بسم الله

دانشکده فنی و مهندسی
گروه مهندسی برق

پایان نامه تحصیلی برای دریافت درجه‌ی کارشناسی ارشد مهندسی برق- مخابرات

چندی سازی برداری وفقی جدیدی برای پارامترهای LPC سیگنال صحبت در حوزه LSF

استاد راهنما:

دکتر سیامک طالبی

مؤلف:

حسن عبداللهی فوزی

تیرماه ۱۳۸۸

قدردانی و سپاس:

سپاس و ستایش خدایی که مرا مستحق دانستن دانست. این رساله کوششی خرد با دانشی ناچیز است که به بهینه سازی چندی ساز برداری ضرائب LPC سیگنال صحبت می‌پردازد. پژوهش حاضر اولین کار تحقیقی در زمینه‌ی بهبود قابل توجه چندی ساز برداری بر اساس یک کتاب کد با اندازه بزرگتر از اندازه معمول می باشد و بدیهی است که دشواری های مضاعف خاص خود را دارد. اما بیش از آن که به بضاعت خویش بنگریم به امید بخشش و راهنمایی خوانندگان عزیز چشم داریم. امیدوارم که این گام اندک، راهگشای قدمهای بعدی در این زمینه و زمینه‌های مشابه آن گردد. در این جا از راهنمایی‌ها و مشورت‌های استاد گرانقدر؛ جناب آقای دکتر سیامک طالبی که چراغی فرا راهم بودند، قدردانی می کنم و از صمیم قلب سپاسی چندباره می گویم. از راهنمایی های برادر عزیزم، مهندس حسین عبداللهی، در تنظیم این پایان نامه صمیمانه تقدیر و تشکر می نمایم.

تقدیم به پدر و مادرم؛

همسر مهربانم؛

خانواده ام که همیشه و در همه حال یاورم بوده اند؛

معلمانم؛

و

مردمانی که آن ها را دوست دارم و با آن ها زندگی می کنم

چکیده:

برای کدینگ سیگنال صحبت، سیگنال صحبت ورودی به منظور اخذ تعدادی پارامتر موجود در طول یک بازه محدود (فریم) پردازش و آنالیز می شود. این پارامترها بصورت مجموعه شاخص های باینری (رشته بیت فشرده شده)، رمزگذاری و چندی سازی می شوند. یکی از پارامترهای اخذ شده‌ی سیگنال صحبت ضرائب مربوط به فیلتر سنتز مرکب[1] می باشد که اطلاعات پوش طیف سیگنال صحبت را به نمایش می گذارد. در نهایت ضرائب این فیلتر که بعنوان ضرائب LPC معروف می باشند چندی سازی شده و به سمت گیرنده فرستاده می شوند. این پایان نامه، به وفقی سازی چندی ساز برداری ضرائب LPC سیگنال صحبت در حوزه LSF و در پهنای باند تلفنی می‌پردازد. همان طور که می دانیم هر چقدر اندازه یک کتاب کد، بزرگتر باشد کیفیت چندی سازی برداری افزایش می یابد. اما این باعث افزایش نرخ بیت و هزینه های محاسباتی (هزینه های جستجو) خواهد شد. چنانچه بتوان با حفظ هزینه های جستجو و نرخ بیت، از یک کتاب کد بزرگتر به چندی سازی برداری بپردازیم، کیفیت خروجی چندی سازی شده تا حد قابل ملاحظه ای بهبود می یابد. در تمامی کارهای وفقی سازی انجام شده تاکنون، از کتاب کدی با اندازه بزرگتر از اندازه معمول استفاده نشده است. بنابراین در این پایان نامه سعی شده است که با استفاده از یک کتاب کد با اندازه بزرگ به بهینه سازی چندی سازی برداری پرداخته شود. به منظور بهره برداری مؤثر از کتاب کد بزرگتر، از ابزار پیش بینی استفاده می کنیم. در عین حال به منظور جلوگیری از افزایش اعوجاج بالا بدلیل کاهش اثرات همبستگی پایین بین فریمی در حضور پیش بینی، از یک چندی سازی Safety Net استفاده خواهیم کرد. در عین حال در دو حالت دیگر ابزار پیش بینی را کنار گذاشته و در یک حالت از یک شبکه عصبی پرسپترون و در حالت دیگر از ایده مرتب سازی کتاب کد بر اساس اعوجاج نیز کمک می گیریم. نتایج شبیه سازی نشان می دهد که یک trade off بین اندازه کتاب کد بزرگ و عملکرد چندی سازی وفقی وجود دارد.

