تعامل انسان و اطلاعات، جلد۴ شماره ۴
http://hii.khu.ac.ir
مقاله اصیل پژوهشی

پیش‌نمون‌سازی سامانه ویبراتو-فهم بازیابی زمزمه-محور اطلاعات موسیقایی برای ابزارهای ارتباطی همراه: مطالعه موردی هارمونیکای کروماتیک

کیوان برنا: گروه علوم کامپیوتر، دانشکده علوم ریاضی و کامپیوتر، دانشگاه خوارزمی، کرج، ایران.
داریوش علیمحمدی: گروه علم اطلاعات و دانش‌شناسی، دانشکده روان‌شناسی و علوم تربیتی، دانشگاه خوارزمی، کرج، ایران، (نویسنده مسئول).
alimohammadi@khu.ac.ir

چکیده

دریافت: ۱۳۹۶/۱۰/۱۳
پذیرش: ۱۳۹۷/۳/۰۷

زمینه و هدف: پژوهش حاضر با هدف پیش‌نمون‌سازی سامانه‌های زمزمه-محور روی گوشی‌های هوشمند انجام شده است.
روش پژوهش: در این تحقیقِ چند-روشی، از فن شبیه‌سازی در مجموعه مدل‌های آمیخته روش پژوهش در عملیات، و نیز روش سندی، به طور همزمان، استفاده شده است. آزمایش این پژوهش بر دو آلبوم هارمونیکای کروماتیک بنا شد. به این منظور، ۲۴ قطعه هوموفونیک با استفاده از نرم‌افزار *هلیوم اُدیو اِسپلیتِر* برش زده شدند. برش‌های برگزیده در محیط نرم‌افزار *سونیک ویژوآلایزرِ* پردازش، و ۱۶۸ سند ایکس اِم اِل (گروه اول) تولید شدند. از سوی دیگر، ۴ سوژه پژوهش همان برش‌ها را همراه با ویبراتو، زمزمه و ضبط کردند. زمزمه‌ها با استفاده از نرم‌افزار *ای اِم آر تو اِم پی تری کانورتُر* تبدیل فرمت شده، و برای دریافت خروجی‌های ایکس اِم اِل به *سونیک ویژوآلایزرِ* تحویل شدند. در این مرحله نیز مجموعه‌ای متشکل از ۶۷۲ سند ایکس اِم اِل (گروه دوم) تولید شد. نرم‌افزار *مَتلَب* بر اساس ۱۶۸ سند گروه اول با استفاده از شبکه عصبی مصنوعی تعلیم داده شد. سپس، ۶۷۲ سند گروه دوم، به *مَتلَب* خورانده، و پردازش شدند. در نهایت، خروجی‌های *مَتلَب* برای هر دو گروه، به وسیله نرم‌افزار *ایمیج کامپیرِر* با یکدیگر مقایسه شدند.
یافته‌ها: یافته‌های پژوهش، شباهت بسیار زیاد (۹۹ درصد) میان خروجی‌های منتج از پردازش قطعات (گروه اول)، و خروجی‌های منتج از پردازش زمزمه‌ها (گروه دوم) را نشان دادند. همچنین، مشخص شد که تغییر جنسیت و مهارت موسیقایی کاربران تفاوتی در نتایج ایجاد نمی‌کند.
نتیجه‌گیری: می‌توان اذعان کرد طراحی سامانه زمزمه-محوری که از طریق تبدیل صوت به اسناد ایکس اِم اِل، و مقایسه اسناد، قطعات موسیقایی را بازیابی می‌کند، راهبرد مناسبی برای توسعه نرم‌افزارهای بازیابی این قطعات روی گوشی‌های هوشمند است.

کلیدواژه‌ها: پیش‌نمون‌سازی، سامانه زمزمه-محور بازیابی اطلاعات موسیقایی، ویبراتو، ابزار ارتباطی همراه، هارمونیکای کروماتیک

مقدمه

علاقه وافر اقشار گوناگون مردم به موسیقی قابل توجه است. در این میان، گروهی از مردم مهارت نواختن ساز را دارند. این گروه، ملودی‌های موجود و بعضاً خود-ساخته را می‌نوازند؛ و به این شکل هیجانات خود را ابراز و تخلیه می‌کنند. اما کسانی که در موسیقی فاقد مهارت هستند، به دلیل ناآگاهی از نحوه نگارش ملودی، مهارت انتخاب سازها، اجرای آهنگ، ضبط صدا، و دیگر مسائل مرتبط، نیاز خود را غالباً از طریق شنیدن آثار موسیقایی برطرف می‌کنند. امروزه، ابزارهای ارتباطی همراه به عنوان مناسب‌ترین دستگاه پخش و شنیدن قطعات موسیقایی، در میان کاربران رواج یافته‌اند.

بهره‌مندی مداوم از موسیقی، ثبت تدریجی اثر در حافظه را سبب می‌شود. اما هر کسی ممکن است در معرض فراموشی قرار بگیرد. این مسئله، یعنی ناتوانی در فراخوانی قطعات موسیقایی از حافظه، نیازمند ارائه راهکاری ساده است. راه‌حل مطلوب، فارغ از میزان دشواری محتمل فناورانه، باید معطوف به ساده‌ترین و طبیعی‌ترین شیوه در دسترس عموم باشد. این شیوه چیزی نیست جز زمزمه کردن. زمزمه[1] به زبان ساده تکرار ملودی بر پایه صدای *هوم*، یا صوت نغمه‌ای بی‌کلام است؛ که با خارج کردن صدا از بینی و بدون باز کردن دهان تولید می‌شود.

در میان سازها، هارمونیکا یا ساز دهنی، به عنوان یک ساز بادی و به خاطر رواج فراوان در تمام کشورها، می‌تواند در پژوهش‌های مرتبط با بازیابی اطلاعات موسیقایی مورد توجه قرار گیرد. زیرا گذشته از صدای زیبا و اثرگذار، و شیوه یادگیری ساده تا کنون در هیچ پژوهش مرتبطی استعداد این ساز در نظام‌های بازیابی اطلاعات موسیقایی به محک تجربه گذاشته نشده است. حال آنکه بنا به نظر آرجانیکُف[2] (۲۰۱۳)،

1 Humming
2 Arjannikov

[Downloaded from system.khu.ac.ir on 2024-05-06]
[DOR: 20.1001.1.24237418.1396.4.4.6.3]

هارمونیکا شایستگی این توجه را دارد. با مسلم دانستن این دو پیش‌فرض که ویبراتو[1] یکی از رایج‌ترین فنون غنی‌سازی صدا در اجراهای موسیقایی است (پانگ[2] و یون[3]، ۲۰۰۵، ص ۱۱۳۵)، و هارمونیکا نیز در مقام یک ساز بادی، ویبراتوِ خاص خود را دارد (اِشمِل ۱۳۸۹الف؛ ۱۳۸۹ب)، می‌توان نقش ویبراتوِ هارمونیکا برای تسهیل پردازش پرسش‌های کاوش مبتنی بر زمزمه در یک پایگاه موسیقایی را بررسی کرد. بر این اساس، پژوهش حاضر یک پرسش اصلی و دو پرسش فرعی را مد نظر قرار داده است.

