

شکل دهی طیف نویز بازسازی در کدگذار ADPCM با پیش بینی پسر و

قاسم علی پور و محمدحسن ساوجی

متناسب با شکل طیف سیگنال صحبت می‌تواند منجر به افزایش کیفیت ادراکی این کدگذار شود.

دو روش معمول برای کاهش اثر خطای بازسازی بر روی کیفیت ادراکی سیگنال بازسازی در کدگذاری ADPCM وجود دارد که روش اول شامل اعمال یک فیلتر وقتی پسین^۵ در خروجی کدگذار است [۵] تا [۷]. در این روش سیگنال بازسازی شده در کدگشا به صورت وقتی متناسب با برآوردی از ویژگی‌های طیفی سیگنال صحبت ورودی فیلتر می‌شود. در روش دوم همراه با کدگذار از یک فیلتر جهت شکل دهی طیف نویز بازسازی استفاده می‌شود [۸] تا [۱۳]. تفکر اصلی در این روش این است که طیف نویز به گونه‌ای شکل دهی شود که شکل طیفی سیگنال ورودی را دنبال کند. در این صورت با توجه به ویژگی پوشاندن نویز^۶ در سیستم شنوایی انسان، نویز بازسازی اثر کمتری بر روی کیفیت ادراکی سیگنال بازسازی شده خواهد داشت. طرح‌های مختلفی برای این هدف پیشنهاد شده است [۸] تا [۱۰] که از این میان می‌توان به طرح پیشنهاد شده در [۱۰] اشاره کرد. این طرح در قالب کدگذار ADPCM با پیش بینی پسر و پیشنهاد و بررسی شده است. در مطالعه اخیر سعی می‌شود که این روش برای اولین بار به کدگذار با پیش بینی پسر و هم تعمیم داده شود.

در بخش دوم روش کلی کدگذاری ADPCM به صورت اجمالی معرفی می‌شود. بخش سوم هم شامل تشریح پیش بینی وقتی پسر و در قالب این کدگذار است. چگونگی اعمال فیلتر شکل دهی نویز و نتایج حاصل به ترتیب در بخش‌های چهارم و پنجم آورده شده‌اند و در انتها هم یک جمع بندی ارائه می‌شود.

۲- روش کدگذاری ADPCM

در یک چندی کننده اسکالر یکنواخت مقدار خطای چندی کردن متناسب با دامنه سیگنال ورودی و تعداد بیت مورد استفاده برای نمایش هر نمونه سیگنال است. به عبارت دیگر در یک مقدار ثابت بیت بر نمونه، افزایش دامنه سیگنال ورودی منجر به افزایش سطح نویز چندی کردن می‌شود. دلیل اصلی بهبود کارایی کدکننده‌های تفاضلی کاهش محدوده پویای سیگنال ورودی به چندی کننده با حذف افزونگی‌های کوتاه مدت^۷ موجود در آن است. به این منظور در روش کدگذاری ADPCM این افزونگی‌ها تا حد ممکن حذف می‌شود [۱] و [۲] که این کار مطابق ساختار کلی شکل ۱ و با چندی کردن وقتی سیگنال خطا یا باقیمانده زیر انجام می‌گیرد

$$e(i) = s(i) - \hat{s}(i) \quad (1)$$

که $\hat{s}(i)$ مقدار پیش بینی شده سیگنال صحبت در لحظه i است.

متناسب با ویژگی‌های غیر ایستان سیگنال صحبت باید الگوریتم‌های وقتی را برای به روز کردن ضرایب فیلتر پیش بینی به کار برد. دو روش کلی به روز کردن این ضرایب عبارتند از پیش بینی وقتی پیشرو^۸ (APF) و

چکیده: تفکر اصلی در روش کدگذاری ADPCM حذف افزونگی‌های موجود در سیگنال صحبت قبل از چندی کردن آن است. علی‌رغم سطح پایین خطای بازسازی، یک ویژگی مهم این کدگذار صاف بودن شکل طیفی سیگنال خطای بازسازی است. در پژوهش جاری سعی می‌شود که با اعمال یک فیلتر تمام صفر به کدگذار ADPCM با پیش بینی پسر و، جهت شکل دهی طیفی نویز بازسازی، کیفیت ادراکی سیگنال بازسازی شده را افزایش داد. این عمل منجر به ایجاد یک تعامل مناسب میان سطح انرژی و شکل طیفی این سیگنال خطا می‌شود. نتایج حاصل بیان گر افزایش کیفیت ادراکی سیگنال بازسازی شده (بر اساس معیار PESQ) در عوض افزایش انرژی خطای بازسازی (بر اساس معیار SNR) است.

کلید واژه: پیش بینی وقتی پسر و، شکل دهی طیفی نویز، کدگذاری صحبت، کدگذار ADPCM، کیفیت ادراکی سیگنال صحبت.

۱- مقدمه

یکی از روش‌های بسیار پر کاربرد در کدگذاری سیگنال صحبت روش شکل موجی^۱ ADPCM است. در این روش، قبل از چندی کردن سیگنال، افزونگی‌های موجود در آن تا حد امکان حذف می‌شود و این کار با چندی کردن سیگنال خطا یا باقیمانده حاصل از پیش بینی انجام می‌گیرد [۱] و [۲]. این گونه الگوریتم‌ها کیفیت بازسازی بالایی را در نرخ بیت‌های متوسط نتیجه می‌دهند. دلیل اصلی بهبود کارایی این گونه کدکننده‌ها کاهش محدوده پویای سیگنال ورودی به چندی کننده (سیگنال باقیمانده) و از آنجا کاهش خطای چندی کردن در یک مقدار ثابت بیت بر نمونه^۳ است. پیش بینی مورد نظر عموماً با یک ترکیب خطی از نمونه‌های قبلی سیگنال صحبت انجام می‌گیرد که در واقع متناظر با مدل کردن فیلتر محافظه صوتی با یک فیلتر خطی FIR است.

