



تخمین فرکانس گام در سیگنال گفتار با استفاده از تبدیل موجک

امین شادروان لاله زاری، جلیل شیرازی

۱- دانشجوی کارشناسی ارشد موسسه آموزش عالی خاوران

۲- استادیار موسسه آموزش عالی خاوران



aminshadravan@yahoo.com

امین شادروان لاله زاری

خلاصه

در این مقاله روشی جدید برای تخمین فرکانس گام بر اساس تبدیل موجک ارائه می شود. فرکانس گام، فرکانس پایه در سیگنال گفتار است. تخمین این فرکانس در اکثر کاربردهای پردازش گفتار ضروری است و از روش های مختلفی برای این منظور استفاده شده است. نتایج بدست آمده از این روش ها برای پردازش سیگنال گفتار مفید است اما کافی نیست. با استفاده از تبدیل موجک در الگوریتم های قبلی، باب جدیدی در آنالیز سیگنال گفتار باز شده است. با تغییر منبع داده ی استفاده شده در هر پژوهش نتایج بدست آمده از هر روش متفاوت می باشد. در این مقاله از بانک اطلاعاتی keele شامل صدای سه زن و سه مرد استفاده شده است. در نهایت از مقایسه ی نتایج بدست آمده از روش پیشنهادی با نتایج بدست آمده از روش AMDF و تعمیم یافته های آن می توان دریافت که روش پیشنهاد شده بر اساس تبدیل موجک عملکرد بهتری نسبت به روش های ذکر شده دارد.

واژه های کلیدی: تخمین فرکانس گام، سیگنال گفتار، تبدیل موجک

۱. مقدمه

روش‌های متعددی برای تخمین فرکانس گام ارائه شده‌است. از آنجا که در روش AMDF^۱ میزان محاسبات کم و دقت آن مناسب است، استفاده از این روش میان پژوهشگران رایج است [1]. مشکلی که در این روش وجود دارد این است که در برخی لحظات دوره‌ی تناوب گام را نصف مقدار واقعی آن یا ضریبی از مقدار واقعی آن تخمین می‌زند. برای رفع این مشکل روش‌های تعمیم‌یافته‌ی آن با عناوین CAMDF^۲ و EAMDF^۳ ارائه شدند [2][3]. اما با وجود بهبود نتایج، تأثیر چندانی بر رفع مشکل نداشتند. در ادامه‌ی روند بهبود نتایج بدست آمده از روش AMDF روش EMDAMDF^۴ ارائه شد که تا حد زیادی این مشکل را حل کرد [4]. در روش empirical mode decomposition سیگنال به تعدادی سیگنال با باند محدود تجزیه می‌شود که به بالا بردن دقت تخمین کمک زیادی می‌کند [5]. اما با وجود این در تخمین فرکانس گام با استفاده از تبدیل موجک، نتایج بدست آمده از تمامی روش‌های ذکر شده دقیق‌تر است. دلایل متعددی مانع تعیین دقیق فرکانس گام سیگنال صوتی می‌شود. اولین دلیل اینکه ساختار شکل‌دهنده بعنوان بیانی از رفتار دینامیکی مجرای عبور جریان هوا در سیستم تولید صوت تأثیر بارزی بر ساختار و شکل موج سیگنال گفتار خواهد داشت. دومین عامل اصلی اختلاف کمی است که می‌تواند مابین سیگنال صوتی صدا دار با دامنه‌ی پایین و سیگنال بی‌صدا وجود داشته باشد. مشکل سوم این است که تعیین دقیق نقاط ابتدایی و انتهایی در یک دوره‌ی تناوب از سیگنال دشوار است [6]. الگوریتم‌های تعیین فرکانس گام سیگنال با استفاده از تبدیل موجک، اغلب این کار را با مشخص کردن لحظات بسته‌شدن دهانه‌ی حنجره در حین عبور جریان هوا از میان آن و در نهایت اندازه‌گیری فواصل زمانی مابین این لحظات انجام می‌دهند [7]. دلیل این امر آن است که در دو مقیاس غالباً متوالی از تبدیل موجک سیگنال، در چنین لحظاتی بیشینه‌های محلی حداکثر میزان همبستگی با یکدیگر را خواهند داشت. در این مقاله ابتدا در بخش دوم روش AMDF و تعمیم‌یافته‌های آن تشریح می‌شوند و روش تخمین فرکانس گام بر اساس تبدیل موجک ارائه می‌شود. در بخش سوم روش پیشنهادی بر اساس تبدیل موجک بیان شده است. نتایج حاصل از آزمایش‌های انجام‌شده بر روی منبع داده‌ی keele در بخش چهارم مقایسه شده‌اند. در نهایت در بخش پنجم نتیجه‌گیری کلی بیان شده‌است.