پرسش اصلی

• سامانه ویبراتو-فهم بازیابی زمزمه-محور اطلاعات موسیقایی برای ابزارهای ارتباطی همراه، تا چه حد خواهد توانست قطعات نواخته شده به وسیله هارمونیکای کروماتیک را بازیابی کند؟

پرسش‌های فرعی

• آیا با تغییر جنسیت کاربران، در توان سامانه ویبراتو-فهم بازیابی زمزمه-محور اطلاعات موسیقایی برای ابزارهای ارتباطی همراه، به هنگام بازیابی قطعات، تفاوتی ایجاد خواهد شد؟

• آیا با تغییر مهارت موسیقایی کاربران، در توان سامانه ویبراتو-فهم بازیابی زمزمه-محور اطلاعات موسیقایی برای ابزارهای ارتباطی همراه، به هنگام بازیابی قطعات، تفاوتی ایجاد خواهد شد؟

پژوهشگر به منظور حفظ بی‌طرفی پیش‌بینی می‌کند که بین عملکرد پیش‌نمون سامانه ویبراتو-فهم بازیابی زمزمه-محور اطلاعات موسیقایی از یک سو، و جنسیت و مهارت موسیقایی کاربران از سوی دیگر، رابطه‌ای وجود ندارد. بنابراین، دو فرض صفر به قرار ذیل شکل می‌گیرند:

• H_0 یک: بین عملکرد پیش‌نمون سامانه ویبراتو-فهم بازیابی زمزمه-محور اطلاعات موسیقایی و جنسیت کاربران، رابطه‌ای وجود ندارد.

• H_0 دو: بین عملکرد پیش‌نمون سامانه ویبراتو-فهم بازیابی زمزمه-محور اطلاعات موسیقایی و مهارت موسیقایی کاربران، رابطه‌ای وجود ندارد.

دو فرضیه پژوهش نیز به شکل ذیل بیان می‌شوند:

• بین عملکرد پیش‌نمون سامانه ویبراتو-فهم بازیابی زمزمه-محور اطلاعات موسیقایی و جنسیت کاربران، رابطه‌ای معنادار وجود دارد.

• بین عملکرد پیش‌نمون سامانه ویبراتو-فهم بازیابی زمزمه-محور اطلاعات موسیقایی و مهارت موسیقایی کاربران، رابطه‌ای معنادار وجود دارد.

روش

پژوهش جاری، از حیث معرفت‌شناسی[4] عینی‌گرا[5] است. در چارچوب نظریه اثبات‌گرایی منطقی[6] طراحی شده است؛ و حواس تجربی را با منطق ریاضی ترکیب می‌کند. در تفسیر نتایج، تابع رویکرد استقرایی[7] است؛ و از نظر امکان استفاده از نتایج، کاربردی به حساب می‌آید. زیرا یافته‌های آن را می‌توان در توسعه فناوری بازیابی اطلاعات موسیقایی از ابزارهای ارتباطی همراه، مؤثر دانست. راهبرد[8] حاکم بر آن، روش‌شناسی[9] پژوهش در عملیات[10] است. پژوهش در عملیات، به زبان ساده، رفتار پدیده‌ها و سامانه‌ها را مطالعه می‌کند تا بتواند تغییرات آنها را برنامه‌ریزی و کنترل کند. این روش، همچنین، ناظر بر استفاده از قوانین علمی برای تصمیم‌گیری است. در هر موقعیت تصمیم‌گیری با دو یا بیش از دو گزینه مواجه می‌شویم که انتخاب هر یک به نتیجه‌ای متفاوت و شاید ناشناخته ختم خواهد شد. هدف کلی روش پژوهش در عملیات گزینش راه‌حلی است که بهترین نتیجه را به بار بیاورد. در حوزه تمرکز این پژوهش، یعنی نظام‌های ذخیره‌سازی و بازیابی اطلاعات، با الگوریتم‌ها و برنامه‌هایی روبه‌رو هستیم که شیوه استنباط و اجرای آنها می‌تواند مجموعه متفاوتی از مدارک متنی و غیرِ متنی را برای کاربر بازیابی نماید. از این رو، لازم است مدل مناسبی از این روش اتخاذ شود. در این تحقیق، فن شبیه‌سازی از مجموعه مدل‌های آمیخته[11] پژوهش در عملیات مورد استفاده قرار گرفته است. همچنین، از آنجا که پژوهش حاضر برای گردآوری اطلاعات مرتبط، روش سندی[12] را نیز همزمان به کار می‌گیرد، ماهیت چند-روشی[13] دارد.

دو آلبوم هوموفونیک هارمونیکا که تا کنون با مجوز وزارت

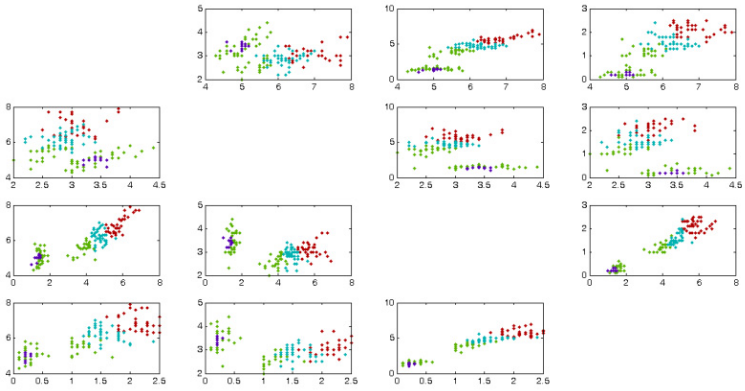
---

[1] Vibrato (ویبراتو به زبان ساده عبارت است از لرزش صدای ساز.)
[2] Pang
[3] Yoon
[4] Epistemology
[5] Objectivist
[6] Logical Positivism
[7] Inductive
[8] Strategy
[9] Methodology
[10] Operations Research
[11] Mixed Models
[12] Documentary Research Method
[13] Multi-Method

[Downloaded from system.khu.ac.ir on 2024-05-06]
[DOR: 20.1001.1.24237418.1396.4.4.6.3]

جدول ۱- آزمون‌ها

		آزمون اول	آزمون دوم
جنسیت		مردان آموزش‌دیده و غیرِ متخصص (۳۳۶ سند ایکس اِم اِل)	آموزش‌دیده‌های مرد و زن (۳۳۶ سند ایکس اِم اِل)
		زنان آموزش‌دیده و غیرِ متخصص (۳۳۶ سند ایکس اِم اِل)	غیرِ متخصص‌های مرد و زن (۳۳۶ سند ایکس اِم اِل)
جمع		۶۷۲ سند ایکس اِم اِل	۶۷۲ سند ایکس اِم اِل
مهارت موسیقایی			
جمع			



نمودار ۱- نتیجه تحلیل مقادیر عددی متغیرهای معتبر ۱۶۸ سند ایکس اِم اِل

فرهنگ و ارشاد اسلامی در بازار موسیقی ایران عرضه شده‌اند، عبارتند از:

- آلبوم *دوست قدیمی*[۱] مشتمل بر ۱۴ قطعه، اثر بابک صفرنژاد؛ و
- آلبوم *دروازه خورشید*[۲] مشتمل بر ۱۲ قطعه، اثر علی معتمدی.

بر اساس جدول تعیین حجم نمونه کِرجسی و مورگان اگر جامعه آماری متشکل از ۲۶ عنصر باشد، نمونه‌ای مشتمل بر ۲۴ عضو می‌تواند آن جامعه را نمایندگی کند. شیوه نمونه‌گیری تصادفی طبقه‌ای، این امکان را فراهم می‌کند که از هر آلبوم ۱ قطعه را تصادفاً کنار گذاشته و نمونه مورد نیاز را به دست آوریم. برای گردآوری اطلاعات، سیگنال بخش‌هایی از ۲۴ قطعه موسیقایی که نمونه آماری این پژوهش را تشکیل می‌دهند، پردازش شدند. علت پردازش تنها یک بخش از هر قطعه آن است که به دلیل ماهیت هوموفونیک قطعات، پژوهشگر در هر قطعه با چند صدا مواجه است. به زبان ساده‌تر، هر چند این دو آلبوم، به عنوان آلبوم‌های هارمونیکا شناخته می‌شوند، اما در اجرای قطعات، تنها از این ساز استفاده نشده است. برای نمونه، پیانو، و ویولن از دیگر سازهایی هستند که صدای آنها در این آلبوم‌ها به وضوح قابل تشخیص است. در نتیجه، پژوهشگر نمی‌توانست کل یک قطعه را پردازش کند. ضمن اینکه پردازش تصادفی یک بخش و تعمیم نتیجه، تنها به نام هارمونیکا نیز منطقی به نظر نمی‌رسد. تحت این شرایط، فقط می‌باید بر بخش‌هایی از قطعات تمرکز می‌شد که سایر سازها، نُتِ هارمونیکا را دنبال می‌کنند. یعنی بخش‌هایی که در آنها هارمونیکا ملودی اصلی قطعه را نواخته و سایر سازها نقش آکورد را داشته‌اند.