علاوه بر سطح نویز بازسازی، کیفیت نهایی سیگنال بازسازی شده متأثر از شکل طیفی آن هم است. در واقع رابطه میان شکل‌های طیفی سیگنال‌های صحبت و خطای بازسازی تأثیر بالایی در کیفیت ادراکی سیگنال بازسازی شده دارد. برای نمونه سطح نویز قابل درک در نقاط فرکانسی که سیگنال صحبت دارای مؤلفه قوی است (از جمله در محل فرمنت^۴ ها) به مراتب بالاتر از نقاط فرکانسی است که در آنها مؤلفه فرکانسی سیگنال صحبت پایین است [۳] و [۴]. با این وجود همچنان که نشان داده می‌شود، یک ویژگی مهم کدگذار ADPCM صاف بودن شکل طیف نویز بازسازی است. لذا شکل دهی طیف سیگنال خطای بازسازی

این مقاله در تاریخ ۹ دی ماه ۱۳۹۱ دریافت و در تاریخ ۲ مهر ماه ۱۳۹۲ بازنگری شد.

قاسم علی پور، گروه مهندسی برق، دانشگاه صنعتی همدان، همدان، کدپستی ۶۵۱۶۹۱۳۴۱۸ (email: alipoor@hut.ac.ir).

محمدحسن ساوجی، دانشکده مهندسی برق و کامپیوتر، اوین، تهران، کدپستی ۱۹۸۳۹۶۳۱۱۳ (email: m-savoji@sbu.ac.ir).

5. Post-Filter
6. Noise Masking
7. Short-Term Redundancy
8. Adaptive Prediction-Forward

1. Adaptive Differential Pulse Code Modulation
2. Dynamic Range
3. Bit per Sample
4. Formant

برآورد $\hat{\sigma}_e$ می‌تواند با میانگین‌گیری بر روی یک قالب سیگنال ورودی انجام شود. در این صورت باز این پارامتر باید جهت بازسازی سیگنال به شکل مناسبی چندی شده و در اختیار کدگشا قرار گیرد. این برآورد در AQB می‌تواند با وزن‌دهی نمایی نمونه‌های قبلی سیگنال باقیمانده چندی شده \hat{e} (که در کدگشا هم موجود است) به صورت زیر انجام شود

$$\hat{\sigma}_e^2 = \alpha \hat{\sigma}_{e_{i-1}}^2 + (1-\alpha) \hat{e}^2(i-1) \quad (3)$$

$0 < \alpha < 1$ پارامتری است که طول حافظه این برآورد را کنترل می‌کند. هر اندازه سیگنال ورودی به چندی‌کننده ناپیستان‌تر باشد مقدار α باید به مقدار صفر نزدیک‌تر شود.

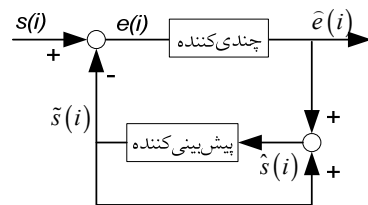
در پژوهش جاری برای چندی‌کردن اسکالر سیگنال باقیمانده در قالب کدگذار ADPCM همواره از روش فوق استفاده می‌شود. در همه آزمایش‌های گزارش‌شده در این پژوهش فرض شده که سیگنال ورودی به چندی‌کننده (سیگنال باقیمانده) دارای توزیع لاپلاسی است. در واقع توزیع‌های یکنواخت، گوسی، گاما و لاپلاسی آزمایش شدند و دیده شد که فرض لاپلاسی بودن توزیع سیگنال باقیمانده برای این چندی‌کننده اسکالر منجر به نتایج نسبتاً بهتری از نقطه نظر کیفیت سیگنال بازسازی شده می‌شود.

۳- پیش‌بینی وقفی - پسرو

معایب اصلی APF عبارتند از تأخیر در کدگذاری و نیاز به بافر کردن داده و نیز ظرفیت اضافی کانال مورد نیاز برای ارسال اطلاعات جانبی. همچنان که پیش از این هم گفته شد، اگر بتوان ضرایب پیش‌بینی کننده بهینه را بر پایه داده چندی‌شده و ارسال شده برآورد کرد می‌توان آنها را در کدگشا بازسازی کرد تا دیگر نیازی به ارسال آنها نباشد. برای این منظور می‌توان از پیش‌بینی وقفی پسرو بر پایه انجام عملیات یکسان پیش‌بینی در کدگذار و کدگشا استفاده کرد [۱] و [۲]. در این روش ضرایب فیلتر پیش‌بینی و نیز چندی‌کننده به صورت متوالی بر اساس سیگنال باقیمانده چندی‌شده (که هم در کدگذار و هم در کدگشا موجود است) به‌روز می‌شوند.