۲. چند نمونه از روش‌های تخمین فرکانس گام

۲.۱. مروری بر روش AMDF و تعمیم‌یافته‌های آن

روش AMDF توسط Ross et al. در سال ۱۹۷۴ ارائه شد که رابطه آن به این صورت است:

^۱ Average magnitude difference function

^۲ Circular AMDF

^۳ Extended AMDF

^۴ Empirical mode decomposition AMDF

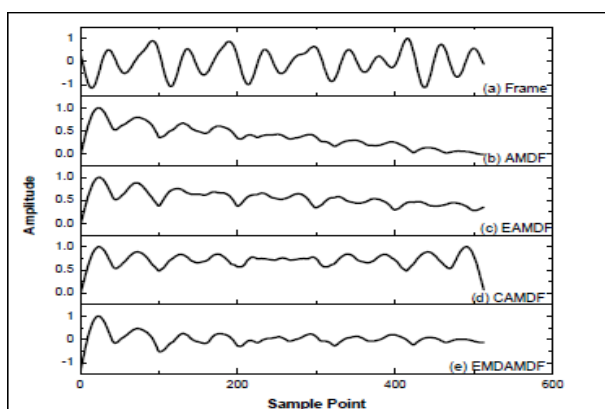
$$D(\tau) = \sum_{n=0}^{N-\tau-1} |x(n) - x(n + \tau)| \quad (1)$$

که در آن $x(n)$ یک قاب سیگنال صدادار است که در یک پنجره‌ی مستطیلی با طول N ضرب شده‌است. τ نمایانگر تاخیر است و محدوده‌ی آن بین صفر تا $N - 1$ می‌باشد. برای یک سیگنال متناوب یا شبه متناوب با دوره‌ی تناوب T_p ، $D(\tau)$ باید دره‌ها را در T_p و ضرائب آن نشان دهد. بطور کلی دوره‌ی تناوب از روش AMDF اینگونه تخمین زده می‌شود:

$$T_p = \arg \min_{\tau \min}^{\tau \max} (D(\tau)) \quad (2)$$

که $\tau \min$ و $\tau \max$ به ترتیب مقادیر بیشینه و کمینه‌ی τ می‌باشند.

از آنجایی که داده‌ها برای محاسبه‌ی $D(\tau)$ در تاخیرهای بالا در رابطه‌ی ۱ کم است، می‌توان روند نزولی را در منحنی تابع دید. بنابراین در پاره‌ای مواقع کمینه‌ها موجب تخمین نادرست گام در روش AMDF می‌شوند. همانطور که در شکل ۱(b) نشان داده شده‌است، بجای دوره‌ی گام صحیح، دو برابر دوره‌ی گام محاسبه می‌شود. در شکل ۱(a) سیگنال صدا مربوط به قاب صدادار سیگنال صدای یک زن است. دوره‌ی گام صحیح در تاخیر ۲۰۳ می‌باشد.