بنابراین، ۲۴ برش چند ثانیه‌ای با استفاده از نسخه ۱/۹/۰/۳۴۳ نرم‌افزار هلیوم اُدیو اِسپلیتر[۳] آماده، و در محیط نرم‌افزار سونیک ویژوآلایزر[۴] پردازش شدند. خروجی‌های ویرایش ۲/۴/۱ نرم‌افزار سونیک ویژوآلایزر که در قالب اسناد ایکس اِم اِل آماده می‌شوند، در ۷ لایه[۵] قابل دریافت هستند. به منظور رعایت اصل بی‌طرفی، هر برش در چارچوب تمام ۷ لایه، پردازش، و در نتیجه ۱۶۸ سند ایکس اِم اِل تولید شدند. در طرف مقابل، ۴ نفر به عنوان سوژه تحقیق مشارکت کردند. یک مرد آموزش‌دیده، یک مرد غیرِ متخصص، یک زن

[۱] . آلبوم دوست قدیمی در سال ۱۳۸۸ به وسیله نشر رهگذر هفت اقلیم به بازار موسیقی ارائه شده است.

[۲] . آلبوم دروازه خورشید در سال ۱۳۹۱ به وسیله موسسه گلچین آوای شرق به بازار موسیقی ارائه شده است.

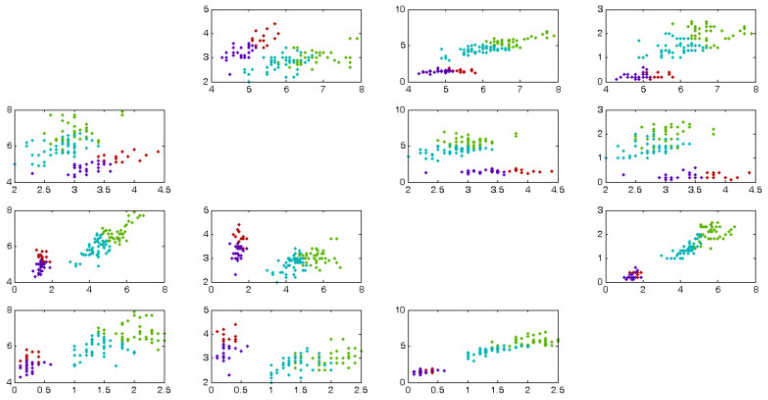
[۳] Helium Audio Splitter

[۴] Sonic Visualiser

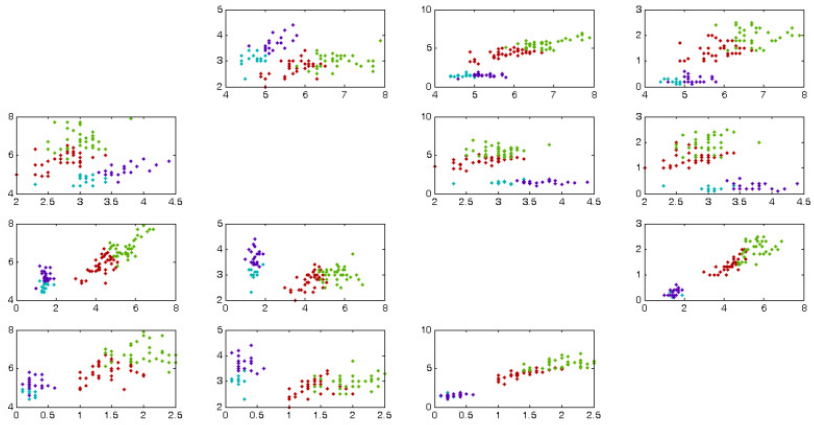
[۵] Layer

[Downloaded from system.khu.ac.ir on 2024-05-06]

[DOR: 20.1001.1.24237418.1396.4.4.6.3]



نمودار ۲- نتیجه تحلیل مقادیر عددی متغیرهای معتبر ۳۳۶ سند ایکس اِم اِل (گروه مردان)

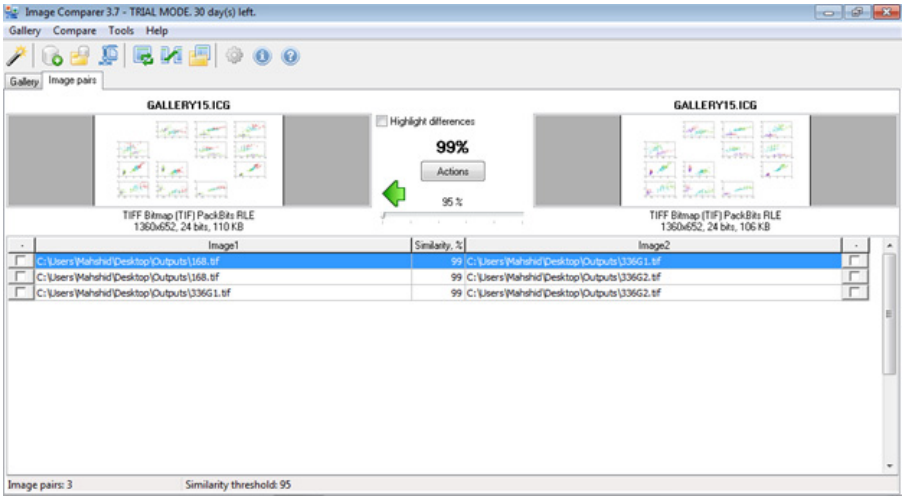


نمودار ۳- نتیجه تحلیل مقادیر عددی متغیرهای معتبر ۳۳۶ سند ایکس اِم اِل (گروه زنان)