کلمات کلیدی: چندی سازی برداری، وفقی سازی، پیش بینی، LSF، Safety Net، شبکه عصبی پرسپترون، کتاب کد ، Split VQ.

[1] formant synthesis filter

فهرست جداول

جدول ۲-۱- مقایسه هزینه حافظه بین VQ مستقیم و MSVQ برای وضوح بیت متفاوت..... ۱۸
جدول ۲-۲- تخصیص بیت کدر FS1015..... ۲۹
جدول ۲-۳- تخصیص بیت کدر FS1016..... ۳۲
جدول ۵-۱- عملکرد روش پیشنهادی اول با نرخ بیت ثابت 25 bit/frame و اندازه های کتاب کد L متفاوت..... ۸۰
جدول ۵-۲- عملکرد روش پیشنهادی دوم با نرخ بیت ثابت 25 bit/frame و اندازه های کتاب کد L متفاوت..... ۸۳
جدول ۵-۳- عملکرد روش پیشنهادی سوم با نرخ بیت ثابت 25 bit/frame و اندازه های کتاب کد L متفاوت ۸۵
جدول ۵-۴- بررسی روش های پیشنهادی به ازای نرخ بیت های متفاوت و $k = 3$ ۸۶
جدول ۵-۵- بررسی اثر کاهش پیچیدگی برروی روش پیشنهادی سوم و برای $k = 3$ و نرخ بیت ۲۵ بیت در هر فریم. ۸۸
جدول ۵-۶- مقایسه روش پیشنهادی ما با روش پیشنهادی در مرجع [۲۴]. ۸۹

فهرست شکل ها

شکل ۲-۱: مفهوم چندی سازی برداری با اتصال سری انکدر و دیکدر. ۷
شکل ۲-۲: نمایش سلول های Voroni و مراکز سلول ها. ۱۰
شکل ۲-۳: دو نیمه کردن یک کلمه رمز. ۱۲
شکل ۲-۴: نمایش سلول های Voroni قبل(سمت چپ) و بعد (سمت راست) از همگرایی. ۱۳
شکل ۲-۵: بردارهای آموزش و خط سیر کلمات رمز بعد از همگرایی. ۱۴
شکل ۲-۶: منحنی مجموع فاصله بر حسب تکرار. ۱۵
شکل ۲-۷: دیکدر چندی سازی برداری k مرحله ای. ۱۷
شکل ۲-۸: انکدر چندی سازی برداری k مرحله ای. ۱۷
شکل ۲-۹: دیکدر 2SplitVQ. ۱۹
شکل ۲-۱۰: انکدر 2SplitVQ. ۱۹
شکل ۲-۱۱: بلوک دیاگرام سیستم رمزگذار سیگنال صحبت. ۲۵
شکل ۲-۱۲: بلوک دیاگرام جامع کدر سیگنال صحبت. ۲۶
شکل ۲-۱۳: بلوک دیاگرام انکدر FS1015. ۲۷
شکل ۲-۱۴: بلوک دیاگرام دیکدر FS1015. ۲۹
شکل ۲-۱۵: مدل تولید سیگنال صحبت CELP. ۳۰
شکل ۲-۱۶: بلوک دیگرام انکدر FS1016. ۳۱
شکل ۲-۱۷: بلوک دیاگرام دیکدر FS1016. ۳۳
شکل ۳-۱: هیستوگرام نرم بردارهای سیگنال صحبت. ۳-۱-الف: قبل از نرمالیزه کردن توسط بهره. ۳۶
۳-۱-ب: بعد از نرمالیزه کردن توسط بهره ده. ۳۷
شکل۳-۲: ساختار پایه forward gain-adaptive VQ. ۳۸
شکل۳-۳: ساختار پایه backward gain-adaptive VQ. ۳۹
شکل۳-۴: چندی سازی چند مرحله ای به همراه کتاب کد پاره وفقی. ۴۰
شکل ۳-۵ : 2-split,2-stage LSP-VQ همراه با کتاب کد پاره وفقی. ۴۰
شکل ۳-۶ : 2-split,2-stage LSP-VQ همراه با کتاب کدوفقی. ۴۱
شکل ۳-۷ : نمایش عملکرد Partition-deletion ۴۲
شکل ۳-۷ : چندی سازی برداری دو مرحله ای وفقی شده با استفاده از الگوریتم LMS. ۴۶
شکل ۴-۱: ترسیم دنباله ای از ۱۰ فرکانس طیف خطی (LSF) در بازه ۲ ثانیه از یک سیگنال صحبت. ۴۷
شکل ۴-۲: چندی سازی برداری پیش بینی کننده. الف: دیکدر. ب: انکدر. ۵۰
شکل ۴-۳: شماتیک پیش بینی یک بعدی. الف: متناظر با فرمول (۴-۶). ب: متناظر با فرمول (۴-۷). پ متناظر با فرمول (۴-۸).