شکل ۱- مقایسه‌ی بین روش AMDF و تعمیم‌یافته‌های آن [4]

برای رفع این مشکل روش CAMDF ارائه شد که رابطه‌ی آن اینگونه است [4]:

$$D_c(\tau) = \sum_{n=0}^{N-1} |x(\text{mod}(n + \tau, N)) - x(n)| \quad (3)$$

که $\text{mod}(n + \tau, N)$ عملگر پیمانه را نشان می‌دهد، به این معنا که $n + \tau$ به پیمانه‌ی N از آنجایی که در این روش تاخیرها یکسان سازی شده‌اند دیگر مشکل روند نزولی در منحنی از بین رفت، اما بدلیل اینکه منحنی CAMDF نسبت به $\tau = N/2$ متقارن است، قادر به تشخیص طول کامل دوره‌ی گام نیست. علاوه بر این دوره‌ی گام و ضرائب آن در این روش چند برابر می‌شوند که می‌تواند خطای جدیدی را موجب شود. با دیدن شکل ۱(d) می‌توان

دریافت که منحنی CAMDF نسبت به $\tau = N/2$ متقارن است و روند نزولی را برطرف کرده است. اما مشکل دو برابر شدن دوره‌ی گام همچنان وجود دارد.

روش EAMDF با دقت بالایی ارائه شده است که رابطه‌ی آن اینگونه توصیف می‌شود [4]:

$$D_e(\tau) = \frac{1}{N-\tau} \sum_{n=-\frac{N}{2}}^{N+\frac{N}{2}-\tau} |x(n) - x(n+\tau)| \quad (4)$$

این روش قادر به حذف روند نزولی و تجمع دوره‌های گام با کاهش تاخیر τ می‌باشد. محاسبات آن با استفاده از سه قاب انجام می‌شود که تابع دیفرانسیلی متفاوت با AMDF واقعی را منجر می‌شود. با وجود اینکه در این روش روند نزولی با موفقیت حذف شده است، اما بدلیل استفاده از سه قاب و بزرگ شدن طول قاب غالباً خطای ایجاد شده، خطای چند برابر شدن دوره‌ی گام است. علاوه بر این خطای بدست آمده از این روش با انجام پیش و پس پردازش‌هایی بدست آمده است. شکل ۱(c) منحنی حاصل از این روش را نشان می‌دهد که همچنان مشکل دو برابر شدن دوره‌ی گام در آن وجود دارد.

روش EMDAMDF با رابطه‌ی زیر توصیف می‌شود [4]:

$$S_{EMDAMDF}(t) = \sum_{n=1}^N C_n(t) \quad (5)$$

در این رابطه $C_n(t)$ تابع AMDF است که با روش EMD اصلاح شده است و N تعداد IMFهاست. IMFها بخش‌های تجزیه شده‌ی یک سیگنال به سیگنال‌های باند محدود می‌باشند. بر خلاف روش AMDF روش EMDAMDF روند نزولی را حذف کرده و در آن طول قاب‌ها غیر قابل تغییر است. به همین دلیل EMDAMDF قادر به تخمین فرکانس گام در هر تاخیری است. خطاهای بوجود آمده در CAMDF و EAMDF نیز در این روش کاهش یافته‌اند. در شکل ۱(e) منحنی حاصل از روش EMDAMDF نشان داده شده است. دوره‌ی گام در تاخیر ۹۲ تخمین زده شده است. در ادامه روشی بر اساس تبدیل موجک ارائه می‌شود و نتایج بدست آمده از آن گواه برتری این روش بر همه‌ی روش‌های منتج از روش AMDF است.

۲,۲ روش تخمین فرکانس گام بر اساس تبدیل موجک

در سال‌های اخیر از تبدیل موجک در کاربردهای پردازش صوت استفاده‌های زیادی می‌شود که موفقیت آمیز هم بوده است. این تبدیل قادر به آنالیز زمان-فرکانس سیگنال صحبت می‌باشد و می‌تواند تغییرات ناگهانی سیگنال را تشخیص دهد. از این رو ابزار مناسبی برای تخمین فرکانس گام در سیگنال گفتار است. در این بخش از مقاله به توصیف تبدیل موجک و روش تخمین فرکانس گام بر پایه‌ی این تبدیل می‌پردازیم.

