ترجمه فا
TarjomeFa.Com

ارائه شده توسط:

سایت ترجمه فا

مرجع جدیدترین مقالات ترجمه شده

از نشریات معتبر

مکان یابی چند-منبعی در محیط های پرانعکاس توسط موسیقی-ریشه و خوشه بندی

چکیده

مکان یابی منابع اکوستیک در محیط های پرانعکاس توسط آرایه های میکروفن, در پردازش سیگنال صوتی, هنوز هم یک کار چالش آور است. در حقیقت, بیشتر مفروضات در مورد مدل های اتخاذ شده رایج در کاربردهای واقعی برآورده نمی شوند. علاوه بر این, در سیستم های عملی, استفاده از معماری های پیچیده و هزینه بر, رایج یا ممکن نیست و به سنکرون سازی دقیق و ترکیب سریع داده های در میان حسگرها نیاز دارد.

در این مقاله, یک رویه چند-مرحله ای نیرومند برای مکان یابی سخنگو (اسپیکر) در اتاق های پرانعکاس معرفی و توصیف شده است. این رویکرد جدید مبتنی بر یک مدل هارمونیک های توزیع شده از تاخیرات زمانی در حوزه فرکانس است و از الگوریتم شناخته شده ROOT-MUSIC (ریشه-موسیقی), بعد از یک پردازش توزیع شده مقدماتی از سیگنال های دریافتی استفاده می کند. سپس موقعیت های نمونه منبع با خوشه بندی برآوردهای TDOA خام تخمین زده می شوند.

ویژگی های اصلی رویکرد پیشنهادی, در مقایسه با راه حل های قبلی, توانمندی ردیابی سخنگوهای متعدد و دقت بالای برآوردگر TDOA شکل بسته هستند.

1. مقدمه

مکان یابی منابع اکوستیک در محیط های پرانعکاس, یک وظیفه مهم در بسیاری از سیستم های خودکار برای نظارت, ویدئو کنفرانس, صحبت با هندزفری [1] است. پارامترهای فضایی به دست آمده در فرآیند موقعیت یابی را می توان در انواع کاربردها استفاده نمود: طنین زدایی صحبت, پیش بینی خطا و تحلیل در دستگاه, اشاره و ردیابی دوربین های تلویزیونی, تایید سخنگو و غیره.

از نقطه نظر پردازش سیگنال, این موضوع, یک راه حل مناسب برای ورودهای متعدد متناظر با سیگنال (ها) و انعکاسات است. سطوح انعکاسی در محیط های بسته معمولاً توسط اطلاعات منابع مجازی مدلسازی می شوند [2] که تعداد آنها معمولاً از اندازه آرایه میکروفن بیشتر می شود. این حقیقت همراه با پهنای باند بسیار بزرگ سیگنال های مورد نظر, تکنیک های پارامتری مورد استفاده در آرایه باند-باریک یا باند-پهن متوسط را در حضور منابع میدان-دور نامناسب می سازد [5][4][3]. به این دلایل, بیشتر رویکردها برای مکان یابی منبع شامل استفاده از تاخیرات زمانی مختلف (تاخیر زمانی ورودی, TDOA, در میان جفت های (دوبل ها) میکروفن های همزمان-واقع شده) می شوند [9][8][7][6]. این فرآیند به بهینه سازی پارامتری پیوسته از سیگنال های جمع آوری شده توسط چند حسگر در یک زمان نیاز دارد.

نوعاً, برآورد TDOA توسط روش های همبستگی متقابل تعمیم یافته انجام می شود [9][6] که برای سادگی و سهولت پیاده سازی, جذاب هستند. در هر صورت, روش های همبستگی متقابل تعمیم یافته, یک مدل منبع-تک را فرض می کنند که در بسیاری از محیط های عملیاتی نوی فراتر از واقعیت است. بنابراین یک مدل متفاوت و راهبردها, برای غلبه بر محدودیت های رویکردهای سنتی نیاز می شوند.

از نقطه نظر طراحی سیستم, کاهش الزامات سنکرون سازی و مسیرهای سیگنال به حداقل, برای کاهش هزینه ها در کاربردهای کنونی, بسیار مهم است.