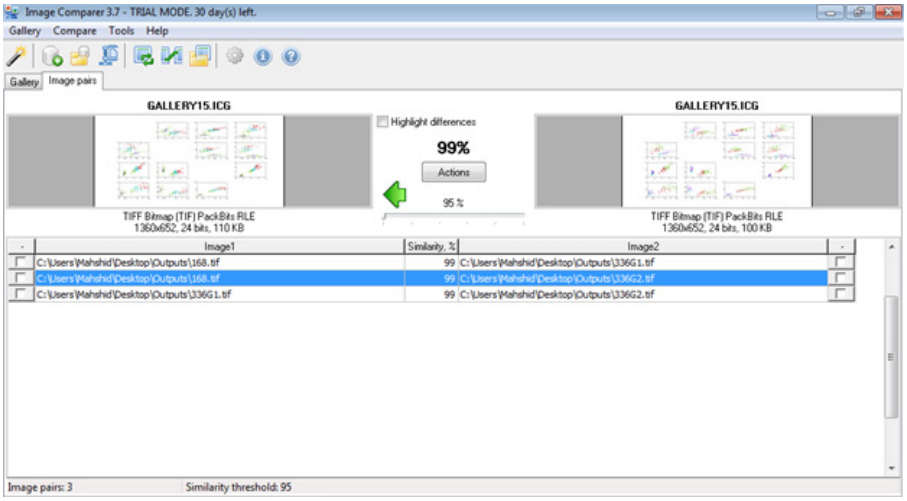
آموزش‌دیده و یک زن غیرِ متخصص، سوژه‌های پژوهش را تشکیل دادند. در این پژوهش، مراد از فرد آموزش‌دیده، صاحب مدرک کارشناسی موسیقی است که در این حوزه فعالیت حرفه‌ای نیز داشته باشد. هر سوژه، ۲۴ برش برگزیده را همراه با ویبره زمزمه کرد. زمزمه‌ها به وسیله گوشی همراه، ضبط، در فرمت اِی اِم آر ذخیره، و با استفاده از نسخه ۱/۴ نرم‌افزار اِی اِم آر تو اِم پی تِری کانورتُر[1] به فرمت اِم پی تِری تبدیل شدند. در مجموع، ۹۶ قطعه زمزمه‌ای تولید شد؛ که با ۲۴ برش اولیه برابری می‌کردند. این ۹۶ قطعه نیز در چارچوب همان ۷ لایه پردازش شده، و ۶۷۲ خروجی ایکس اِم اِل را به دست دادند. ۱۶۸ سند ایکس اِم اِلِ گروه اول وارد ویرایش سال ۲۰۱۲ نرم‌افزار مَتلَب شدند. از میان متغیرهای موجود در اسناد ایکس اِم اِل، چهار متغیر *model id*، *sampleRate*، *dataset id* و *layer id* ویژگی‌های صوتی قطعات موسیقایی را بازنمایی می‌کنند. مقادیر عددی این متغیرها استخراج و بر اساس چهار آرایه در مَتلَب وارد شدند. بعد از درون‌دهی متغیرها، برنامه لازم برای ایجاد شبکه مصنوعی و آموزش دادن مَتلَب، نوشته، و نرم‌افزار بر اساس متغیرها تعلیم داده شد. در مرحله بعد، ۶۷۲ سند ایکس اِم اِلِ گروه دوم به عنوان درون‌داد به مَتلَب خورانده شدند. سپس، از همان برنامه برای درون‌دهی مقادیر عددی متغیرهای معتبر ۶۷۲ سند ایکس اِم اِل گروه دوم، و آموزش مجدد مَتلَب تبعیت شد. در این مرحله، درون‌دادها چهار بار بر اساس جنسیت (آزمون اول)، به تفکیک مرد (آموزش‌دیده و غیرِ متخصص)، و زن (آموزش‌دیده و غیرِ متخصص)، و نیز بر اساس مهارت موسیقایی (آزمون دوم)، به تفکیک آموزش‌دیده (مرد و زن)، و غیرِ متخصص (مرد

[1] AMR to MP3 Converter

[Downloaded from system.khu.ac.ir on 2024-05-06]
[DOR: 20.1001.1.24237418.1396.4.4.6.3]



شکل ۱- تشابه زمزمه مردان با قطعات موسیقایی



شکل ۲- تشابه زمزمه زنان با قطعات موسیقایی

و زن) به نرم‌افزار خورانده شدند (جدول ۱).

در نهایت، خروجی‌های دو مرحله تعلیم، به منظور تعیین میزان مشابهت قطعات و شبیه‌سازی عملکرد پیش‌نمون سامانه زمزمه-محور بازیابی اطلاعات موسیقایی به وسیله ویرایش ۳/۷ نرم‌افزار ایمِیج کامپیرر[1] با یکدیگر مقایسه شدند. منطق حاکم بر استفاده از نرم‌افزار ایمِیج کامپیرر، قانون پایستگی انرژی است که بر اساس آن انرژی از بین نمی‌رود؛ بلکه از شکلی به شکل دیگر تبدیل می‌شود. بنابراین، می‌توان تصویر را با صوت هم‌ارزش فرض کرد. همچنان که در فن نمایه‌سازی به عنوان یکی از راهبردهای ذخیره‌سازی و بازیابی اطلاعات، کلیدواژه‌ها به مثابه جایگزین و بازنمون متون، اسناد و مدارک متنی و غیرِ متنی، حامل همان انرژی اصیل در انتقال معنا فرض می‌شوند.

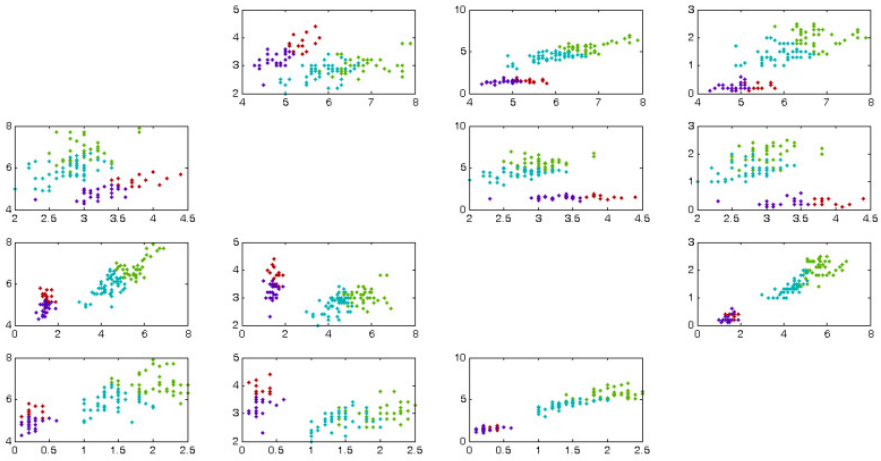
یافته‌ها

برای پاسخگویی به پرسش اصلی پژوهش باید ابتدا به پرسش‌های فرعی پاسخ داد. نمودارهای[2] ۱، ۲، و ۳ به ترتیب، نتیجه تحلیل مقادیر عددی متغیرهای معتبر ۱۶۸ سند ایکس اِم اِل گروه اول (نمودار ۱)، و نیز نتیجه تحلیل مقادیر عددی متغیرهای معتبر ۶۷۲ سند ایکس اِم اِل گروه دوم در چارچوب آزمون اول (نمودارهای ۲ و ۳) به وسیله نرم‌افزار متَلَب را نشان می‌دهند.

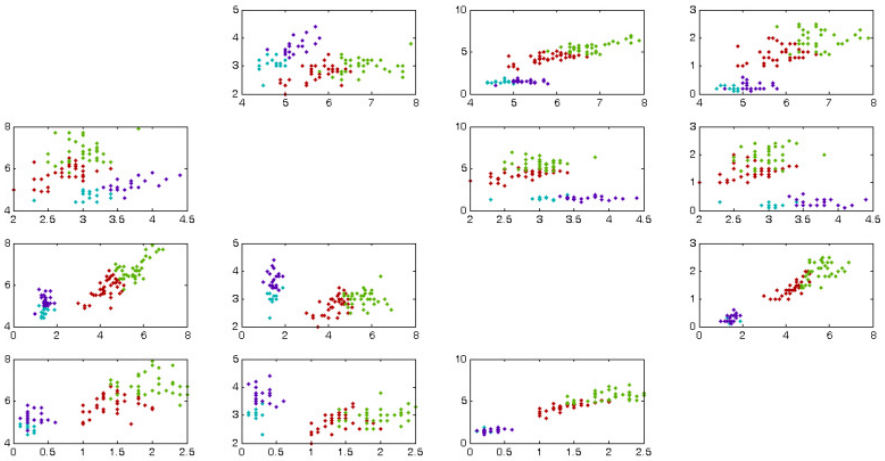
[1] Image Comparer

[2] Plots

[Downloaded from system.khu.ac.ir on 2024-05-06]
[DOR: 20.1001.1.24237418.1396.4.4.6.3]



نمودار ۴- نتیجه تحلیل مقادیر عددی متغیرهای معتبر ۳۳۶ سند ایکس اِم اِل (گروه آموزش‌دیده‌ها)



نمودار ۵- نتیجه تحلیل مقادیر عددی متغیرهای معتبر ۳۳۶ سند ایکس اِم اِل (گروه غیرِ متخصصان

این نمودارها در ظاهر هیچ تفاوتی با هم ندارند. تحلیل‌های مکمل نرم‌افزار ایمِیج کامپیرِر نیز آشکار می‌کنند که مقدار شباهت بسیار بالاست (شکل‌های ۱ و ۲).