توابع موجک از باز، بسته و جابجا شدن زمانی یک تابع اولیه مانند ψ تولید می‌شوند که به این صورت توصیف می‌شود:

$$\psi_{a,b}(t) = |a|^{-\frac{1}{2}} \psi\left(\frac{t-b}{a}\right) \quad (6)$$

در واقع در تبدیل موجک هر تابعی مانند $f(t)$ به صورت یک ترکیب خطی از توابع مقیاس و موجک بیان می شود. در حالت کلی پارامترهای مربوط به مقیاس بندی و جابجایی زمانی به این صورت گسسته سازی می شوند:

$$a = a_0^j, \quad b = kb_0 a_0^j \quad (7)$$

که در آن j و k حوزه ی اعداد صحیح را شامل می شوند. بنابراین خانواده ی موجک جدید با پارامترهای گسسته سازی شده به صورت زیر قابل بیان خواهد بود:

$$\psi_{j,k}(t) = a_0^{-\frac{j}{2}} \psi(a_0^{-j} t - kb_0) \quad (8)$$

که $\psi_{j,k}(t)$ در آن موجک مادر، در تبدیل موجک گسسته نامیده می شود. اگر چه این تبدیل را موجک گسسته می نامیم اما هنوز متغیر زمانی آن مقداری پیوسته است. ضرائب تبدیل نیز به این صورت قابل بیان هستند:

$$d_{j,k} = \langle f_w(t), \psi_{j,k}(t) \rangle = \left(\frac{1}{a_0^{\frac{j}{2}}} \right) \int f_w(t) \psi(a_0^{-j} t - kb_0) dt \quad (9)$$

بدین ترتیب هر تابعی مانند $f_w(t)$ بصورت ترکیب خطی از توابع تولید شده توسط موجک مادر قابل بیان خواهد بود:

$$f_w(t) = \sum_j \sum_k \langle f_w(t), \psi_{j,k}(t) \rangle \psi_{j,k}(t) \quad (10)$$

به بیان دیگر هر تابعی را می توان به صورت ترکیب خطی تعداد معدودی (L) از رزولوشن های توابع مقیاس و موجک با پارامترهای $a_0 = 2$ و $b_0 = 1$ بیان نمود:

$$f_w(t) = \sum_{k=-\infty}^{\infty} C_{L,k} 2^{-\frac{L}{2}} \varphi(2^{-L} - k) + \sum_{j=1}^L \sum_{k=-\infty}^{\infty} d_{j,k} 2^{-\frac{j}{2}} \psi(2^{-j} - k) \quad (11)$$

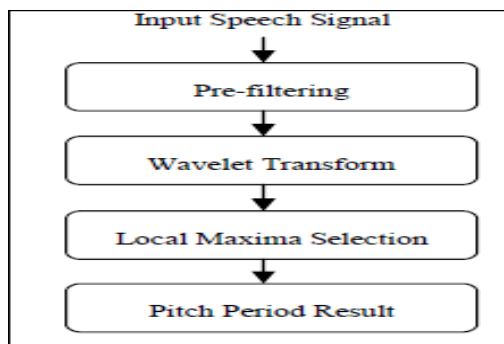
این در حالیکه ضرائب تابع مقیاس به این صورت توصیف می شوند:

$$C_{L,k} = \langle f_w(t), \varphi_{L,k}(t) \rangle = \int f_w(t) 2^{-L/2} \varphi(2^{-L} t - k) dt \quad (12)$$

$$\varphi_{L,k}(t) = 2^{-L/2} \varphi(2^{-L} t - k) \quad (13)$$

۲,۲,۱ روش تخمین فرکانس گام اصلاح شده بر اساس تبدیل موجک

در سال ۲۰۱۰ با استفاده از الگوریتمی که در شکل ۲ آمده است، تخمین فرکانس گام انجام شده است. در این الگوریتم از تکنیک Pre-Filtering استفاده شده است [8].