شناخته‌شده‌ترین الگوریتم‌های وقفی متوالی برای تطبیق ضرایب فیلتر عبارتند از الگوریتم‌های کمترین میانگین مربع $(LMS)^6$ و کمترین مربع‌های بازگشتی $(RLS)^7$ [۱۴] و [۱۵]. الگوریتم LMS در مقایسه با الگوریتم RLS از نظر محاسباتی مؤثرتر و از نظر عددی پایدارتر است ولی علاوه بر وابستگی رفتار همگرایی و ویژگی‌های حالت دائم الگوریتم LMS به ویژگی‌های سیگنال ورودی، نرخ همگرایی آن نیز پایین‌تر است. با این وجود سادگی و پایداری بیشتر الگوریتم LMS منجر به کاربرد بیشتر آن در مبحث فیلترهای وقفی شده است.

الگوریتم LMS یک عضو مهم از خانواده الگوریتم‌های وقفی بر پایه گرادیان^۸ است. می‌توان این الگوریتم را به صورت یک حل تقریبی از یک مسأله بهینه‌سازی آماری در نظر گرفت که هدف آن رسیدن به کمینه میانگین مربع خطا^۹ (mmse) است. در تقریب LMS نرمالیزه (NLMS) مورد استفاده در قالب کدگذار ADPCM پسرو، در هر لحظه بردار ضرایب فیلتر پیش‌بینی $\mathbf{w}_i = [w_i(1) \ w_i(2) \ \dots \ w_i(P)]^T$ با استفاده از



شکل ۱: شمای کلی کدگذار ADPCM.

پسرو^۱ (APB)، در APF برآورد پارامترها با توجه به نمونه‌های سیگنال صحبت ورودی انجام می‌گیرد که در کدگشا موجود نیستند. پس به منظور بازسازی سیگنال صحبت لازم است که ضرایب فیلتر پیش‌بینی همراه با سیگنال خطای چندی‌شده به سمت کدگشا فرستاده شوند. ولی در APB پارامترها به صورت متوالی و با توجه به نمونه‌های قبلی سیگنال بازسازی شده‌ای به‌روز می‌شوند که در کدگشا نیز موجود هستند. پس می‌توان پارامترهای پیش‌بینی را به صورت محلی در کدگشا برآورد کرد تا دیگر نیازی به ارسال اطلاعات اضافی نباشد [۱] و [۲].

مطابق ساختار کلی شکل ۱ در هر دو حالت پیش‌بینی پسرو و پیشرو، فیلتر پیش‌بینی درون یک حلقه بازخورد بر روی سیگنال بازسازی شده اعمال می‌شود. این گونه پیاده‌سازی در شمای پسرو الزامی است تا بتوان عملیات به‌روز کردن را در گیرنده و فرستنده بر اساس سیگنال‌های مشابه (سیگنال‌های بازسازی شده در هر دو طرف) انجام داد. با این وجود در صورت استفاده از پیش‌بینی پیشرو می‌توان عملیات تحلیل را مستقیماً بر روی سیگنال صحبت ورودی اعمال کرده و متعاقب آن (بدون ایجاد حلقه بازخورد) سیگنال باقیمانده را چندی کرد. با این وجود همچنان که در ادامه دیده می‌شود استفاده از ساختار بازخورددار شکل ۱ منجر به حداقل شدن سطح نویز چندی‌شدن می‌شود.

به منظور کاهش خطای چندی‌کردن باید یک چندی‌کننده وقفی را برای چندی‌کردن سیگنال خطا به کار برد. در بررسی حاضر از یک چندی‌کننده بی‌حافظه^۲ و لحظه‌ای^۳ استفاده شده که در آن، چندی‌کننده در هر لحظه وابسته به نمونه‌های قبلی و بعدی نیست. در این حالت عبارت چندی‌کننده وقفی می‌تواند بیان‌کننده یک چندی‌کننده با ساختار ثابت و اندازه پله متغیر باشد. مسأله اصلی تطبیق اندازه پله Δ_i به واریانس متغیر سیگنال ورودی به چندی‌کننده (که در کاربرد مورد نظر ما سیگنال باقیمانده می‌باشد) است. در حقیقت مقدار بهینه Δ_{opt} از طریق یک پارامتر ρ که تنها وابسته به تابع چگالی احتمال (pdf) ورودی و bps است، متناسب با انحراف معیار سیگنال ورودی تعیین می‌گردد. مقدار این پارامتر را می‌توان برای یک pdf و bps خاص به صورت عددی بهینه کرد [۱]. بنابراین کارکرد یک چندی‌کننده وقفی را می‌توان به شکل زیر بیان کرد (ρ پارامتر ثابت است)

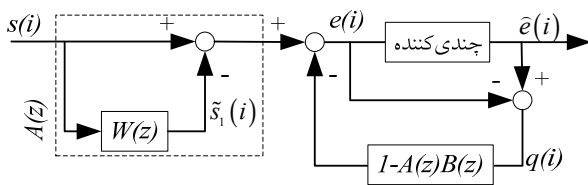
$$\Delta_i = \rho \hat{\sigma}_e \quad (2)$$

$\hat{\sigma}_e^2$ برآوردی از واریانس سیگنال باقی‌مانده e در لحظه i است که برای ورودی ناپیستان متغیر با زمان است. مطابق رابطه بالا مسأله اصلی در چندی‌کردن وقفی، برآورد وقفی و پیوسته $\hat{\sigma}_e$ است.