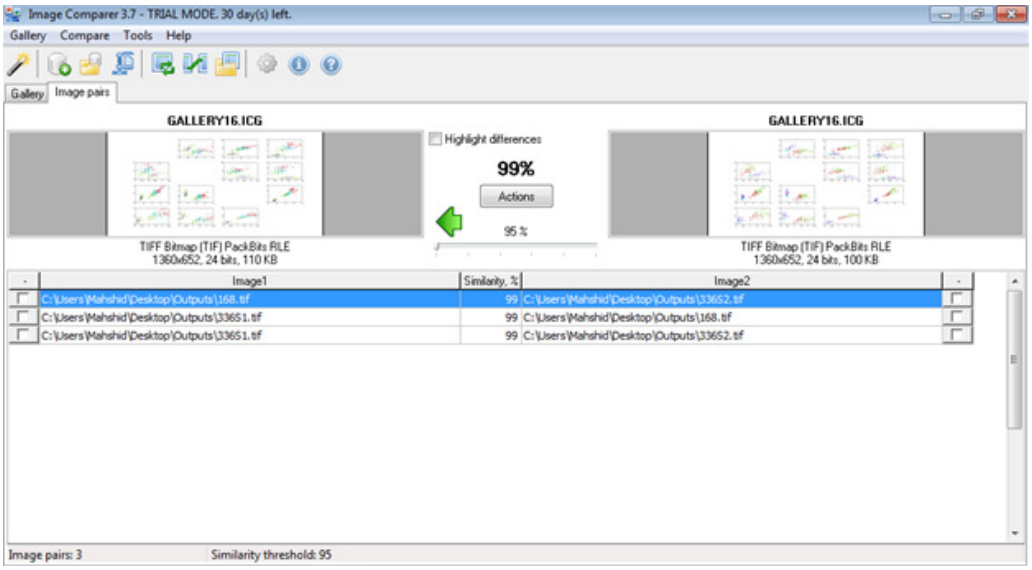
شباهت بسیار زیاد (در حد ۹۹ درصد) میان نمودارها که از طریق مقایسه و تطبیق نقاط رنگی به دست آمده، نشان می‌دهد که با تغییر جنسیت کاربران، در عملکرد پیش‌نمون سامانه ویبراتو-فهم بازیابی زمزمه-محور اطلاعات موسیقایی برای ابزارهای ارتباطی همراه، به هنگام بازیابی قطعات، تفاوتی ایجاد نخواهد شد. بنابراین، پاسخ نخستین پرسش فرعی منفی است.

نمودارهای ۴ و ۵، نتیجه تحلیل مقادیر عددی متغیرهای معتبر ۶۷۲ سند ایکس اِم اِل گروه دوم در چارچوب آزمون دوم را نشان می‌دهند.

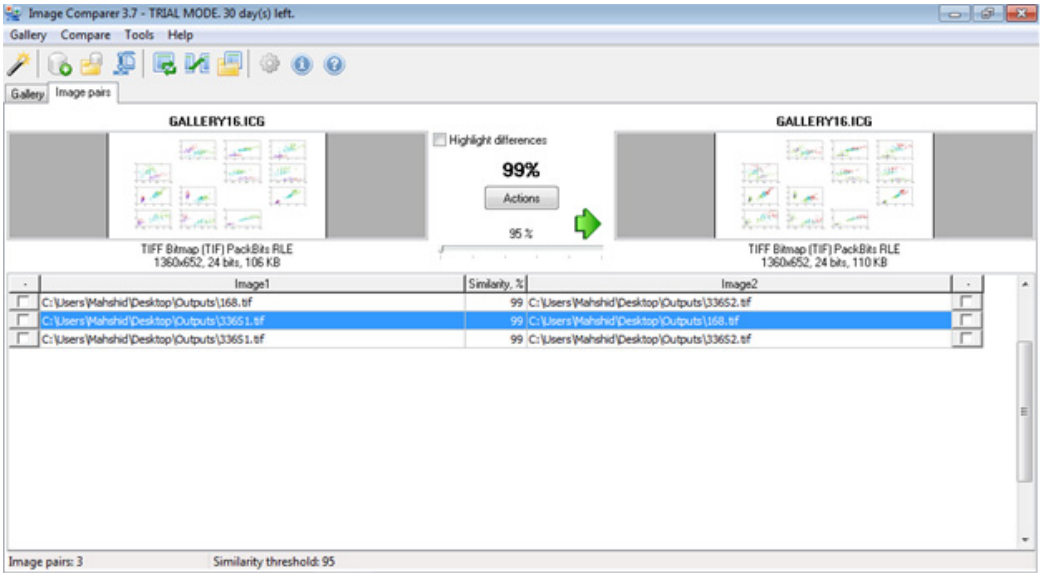
شکل‌های ۳ و ۴ نیز، مقایسه نمودارهای ۴ و ۵ با نمودار ۱ در محیط نرم‌افزار ایمِیج کامپیرِر را نشان می‌دهند.

به این ترتیب، روشن می‌شود که حتی با تغییر مهارت موسیقایی کاربران، فرآیند شبیه‌سازی عملکرد پیش‌نمون سامانه ویبراتو-فهم بازیابی زمزمه-محور اطلاعات موسیقایی برای ابزارهای ارتباطی همراه، به هنگام بازیابی قطعات، تفاوتی را نشان نخواهد داد. زیرا در این مورد نیز شباهت ۹۹ درصدی وجود دارد. در نتیجه، می‌توان گفت که پاسخ دومین پرسش فرعی پژوهش نیز منفی است. در مجموع، و در پاسخ به پرسش اصلی پژوهش حاضر می‌توان گفت که با شبیه‌سازی عملکرد پیش‌نمون سامانه ویبراتو-فهم بازیابی زمزمه-محور اطلاعات موسیقایی برای ابزارهای ارتباطی همراه، امکان بازیابی قطعات نواخته شده به وسیله هارمونیکای کروماتیک تا حد بسیار زیادی فراهم شد. همچنین، یافته‌های پژوهش نشان دادند که نمی‌توان به وجود

[Downloaded from system.khu.ac.ir on 2024-05-06]
[DOR: 20.1001.1.24237418.1396.4.4.6.3]



شکل ۳- تشابه زمزمه آموزش‌دیده‌ها با قطعات موسیقایی



شکل ۴- تشابه زمزمه غیرِ متخصصان با قطعات موسیقایی

رابطه معنادار بین مؤلفه‌های جنسیت و مهارت موسیقایی، و سامانه باور داشت. بنابراین، فروض صفر پژوهش جاری، تأیید، و در نتیجه، فرضیه‌های آن رد می‌شوند. در نتیجه، هیچ آزمونی روی فرضیه‌ها اجرا نشد.

