بررسی کارایی رابط کاربر صوتی در سامانه‌های اطلاعات مکانی تحت وب

امیدرضا عباسی*[۱]، علی اصغر آل‌شیخ[۲]

[۱] دانشجوی دکتری سیستم‌های اطلاعات مکانی - دانشکده مهندسی نقشه‌برداری - دانشگاه صنعتی خواجه نصیرالدین طوسی
oabbasi@mail.kntu.ac.ir

[۲] استاد دانشکده مهندسی نقشه‌برداری - دانشگاه صنعتی خواجه نصیرالدین طوسی
alesheikh@kntu.ac.ir

(تاریخ دریافت شهریور ۱۳۹۶، تاریخ تصویب آذر ۱۳۹۶)

چکیده

امروزه نقشه‌ها به طور گسترده به منظور ناوبری در هنگام رانندگی یا دوچرخه‌سواری استفاده می‌شوند. در این موارد، برقراری ارتباط کاربر و نقشه با استفاده از دست غیرممکن و یا خطرناک است. علاوه بر این، افراد کم بینا در دیدن نشان‌ها و علائم کوچک مورد استفاده در صفحه مشکلاتی دارند. با توجه به این که گفتگو یک روش تعامل طبیعی در انسان است، ورودی‌های صوتی می‌توانند از پیچیدگی رابط‌های کاربر سیستم‌های اطلاعات مکانی عمومی بکاهند. از این رو، توسعه یک سامانه اطلاعات مکانی کاربرمحور دارای قابلیت تعامل با افراد کم‌بینا یا رانندگان و دوچرخه‌سواران ضروری می‌نماید. در این مقاله، ابتدا به بررسی سامانه‌های ناوبری موجود پرداخته‌شده است. سپس، نمونه اولیه یک سامانه اطلاعات مکانی تحت وب متن باز با رابط کاربر صوتی توسعه داده شده است. در این سامانه، کاربر بدون نیاز به استفاده از دست و تنها با استفاده از فرمان‌های صوتی به ناوبری نقشه می‌پردازد. به منظور برقراری ارتباط با موتور بازشناسی گفتار در مرورگر از کتابخانه annyang استفاده شده است. از مزایای این رابط کاربر نسبت به موارد مشابه پیشین می‌توان به عدم استفاده از دست، سازگاری با و انطباق بر فناوری‌های روز دنیای وب، عدم نیاز به تجهیزات سخت‌افزاری اضافی و در نتیجه هزینه اندک اشاره کرد. علاوه بر این، قابلیت استفاده از رابط کاربر گرافیکی نیز در این سامانه فراهم شده است. بدین ترتیب، امکان استفاده همزمان از رابط کاربر صوتی و رابط کاربر گرافیکی تعبیه شده است.

واژگان کلیدی: سامانه اطلاعات مکانی تحت وب، رابط کاربر صوتی، بازشناسی گفتار، رابط کاربر چندحالته، تعامل انسان-رایانه

* نویسنده رابط

۱- مقدمه

فضای تعامل میان کاربر و ماشین، رابط کاربر[۱] نام دارد. این عبارت به طور گسترده برای ارتباط میان انسان و هر نوع ماشین به کار می‌رود. امروزه یکی از شایع‌ترین ماشین-های مورد استفاده توسط انسان رایانه‌ها هستند. تعامل میان انسان و رایانه (HCI[۲]) یکی از حوزه‌های تحقیقاتی میان-رشته‌ای است که به مطالعه روش‌های استفاده انسان از رایانه و طراحی آن‌ها می‌پردازد. تعامل کاربر با رایانه از طریق صوت یکی از انواع رابط‌های کاربری است [۱, ۲].

سامانه‌های اطلاعات مکانی تحت وب قادر به گردآوری داده‌ها و لایه‌های مکانی از منابع مختلف و نمایش آن در یک محیط نرم‌افزاری بصری هستند. این محیط نرم‌افزاری غالبا مرورگرها[۳] هستند و محیط انتقال داده‌ها و لایه‌ها از طریق اینترنت صورت می‌گیرد. فراهم آوردن یک رابط کاربرپسند جهت دسترسی آسان به اطلاعات مکانی از طریق مرورگر ضروری است. با این وجود، سامانه‌هایی که تا کنون طراحی شده‌اند، غالبا فاقد سازوکارهایی برای تعامل با کاربران خاص یا وضعیت‌های خاص هستند. برای مثال، افرادی که مشکلاتی در بینایی خود دارند یا چشمان آن‌ها ضعیف است، قادر به مشاهده نقشک‌ها[۴]، که معمولا اندازه‌ای کوچک دارند، نیستند. همچنین، فرض کنید فردی با استفاده از تلفن همراه خود قصد استفاده از یک سامانه اطلاعات مکانی به منظور ناوبری در هنگام دوچرخه سواری دارد. در این صورت، تعامل این کاربر با صفحه کلید مجازی یا استفاده از قلم از طریق استفاده از دست غیرممکن و یا خطرناک است. یکی از رهیافت‌های ممکن در رفع این چالش استفاده از رابط‌های کاربر صوتی[۵] است [۳]. رابط-های کاربر صوتی امکان تعامل کاربر با رایانه از طریق فرمان‌های صوتی را فراهم می‌کنند. از آنجا که گفتگو یک روش طبیعی جهت برقراری تعامل بین انسان‌ها به‌شمار می‌رود، رابط کاربر صوتی می‌تواند از پیچیدگی رابط‌های سامانه‌های اطلاعات مکانی عمومی بکاهد [۴, ۵].