همانند پیش‌بینی وقفی، چندی‌کننده وقفی را هم می‌توان به دو دسته کلی پیشرو^۴ (AQF) و پسرو^۵ (AQB) دسته‌بندی کرد. در AQF

5. Adaptive Quantization - Backward
6. Least Mean Square
7. Recursive Least Squares
8. Gradient - Based
9. Minimum Mean Squared Error

1. Adaptive Prediction - Backward
2. Memoryless
3. Instantaneous
4. Adaptive Quantization - Forward



شکل ۳: ساختار کدکننده ADPCM همراه با فیلتر شکل‌دهی طیف نویز.

از ساختار مستقیم بدون بازخورد که پیش از این توصیف شد و یک نسخه فیلترشده از نویز چندی‌کردن. به عبارتی دیگر در لحظه i

$$e(i) = s(i) - \sum_{k=1}^P w_i(k)s(i-k) - \sum_{k=1}^P w_i(k)q(i-k) \quad (10)$$

لذا با توجه به (۱۰) دیده می‌شود که نمونه‌های گذشته نویز چندی‌کردن هم در محاسبه سیگنال باقیمانده دخیل هستند. می‌توان کدگذار ADPCM را به صورت معادل شکل ۲ هم نمایش داد. این ساختار به خوبی تفاوت میان سیگنال باقیمانده در دو روش پیاده‌سازی مستقیم و پیاده‌سازی همراه با بازخورد و نیز اثر نویز چندی‌کردن بر روی ساختار اخیر را نشان می‌دهد.

با توجه به (۷) و (۹) می‌توان نشان داد که

$$\hat{E}(z) = A(z)S(z) + A(z)Q(z) \quad (11)$$

که $\hat{E}(z)$ ، $S(z)$ و $Q(z)$ به ترتیب تبدیل z سیگنال‌های باقیمانده چندی‌شده، صحبت ورودی و خطای چندی‌کردن می‌باشد. به عبارت دیگر سیگنال باقیمانده چندی‌شده برابر است با سیگنال باقیمانده حاصل از اعمال مستقیم فیلتر پیش‌بینی به سیگنال تمیز به اضافه یک بخشی که خود حاصل فیلتر کردن نویز چندی‌کردن با فیلتر آنالیز $A(z)$ است. لذا با توجه به (۶) می‌توان نوشت

$$\hat{S}(z) = S(z) + Q(z) \quad (12)$$

از آنجا که چندی‌کننده مورد استفاده جهت کمینه‌کردن نویز چندی‌کردن بهینه می‌شود، ساختار ADPCM شکل ۱ منجر به بیشینه‌کردن نسبت سیگنال به نویز (SNR) سیگنال بازسازی شده می‌شود. با این وجود از آنجایی که نویز چندی‌کردن دارای یک طیف صاف است، طیف سیگنال خطای بازسازی شده هم صاف خواهد بود. از طرفی دیگر علاوه بر سطح انرژی نویز بازسازی، شکل طیفی آن هم تأثیر بالایی در کیفیت ادراکی سیگنال بازسازی شده دارد. روش‌های مختلفی برای مقابله با این پدیده شناخته شده در کدکننده‌های ADPCM وجود دارند [۵] تا [۱۰]. از جمله در [۱۰] روشی برای کنترل شکل طیفی سیگنال خطای بازسازی در شمای پیشرو پیشنهاد شده است. در ادامه پس از تشریح این روش آن را به شمای پسر هم تعمیم خواهیم داد.

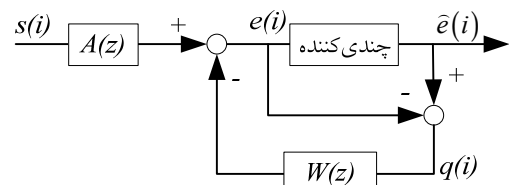
در صورت استفاده از یک فیلتر $B(z)$ برای شکل‌دهی طیف سیگنال خطای بازسازی، سیگنال بازسازی شده را می‌توان به صورت زیر مدل کرد

$$\hat{S}(z) = S(z) + B(z)Q(z) \quad (13)$$

با توجه به (۶)، (۷) و (۱۳) می‌توان نوشت

$$E(z) = A(z)S(z) - [1 - A(z)B(z)]Q(z) \quad (14)$$

این ساختار در شکل ۳ نمایش داده شده است. تنها تفاوت میان این ساختار و ساختار شکل ۲ (که خود معادل ساختار اصلی شکل ۱ است) در وجود فیلتر شکلی‌دهی طیف نویز $B(z)$ است.



شکل ۲: شمای معادل یک کدکننده ADPCM.

الگوریتم تندترین شیب^۱ و با کمینه‌کردن مربع مقدار لحظه‌ای سیگنال باقیمانده چندی‌شده $(\hat{e}^T(i))$ به صورت زیر به روز می‌شود [۱۴] و [۱۵]

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \frac{\gamma\mu}{\hat{\sigma}_{s_i}} \mathbf{s}_i \hat{e}(i) \quad (4)$$

بردار ورودی در لحظه i $\mathbf{s}_i = [\hat{s}(i-1) \ \hat{s}(i-2) \ \dots \ \hat{s}(i-P)]^T$ است. همچنین $\hat{\sigma}_{s_i}^T$ برآوردی از توان سیگنال ورودی به فیلتر در لحظه i می‌باشد که در هر لحظه به صورت زیر به روز می‌شود

$$\hat{\sigma}_{s_i}^T = \beta \hat{\sigma}_{s_{i-1}}^T + (1-\beta) \hat{s}(i-1)^T \quad (5)$$

β عامل تعیین‌کننده حافظه در برآورد مذکور است. می‌توان گفت که در الگوریتم NLMS (که از این پس برای سادگی آن را LMS می‌نامیم) در هر لحظه پارامتر همگرایی متناسب با یک برآورد از انرژی سیگنال ورودی تغییر می‌کند.