شکل ۲- الگوریتم تخمین فرکانس گام [8]

منبع داده‌ی استفاده شده در این پژوهش از سیگنال صدای مربوط به چند عبارت و جمله‌ی چینی گفته شده توسط سه زن و دو مرد بدست آمده است. در نهایت درصد خطای بدست آمده با استفاده از این روش با روش تبدیل موجک وزن دهی شده در جدول ۱ مقایسه شده است.

جدول ۱- درصد خطاهای بدست آمده [8]

SNR	Pitch Period Error Rate	
	<i>The Proposed Method</i>	<i>Weighted Wavelet method</i>
Clean	0.23%	0.95%
20	0.65%	1.22%
10	1.35%	3.60%
5	2.87%	5.78%

۲,۲,۲ تخمین فرکانس گام با استفاده از نسل دوم تبدیل موجک

در این روش که در سال ۲۰۱۰ ارائه شده است، ابتدا سیگنال به دو بخش شامل نمونه های زوج و فرد تقسیم می شود. سپس مراحل تخمین فرکانس گام انجام می شود [9]. نتایج حاصل از این روش که بر روی سیگنال های مربوط به تلفظ حروف /a/ و /i/ انجام شده است، در جدول ۲ نشان داده شده است.

جدول ۲- میانگین درصد خطاهای بدست آمده [9]

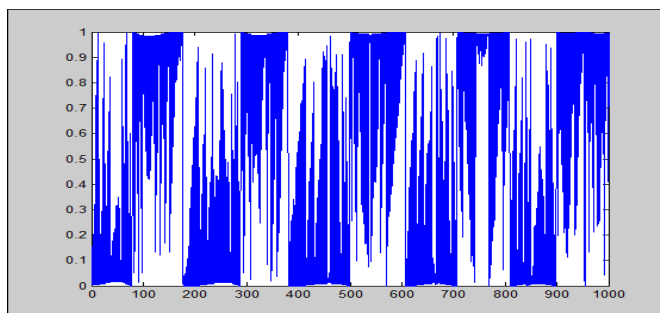
SIGNAL TO NOISE RATIO (DB)	PERCENTAGE OF MEAN ESTIMATION ERROR
Noiseless	0%
25	0.3%
15	1.2%
3	1.8%
0	2.1%

۳. روش پیشنهادی تخمین فرکانس گام بر اساس تبدیل موجک

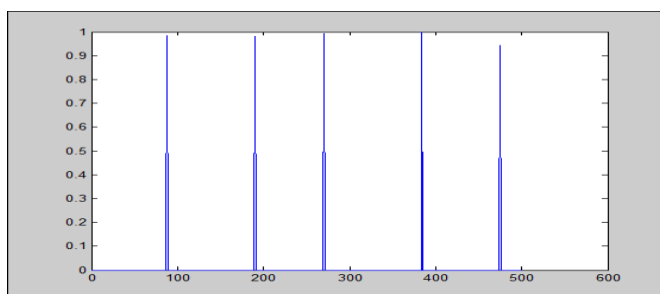
مراحل پیاده سازی الگوریتم تخمین فرکانس گام بر اساس تبدیل موجک به شرح زیر خلاصه می شود [10]:

بخشی از سیگنال، حداقل شامل دو تناوب توسط پنجره قطعه بندی می شود. در مرحله ی بعد، ضرائب تابع مقیاس برای دو مقیاس متوالی i و $i+1$ محاسبه می شود. در ضرائب بدست آمده در هر مقیاس، تمامی بیشینه های محلی که دارای دامنه ای بالاتر از مقدار آستانه ی مطلوب (معمولاً ۰/۷ برابر بزرگترین دامنه ی قطعه) هستند ذخیره می شوند. در بین نقاط ذخیره شده، از هر چندین نقطه ی متوالی که دارای فاصله ای کمتر از حداقل دوره ی تناوب گام در صدای انسان می باشند، نقطه ی مربوط به بزرگترین دامنه حفظ گشته و سایر نقاط مجاور آن حذف می شوند. در هر مرحله ای از محاسبه ی ضرائب تابع مقیاس اگر نقاط بیشینه ی واجدالشرايط در دو مقیاس متوالی در محل های یکسانی نسبت به یکدیگر واقع شده باشند، آنگاه متوسط فواصل مابین این نقاط بعنوان تخمینی از دوره ی تناوب گام قطعه ی موردنظر برگزیده می شود. عمل محاسبه ی ضرائب تابع مقیاس حداکثر تا پنج مرحله ادامه می یابد و در صورتی که تا پایان این عمل، ضرائب تابع مقیاس در هیچ دو مقیاس متوالی شرط فوق را برآورده نکنند، قطعه ی موردنظر بی صدا محسوب خواهد شد.

در این الگوریتم از تابع موجک Haar استفاده شده است. همچنین در الگوریتم ارائه شده برخلاف روش تخمین فرکانس گام بر اساس نسل دوم تبدیل موجک، سیگنال به دو بخش زوج و فرد تبدیل نشده است و بطور کامل در نظر گرفته می شود. سیگنال نیز با قاب های ۲۶/۵ میلی ثانیه ای با گام ۱۰ میلی ثانیه قطعه بندی می شود. نمودار یک قاب و ضرائب نقاط بیشینه ی واجدالشرايط تابع مقیاس آن به ترتیب در شکل های ۳ و ۴ نشان داده شده است.



شکل ۳- یک قطعه‌ی ۲۶/۵ میلی‌ثانیه‌ای از سیگنال صدای یک زن



شکل ۴- ضرائب نقاط پیشینه‌ی واجدالشرايط تابع مقیاس قطعه‌ی نشان داده‌شده در شکل ۳

۴. نتایج آزمایش‌های انجام شده

در این بخش نتایج حاصل از پیاده‌سازی دو روش مبتنی بر تبدیل موجک و AMDF ارائه شده‌است. مجموعه‌ی داده‌های مورد بررسی از بانک اطلاعاتی keele شامل صدای سه زن و سه مرد انتخاب شده‌است. شایان ذکر است که ملاک موجود جهت تعیین و ارزیابی میزان دقت هر یک از روش‌های گفته‌شده جهت تخمین فرکانس گام هریک از نمونه‌ها، مقادیری است که همراه با مجموعه‌ی داده‌ها در اختیار قرار داده شده‌است. ضبط فایل‌ها تحت فرمت SAM، ۱۶ بیتی، استریو و با فرکانس نمونه‌برداری ۲۰ کیلوهرتز انجام پذیرفته‌است [11]. نمونه‌های استفاده شده شامل سه زن با عناوین $(F_1-F_2-F_3)$ و سه مرد با عناوین $(M_2-M_3-M_4)$ می‌باشند. با توجه به تعریف Rabiner [12] اگر اختلاف دوره‌ی تناوب گام بدست آمده از روش پیشنهادی با ارزش‌های صحیح موجود در مرجع داده کمتر از ۱ میلی‌ثانیه باشد، می‌توان از آن صرف نظر کرد و خطای بدست آمده GPE^۶ نام دارد. خطاها در جداول ۳ و ۴ به صورت GPE % ارائه شده‌است.