امروزه با پیشرفت فناوری‌ها در زمینه پردازش زبان طبیعی و زبان‌شناسی رایانشی[۶]، روش‌ها و فنونی توسعه یافته که رایانه‌ها را قادر می‌سازند تا زبان کلامی را تشخیص دهند و آن را به متن تبدیل کنند. پیاده‌سازی این فناوری در مرورگرها در حال تکمیل است [۶]. از این رو، فرصت‌های مناسبی جهت پیاده‌سازی رابط‌های کاربر صوتی در نرم‌افزارهای تحت وب در حال ایجاد است.

در این مقاله نمونه اولیه یک سامانه اطلاعات مکانی تحت وب با رابط کاربر صوتی طراحی و پیاده‌سازی شده است. در این سامانه، کاربر قادر است تا با استفاده از فرمان‌های صوتی به ناوبری نقشه بپردازد و لایه‌ها را خاموش و روشن کند. این سامانه از محیط توسعه متن باز MapServer بر روی سرور Apache بهره می‌برد. همچنین به منظور برقراری ارتباط با موتور بازشناسی گفتار[۷] در نرم‌افزار مرورگر از کتابخانه annyang استفاده شده است. علاوه بر این، به منظور رفع کاستی‌های یک رابط کاربر صوتی نظیر عدم توانایی بازشناسی گفتار در محیط‌های شلوغ، یک رابط کاربر گرافیکی نیز برای این سامانه طراحی شده است. لازم به ذکر است که برای تغییر وضعیت از یک رابط به رابط دیگر نیاز به فعال‌سازی وجود ندارد. در واقع، هر دو رابط به طور همزمان فعال و قابل استفاده هستند. از این رو، این سامانه دارای رابط کاربر چندحالته[۸] است. پژوهشگران برای رابط‌های کاربر چندحالته مزایایی نظیر دقت بیشتر و کاربرپسند بودن را نسبت به رابط‌های کاربر رایج برشمرده‌اند [۲, ۷].

ساختار مقاله پیش‌رو بدین صورت است: بخش ۲ به مرور تحقیقات پیشین در این زمینه می‌پردازد. سپس انواع رابط‌های کاربر معرفی و رابط کاربر صوتی شرح داده خواهد شد. بخش ۴ پیش‌نیازهای لازم جهت پیاده‌سازی سامانه را بررسی خواهد کرد. در بخش ۵ به جزئیات پیاده-سازی و معماری سیستم پرداخته خواهد شد. در نهایت بخش ۶ به نتیجه‌گیری می‌پردازد و محدودیت‌های پیش-روی تحقیق را بیان می‌کند.

۲- تحقیقات پیشین

یکی از روش‌های قدیمی ایجاد ارتباط با کاربران نابینا استفاده از نقشه‌های لمسی[۹] است. در این نقشه‌ها، ویژگی-ها و عوارض به صورت فیزیکی قابل لمس هستند. بدین

۱ User Interface (UI)
۲ Human-Computer Interaction (HCI)
۳ Browsers
۴ Icons
۵ Voice User Interface (VUI)
۶ Computational linguistics
۷ Speech recognition engine
۸ Multi-modal user interface
۹ Tactile maps

ترتیب، کاربر با لمس نقشه توسط انگشتان خود پی به وجود عوارض می‌برد. برخی پژوهشگران این روش را با استفاده از قابلیت‌های صوتی ارتقاء داده‌اند [۸, ۹]. با این حال، تولید نقشه‌های فیزیکی هزینه‌بر است. علاوه بر این، کاربران برای درک نقشه‌ها باید آموزش‌های لازم را ببینند.