۴- شکل‌دهی نویز بازسازی

در این بخش پس از معرفی روش پیشنهادشده در قالب کدگذار ADPCM با پیش‌بینی پیشرو برای شکل‌دهی طیفی سیگنال خطای بازسازی، چگونگی تعمیم این روش به کدگذار ADPCM با پیش‌بینی پسر و تشریح می‌شود.

الف) اعمال فیلتر شکل‌دهی نویز به کدگذار پیشرو

با توجه ساختار کلی کدگذار ADPCM (شکل ۱) می‌توان گفت که سیگنال بازسازی شده $\hat{s}(i)$ حاصل عبور سیگنال باقیمانده چندی‌شده $\hat{e}(i)$ از فیلتر بازسازی $1/A(z)$ است، یعنی

$$\hat{S}(z) = \frac{\hat{E}(z)}{A(z)} \quad (6)$$

فیلتر $A(z) = 1 - W(z)$ تحلیل است. سیگنال باقیمانده چندی‌شده خود برابر است با مجموع سیگنال باقیمانده و خطای چندی‌کردن

$$\hat{E}(z) = E(z) + Q(z) \quad (7)$$

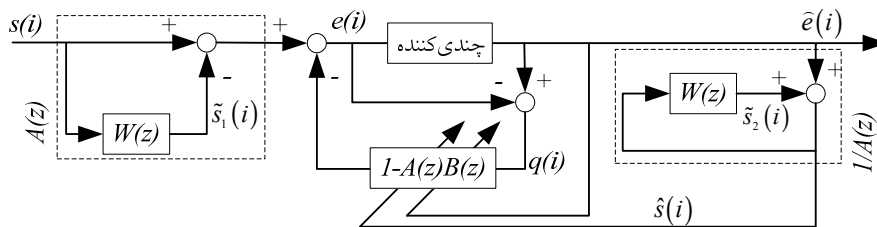
با توجه به (۶) و نیز شکل ۱ می‌توان نوشت

$$E(z) = S(z) - W(z)\hat{S}(z) = S(z) - \frac{1-A(z)}{A(z)}\hat{E}(z) \quad (8)$$

لذا با توجه به (۷) می‌توان نوشت

$$E(z) = A(z)S(z) + [A(z) - 1]Q(z) \quad (9)$$

از طرفی دیگر $A(z)S(z)$ سیگنال باقیمانده ناشی از اعمال فیلتر پیش‌بینی به سیگنال صحبت تمیز است. لذا سیگنال باقیمانده در کدگذار ADPCM برابر خواهد بود با مجموع سیگنال باقیمانده ناشی از اعمال مستقیم فیلتر پیش‌بینی به سیگنال صحبت تمیز (سیگنال باقیمانده حاصل



شکل ۴: ساختار کدگذار ADPCM پسر و همراه با فیلتر شکل دهی طیف نويز.

$aq = [\cdot \cdot]^T$, $\tilde{\sigma}_{s_i}$ = a small positive number
while i

$$\tilde{s}_1(i) = \sum_{k=0}^p w_1(k)s(i-k)$$

$$wq = \sum_{k=0}^p w_1(k)q(i-k)$$

$$e(i) = s(i) - \tilde{s}_1(i) - wq + \sum_{k=0}^p b_1(k)aq(i-k)$$

$$\hat{e}(i) = Q\{e(i)\}$$

$$q(i) = e(i) - \hat{e}(i)$$

$$aq = [q(i) - wq \quad aq(i)]^T$$

$$\tilde{s}_2(i) = \sum_{k=0}^p w_2(k)\hat{s}(i-k)$$

$$\hat{s}(i) = \tilde{s}_2(i) + \hat{e}(i)$$

$$\tilde{\sigma}_{s_i}^2 = \beta \tilde{\sigma}_{s_{i-1}}^2 + (1-\beta)\hat{s}(i)^2$$

$$a_{i+1}(k) = a_i(k) + \frac{\gamma \mu}{\tilde{\sigma}_{s_i}} \hat{e}(i)\hat{s}(i-k)$$

شکل ۵: چگونگی اعمال فیلتر شکل دهی نويز به کدگذار ADPCM با پیش بینی پسر و بر اساس الگوریتم وقتی LMS.

تنها تفاوت در چگونگی به روز کردن ضرایب فیلتر است. در پیش بینی پسر و ضرایب فیلتر پیش بینی به صورت پسر و با استفاده از نمونه های قبلی سیگنال بازسازی شده به روز می شوند. برای تشریح بهتر این رابطه می توان ساختار مفصل تر شکل ۴ را برای این حالت در نظر گرفت. همچنان که در این شکل نشان داده شده است، در پیش بینی پسر و ضرایب فیلتر پیش بینی (و همچنین خود چندی کننده) بر اساس مقادیر گذشته سیگنال باقیمانده چندی شده و سیگنال بازسازی شده به روز می شوند. بازسازی سیگنال هم مستلزم انجام پیش بینی از روی مقادیر قبلی سیگنال بازسازی شده است. به عبارت دیگر علاوه بر پیش بینی که در خلال عملیات تحلیل بر مبنای سیگنال ورودی تمیز انجام می شود، پیاده سازی این کدگذار مستلزم انجام عملیات پیش بینی از روی سیگنال بازسازی شده نیز می باشد. در ساختار شکل ۴ این دو سیگنال پیش بینی به ترتیب با \tilde{s}_1 و \tilde{s}_2 نشان داده شده اند. در حالت کلی این دو سیگنال متفاوت از هم هستند ولی برای حالت $B(z) = 1$ که معادل ساختار اصلی بدون فیلتر شکل دهی طیفی است، این دو سیگنال یکی خواهند بود.