بحث و نتیجه‌گیری

تِریپاتی[۱] و دیگران (۲۰۰۹، ص ۳۷۳)، به درستی یادآور شده‌اند که بنا کردن سامانه بازیابی اطلاعات موسیقایی بر تجزیه نُت‌ها و تشخیص زیری و بمی، حتی در قطعات مونوفونیک (تک‌صدایی) نیز می‌تواند به نتایج گمراه‌کننده ختم شود. به علاوه، کیم[۲] (۲۰۱۰، ص ۱۱)، اذعان کرده است که موسیقی هوموفونیک (چندصدایی) از این نظر دشواری افزون‌تری را ایجاد می‌کند. با درک این حقیقت، پردازش سیگنال و گرفتن خروجی ایکس اِم اِل در دستور کار قرار گرفت. همچنین، تنها آن بخش از ملودی هر قطعه تجزیه شد که بر پایه صدای هارمونیکا نوشته و نواخته شده بود. در نقطه مقابل، از آزمودنی‌ها نیز خواسته شد که بخش منتخب را زمزمه کنند. به باور دِ لا بَندِرا[۳] و دیگران (۲۰۱۱، صص ۴۹-۵۰)، گزینش برش اصلی یک قطعه موسیقایی، صدای اصلی را غنی و پردازش سیگنال را میسرتر خواهد کرد. از سوی دیگر، پاردُ[۴] و بیرمنگام[۵] (۲۰۰۳، ص ۱۰۷)، استفاده از قطعاتی معدود و

۱ Tripathy
۲ Kim
۳ de la Bandera
۴ Pardo
۵ Birmingham

[Downloaded from system.khu.ac.ir on 2024-05-06]
[DOR: 20.1001.1.24237418.1396.4.4.6.3]

آزمودنی‌هایی اندک شمار را برای قضاوت کم هزینه درباره کارایی یک سامانه و نه الزاماً کیفیت رتبه‌بندی نتایج آن، مناسب دانسته‌اند. بر این اساس، تحلیل ۲۴ قطعه موسیقایی با مشارکت تنها ۴ نفر در این پژوهش، قابل توجیه است. در کنار مزایای اسناد ایکس اِم اِل باید پذیرفت که در جریان تحلیل اصوات موسیقایی زمزمه شده، بخشی از صدا که بسامد پایینی دارد امکان بازنمایی را نمی‌یابد. در حال حاضر این واقعیت را به عنوان ضعف پردازش سیگنال می‌توان مورد توجه قرار داد.

گزارش شده است که مردان نسبت به زنان، و موسیقی‌دانان نسبت به افراد غیرِ متخصص، درک موسیقایی بهتری دارند (اونال[۱] نارایانان و چِو[۲] ۲۰۰۴، ص ۱۱۸؛ بِرآن[۳] ۱۹۹۱؛ بِرآن و وان[۴] ۱۹۹۳، ص ۱۸۶۰؛ شیه[۵] ژانگ[۶] و کو[۷] ۱۹۹۹؛ کُرس[۸] و لِویس[۹] ۱۹۵۰، ص ۸۳). با وجود این، یافته‌ها نشان داد که رفتار سامانه مفروض در حالت استفاده از سوژه‌های غیرِ متخصص نسبت به زمان مشارکت سوژه‌های متخصص دچار افت کیفی محسوس نمی‌شود. همچنین، تفاوت معناداری بین بازدهی مردان و زنان مشاهده نشد. از دیگر نکات تأمل‌برانگیز آن است که طول مناسب پرسش کاوش در واحد زمان ۸ ثانیه دانسته شده است (کُسوگی؛[۱۰] ۲۰۰۴، ص ۲۴). در پژوهش جاری، میانگین زمانی پرسش‌های کاوش ۸/۲ ثانیه بود. با وجود این تفاوت، نشان داده شد که اگر آزمایش سامانه مبتنی بر خروجی‌های ایکس اِم اِل طراحی شود، تغییر محسوسی در نتایج بازیابی شده ایجاد نخواهد شد.

همچنین، اذعان شده است که کاربر ممکن است به هر شکل ممکن و به طور سلیقه‌ای ملودی یک قطعه را زمزمه کند (کُسوگی، ۲۰۰۴، صص ۱۵-۱۶)؛ و در ناخودآگاهش خود را از این نظر آزاد می‌داند. در این پژوهش نیز، با وجود اینکه پیش از آزمون درباره شیوه زمزمه کردن به سوژه‌ها تذکر داده شده بود، اما به هر حال نمی‌توان ادعا کرد که زمزمه‌ها با کیفیت بالا عرضه شده و عاری از هر گونه خطا بوده‌اند. زیرا ابهام در زمزمه کاربران، یک نقیصه ذاتی است. شاید بتوان بخشی از مشکل را به ماهیت هوموفونیک قطعات موسیقایی و ادراک سوژه‌های تحقیق از آنها نسبت داد. ژو و شاشا (۲۰۰۳الف، ص ۴؛ ۲۰۰۳ب، ص ۱۸۴)، تأکید کرده‌اند که از هر ده هزار نفر، تنها یک نفر می‌تواند زیری و بمی را به صورت کاملاً صحیح ادا کند. اونال، نارایانان و چِو (۲۰۰۴، ص ۱۱۴)، پا را از این هم فراتر نهاده و این رقم را به کمتر از یک صدم درصد کاهش داده‌اند. در تفسیری واقع‌بینانه، همان یک نفر هم بسته به جنسیت، سن و حالت روحی ممکن است به درستی از عهده این کار برنیاید. این بدان معنا است که کیفیت بازتولید صدا در قالب زمزمه، تنها از آشنایی زمزمه‌گر با موسیقی متأثر نمی‌شود. به باور اونال و دیگران (۲۰۰۵، ص ۴۷۷) کاربران، ملودی‌های آشنا را بهتر زمزمه می‌کنند. اما باید این حقیقت را نیز در نظر داشت که هر چه آشنایی کمتر باشد، عدم اطمینان حاصل از تشابه صدای زمزمه شده با اصل آهنگ، افزایش می‌یابد. به این ترتیب، حتی شناخته شده بودن ملودی نیز ممکن است کمک چندانی به بهبود بازیابی آن نکند. زیرا مهارت کاربر در زمزمه عامل تعیین‌کننده است (اونال و دیگران ۲۰۰۳، صص ۱-۲؛ چِن[۱۱] ۲۰۱۰، ص ۱؛ کائو[۱۲] و دیگران، ۲۰۱۳، ص ۱؛ ما[۱۳] ۲۰۱۰، ص ۱؛ یانگ، چِن و وانگ، ۲۰۱۰، ص ۱۱).

گِرسیک[۱۴] (۲۰۰۵، ص ۱)، کیفیت تجهیزات ضبط زمزمه را نیز مؤثر دانسته است. شاید به همین دلایل باشد که ایرادها باقی و اجتناب‌ناپذیر (هاس[۱۵] و پُلاستری[۱۶] ۲۰۰۱، ص ۶۶، دوک[۱۷] و دیگران، ۲۰۰۸، ص ۱۰۹۰؛ سوزوکی[۱۸] و دیگران، ۲۰۰۹، ص ۴۹۸) دانسته شده‌اند. از این رو، ضبط زمزمه‌ها در این پژوهش به وسیله گوشی همراه انجام شد تا رفتار سامانه مفروض بازیابی زمزمه-محور اطلاعات موسیقایی در طبیعی‌ترین شرایط ممکن شبیه‌سازی شود. به هر حال، نمی‌توان انکار کرد که اگر فرآیند ضبط در یک استودیوی مجهز به اتاق آکوستیک و تجهیزات صدابرداری با کیفیت انجام می‌شد، امکان به حداقل رساندن خشه[۱۹] و تولید اسناد ایکس اِم اِل غنی‌تر فراهم می‌بود. اما در آن صورت، پژوهشگر نمی‌توانست یافته‌های تحقیق را به شکلی منطقی و موجه به