پرنته و بیشاپ یک سامانه اطلاعات مکانی به نام BATS[۱] برای کاربران با مشکلات بینایی توسعه داده‌اند. رابط کاربر BATS از یک ابزار نشان‌گر[۲] برای ورودی و از قابلیت‌های صوتی و لمسی جهت خروجی استفاده می‌کند. در این سامانه، کاربر با استفاده از یک موشواره[۳]، توپک[۴] و یا اهرمک بازی[۵] به ناوبری نقشه می‌پردازد. هنگامی که نشان‌گر بر روی مرزها یا بر مرکز شهرها حرکت می‌کند، ابزار نشان‌گر که در دست کاربر است، می‌لرزد. علاوه بر این، هنگامی که نشان‌گر بر روی نقشک‌ها قرار می‌گیرد، نام آن نقشک از طریق بلندگو خوانده می‌شود. بدین ترتیب، کاربر متوجه خواهد شد که نشان‌گر در هر لحظه در کجای صفحه قرار دارد. مشکل عمده این رابط کاربر، پیچیده بودن مسئله درک موقعیت نشان‌گر در فضای صفحه توسط کاربر است. به نظر می‌رسد، بسیاری از ارتباطاتی که کاربر با نقشک‌های چنین سامانه‌ای برقرار می‌سازد، اتفاقی باشد. برای مثال، اگر کاربر به دنبال یک نقشک خاص در صفحه باشد باید نشان‌گر را به طور تصادفی در صفحه بچرخاند تا بر روی نقشک موردنظر قرار بگیرد [۱۰].

وانگ و همکاران جهت رفع مشکل مذکور، یک روش دسترسی بلادرنگ خودکار ارائه کرده‌اند که در آن کاربران کم بینا قادر به ناوش نقشه هستند. در این روش، نقشه با استفاده از یک چاپگر لمسی بر روی کاغذ چاپ می‌شود. به طور همزمان یک اسکنر تصویر نقشه را بر روی صفحه نمایش نشان می‌دهد. بدین ترتیب، کاربر با حس لامسه انگشتان در هر لحظه در می‌یابد که در کدام بخش از فضای تصویر قرار دارد. علاوه بر این، ویژگی‌های هرعارضه توسط بلندگوها به کاربر اعلام می‌شود. با این وجود، چنین سامانه‌ای بسیار هزینه‌بر است. از طرفی، کاربران این سیستم ملزم به استفاده از دست در هنگام کار با نقشه هستند. بنابراین، استفاده در هنگام رانندگی یا دوچرخه-سواری ممکن نیست. شکل ۱ کاربر این سامانه را در حال ناوش نقشه نشان می‌دهد [۹].

شکل ۱- سامانه GIS توسعه یافته توسط Wang و همکاران [۹] با رابطه کاربر لمسی - حسی - صوتی

لاورنس و همکاران [۱۱] سامانه‌ای شبیه به سامانه BATS برای نمایش نقشه محوطه دانشگاه Oregon توسعه داده‌اند که در آن از فناوری Flash استفاده شده است. علاوه بر این، در این سامانه، یک شیوه‌نامه جدید شبیه به CSS[۶] تولید شده است تا بتوان از متغیرهای بصری، صوتی و حسی به طور همزمان استفاده کرد. علاوه بر مشکلاتی که برای سامانه BATS ذکر شد، برای این سامانه می‌توان به موانعی همچون استفاده از فناوری Flash اشاره کرد. برای استفاده از این فناوری باید افزونه مناسب بر روی مرورگر نصب گردد. علاوه بر این، بسیاری از مرورگرهای همراه به علت مشکلات امنیتی استفاده از این فناوری را پیشنهاد نمی‌کنند [۱۲].

ویلسون و همکاران یک سامانه اطلاعات مکانی همراه توسعه داده‌اند که در آن از رابط کاربر چندحالته استفاده کرده اند. در این سامانه کاربر قادر به ناوش نقشه و انجام پرسش و پاسخ از طریق رابط صوتی و گرافیکی است. با این حال، این سامانه بر روی سکوی همراه آزمایش شده است. همچنین، برای استفاده از این سامانه باید نرم‌افزار آن بر روی تلفن همراه نصب شود [۴]، در حالی که سامانه ارائه شده در مقاله حاضر بر بستر وب طراحی شده است و نیاز به نصب هیچ‌گونه نرم‌افزار اضافه‌ای نیست.

راشرت و همکاران یک سامانه GIS رومیزی[۷] جهت پشتیبانی مدیریت بحران توسعه داده‌اند. این سامانه از یک

۱ The Blind Audio Tactile Mapping System (BATS)
۲ Pointing device
۳ Mouse
۴ Trackball
۵ Joystick

۶ Cascading Style Sheet (CSS)
۷ Desktop GIS

رابط کاربر صوتی و یک رابط کاربر با قابلیت شناسایی حالت بدن استفاده می‌کند. برای استفاده از این سامانه یک صفحه نمایش‌گر بزرگ مورد نیاز است تا بتواند حالات بدن را به طور همزمان برای کاربران در نظر بگیرد. از طرفی این سامانه بر بستر رومیزی توسعه داده شده است. علاوه بر این، برای تشخیص حالت بدن باید از یک سنسور یا دوربین استفاده کرد [۱۳]. فوهرمان و همکاران نیز سامانه‌ای مشابه با سامانه فوق توسعه داده‌اند [۱۴].