به منظور تعمیم الگوریتم شکل دهی نويز به کدگذار پسر و، بخش دوم سمت راست (۱۴) به صورت زیر بازنویسی می شود

$$\begin{aligned} [1 - A(z)B(z)]Q(z) &= [1 - (1 - W(z))(1 + \bar{B}(z))]Q(z) \\ &= [W(z) - \bar{B}(z)A(z)]Q(z) \\ &= W(z)Q(z) - \bar{B}(z)A(z)Q(z) \end{aligned} \quad (19)$$

که در رابطه فوق $\bar{B}(z) = \sum_k b(k)z^{-k}$ است. با توجه به رابطه اخیر و در صورت استفاده از الگوریتم وقتی LMS می توان ساختار کدگذار ADPCM با پیش بینی پسر و همراه با فیلتر شکل دهی طیف نويز را به صورت شکل ۵ خلاصه کرد.

یک محدودیت مهم بر روی فیلتر $B(z)$ این است که باید پاسخ ضربه $b(k)$ برای $k=0$ برابر با ۱ باشد، یعنی $b(0)=1$. به عبارتی دیگر فیلتر باید به گونه ای باشد که $[1 - A(z)B(z)]$ فیلتری باشد که تنها بر روی مقادیر گذشته نويز عمل کند، زیرا مقدار $q(i)$ تنها زمانی مشخص می شود که مقدار $e(i)$ محاسبه شود که خود مستلزم معلوم بودن فیلتر شکل دهی طیفی نويز است. بنابراین فیلتر $B(z)$ باید دارای ساختار کلی زیر باشد

$$B(z) = 1 + \sum_k b(k)z^{-k} \quad (15)$$

به علاوه طراحی $B(z)$ باید وابسته به ملاحظات ادراکی که هدف اصلی اعمال این فیلتر هم است، باشد.

دو نمونه خاص از این فیلتر عبارتند از $B(z) = 1$ و $B(z) = 1/A(z)$. [۱۰]. نمونه اول معادل ساختار اصلی ADPCM با پیاده سازی بازخوردی عملیات تحلیل (شکل ۱) است. همچنان که پیش از این هم گفته شد این ساختار هر چند منجر به بیشینه کردن SNR می شود ولی شکل طیفی سیگنال نويز بازسازی آن صاف است. نمونه دوم هم معادل همان ساختار شکل ۲ با حذف مسیر بازخورد حول چندی کننده یا به عبارتی دیگر معادل اعمال مستقیم فیلتر تحلیل به سیگنال ورودی است. در این صورت طیف سیگنال بازسازی شده هم شکل طیف خود سیگنال صحبت است، ولی SNR آن طبیعتاً پایین تر است.

با توجه به نکات فوق لازم است که از یک فیلتر بینابینی استفاده کرد که بتواند یک تعامل مناسب میان انرژی و شکل طیفی سیگنال خطای بازسازی ایجاد کند. در [۱۰] بیان شده است که هیچ کدام از فیلترهای تمام قطب بررسی شده این هدف را برآورده نکردند. برای این منظور یک فیلتر تمام صفر درجه دو پیشنهاد شده و برای یافتن ضرایب بهینه این فیلتر در هر لحظه، نخست ضرایب خودهمبستگی ضرایب فیلتر $w_i(k)$ به صورت زیر محاسبه می شوند

$$\rho_m = \sum_{k=0}^{P-|m|} w_i(k)w_i(k+|m|) \quad , \quad 0 \leq m \leq 2 \quad (16)$$

سپس ضرایب $b(k)$ با حل دستگاه معادلات خطی زیر حاصل می شوند

$$\sum_{m=1}^2 b(m)\rho_{k-m} = -\rho_k \quad (17)$$

حل این دستگاه به صورت زیر خواهد بود

$$\begin{aligned} b(1) &= \frac{\rho_1(\rho_2 - \rho_1)}{\rho_2^2 - \rho_1^2} \\ b(2) &= \frac{\rho_1^2 - \rho_2\rho_1}{\rho_2^2 - \rho_1^2} \end{aligned} \quad (18)$$

ب) تعمیم فیلتر شکل دهی نويز به کدگذار پسر و

در صورت استفاده از پیش بینی پسر و تمامی مباحث فوق معتبر هستند،

همچنان که نشان داده شد در صورت استفاده از یک چندی‌کننده بهینه، هر چند انرژی نوین بازسازی در این گونه کدگذارها کمینه است ولی با این وجود شکل طیفی آن صاف است. لذا در صورت استفاده مناسب از یک فیلتر جهت شکل‌دهی طیفی سیگنال خطای بازسازی متناسب با شکل طیفی سیگنال ورودی، می‌توان کیفیت ادراکی سیگنال بازسازی شده را افزایش داد. برای این منظور از یک فیلتر تمام‌صفر که پیش از این در قالب کدگذار ADPCM با پیش‌بینی وفقی پیشرو پیشنهاد شده بود، استفاده شد. نتایج به دست آمده نشان دادند که در صورت استفاده از این فیلتر شکل‌دهی طیفی، در عوض بالا رفتن انرژی خطای بازسازی (بر حسب معیار SNR) می‌توان کیفیت ادراکی آن (بر حسب معیار PESQ) را افزایش داد. این نتیجه‌گیری کلی منطبق بر انتظاری است که از این گونه فیلترها وجود دارد.

