

1- مقدمه

پایان بحث وکودرها

وکدر Multi-Band Excitation LP

2- وکودر MBE

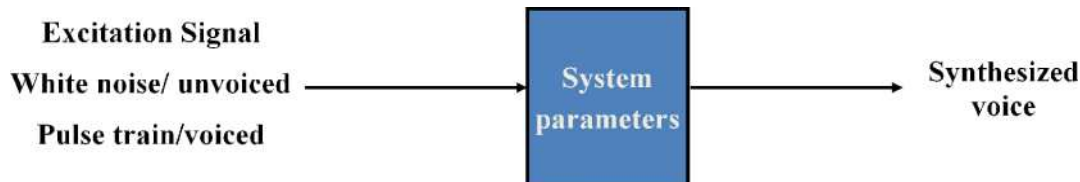
همان طور که می دانیم در وکودرها، گفتار قبل از هر پردازشی به پنجره هایی تقسیم می شود.

سپس پارامترهای تحریک و پارامترهای سیستم برای آن قسمت ها محاسبه می شوند.

- پارامترهای تحریک شامل: شناسایی صدا دار بودن/نبودن و پریود گام
- پارامترهای سیستم شامل: پوش طیفی یا همان پاسخ ضربه سیستم

سپس این اطلاعات به دست آ «ده ارسال می شوند.

در تصویر 1 مشاهده می کنید که در سمت گیرنده، سیگنال تحریک از پارامترها ساخته می شود و سپس با عبور آن از فیلتر، گفتار سنتز شده حاصل می شود.



تصویر 1 – فرآیند تولید گفتار سنتز شده

معمولاً این وکودرهای ساده کیفیت پایینی دارند:

- مدل ها گفتار محدودیت های اساسی دارد.
- تخمین پارامترها ممکن است دقیق نباشد.
- عدم قادر بودن قطار ضربه /نویز سفید برای تولید همه صداها: گفتار سنتز شده با پالس پریودیک تقریباً حالت وزوز دارد و تحریک کاملاً نویز کیفیت بدی دارد.

برای رفع این خاصیت وزوز داشتن صدا استفاده از تحریک هایی است که مخلوطی از پالس پریودیک و نویز هستند.

در این وکودرها پالس های پریودیک و نویز ها با یک نسبت خاصی با هم ترکیب می شوند و این نسبت است که به سمت گیرنده ارسال می شود تا سیگنال تحریک ساخته شود.

به دلیل ایستا بودن سیگنال گفتار در یک فریم، یک پنجره به سیگنال اعمال می شود

$$s_w(n) = w(n)s(n)$$

تبدیل فوری به قطعه پنجره شده $s_w(\omega)$ را می توان به صورت ضرب پوش طیف $H_w(\omega)$ و طیف تحریک $|E_w(\omega)|$ در نظر گرفت (فرمول 1).

$$\hat{s}_w(\omega) = H_w(\omega) |E_w(\omega)| \quad \text{فرمول 1}$$

در اغلب مدل ها $H_w(\omega)$ نسخه صاف شده از طیف اصلی $s_w(\omega)$ می باشد.

پوش طیف باید به صورت دقیق بازنمایی شود.

می توان با اضافه کردن تصمیم های صدا دار بودن/نبودن به صورت وابسته به فرکانس کیفیت را افزایش داد.

در مدل های ساده پیشین، طیف تحریک تماماً بوسیله فرکانس گام و تصمیم صدا دار بدون برای کل طیف ساخته می شود.

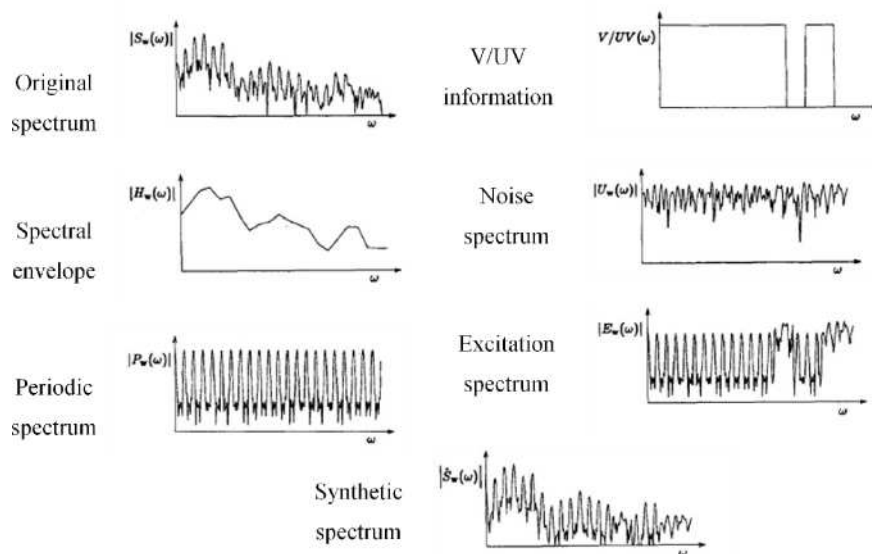
«در مدل MBE طیف تحریک بوسیله فرکانس گام و تصمیم های صدا دار بودن وابسته به فرکانس ساخته می شود.»

در کل برای یک طیف پیوسته به تعداد خیلی زیادی تصمیم صدا دار بودن نیاز داریم.

برای کاهش تعداد بیت های مورد نیاز طیف را به چندین باند فرکانسی تقسیم می کنیم و برای هر باند یک پارامتر باینری صدا دار بودن در نظر گرفته می شود.

تفاوت MBE با مدل های دیگر این است که تعداد باند ها معمولاً زیاد و در حدود 20 در نظر گرفته می شود.