جانستون و همکاران یک نرم‌افزار همراه به نام Multimodal Virtual Assistant (MVA) توسعه داده‌اند که قادر به تعامل با کاربر از طریق حالت بدن، صدا، لمس صفحه و موقعیت تلفن همراه است. از این رو، این نرم‌افزار از نوع بافت‌آگاه به‌شمار می‌رود. در این سانامه، کاربر می‌تواند به جستجوی مکان‌های موردنظر خود در حوالی یک نقطه بر روی یک نقشه پویا بپردازد [۱۵].

علاوه بر این، شرکت‌های تجاری نیز سعی در به‌کارگیری رابط کاربر صوتی در محصولات عمومی خود شده‌اند. به عنوان نمونه، شرکت .Apple Inc در سال ۲۰۱۱ دستیار شخصی هوشمند خود به‌نام Siri را بر روی تلفن همراه هوشمند iPhone 4S معرفی و راه‌اندازی نمود. این نرم‌افزار قابلیت بهبود شناسایی گفتار از طریق الگوریتم‌های پیچیده یادگیری را داراست. با این وجود، این محصول کاملا در انحصار شرکت Apple می‌باشد و تنها بر روی مجموعه سیستم‌های عامل تولید این شرکت نظیر iOS، watchOS، macOS و tvOS قابل استفاده است. شرکت سامسونگ نیز در سال ۲۰۱۷ محصولی مشابه به‌نام Bixby را در تلفن‌های هوشمند S8 و S8+ معرفی کرد.

جدول۱ خلاصه‌ای از انواع رابط‌های کاربر و مزایا و معایب هر یک را نشان می‌دهد.

جدول۱- مروری بر رابط‌های کاربر در سامانه‌های GIS

نوع رابط کاربر	پژوهشگران	مزایا	معایب
لمسی - حسی- صوتی (خروجی)	یاکوبسون [۸] ونگ و همکاران [۹]	برقراری ارتباط مستقیم کاربر با صفحه نمایش	هزینه زیاد، نیاز به تجهیزات سخت افزاری خاص، نیاز به آموزش، استفاده از دست
حسی - صوتی (خروجی)	پرنته و بیشاپ [۱۰]	عدم نیاز به آموزش	هزینه نسبتا زیاد، استفاده از دست، عدم ارتباط مستقیم کاربر با صفحه نمایش
حسی - صوتی	لاورنس و همکاران [۱۱]	عدم نیاز به آموزش، عدم نیاز به استفاده از دست	هزینه نسبتا زیاد، عدم ارتباط مستقیم کاربر با صفحه نماش، استفاده از فناوری قدیمی Flash، نیاز به نصب افزونه بر روی مرورگر
گرافیکی - صوتی	ویلسون و همکاران [۴]	عدم نیاز به آموزش، عدم نیاز به استفاده از دست، هزینه پایین	نصب نرم افزار بر روی تلفن همراه
گرافیکی - صوتی - حالت بدن	راشرت و همکاران [۱۳] فوهرمان و همکاران [۱۴] جانستون و همکاران [۱۵]	استفاده همزمان چند کاربر	نیاز به آموزش حالات بدن، نیاز به سخت افزار تشخیص ژست
صوتی گرافیکی	Apple Inc. Samsung Electronics	قابلیت یادگیری	انحصاری (یا تجاری) بودن، وابسته به سیستم‌عامل یا دستگاه

۳- رابط کاربر صوتی

رابط کاربر در رایانه‌ها از منظر ماهیت به دو نوع سخت‌افزاری و نرم‌افزاری تقسیم می‌شوند. موشواره، صفحه‌کلید[۱] و بلندگو[۲] نمونه‌هایی از رابط‌های سخت‌افزاری و گزینگان‌ها[۳]، پنجره‌ها[۴] و نقشک‌ها نمونه‌هایی از رابط‌های نرم‌افزاری به‌شمار می‌روند. از منظر عملکرد، رابط‌های کاربر انواع مختلفی دارند. رابط کاربر گرافیکی (GUI[۵])، خط فرمانی (CLI[۶]) و صفحه نمایش لمسی[۷] از رابط‌های کاربر رایج به شمار می‌روند [۵].

۱ Keyboard
۲ Speaker
۳ Menus
۴ Window
۵ Graphical User Interface (GUI)
۶ Comman Line Interface (CLI)
۷ Touchscreen

وضعیت‌هایی مانند دوچرخه‌سواری و رانندگی وجود دارند که در آن‌ها اطلاعات موجود در یک نقشه به راحتی برای کاربر در دسترس نیستند و یا دسترسی به آن‌ها با خطر همراه است [۷]. همچنین، با توجه به این که گفتار زبان تعامل طبیعی انسان است، استفاده از رابط کاربر صوتی می‌تواند از پیچیدگی‌های تعامل با نقشه‌ها بکاهد [۴]. رابط کاربر صوتی با بهره‌گیری از فناوری بازشناسی گفتار برای کاربر امکان کنترل نقشه بدون استفاده از دست و یا چشم را فراهم می‌سازد. بدین ترتیب، کاربر می‌تواند با بیان فرمان‌های صوتی از پیش‌شناخته‌شده به تعامل با نقشه بپردازد. از طرفی، رایانه نیز می‌تواند از طریق فناوری تبدیل متن به گفتار با کاربر صحبت کند. با این حال، استفاده از این نوع رابط کاربر در محیط‌های شلوغ و پرصدا می‌تواند در برقراری ارتباط با سامانه اختلال ایجاد نماید [۱۶]. از دیگر مزایای رابط کاربر صوتی، استفاده همزمان چند کاربر بدون نیاز به فراهم آوردن سخت‌افزار بیشتر است [۱۳]. تحقیقات نشان‌داده‌اند که استفاده از رابط کاربر صوتی باعث افزایش سه برابری سرعت برقراری ارتباط به سامانه خواهد شد [۱۷].

