ارائه مدل پیش‌بینی سری زمانی آشوبی برای کد کردن سیگنال صوتی بر پایه برنامه‌سازی ژنتیک بهینه شده با الگوریتم فاخته

سمیه احمدی

دانشگاه سمنان ، S_ahmadi@semnan.ac.ir

چکیده:

در این مقاله یک روش حل برای مدل پیش‌بینی سری زمانی آشوبی ارائه شده است. فضای فاز برپایه خصوصیات سیگنال صوتی آشوبی بازسازی شده و برنامه‌سازی ژنتیک[1] برای حل مدل پیش‌بینی سری زمانی آشوبی روی فضای فاز با بعد جاسازی[2] m و تاخیر زمانی τ معرفی شده است. سپس مدلهای پیش‌بینی سری زمانی آشوبی سیگنال صوتی ساخته می‌شوند. با استانداردسازی پردازش این مدلها و بهینه‌سازی پارامترهای آنها برپایه الگوریتم فاخته، مدل کد کننده سری زمانی آشوبی سیگنال صوتی با توانایی تعمیم خاص بدست می‌آید. در پایان نتایج آزمایش نشان می‌دهند که روش پیشنهادی می‌تواند مدلهای سری زمانی آشوبی موثرتری با دقت کدکننده‌گی بهتر نسبت به روشهای قبلی مانند کدکننده پیش‌بین خطی[3] و همچنین برنامه‌سازی ژنتیک مبتنی بر الگوریتم ازدحام ذرات[4] ارائه داده و بدست دهد.

۱- مقدمه

کدکننده پیش‌بین خطی زمینه پردازش سیگنال صوتی کاربرد وسیع و گسترده‌ای داشته و برپایه همبستگی[5] بین نمونه‌های صوتی با استفاده از مدل خطی با p مقدار نمونه گذشته برای پیش‌بینی نمونه‌های حال و آینده بکارگرفته می‌شود. سپس تعداد زیادی از مطالعات روی خصوصیات غیرخطی سیگنال صوتی انجام شده و به این نتیجه دست یافتند که سیگنال صوتی و دنباله گسسته سیگنال صوتی یک فرآیند خطی بسیار پیچیده بوده و خواص آشوبی آشکاری دارد[۱-۴]. بنابراین غیرخطی بودن سیگنال صوتی استفاده از مدلهای غیرخطی را برای تجزیه و تحلیل آن ایجاب می‌کند. مطالعه مدلهای غیرخطی برای کد کردن سیگنال صوتی یکی از مهمترین مسایل پردازش سیگنال صوتی شده و بکارگیری تئوری آشوب برای مطالعه سیگنال صوتی باعث دستیابی به یافته‌های جدید در مورد سیگنال صوتی شده است[۴-۶]. ثابت شده است که سیگنال صوتی بازسازی شده برای تجزیه و تحلیل خصوصیات عمیق و دقیق آن کمک شایانی می‌کند[۶]. بنابراین مقدمه‌ای از تئوری تجزیه و تحلیل سری زمانی آشوبی علاوه بر اینکه ارزش عملی و تئوری بسیار بالایی برای مطالعه روی خصوصیات آشوبی و روشهای پردازش سیگنال صوتی دارد، میتوان از این تئوری برای ساخت مدلهای پردازش سیگنال صوتی غیرخطی استفاده کرد.

مدلهای شبکه عصبی از عمومی‌ترین روشهای مورد استفاده پردازش سیگنال صوتی غیرخطی می‌باشند[۷-۹]. همچنین از شبکه عصبی در تشخیص صوت[۱۰] و انتقال صوت [۱۱] استفاده می‌شود.

تایسن (Thyssen) و همکارانش بنا را براین گذاشته بودند که سیگنال صوتی از ترکیب توابع خطی و غیرخطی تشکیل شده است. آنها از مدلهای پیش‌بین خطی برای مدلسازی قسمت خطی استفاده کرده و باقیمانده سیگنال که خواص کاملا غیرخطی دارد توسط شبکه‌های عصبی قابل مدل کردن بود. ترکیب مدلهای خطی و غیرخطی دقت مدلسازی را افزایش می‌داد. آقای لوکا (Atthew K. Luka) یک سیستم تشخیص صوت نژادی محلی ارائه دادند که از ضرایب کپستروم فرکانس Mel. سیگنال صوتی که از بانکهای فیلتری MEL بدست می‌آمدند، استفاده کرده و بعنوان ورودی به شبکه عصبی چند لایه اعمال می‌کردند. آنها الگوریتمهای پس انتشار خطای گرادیان مزدوج[6] را با شبکه عصبی برای بهبود سرعت همگرایی مخلوط کردند که توانایی تشخیص خوبی را نشان می‌داد[۱۳]. با بکارگیری تئوری آشوب برای سیگنال صوتی، لی و تانگ (Lee and Tong) [۱۴] و آینا(Aina) و همکارانش[۱۵] شبکه‌های عصبی شعاعی و خطی محلی رایج را برای تجزیه و تحلیل سریهای زمانی بکار برده و به دو نوع پیش‌بین کد صوتی غیرخطی دست یافتند. نتایج بدست

[1] Genetic Programming
[2] Embedding Dimension
[3] Linear Predictive Coding
[4] Particle Swarm Optimization(PSO)
[5] Correlation
[6] Conjugate gradient back propagation algorithm

این مقاله طبق مجوز شماره ۹۳/۱۷/۸۰۵ مرکز منطقه ای اطلاع رسانی علوم و فناوری RICeST زیر نظر وزارت علوم، تحقیقات و فناوری در آن مرکز نمایه و منتشر شده است.