مراجع

- [1] N. S. Jayant and P. Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video*, New Jersey: Prentice Hall, 1984.
- [2] J. R. Deller et al., *Discrete-Time Processing of Speech Signals*, New York: Macmillan, 1993.
- [3] N. Jayant, J. Johnston, and R. Safranek, "Signal compression based on models of human perception," *Proceeding of the IEEE*, vol. 81, no. 10, pp. 1385-1422, Oct. 1993.
- [4] T. Painter and A. Spanias, "Perceptual coding of digital audio," *Proceeding of the IEEE*, vol. 88, no. 4, pp. 451-515, Apr. 2000.
- [5] J. H. Chen and A. Gersho, "Adaptive postfiltering for quality enhancement of coded speech," *IEEE Trans. Speech and Audio Processing*, vol. 3, no. 1, pp. 59-70, Jan. 1995.
- [6] D. Malah and R. V. Cox, "A generalized comb filter technique for speech enhancement," in *Proc. IEEE ICASSP*, vol. 7, pp. 160-163, May 1982.
- [7] V. Ramanaoorthy, N. S. Jayant, R. V. Cox, and M. M. Sondhi, "Enhancement of ADPCM speech coding with backward-adaptive algorithms for postfiltering and noise feedback," *IEEE J. Select. Areas Commun.*, vol. 6, no. 2, pp. 364-382, Feb. 1988.
- [8] B. S. Atal and M. R. Schroeder, "Predictive coding of speech and subjective error criteria," *IEEE Trans. Acoust. Speech, Signal Process.*, vol. 27, no. 2, pp. 247-254, 1979.
- [9] M. R. Schroeder, B. S. Atal, and J. L. Hall, "Optimizing digital speech coders by exploiting masking properties of the human ear," *J. of Acoust. Soc. Am.*, vol. 66, no. 6, pp. 1647-1652, 1979.
- [10] J. Makhoul and M. Berouti, "Adaptive noise spectral shaping and entropy coding in predictive coding of speech," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 27, no. 1, pp. 63-73, Feb. 1979.
- [11] A. Borowicz and A. Petrovsky, "Signal subspace approach for psychoacoustically motivated speech enhancement," *Speech Communication*, vol. 53, no. 2, pp. 210-219, Feb. 2011.
- [12] E. Ravelli et al., "Joint Optimization of Base and Enhancement Layers in Scalable Audio Coding," *IEEE Trans. Acoust. Speech, Signal Process.*, vol. 21, no. 4, pp. 711-724, Apr. 2013.
- [13] ITU-T, "Recommendation G.711.1, wideband embedded extension for G.711 pulse code modulation," in *ITU-T Recommendation*, Ed. Geneva, Switzerland, 2008.
- [14] S. Haykin, *Adaptive Filter Theory*, Pearson Education India, 2008.
- [15] A. Sayed, *Fundamentals of Adaptive Filtering*, New York, 2003.
- [16] TIMIT, *DARPA TIMIT-Acoustic-Phonetic Continuous Speech Corpus*, [Online].
- [17] ITU-T, "P.862: perceptual evaluation of speech quality (PESQ): an objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech Codexes," in *ITU-T Recommendation*, Ed. Geneva, Switzerland, 2001.

قاسم علی‌پور در سال ۱۳۸۱ مدرک کارشناسی مهندسی مخابرات خود را از دانشگاه تبریز و در سال ۱۳۸۴ مدرک کارشناسی ارشد مهندسی الکترونیک خود را از دانشگاه شهید بهشتی تهران دریافت نمود. وی از مهرماه سال ۱۳۸۴ به مدت دو سال به عنوان کارشناس ارشد ابزار دقیق در مجتمع صنایع پتروشیمی کرمانشاه فعالیت داشته و پس از آن به دوره دکتری مهندسی الکترونیک دانشگاه شهید بهشتی تهران وارد گردید و در سال ۱۳۹۱ موفق به اخذ درجه دکترا در مهندسی الکترونیک از دانشگاه مذکور گردید. نامبرده از سال ۱۳۹۱ به عنوان استادیار در گروه مهندسی برق دانشگاه صنعتی همدان

جدول ۱: نتایج حاصل از اعمال فیلتر شکل‌دهی طیف نوین به کدگذار پسر با الگوریتم وفقی NLMS.

		bps = ۳	bps = ۴	bps = ۵	bps = ۶
NLMS					
(الف)	PESQ	۳,۵۵۹۰	۳,۹۴۱۶	۴,۱۷۵۱	۴,۲۱۳۷
	SNR (dB)	۱۷,۰۸۴۲	۲۱,۱۰۳۶	۲۴,۸۴۲۴	۲۷,۴۳۱۴
NLMS-NS					
(ب)	PESQ	۳,۶۰۹۵	۳,۹۶۱۷	۴,۱۸۷۸	۴,۳۰۱۰
	SNR (dB)	۱۶,۴۶۱۰	۲۰,۴۵۴۸	۲۴,۱۴۰۶	۲۷,۶۵۷۴

