام در استودیو یا اتاقِ صامت می‌توان ضبط کرد. سپس با ترکیبِ آن با اکوستیکِ سالنِ موردِ نظر، صدایی به‌دست آورد که حسِ حضور در سالنِ موردِ نظر را القاء کند.

[1] Spatial Impulse Response Rendering

بنابراین به‌دست آوردنِ اکوستیکِ سالن بسیار مهم است که در این پایان‌نامه به آن پرداخته شده است. روش‌های متداول برای دریافتِ اکوستیکِ سالن، استفاده از میکروفنِ اضافی در هنگامِ ضبط برای دریافتِ بازتاب‌ها، استفاده از میکروفن‌گذاریِ چندکانالی [۱] و نیز استفاده از پاسخِ ضربه است. در فصلِ ۲ مشاهده می‌شود که استفاده از روش‌های میکروفن‌گذاریِ چندکانالی مشکلاتِ فراوانی دارند و هنوز تحقیقاتِ فراوانی در این خصوص در حالِ انجام است. در این پایان‌نامه از پاسخِ ضربه به شیوه‌ای جدید استفاده شده است که همان رندرینگِ پاسخ ضربه‌ی مکانی است [۲]. در این روش، با توجه به نحوه‌ی مکان‌یابیِ انسان و استفاده از فیزیکِ میدانِ صوتی، پاسخ ضربه‌ای از سالن ساخته می‌شود که در آن تمامیِ شاخص‌های موردِ نیاز برای مکان‌یابیِ منابع و به‌وجود آمدنِ حسِ آن فضا وجود داشته باشد. در این راستا لازم است سه موضوعِ مختلف در نظر گرفته شوند که عبارتند از:

۱- فیزیکِ میدانِ صوتی،

۲- ادراکِ انسان،

۳- روش‌های بازتولیدِ صدا.

این موارد به عنوانِ زیربنای موضوعِ پایان‌نامه تحتِ عنوانِ مبانیِ نظریِ تحقیق در فصلِ ۲ موردِ بررسی قرار گرفته‌اند. بر اساسِ این نظریه‌ها، رندرینگِ پاسخ ضربه‌ی مکانی در فصلِ ۳ فرمول‌بندی و تشریح شده است. یافته‌های حاصل از این روش در فصلِ ۴ ارائه گردیده است و در فصلِ ۵ جمع‌بندی و پیشنهاداتی برای کارهای آتی بیان شده است.

۱-۱- طرحِ مسأله

معمولاً مهندسِ صدا، ضبطِ منبعِ صوتی یا به‌طورِ کلی ضبطِ برنامه (ارکستر موسیقی، نمایش‌نامه، سخنرانی) را در سالنی را انجام می‌دهد که هم‌زمان صدای منبعِ اصلی و اکوستیکِ فضا را دریافت کند. غالباً استفاده از چنین سالن‌هایی برای ضبطِ برنامه‌های بزرگ، امکان ندارد و برای کنترلِ بیشترِ منابعِ صدا، لازم است که از روش‌های میکروفن‌گذاریِ نزدیک استفاده شود و سپس در مرحله‌ی میکسِ برنامه، صحنه‌ی مصنوعی از فضا ساخته شود [۳].

در روش‌های ضبطِ برنامه با استفاده از میکروفن‌گذاریِ نزدیک، چندین میکروفن، نزدیکِ منبعِ صدا گذاشته می‌شوند که سیگنالِ صدای تقریباً خشک[۱] و بدون بازتابِ اتاق، به‌دست آید. در مرحله‌ی

[۱] به صدای مستقیمِ دریافت شده از منبعِ صوتی که بازتاب‌های اتاق بر آن تأثیر نداشته است، صدای خشک (dry)، گفته می‌شود. مثلاً صدای ضبط شده در استودیو یا اتاقِ صامت، صدای خشک است.