⁶ Gross pitch error

جدول ۳- GPE% بدست آمده از الگوریتم‌های مختلف AMDF و روش پیشنهادی برای سیگنال صدای سه زن

	F ₁	F ₂	F ₃
AMDF	۲۲/۶۶	۱۱/۹۳	۱۳/۱۱
CAMDF	۹/۳۴	۵/۷۳	۷/۷۵
EAMDF	۷/۵۱	۴/۵۸	۵/۰۳
EMDAMDF	۶/۰۷	۳/۸۴	۴/۶۳
روش پیشنهادی	۲/۶۴	۲/۹۰	۱/۶۷

جدول ۴- GPE% بدست آمده از الگوریتم‌های مختلف AMDF و روش پیشنهادی برای سیگنال صدای سه مرد

	M ₂	M ₃	M ₄
AMDF	۹/۹۲	۲۱/۳۱	۱۹/۵۱
CAMDF	۷/۳۲	۲۲/۰۴	۱۷/۸۷
EAMDF	۳/۰۸	۱۱/۳۳	۹/۴۲
EMDAMDF	۲/۸۱	۹/۱۰	۸/۳۵
روش پیشنهادی	۵/۵۲	۵/۵۲	۷/۳۹

همانطور که در جداول ۳ و ۴ پیداست عملکرد الگوریتم پیشنهادی بر پایه‌ی تبدیل موجک از روش‌های AMDF و تعمیم‌یافته‌های آن بهتر است. خطای حاصل از تخمین فرکانس گام برای سیگنال F₁ با روش پیشنهادی بر اساس تبدیل موجک ۲/۶۴ درصد می‌باشد که نسبت به خطای بدست آمده از روش‌های دیگر به نحو قابل ملاحظه‌ای کمتر است. پیداست که خطای حاصل از تخمین فرکانس گام برای سیگنال F₂ با روش پیشنهادی بر اساس تبدیل موجک ۲/۹۰ درصد می‌باشد که نسبت به خطای بدست آمده از روش EMDAMDF تقریباً ۱ درصد کمتر است و نسبت به روش‌های دیگر با اختلاف حداقل ۱/۵ درصد کمتر است. خطای حاصل از تخمین فرکانس گام برای سیگنال F₃ با روش پیشنهادی بر اساس تبدیل موجک ۱/۶۷ درصد می‌باشد که کمترین میزان خطا در بین تمامی سیگنال‌های مورد آزمایش می‌باشد و نسبت به روش‌های دیگر با اختلاف حداقل ۲/۵ درصد کمتر است. با نگاهی اجمالی به نتایج بدست آمده در جداول ۳ و ۴ این مطلب روشن می‌شود که بطور کلی خطاهای بدست آمده از روش‌های انجام شده برای سیگنال‌های مربوط به زنان کمتر از سیگنال‌های مربوط به مردان است. همانطور که در جدول ۴ مشاهده می‌شود، خطای حاصل از تخمین فرکانس گام برای سیگنال M₂ با روش پیشنهادی بر اساس تبدیل موجک ۵/۵۲ درصد می‌باشد که این میزان خطا در مقایسه با روش EMDAMDF ۲/۷۱ درصد و در مقایسه با روش EAMDF ۲/۴۴

درصد بیشتر می باشد اما نسبت به روش های دیگر با اختلاف حداقل ۱/۵ درصد کمتر است. خطای حاصل از تخمین فرکانس گام برای سیگنال M_3 با روش پیشنهادی بر اساس تبدیل موجک ۵/۵۲ درصد می باشد که این میزان خطا در مقایسه با روش EMDAMDF ۳/۵۸ درصد، در مقایسه با روش EAMDF ۵/۸۱ درصد، در مقایسه با روش CAMDF ۱۶/۵۲ درصد و در مقایسه با روش AMDF ۱۵/۷۹ درصد کمتر می باشد. آنگونه که پیداست برای سیگنال M_4 خطای بدست آمده از روش پیشنهادی بر اساس تبدیل موجک نسبت به دیگر روش ها کمتر است. در جدول ۴، خطای حاصل از تخمین فرکانس گام برای این سیگنال با روش پیشنهادی بر اساس تبدیل موجک ۷/۳۹ درصد می باشد که این میزان خطا در مقایسه با خطاهای بدست آمده برای سیگنال های دیگر با همین روش از بیشترین مقدار برخوردار است. در عین حال، خطا برای این سیگنال با روش پیشنهادی بر اساس تبدیل موجک نسبت به روش های دیگر از کمترین میزان برخوردار است.