۴- بازشناسی گفتار

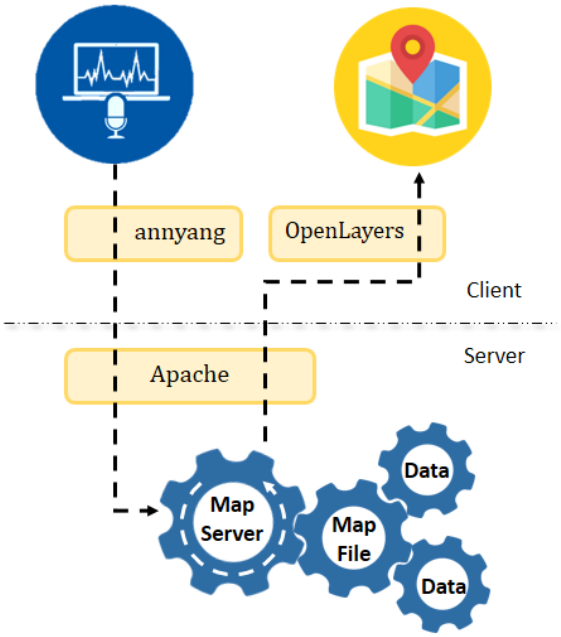
بازشناسی گفتار در حوزه‌های مختلفی نظیر رابط کاربر صوتی، شماره‌گیری صوتی[۱] [۱۸]، کنترل لوازم خانگی[۲] [۱۹]، جستجو و ورود داده [۲۰] کاربرد دارد. بازشناسی گفتار از زیرمجموعه‌های شاخه زبان‌شناسی رایانشی به شمار می‌رود. اخیرا این فناوری در برخی مرورگرها پیاده‌سازی شده است. با استفاده از رابط برنامه-نویسی کاربردی (API[۳]) WebSpeech می‌توان از داده‌های صوتی در نرم‌افزارهای تحت وب استفاده کرد. این API از دو بخش SpeechSynthesis و SpeechRecognition تشکیل شده است. وظیفه SpeechSynthesis تبدیل متن به صوت است. در مقابل، بخش SpeechRecognition وظیفه شناسایی گفتار حاصل از یک ورودی صوتی نظیر میکروفون را بر عهده دارد [۲۱]. هنگامی که یک واژه یا عبارت با موفقیت تشخیص داده شود، یک رشته متن به عنوان نتیجه بازگردانده می‌شود و به تناسب عکس‌العملی اجرا خواهد شد. در این مقاله جهت برقراری ارتباط با سرویس بازشناسی گفتار مرورگر از کتابخانه annyang استفاده شده است. annyang یک کتابخانه متن باز سبک به زبان جاوااسکریپت[۴] است و به طور مستقیم به موتور بازشناسی گفتار مرورگر متصل می‌شود [۲۲].

۵- پیاده‌سازی

در سامانه GIS پیاده‌سازی شده در این مقاله از Apache به عنوان وب سرور و از نرم‌افزار سمت سرور MapServer جهت افزودن قابلیت‌های مکانی استفاده شده است. به منظور نمایش نقشه از کتابخانه سمت کاربر OpenLayers استفاده شده است. همچنین، کتابخانه annyang به عنوان افزونه به صفحه HTML اضافه شده تا قابلیت پردازش گفتار فراهم شود. باید توجه داشت که پردازش و تبدیل گفتار به متن توسط موتور بازشناسی گفتار مرورگر و در فضای ابر[۵] انجام می‌شود، بنابراین سرعت بیشتر در اتصال به شبکه اینترنت موجب افزایش سرعت عملکرد سامانه خواهد شد.