---

۱ Unal
۲ Chew
۳ Brown
۴ Vaughn
۵ Shih
۶ Zhang
۷ Kuo
۸ Corso
۹ Lewis
۱۰ Kosugi
۱۱ Chen
۱۲ Kao
۱۳ Ma
۱۴ Gersic
۱۵ Haus
۱۶ Pollastri
۱۷ Duc
۱۸ Suzuki
۱۹ Noise

[Downloaded from system.khu.ac.ir on 2024-05-06]
[DOR: 20.1001.1.24237418.1396.4.4.6.3]

شرایط معمول تعمیم بدهد. در شرایط عادی، کاربری که قصد داشته باشد با استفاده از گوشی هوشمند متصل به وب، قطعه موسیقایی مورد نظر خود را از طریق زمزمه کاوش نماید، لزوماً در فضایی آرام و بی‌تنش پرسش کاوش را عرضه نمی‌کند. ممکن است افراد دیگری در نزدیکی او حضور داشته باشند؛ یا اینکه خود کاربر مستقلاً یا باواسطه در حال حرکت باشد. در هر صورت، او در محیطی آکنده از صداهای طبیعی و غیرِ طبیعی مزاحم کاوش خواهد کرد؛ دقیقاً مشابه وضعیتی که در آن بسیاری از ما اقدام به مکالمه تلفنی می‌کنیم. ضبط زمزمه با گوشی همراه این مزیت را داشت که شرایط آزمون را تا بیشترین حد ممکن به شرایط واقعی پیاده‌سازی سیستم نزدیک کرد.

این پژوهش به طور مشخص دو پیشنهاد را عرضه می‌کند که مجریان برنامه‌های آموزشی و پژوهشی می‌توانند از آنها در جهت تقویت حوزه بازیابی اطلاعات موسیقایی در ایران استفاده کنند:

• طراحی و توسعه سامانه یا نرم‌افزارهای کاربردی‌ای که از طریق تحلیل قطعات موسیقایی و تبدیل آنها به اسناد ایکس اِم اِل امکان ذخیره‌سازی و بازیابی آنها را فراهم کنند؛ و

• معرفی هر چه بیشتر روش معرفی شده در این پژوهش به عنوان روشی بهینه در حوزه بازیابی قطعات موسیقایی از طریق توجه به مقوله بازیابی اطلاعات موسیقایی در برنامه‌های درسی رشته‌های علم اطلاعات و دانش‌شناسی، علوم رایانه، و مهندسی رایانه (گرایش نرم‌افزار)؛ توجه به کاربردهای رایانه در برنامه‌های درسی رشته موسیقی؛ تأسیس گروه‌های تحقیقاتی بازیابی اطلاعات موسیقایی؛ و توسعه دانش تخصصی محققان از طریق شرکت در همایش‌ها و مشارکت در تولید دانش بین‌المللی در زمینه بازیابی اطلاعات موسیقایی.

آنچه به عنوان پیشنهاد برخاسته از پژوهش، قابلیت طرح را دارد، منحصر به فهرستی از تحقیقات می‌شود که همچون اثر حاضر، توجه به خصیصه ویبراتو در آنها، در اولویت باشد. برخی از زمینه‌های مستعد، تا آنجا که پژوهشگر دریافته است، عبارتند از:

• پیش‌نمون‌سازی سامانه ویبراتو-فهم بازیابی زمزمه-محور اطلاعات موسیقایی مبتنی بر تحلیل اسناد ایکس اِم اِل: مطالعه موردی آلبوم‌های مونوفونیک هارمونیکا؛

• پیش‌نمون‌سازی سامانه ویبراتو-فهم بازیابی زمزمه-محور اطلاعات موسیقایی مبتنی بر استاندارد میدی: مطالعه موردی آلبوم‌های مونوفونیک هارمونیکا؛

• پیش‌نمون‌سازی سامانه ویبراتو-فهم بازیابی زمزمه-محور اطلاعات موسیقایی مبتنی بر استاندارد میدی: مطالعه موردی آلبوم‌های هوموفونیک هارمونیکا؛

• پیش‌نمون‌سازی سامانه ویبراتو-فهم بازیابی زمزمه-محور اطلاعات موسیقایی مبتنی بر کُدِ پارسونز: مطالعه موردی آلبوم‌های مونوفونیک هارمونیکا؛

• پیش‌نمون‌سازی سامانه ویبراتو-فهم بازیابی زمزمه-محور اطلاعات موسیقایی مبتنی بر کُدِ پارسونز: مطالعه موردی آلبوم‌های هوموفونیک هارمونیکا؛

• پیش‌نمون‌سازی سامانه ویبراتو-فهم بازیابی زمزمه-محور اطلاعات موسیقایی مبتنی بر پردازش سیگنال برگرفته از بسامد: مطالعه موردی آلبوم‌های مونوفونیک هارمونیکا؛ و

• پیش‌نمون‌سازی سامانه ویبراتو-فهم بازیابی زمزمه-محور اطلاعات موسیقایی مبتنی بر پردازش سیگنال برگرفته از بسامد: مطالعه موردی آلبوم‌های هوموفونیک هارمونیکا.

باید خاطر نشان ساخت که اجرای این پژوهش با محدودیت‌های ویژه‌ای نیز مواجه بوده است؛ به طوری که آن را تحت تأثیر قرار داده است. نقیصه نخست به بستر عرضه گزارش‌ها مربوط می‌شود. گزارش تعداد زیادی از پروژه‌های این حوزه در همایش‌های مرتبط عرضه شده است. مقالات همایش‌ها، ماهیتاً، نخستین گزارش از یک پروژه هستند. روشن است که محققان می‌توانند با تکمیل کار خود سامانه بهتری را توسعه بدهند. از این رو، اگر فرض بر این باشد که گزارش‌های تکمیلی بعداً در مجلات تخصصی عرضه خواهند شد، منطقاً می‌توان پذیرفت که نخستین گزارش حاوی ابهام‌هایی است. دست کم می‌توان ادعا کرد که نخستین گزارش به اعتبار اینکه هنوز نقائص سامانه به طور کامل برطرف نشده است، کمبودهای گمراه‌کننده دارد. در مسیر اجرای پژوهش حاضر، محقق از اینگونه آسیب‌ها به دور نبوده؛ و الگوبرداری از روش‌ها و استنباط از نتایج معرفی شده در متون، به طور ناخواسته به وی جهت داده است. دومین عاملی که فضای کار را محدود کرد، ماهیت تجاری بسیاری از پروژه‌های اجرا شده در این حوزه بوده است. تعداد قابل توجهی از تحقیقات روی سامانه‌های زمزمه-محور را شرکت‌های تجاری تولیدکننده تجهیزات الکترونیکی و رایانه‌ای اجرا کرده‌اند. این شرکت‌ها با هدف انتفاع مادی وارد حوزه مورد بحث شده‌اند. از این رو، در مرحله گزارش، تصویر کاملی از فرآیندها عرضه نکرده‌اند. بنا کردن یک پژوهش روی گزارش‌های ناقصی که هر کدام گوشه‌هایی از فرآیند را ناگفته و نانوشته رها کرده‌اند، دشواری فراوانی را به دنبال دارد.

[Downloaded from system.khu.ac.ir on 2024-05-06]

[DOR: 20.1001.1.24237418.1396.4.4.6.3]

Kim, J. (2010). *HMM-based note onset detection of natural humming for query by humming systems*. Master's Dissertation, Department of Electrical Engineering, School of Electrical Engineering, Korea Advanced Institute of Science and Technology (KAIST), South Korea.

Kosugi, N. (2004). *A study on content-based music retrieval with humming*. PhD Thesis, Keio University, Japan. Retrieved January 3, 2018. http://www.brl.ntt.co.jp/people/nao/doc/d_final_softbook.pdf.