در این کار, ما یک راهبردی سه-مرحله ای جدید را برای مکانیابی نیرومند چندین سخنگو در اتاق های پرانعکاس پیشنهاد می دهیم. اولین مرحله شامل پیش-سفیدسازی داده ها با استفاده از کدگذاری پیش بینی خطی (LPC) می شود. آثار پیش سفیدسازی سیگنال برای تولید یک غلظت تقریبی از تابع احتمال (تحت یک فرض گاوسی ساده کننده) [7] و کاهش آثار انعکاس است (مثلاً تعداد TDOA چشمگیر که باید برآورد شوند).

در مرحله دوم, TDOAها برای مسیر مستقیم و انعکاسات اولیه (قوی ترین) توسط یک رویکرد پارامتری شکل-بسته بر اساس الگوریتم ریشه-موسیقی برآورد می شوند[12].

در نهایت, مرحله سوم, محتمل ترین موقعیت سخنگوها را با استفاده از خوشه بندی در فضا می یابد که در میان تمام مکان های تخمین زده شده انجام می شود. چگال ترین خوشه ها به عنوان سخنگوهای نماینده انتخاب می شوند و در

نتیجه بیشتر تشخیص های غلط تولید شده توسط برون هشته ها حذف می شود (منابع مجازی, ابهامان مکانی, نویز ضربه ای و غیره).

2.4 خوشه بندی منبع

از TDOAهای تولید شده برای هر جفت دوتایی ها, یک موقعیت منبع نماینده توسط الگوریتم های هندسی کارآمد موجود در نوشته ها [18] محاسبه می شود. برآوردهایی که درون مرزهای اتاق قرار می گیرند در فضا با استفاده از الگوریتم فازی سریع خوشه بندی می شوند [19].

ماکزیمم تعداد Lp برآوردها که می توان به یک سخنگوی تک نسبت داد, به آسانی از هندسه آرایه قابل محاسبه است [18]. بیشتر برآوردهای باقیمانده توسط جفت نمودن TDOAها تولید می شوند که به منابع اکوستیک مختلف اشاره می کند و هیچ معنای فیزیکی ندارد. دیگر مکان های ناصحیح از ابهامات هندسی آرایه, انعکاس و پراش ها ناشی می شود.

در حالیکه برآوردهای نادرست معمولاً خوشه های پراکنده شامل چند نقطه را تولید می کنند, خوشه های چگال دارای یک شکل بیضوی تقریبی, حول سخنگوها تشکیل می شوند. مرکز ثقل خوشه های جمع آوری کننده تعدادی از عناصر نزدیک به Lp در نهایت به عنوان مکان های سخنگو انتخاب می شوند. عملکرد این کار را می توان توسط گسترش فاز خوشه بندی روی چارچوب های زمانی پشت سر هم بهبود بخشید.

4. خلاصه

یک رویکرد جدید برای مکان یابی سخنگو در محیط های پرانعکاس معرفی و توصیف شده است. عملکرد بهتر روش پیشنهادی با توجه به فنون سنتی در شبیه سازی رایانه ای واقعی نشان داده شده است, حتی در حضور چندین سخنگو. هزینه محاسباتی رویه جدید که بر اساس بلوک های ساختمانی استاندارد و بهینه سازی شده است (SVD ,LPC, FFT اندازه-کوچک, یابنده های ریشه چندجمله ای) توسط پردازنده های DSP مدرن قابل پرداخت است و وابسته به پیاده سازی خاص می باشد. خوشه بندی فازی نهایی را می توان در زبان های سطح-بالا روی ایستگاه های کاری هدف-کلی پیاده سازی نمود. در نهایت, الگوریتم پیشنهادی توسط درجه بالایی از انعطاف پذیری مشخص می شود و به آنی هر سفیدسازی ارتقا یافته, برآورد TDOA و فنون خوشه بندی را ترکیب می کند.

ترجمه فا
TarjomeFa.Com

این مقاله، از سری مقالات ترجمه شده رایگان سایت ترجمه فا میباشد که با فرمت PDF در اختیار شما عزیزان قرار گرفته است. در صورت تمایل میتوانید با کلیک بر روی دکمه های زیر از سایر مقالات نیز استفاده نمایید:

✔ لیست مقالات ترجمه شده

✔ لیست مقالات ترجمه شده رایگان

✔ لیست جدیدترین مقالات انگلیسی ISI

سایت ترجمه فا ؛ مرجع جدیدترین مقالات ترجمه شده از نشریات معتبر خارجی