در این سامانه قابلیت‌های اساسی ناوشی GIS نظیر بزرگ‌نمایی و کوچک‌نمایی، ناوبری نقشه و خاموش و روشن کردن لایه‌ها گنجانده شده است. فرایند درخواست نقشه از سمت کاربر تا بازگرداندن نقشه از سمت سرور در شکل۲ نشان داده شده است. مطابق شکل، ابتدا کاربر با استفاده از فرمان‌های صوتی نظیر go left و zoom in درخواست خود را وارد می‌کند. سپس این فرامین به‌عنوان ورودی توابع کتابخانه annyang شناخته شده و به متن تبدیل می‌گردد. توابع موردنظر فراخوانی و درخواست وارد وب سرور می‌شود. در ادامه، وب سرور آن را به سمت نرم-افزار MapServer هدایت می‌کند. نرم افزار MapServer در هسته خود از یک MapFile استفاده می‌کند که در آن مشخصات داده‌های مورد نظر نظیر سیستم تصویر، محدوده داده و نحوه نمایش فراهم شده است. خروجی حاصل از این نرم افزار به صورت یک سرویس نقشه WMS[۶] یا WFS[۷] به کاربر بازگردانده می‌شود.

۱ Voice dialling
۲ Home appliance control
۳ Application Programming Interface (API)
۴ JavaScript
۵ Cloud
۶ Web Map Service (WMS)
۷ Web Feature Service (WFS)



شکل۲- معماری سامانه WebGIS با رابط کاربر صوتی

در کتابخانه annyang فرمان‌های صوتی به صورت یک شیء معرفی می‌شوند. در این مقاله، به منظور ایجاد قابلیت بازشناسی گفتار برای عملگرهای مختلف سامانه از روش‌های زیر استفاده شده است:

- در حالتی که فرمان تنها دارای یک واژه است، می‌توان از متغیرهای دارای نام استفاده کرد. این متغیر می‌تواند در هر جای فرمان قرار بگیرد. به عنوان نمونه، برای خروجی گرفتن از یک گزارش ماهانه می‌توان از عبارت زیر استفاده کرد:

```
>> calculate :month
stats:reportFunction()
```

در این مثال، اگر کاربر قصد گرفتن خروجی گزارش مربوط به ماه آبان را داشته باشد، کافی است عبارت calculate Aban stats را ذکر کند. بدین ترتیب، کتابخانه از month به عنوان یک متغیر استفاده می‌کند که مقدار Aban را در خود ذخیره کرده است.

- در صورتی که نیاز به شناسایی یک فرمان متشکل از چند واژه باشد، می‌توان از علامت * استفاده نمود. به عنوان نمونه عبارت زیر هر چه پس از واژه for ذکر شود را بر می‌گرداند.

```
>> search for
*searchInput:searchFunction()
```

- در صورتی که یک واژه به صورت اختیاری در فرمان وجود داشته باشد، از علامت پرانتز استفاده می‌شود. برای مثال، با استفاده از عبارت زیر می‌توان تمام لایه‌ها را خاموش کرد یا در صورت ذکر شدن نام لایه، تنها آن لایه خاص خاموش شود.

```
>> turn off (layer)
layer:turnOffFunction()
```

در صورتی که کاربر به جای عبارت داخل پرانتز نام لایه موردنظر را ذکر کند، آن لایه خاموش می‌شود، اما اگر واژه‌ای به جای آن نگوید، تمام لایه‌ها خاموش می‌شوند.

علاوه بر این، برای ایجاد فرمان‌های پیشرفته‌تر می‌توان از عبارات باقاعده[۱] نیز بهره برد. برای مثال، در عبارت زیر تنها چهار ماه مشخص‌شده قابل فهم هستند؛ این در حالی است که در صورت استفاده از متغیرهای دارای نام هر واژه‌ای ممکن بود به تابع مورد نظر فرستاده شود.

```
>> 'calculate :quarter stats':
   {'regexp': /^calculate
   (Mehr|Aban|Azar|Dey) stats$/,
   'callback': calculateFunction}
```

شکل۳ نمایی از سامانه توسعه داده شده را نشان می‌دهد. در این سامانه قابلیت‌های اساسی یک WebGIS نظیر pan، zoom و خاموش و روشن کردن لایه‌ها گنجانده شده است. با این حال، قابلیت‌های دیگری نظیر پرسش و پاسخ نیز با استفاده از همین الگوها قابل پیاده‌سازی هستند.