میکس، از پنینگِ دامنه[1] برای قرار دادن این سیگنال‌ها در جهت‌های دلخواه استفاده می‌شود. حسِ مکانی با کمکِ دستگاه‌های جانبی مثلِ ریوربریتورها[2] و یا با اضافه کردنِ سیگنال‌هایی از میکروفن‌های اضافی دیگر، که دورتر از منبع در سالنِ ضبط قرار گرفته‌اند، خلق می‌شود [۳]. بدین ترتیب حسِ فضا و سالن به صورتِ مصنوعی ساخته می‌شود.

بنابراین با استفاده از ریوربریتورهای کانوالو کننده، این امکان فراهم شده است تا ضبط و اجرای برنامه را در هر سالن، شبیه سازی کرد. اساسِ کارِ ریوربریتورها، استفاده از پاسخ ضربه‌ی اتاق است. با اندازه‌گیری پاسخ ضربه‌ی سالن‌های مختلف و بارگزاریِ آن در ریوربریتور، صحنه‌ی مصنوعی از این سالن‌ها را می‌توان به‌وجود آورد.

رندرینگِ پاسخ ضربه‌ی مکانی، روشِ نوینی است که با استفاده از پاسخ اندازه‌گیری شده از اتاق، درنظرگرفتن نحوه‌ی ادراکِ فضا توسطِ انسان و سامانه‌ی بلندگوییِ مقصد، پاسخ ضربه‌ای می‌سازد که با بارگزاری آن در ریوربریتورها و کانوالو کردنِ صدای خشکِ منبعِ صوتی با آن، حسِ فضا و سالن، در مقصد بازتولید می‌شود. این پاسخ ضربه، چندکانالی است و تعداد کانال‌های آن محدود نیست. بدین ترتیب می‌توان صدای سه بعدی را در هر فضایی بازتولید کرد.

۱-۲- ضرورت و اهمیتِ تحقیق

بازتولیدِ اکوستیکِ هر فضای بسته در فضای بسته‌ی دیگر، با روشی ساده و کم‌هزینه یکی از ضرورت‌های جوامع امروزی است که دسترسیِ مجازی به هر محیطی را امکان‌پذیر می‌سازد. بدین ترتیب بدون نیاز به حضورِ واقعی در یک سالن یا فضای بسته، حسِ حضور در آن را با استفاده از روشِ رندرینگِ پاسخ ضربه‌ی مکانی می‌توان در فضای دیگری به وجود آورد. اهمیت این روش در ساده‌بودن روشِ پردازشیِ مورد نیاز در آن، عملیاتی بودن انجامِ روش و وابسته نبودنِ آن به ساختارِ بلندگویی خاص می‌باشد. همچنین بازتولیدِ صدا به صورت سه‌بعدی و نیز فراگیر در هر فضایی با استفاده از این روش امکان‌پذیر می‌باشد

[1] تقسیمِ دامنه‌ی سیگنال، بین کانال‌ها را پنینگِ دامنه گویند.

[2] ریوربریتور(Reverbrator) دستگاهی است که واخنش را به صورت مصنوعی، به‌گونه‌های مختلف تولید می‌کند.

۱-۳- اهدافِ تحقیق:

در تحقیق انجام شده در این پایان نامه دستیابی به اهدافِ ذیل مدِ نظر بوده است:
۱- تحقیقاتِ جامعی در خصوصِ نحوه‌ی مکان‌یابیِ انسان انجام شود، که این هدف در فصلِ ۲ محقق شده است.
۲- پاسخِ چندکانالیِ سالنِ آمفی‌تئاترِ دانشکده‌ی صدا و سیما اندازه‌گیری شود، که این هدف برای اولین بار در ایران به روشی نوین انجام شد که در فصلِ ۳ تشریح شده است.
۳- پاسخ های به‌دست آمده از اندازه‌گیری و رابطه‌ی آن‌ها با ادراکِ مکانیِ انسان توصیف شود که دستیابی به این هدف در فصلِ ۴ آمده است و تشریح شده است.
۴- بر اساسِ داده‌های به‌دست‌آمده از اندازه‌گیری و روشِ پردازشی رندرینگِ پاسخ ضربه‌ی مکانی، پاسخِ چند کاناله‌ی مناسب، برای بازتولیدشدن توسطِ سامانه‌ی بلندگویی چندکاناله، ساخته شود که دستیابی به این هدف در فصلِ ۴ انجام شده است.