۵. نتیجه گیری

در این مقاله، ابتدا دو روش مبتنی بر AMDF و تبدیل موجک مورد بررسی قرار گرفتند. پس از ارائه نتایج بدست آمده، مشاهده شد که هر دو روش در تخمین فرکانس گام از دقت بالایی برخوردارند. روش های CAMDF و EAMDF نتایج روش AMDF را کمی بهبود بخشیده اند اما تأثیر روش EMDAMDF بر بهبود نتایج بیشتر از این دو روش بوده است. این در حالیست که روش تبدیل موجک پیشنهادی بجز در سیگنال M_2 در بقیه سیگنال ها از دقت بالاتری در تخمین فرکانس گام نسبت به سایر روش های ارائه شده برخوردار است. در پایان لازم به ذکر است که نتایج بدست آمده از این روش ها مربوط به سیگنال های بدون نویز می باشد. برای پی بردن به تأثیر روش تبدیل موجک پیشنهادی بر روی سیگنال های نویزی در تخمین فرکانس گام، می توان نتایج بدست آمده از روش های ارائه شده بر روی سیگنال های نویزی با مقادیر SNR متفاوت را، با نتایج این روش مقایسه کرد. این ایده بعنوان گام بعدی در این زمینه پیشنهاد می شود.

۶. مراجع

- [1] M. J. Ross, H. Shaffer, and R. Freudberg, et al. "Average magnitude difference function pitch extractor" IEEE Trans. Acoustics, Speech Signal Processing, vol 22, pp. 353-362, 1974.
- [2] W. Zhang, G. Xu, Y. Wang, "Pitch estimation based on circular AMDF" Proceedings of IEEE ICASSP, pp. 341-344, 2002.
- [3] Ghulam Muhammad, "Noise Robust Pitch detection Based on Extended AMDF" Proceedings of IEEE ISSPIT, pp. 133-138, 2008.
- [4] Yuan Zong; Yumin Zeng; Mengchao Li; Rui Zheng. "Pitch detection using EMD-based AMDF" Intelligent Control and Information Processing (ICICIP) pp. 594-597, 2013.
- [5] N. E. Huang, S. Zheng, and S. R. Long et al. "The empirical mode decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis" Proceedings of Royal Society, pp. 903-995, 1998.

- [6] John.R.Deller, John.H.L.Hanson, John.G.Proakis. "Discrete-Time Processing of Speech Signal " An IEEE Classic Reissue, IEEE Signal Processing Society, 2000.
- [7] M.S. Obaidat, C. Lee, B. Sadoun, D. Nelson. "Estimation of pitch period of speech signal using a new dyadic wavelet algorithm" Information Sciences 119, 21-39, 1999.
- [8] Runshen Cai ; Yaoting Zhu ; Shaoqiang Shi. "A Modified Pitch Detection Method Based on Wavelet Transform" Multimedia and Information Technology (MMIT), 2010 Second International Conference, vol.2, pp. 246 – 249, 2010.
- [9] Ghazvini, M.; Movahedinia, N; Vafaei, A. "pitch period detection using second generation wavelet transform" Signal and Electronic Systems (ICSES) pp. 53-56, 2010.
- [10] Ergun Erc, elebi. "Second generation wavelet transform-based pitch period estimation and voiced/unvoiced decision for speech signals" Applied Acoustics 64, 25–41, 2003.
- [11] F. Plante. "A pitch extraction reference database" Proceedings of Eurospeech, pp.837-840, 1995.
- [12] L. R. Rabiner, M. J. Cheng, and C. A. McGonegal. "A comparative performance study of several pitch detection algorithms" IEEE Trans. Acoustics, Speech and Signal Processing, vol. 24, pp. 399-417, 1976.