Ma, Y. (2010). *Query By Humming Using HMM-based Approach*. The Ohio State University. Retrieved January 3, 2018. http://web.cse.ohio-state.edu/~may/Project_writeup_CSE_788J04_YiMa.pdf.

Pang, H. S., and Yoon, D. H. (2005). Automatic detection of vibrato in monophonic music. *Pattern Recognition*, 38 (7): 1135-1138.

Pardo, B., and Birmingham, W. P. (2003). Query by Humming: How good can it get? *Proceedings of the Workshop on Music Information Retrieval, Toronto, Canada*, pp. 107-109. Retrieved January 3, 2018. http://www.music-ir.org/evaluation/wp3/wp3_ pardo_query.pdf.

Shih, H-H., Zhang, T., and Kuo, C.-C. J. (1999). Real-time retrieval of songs from musical databases with query-by-humming. *Presented at the International Symposium for Multimedia Signal Processing, Taipei, Taiwan*.

Suzuki, M. and et al. (2009). Novel Tonal Feature and Statistical User Modeling for Query-by-Humming. *Information and Media Technologies*, 4 (2): 498-508. Retrieved January 3, 2018. https://www.jstage.jst.go.jp/article/imt/4/2/4_2_498/_pdf.

Tripathy, A. K. and et al. (2009). Query by Humming System. *International Journal of Recent Trends in Engineering*, 2 (5): 373-379. Retrieved January 3, 2018. http://www.academypublisher.com/ijrte/vol02/ no05/ijrte0205373379.pdf.

Unal, E. and et al. (2003). Creating data resources for designing user-centric front-ends for query by humming systems. In *Proceedings of the 5th ACM SIGMM International Workshop on Multimedia Information Retrieval, Berkeley (California), United States of America*. Retrieved January 3, 2018. http://mcl.usc.edu/wp-content/uploads/2014/01/200510-Creating-data-resources-for-designing-user-centric-front-ends-for-query-by-humming-systems.pdf.

Unal, E. and et al. (2005). Creating data resources for designing usercentric frontends for query-by-humming systems. *Multimedia Systems*, 10 (6): 475-483.

Unal, E., Narayanan, S. S., and Chew, E. (2004). A statistical approach to retrieval under user-dependent uncertainty in query-by-humming systems. In *Proceedings of the 6th ACM SIGMM In-*

References

Arjannikov, T. (2013). *Re: MIR and Harmonica*. E-mail to Dariush Alimohammadi. 6 November 2013.

Brown, J. C., and Vaughn, K. (1993). Pitch center of musical sounds with vibrato. *Journal of the Acoustical Society of America*, 94 (3): 1860.

Brown, S. F. (1991). Determination of Location of Pitch within a Musical Vibrato. *Bulletin of the Council for Research in Music Education*, 108: 15-30.

Chen, R. (2010). Fight Against Variant Tempo: Query by Humming and Clapping. *Presented at the Eleventh International Society for Music Information Retrieval Conference (ISMIR 2010), Utrecht, Netherlands*. Retrieved January 3, 2018. from http://ismir2010.ismir.net/proceedings/late-breaking-demo-10.pdf.

Chmel, F. (1389a). Master Class: Second Part. Translated by Delta Frik. *Harmonica*. Retrieved January 3, 2018. http://iranharmonica.ir/Articles.php?ID=0604. [Persian]

Chmel, F. (1389b). Master Class: Third Part. Translated by Delta Frik. *Harmonica*. Retrieved January 3, 2018. http://www.iranharmonica.ir/ Articles.php?ID=0605. [Persian]

Corso, J. F., and Lewis, D. (1950). Preferred Rate and Extent of the Frequency Vibrato. *Journal of the Acoustical Society of America*, 22 (1): 83.

de la Bandera, C. and et al. (2011). Humming Method for Content-based Music Information Retrieval. *Proceedings of the Twelfth International Society for Music Information Retrieval Conference, (ISMIR 2011), Miami (Florida), United States of America*, pp. 49-54. Retrieved January 3, 2018. http://ismir2011.ismir.net/papers/PS1-2.pdf.

Duc, T. N. T. and et al. (2008). Continuous Pitch Contour as an Improvement Feature for Music Information Retrieval by Humming/Singing. In *Proceedings of the 10th Pacific Rim International Conference on Artificial Intelligence: Trends in Artificial Intelligence, Hanoi, Vietnam*, pp. 1086-1091.

Gersic, T. (2005). Diatonic Key Finding in a Query by Humming System. *Gersic*. Retrieved January 3, 2018. http://gersic.com/pdf/gersic_keyfinding.pdf.

Haus, G., and Pollastri, E. (2001). An Audio Front End for Query-by-Humming Systems. In *Proceedings of the Second International Society for Music Information Retrieval Conference, (ISMIR 2001), Bloomington (Indiana), United States of America*, pp. 65-72. Retrieved January 3, 2018. http://pdf.aminer.org/000/439/488/an_audio_front_end_for_query_by_humming_systems.pdf.

Kao, W-T. and et al. (2013). A Two-stage Query by Singing/Humming System on GPU. *Presented at the Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA), Municipality, Taiwan*.

[Downloaded from system.khu.ac.ir on 2024-05-06]

[DOR: 20.1001.1.24237418.1396.4.4.6.3]

ternational Workshop on Multimedia Information Retrieval, New York (New York), United States of America. pp. 113-118.

Yang, X., Chen, Q., and Wang, X. (2010). A Novel Approach Based on Fault Tolerance and Recursive Segmentation to Query by Humming. *Internatonal Journal of Hybrid Information Technology*, 3 (3): 1-13. Retrieved January 3, 2018. http://www.sersc.org/journals/IJHIT/vol3_no3_2010/1.pdf.

[Downloaded from system.khu.ac.ir on 2024-05-06]

[DOR: 20.1001.1.24237418.1396.4.4.6.3]

Human Information Interaction. Vol. 4, No. 4
http://hii.khu.ac.ir
Original Article

Prototyping a Vibrato-Aware Query-By-Humming (QBH) Music Information Retrieval System for Mobile Communication Devices: Case of Chromatic Harmonica

Keivan Borna: PhD, Kharazmi University,Karaj, Iran.
Dariush Alimohammadi: Assistant Professor of Knowledge and Information Studies, Kharazmi University,Karaj, Iran (Corresponding author). alimohammadi@khu.ac.ir

Abstract

Background and Aim: The current research aims at prototyping query-by-humming music information retrieval systems for smart phones.
Methods: This multi-method research follows simulation technique from mixed models of the operations research methodology, and the documentary research method, simultaneously. Two chromatic harmonica albums comprised the research population. To achieve the purpose of research, 24 homophonic tracks were splitted by using *Helium Audio Splitter* software. The splits were processed by *Sonic Visualiser* software; and 168 XML documents were produced. On the other hand, 4 research participants hummed and recorded splits. Hummed tracks were converted by using *AMR to MP3 Converter* software, processed by *Sonic Visualiser*, and resulted in 672 XML documents. *MATLAB* software was learned by the first group of XML documents (168), and then, processed the second group of XML documents (672) for providing desirable outputs. Outputs were compared by using *Image Comparer* software.
Results: Findings indicated a high degree of similarity (99%) between outputs of two groups of XML documents. It has also been found that the gender and the music skill do not have any impact on the results.
Conclusion: It could be acknowledged that designing query-by-humming systems based on converting audio to XML documents, and document matching, is an appropriate strategy towards developing music retrieval applications for smart phones.

Keywords: Simulation, Query-by-humming music information retrieval system, Vibrato, Mobile communication device, Chromatic harmonica

[Downloaded from system.khu.ac.ir on 2024-05-06]

[DOR: 20.1001.1.24237418.1396.4.4.6.3]