۵- نتایج

آزمایش‌های گزارش شده در این مقاله با میانگین‌گیری بر روی ۴۸ فایل صحبت (شامل ۲۴ نمونه ادا شده توسط زن و ۲۴ نمونه ادا شده توسط مرد) از پایگاه داده TIMIT [۱۶] به دست آمده‌اند که هر کدام از این فایل‌ها شامل یک جمله کامل در زبان انگلیسی است. این سیگنال‌ها در اصل با نرخ ۱۶ kHz نمونه‌برداری شده‌اند که در آزمایش‌های اخیر نرخ نمونه‌برداری آنها به ۸ kHz کاهش داده شده است. این کار با اعمال یک فیلتر پایین‌گذر درجه ۲۰ انجام شده و سیگنال‌ها به صورت اسکالر و یکنواخت با ۱۶ بیت چندی شده‌اند.

نتایج آزمایش‌های ما نشان دادند که کدگذار برای $B(z) = 1/A(z)$ (ساختار حلقه باز) ناپایدار است و در واقع در این حالت بالا رفتن سطح نوین بازسازی منجر به ناپایدار شدن کدگذار می‌شود. با این وجود در صورت استفاده از فیلتر تمام‌صفر درجه دو که پیش از این تشریح شد، کدگذار پایدار است. نتایج حاصل از این فیلتر تمام‌صفر و نیز ساختار اولیه بدون فیلتر شکل‌دهی طیفی $(B(z) = 1)$ بر اساس دو معیار SNR و PESQ و برای چند مقدار bps در جدول ۱ با هم مقایسه شده‌اند. در جدول ۱ ساختار اولیه بدون شکل‌دهی نوین و ساختاری که در آن از فیلتر تمام‌صفر مذکور برای شکل‌دهی طیفی نوین استفاده شده است به ترتیب با NLMS و NLMS-NS مشخص شده‌اند. اندازه‌گیری معیار PESQ هم مطابق استاندارد ITU-T P.862 [۱۷] انجام شده که همبستگی بالایی با معیار ذهنی mos^۱ دارد. همچنان که دیده می‌شود استفاده از فیلتر شکل‌دهی طیفی نوین به صورت میانگین منجر به افزایش در معیار PESQ در عوض کاهش در معیار SNR می‌شود. این نتیجه با توجه به نقش فیلتر شکل‌دهی طیفی که پیش از این تشریح شد، قابل توجیه و منطقی است. همچنین نظر به افزایش ناچیز در پیچیدگی عملیاتی الگوریتم ناشی از افزودن فیلتر شکل‌دهی نوین، این مقدار تغییر در نتایج قابل قبول است. به علاوه لازم به ذکر است که فیلتر معرفی شده برای شکل‌دهی نوین به صورت پی‌درپی^۲ پیاده‌سازی می‌شود و لذا افزودن آن منجر به تغییر در تأخیر الگوریتمی کدگذار نمی‌شود. در واقع تأخیر الگوریتمی کدگذار حاصل هنوز برابر با یک نمونه، یا ۱۲۵ میکروثانیه (در فرکانس نمونه‌برداری ۸ kHz) است.

۶- جمع‌بندی

هدف از این پژوهش بررسی روشی برای افزایش کیفیت ادراکی سیگنال بازسازی شده در کدگذار ADPCM با پیش‌بینی وفقی پسر بود.

1. Mean Opinion Score
2. Sequential

محمدحسن ساوجی تحصیلات خود را در مقطع کارشناسی و کارشناسی ارشد مهندسی برق در دانشگاه صنعتی شریف در سال‌های ۱۳۵۱ و ۱۳۵۳ به اتمام رساند. وی در سال ۱۹۷۹ از انستیتو پلی تکنیک Grenoble فرانسه مدرک دکترای مهندسی خود را در رشته الکترونیک و مخابرات دریافت نمود و در سال ۱۹۸۱ در دانشگاه آکسفورد انگلستان دوره فوق دکترا گذراند. ایشان بین سال‌های ۱۹۸۱ تا ۱۹۹۵ در دانشگاه‌ها و مراکز تحقیقاتی مختلف اروپا مشغول تدریس و تحقیق بود. آخرین مسئولیت ایشان در خارج از کشور استاد و مدیر گروه پردازش سیگنال در دانشگاه Santander کشور اسپانیا بوده است. دکتر ساوجی از سال ۱۳۸۰ تاکنون با سمت استاد الکترونیک و مخابرات در دانشگاه شهید بهشتی به فعالیت خود ادامه می‌دهد. برخی از زمینه‌های تحقیقاتی مورد علاقه ایشان عبارتند از: پردازش سیگنال، پردازش تصویر و ویدئو، پردازش سیگنال صحبت و فیلترهای وقتی و غیرخطی.

مشغول به فعالیت است. زمینه‌های تحقیقاتی مورد علاقه‌ی نامبرده عبارتند از: استفاده از روش‌های آماری و ایده‌های نوین پردازش سیگنال دیجیتال در پردازش سیگنال‌های صحبت و تصویر به‌ویژه برای مقاصد فشرده‌سازی، کدگذاری و نويزدایی و نیز پیاده‌سازی سخت‌افزاری الگوریتم‌های مرتبط.

Archive of SID