۶- نتیجه‌گیری

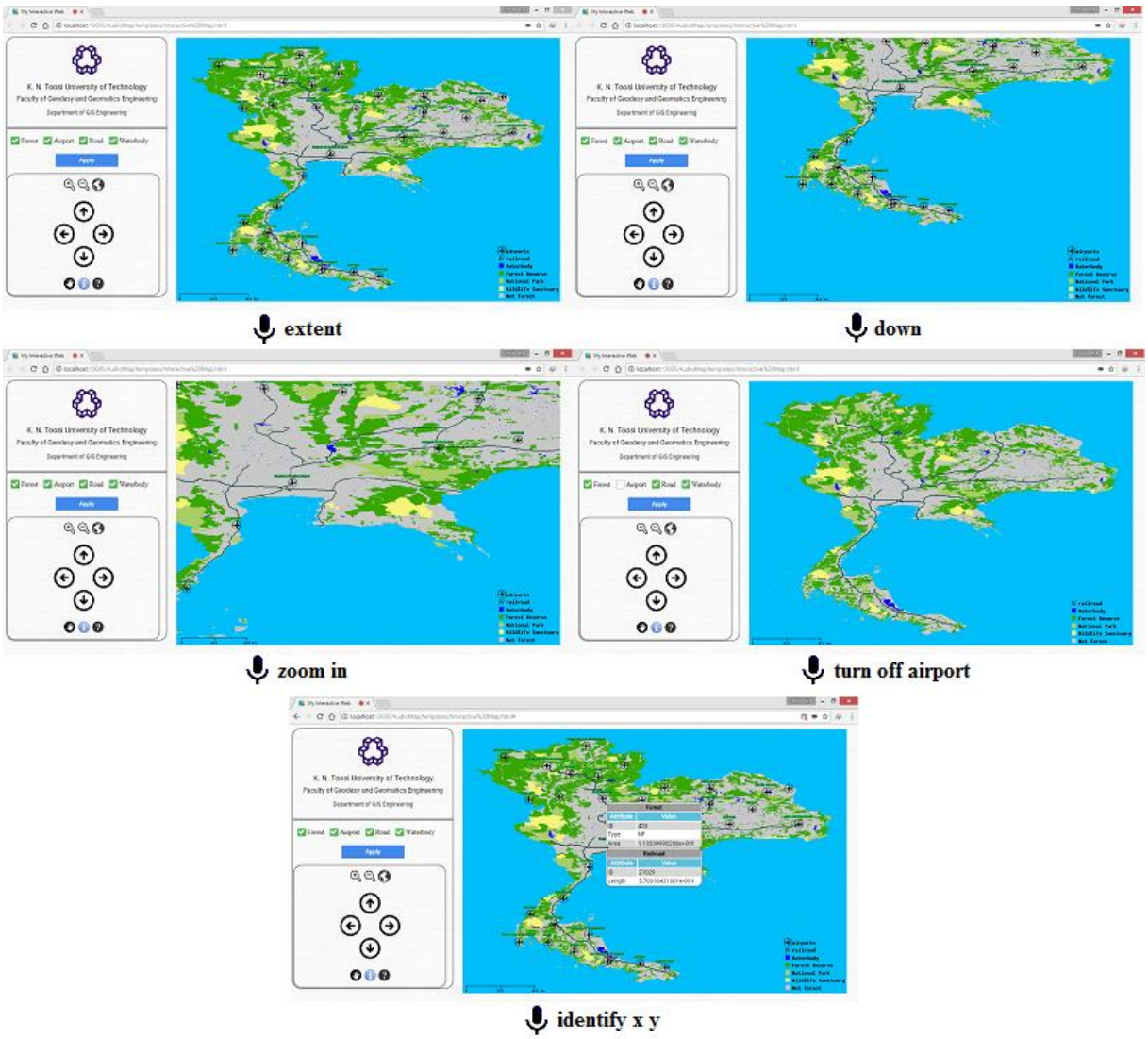
نقشه‌ها روش نمایش بصری داده‌های مکانی هستند. با افزایش حجم و پیچیدگی داده‌های مکانی، طیف کاربران آن‌ها نیز گسترش می‌یابد. از این رو، استفاده از دیگر روش‌های برقراری ارتباط با کاربر نظیر روش‌های شنیداری، گفتاری و یا لامسه نیز اهمیت می‌یابد. علاوه بر این، مواقعی وجود دارد که در آن‌ها کاربر مشغول کار دیگری نظیر رانندگی یا دوچرخه‌سواری است. همچنین ممکن است کاربر دچار ضعف بینایی باشد و کار با نقشک‌ها و علائم کوچک برای او مقدور نباشد.

در این مقاله ابتدا به بررسی رابط‌های کاربر پرداخته شده و پس از آن یک سامانه اطلاعات مکانی تحت وب با استفاده از نرم‌افزارهای متن باز توسعه داده شده است که در آن کاربر قادر به برقراری ارتباط با نقشه از طریق

[۱] Regular expressions

فرمان‌های صوتی است. این سامانه از کتابخانه annyang جهت برقراری با موتور بازشناسی گفتار مرورگر استفاده می‌کند. همچنین یک رابط کاربر گرافیکی برای این سامانه تعبیه شده است که در صورت لزوم، بدون انجام هیچ عمل اضافه‌ای، کاربر می‌تواند به استفاده از دیگر رابط‌های کاربر بپردازد. با توجه به این که سامانه توسعه‌یافته در این مقاله با استفاده از نرم‌افزارهای متن باز ایجاد شده است، هزینه تولید رابط کاربر صوتی نیز کاهش یافته است.

رابط کاربر صوتی به طور ویژه در سامانه‌های ناوبری سه بعدی داخل ساختمان‌ها و فناوری واقعیت مجازی بسیار سودمند خواهد بود. با این وجود، رابط کاربری صوتی با چالش‌هایی روبرو است که باید مرتفع گردند. برای نمونه، در یک محیط شلوغ، سامانه باید بین فرمان-های صوتی و صداهای پس‌زمینه تمایز قائل شود؛ در غیر این صورت، ورودی‌های اشتباه ثبت خواهد شد و عکس-العمل‌های نامناسبی اجرا می‌شود.