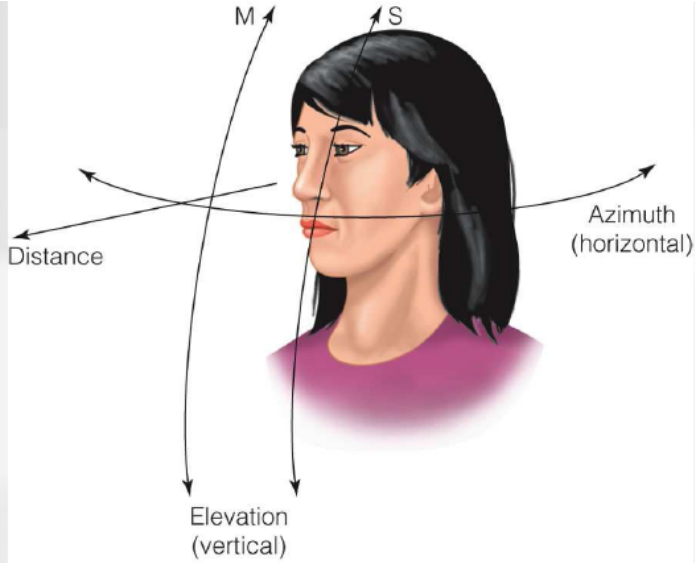
۱-۴- سؤالاتِ تحقیق:

سؤالاتِ اصلیِ انجام شده در این تحقیق عبارت‌اند از:
آیا می‌توان فضایِ اکوستیکیِ محیطی را به فضای دیگری منتقل کرد، بدون این‌که فضای اکوستیکیِ دو محیط، شباهتی به هم داشته باشند؟
چگونه انسان فضای اکوستیکیِ یک محیط را از محیطِ دیگر متمایز می‌کند؟
انسان صدا را چند بعدی درک می‌کند؟
آیا سامانه‌های طراحی شده‌ی چندکانالیِ امروزی، جهتِ صدارسانی به صورتِ چندکانالی واقعاً صدا را به صورتِ طبیعی(همان‌گونه که انسان درک می‌کند) به محیط عرضه می‌کنند؟
سامانه‌های ذکر شده چه میزان قابلیتِ اجرایی شدن را در همه‌ی شرایط دارا هستند؟
آیا می‌توان سامانه‌ای طراحی کرد که با درنظرگرفتنِ نحوه‌ی ادراکِ انسان و استفاده از واقعیت‌های فیزیکیِ موجود، قابلیتِ اجرایی شدنِ آن را ساده‌تر کند؟

۱-۵- فرضیاتِ تحقیق

در انجامِ این پایان‌نامه از دانشِ اکوستیک‌روانی و نیز فیزیکِ صوت جهتِ به‌دست آوردنِ روشی برای تولیدِ صدای سه بعدی و بازتولیدِ اکوستیکِ فضای بسته استفاده شده است.

۱-۶- تعریفِ مفاهیم



شکل۱-۱) نمایشِ آزیموت، ارتفاع و فاصله [۴].

•صدای مکانی در این پایان نامه بر اساسِ سیستمِ مختصاتِ کارتزین توصیف شده است. بر این اساس محور افقی(x)، آزیموت[1] که همان موقعیتِ راست و چپِ شنونده است، محورِ عمودی(z)، ارتفاع[2] که همان جهتِ بالا و پایین برای شنونده است و محورِ (y) فاصله[3] نامیده شده است (شکلِ ۱-۱).

•SIRR: رندرینگِ پاسخ ضربه‌ی مکانی، پردازشِ پاسخِ اندازه‌گیری شده‌ی اتاق برای به‌دست آوردنِ پاسخ‌ضربه‌ی مکانی از آن می‌باشد.

[1] Azimuth
[2] Elevation
[3] Distance