در تصویر 2 خلاصه این عمل را مشاهده می کنید.



تصویر 2 - خلاصه ایجاد سیگنال تحریک در وکودر MBE

پارامترهای MBE عبارتند از:

- پوش طیف
- فرکانس گام
- اطلاعات صدا دار بودن هر باند
- برای فریم های صدا دار، فاز آن باند

پارامترهای پوش طیف بوسیله ضرایب پیشگویی خطی محاسبه می شوند.

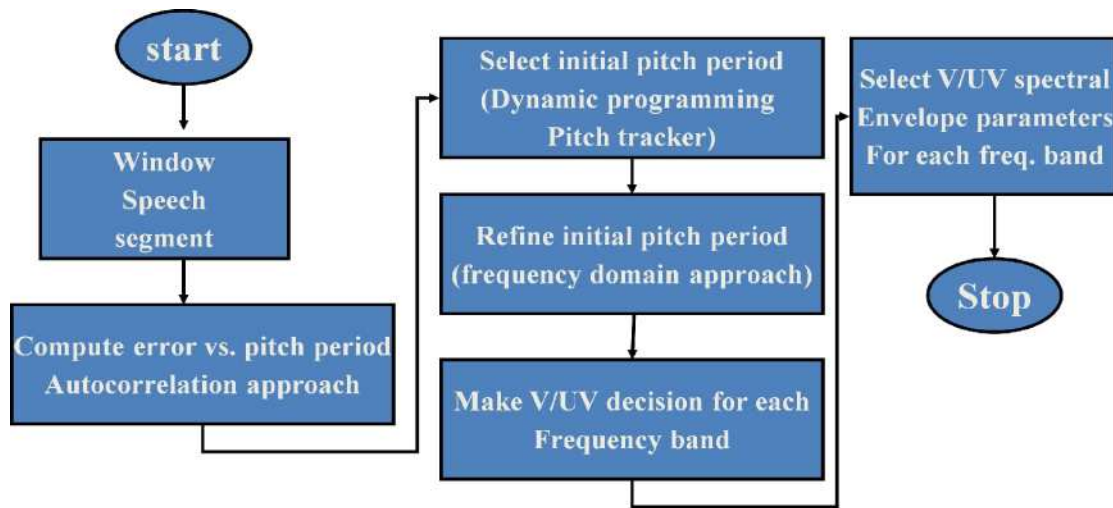
در مدل های ساده معمولاً پارامترهای پوش طیف و پارامترهای تحریک به صورت کاملاً مستقل محاسبه می شوند.

ولی در وکودر MBE این دو به صورت همزمان محاسبه می شوند به این صورت که سعی می کنند سیگنال ستنز شده از لحاظ میانگین مربعات خطا کمترین فاصله را با سیگنال اصلی داشته باشد.

کل مراحل تخمین به دو گام اصلی تقسیم شده اند:

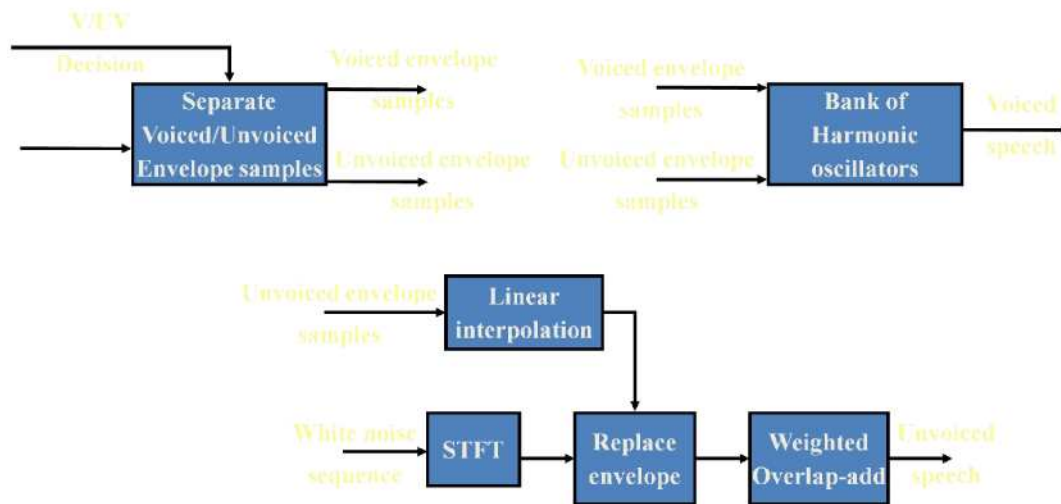
1. در گام اول پریرود گام و پارامترهای پوش طیف تخمین زده می شوند به طوری که خطای بین طیف اصلی و طیف ستنز شده کمینه شود.
2. سپس تصمیم گیری های صدا دار بودن باند ها گرفته می شود.

در تصویر 3 بلوک دیاگرام این وکودر آمده است.



تد

در



تصویر 4 – بلوک دیاگرام کدگذاری MBE

در تصویر 5 جدول تخصیص بیت MBE را مشاهده می کنید.

Parameter	Bits
Fundamental Frequency	9



Magnitude Harmonic	139-94
Harmonic Phase	0-45
Voiced/Unvoiced Bits	12
Total	160

تصویر 5 – جدول تخصیص بیت MBE

4 – خلاصه و نتیجه گیری

در این فصل بحث وکودرها را به پایان رساندیم.

وکودر MBE را بیان کردیم.

6 – منابع درس

- 1- Rabiner, "Fundamentals of Speech Recognition"
- 2- Huang, Acero, "Spoken Language Processing"
- 3- Deller, "Discrete-time processing of speech signals"