شکل۳- نمایی از حالات مختلف سامانه اطلاعات مکانی توسعه یافته با استفاده از رابط کاربر صوتی به همراه فرمان‌های صوتی مربوط

مراجع

[1] C. Pearl, (2016), Designing Voice User Interfaces: Principles of Conversational Experiences: O'Reilly.

[2] R. Edsall, (2017), "Interface Modalities," The International Encyclopedia of Geography.

[3] C. Magnusson, K. Tollmar, S. Brewster, T. Sarjakoski, T .Sarjakoski, and S. Roselier, (2009), "Exploring future challenges for haptic, audio and visual interfaces for mobile maps and location based services," in Proceedings of the 2nd International Workshop on Location and the Web, p. 8.

[4] D. Wilson, J. Doyle ,J. Weakliam, M. Bertolotto, and D. Lynch, (2007), "Personalised maps in multimodal mobile GIS," International Journal of Web Engineering and Technology, vol. 3, pp. 196-216.

[5] M. F. Worboys and M. Duckham, (2004), GIS: a computing perspective: CRC press.

[6] Web Speech API (2017) - Mozilla Development Network.

[7] E. Freeman, G. Wilson, D.-B. Vo, A. Ng, I. Politis, and S. Brewster, (2017), "Multimodal feedback in HCI: haptics, non-speech audio, and their applications," in The Handbook of Multimodal-Multisensor Interfaces, O. Sharon, Bj, S. rn, R. C. Philip, S. Daniel, P. Gerasimos, et al., Eds., ed: Association for Computing Machinery and Morgan & Claypool, pp. 277-317.

[8] R. D. Jacobson, (1998), "Navigating maps with little or no sight: An audio-tactile approach," Proceedings of Content Visualization and Intermedia Representations, pp. 95-102.

[9] Z. Wang, B. Li, T. Hedgpeth, and T. Haven, (2009), "Instant tactile-audio map: enabling access to digital maps for people with visual impairment," in Proceedings of the 11th international ACM SIGACCESS conference on Computers and accessibility, pp. 43-50.

[10] P. Parente and G. Bishop, (2003), "BATS: the blind audio tactile mapping system," in Proceedings of the ACM Southeast regional conference, pp. 132-137.

[11] M. M. Lawrence, N. Martinelli, and R. Nehmer, (2009), "A Haptic Soundscape Map of the University of Oregon," Journal of Maps, vol. 5, pp. 19-29.

[12] R. K. Ramasamy, N. A. Rosly, I. Azizan, and F.-F. Chua, (2017), "CLOUD-BASED WEB SERVICE COMPOSITION USING ACTION SCRIPT," in Proceedings of the 6th International Conference on Computing and Informatics (ICOCI17), Kuala Lampur, Malaysia.

[13] I. Rauschert, P. Agrawal, R. Sharma, S. Fuhrmann, I. Brewer, and A. MacEachren, (2002), "Designing a human-centered, multimodal GIS interface to support emergency management," in Proceedings of the 10th ACM international symposium on Advances in geographic information systems, pp. 119-124.

[14] S. Fuhrmann, A. MacEachren, J. Dou, K. Wang, and A. Cox, (2005), "Gesture and speech-based maps to support use of GIS for crisis management: A user study," AutoCarto 2005.

[15] M. Johnston, J. Chen, P. Ehlen, H. Jung, J. Lieske, A. M. Reddy, (2014) "MVA: The Multimodal Virtual Assistant," in SIGDIAL Conference, 2014, pp. 257-259.

[16] A. Hunter and N. El-Sheimy, (2004), "Mobile GIS and speech recognition," Photogrammetric Engineering & Remote Sensing, vol. 70, pp. 851-860.

[17] S. Ruan, J. O. Wobbrock, K. Liou, A. Ng, and J. Landay, (2016), "Speech Is 3x Faster than Typing for English and Mandarin Text Entry on Mobile Devices," arXiv preprint arXiv:1608.07323.

[18] R. Graham and C. Carter, (2001), "Voice dialling can reduce the interference between concurrent tasks of driving and phoning," International Journal of Vehicle Design, vol. 26, pp. 30-47.

[19] I. Potamitis, K .Georgila, N. Fakotakis, and G. K. Kokkinakis, (2003), "An integrated system for smart-home control of appliances based on remote speech interaction," in INTERSPEECH, 2003.

[20] B. Peters, G. Philip, F. Smith, and D. Crookes, (1986), "Online searching using speech as a man/machine interface," Information processing & management, vol. 25, pp. 391-406.

[21] J. Adorf, (2013), "Web speech api," KTH Royal Institute of Technology.

[22] (https://github.com/TalAter/annyang/blob/master/docs/README.md). API Reference .