شکل ۴-۴: پارامترهای LSF نشان داده شده در فریم های متوالی و طرح پیش بینی ۲ بعدی..... ۵۱
شکل ۴-۵: طرح پیش بینی دو بعدی بهبود یافته..... ۵۴
شکل ۴-۶: سایر طرح های پیش بینی دو بعدی ارائه شده. ۵۴
شکل ۴-۷ : مفهوم Safety Net..... ۵۶
شکل ۴-۸: عملکرد یک SN-PVQ بیست بیتی بعنوان تابعی از ترکیب بین بردارها در چندی سازی حافظه دار و چندی سازی Safety Net برحسب میانگین اعوجاج طیفی..... ۵۹
شکل ۴-۹: شماتیک الگوریتم وفقی سازی پیشنهادی اول..... ۶۱
شکل ۴-۱۰: ساختار Split VQ استفاده شده در کتاب کدهای P، F و L..... ۶۱
شکل ۴-۱۱: فرضیات اسم گذاری کتاب کد های F و P..... ۶۲
شکل ۴-۱۲: نحوه دسته بندی زیربردارهای کتاب کد L بر اساس شرط نزدیک ترین همسایگی نسبت به هر زیر بردار کتاب کد F..... ۶۳
شکل ۴-۱۳: ساختار شبکه عصبی پرسپترون استفاده شده..... ۶۷
شکل ۴-۱۴: شماتیک الگوریتم وفقی سازی پیشنهادی دوم. ۶۷
شکل ۴-۱۵: توزیع هیستوگرام شاخص کتاب کد split VQ ثابت برای سه زیر بردار Split ۸ بیتی. ۶۸
شکل ۴-۱۶: توزیع هیستوگرام شاخص کتاب کد split VQ دینامیکی برای سه زیر بردار Split ۸ بیتی فقط برای توزیع شاخص ۱ تا ۵۰..... ۶۹
شکل۴-۱۷: شماتیک روش پیشنهادی سوم. ۷۰
شکل ۵-۱: شبیه ساز طراحی چندی سازی برداری..... ۷۳
شکل ۵-۲: شبیه ساز کدر CELP سیگنال صحبت..... ۷۸
شکل ۵-۳: آنالیز LPC..... ۷۹
شکل ۵-۴: منحنی میانگین اعوجاج طیفی بر حسب k (اندازه کتاب کد L) برای روش پیشنهادی اول..... ۸۱
شکل ۵-۵: شماتیک ایده چندی سازی حافظه دار (سمت راست) و بدون حافظه (سمت چپ)..... ۸۱
شکل ۵-۶: ساختار شبکه های عصبی استفاده شده در روش پیشنهادی دوم..... ۸۲
شکل ۵-۷: منحنی میانگین اعوجاج طیفی بر حسب k (اندازه کتاب کد L) برای روش پیشنهادی دوم..... ۸۴
شکل ۵-۸: منحنی میانگین اعوجاج طیفی بر حسب k (اندازه کتاب کد L) برای روش پیشنهادی سوم..... ۸۵
شکل ۵-۹: منحنی میانگین اعوجاج بر حسب نرخ بیت برای روش های پیشنهادی و MLSVQ..... ۸۷

فهرست مطالب

فصل اول: پیش درآمد ۱
۱-۱- مقدمه ۲
۱-۲- رئوس مطالب پایان نامه ۳
فصل دوم: مفاهیم اولیه ۴
۲-۱- مقدمه ۵
۲-۲- مفهوم چندی سازی برداری ۵
۲-۲-۱- وضوح ۶
۲-۲-۲- سلول ۶
۲-۲-۳- انکدر و دیکدر ۷
۲-۲-۴- فاصله یا معیار اعوجاج ۷
۲-۲-۵- شرط نزدیکترین همسایگی ۸
۲-۲-۶- شرط مرکز ثقل (Centriod) ۸
۲-۲-۷- مجموعه مرزی ۹
۲-۳- الگوریتم های طراحی چندی سازی برداری ۹
۲-۳-۱- الگوریتم Lloyd ۱۰
۲-۳-۲- الگوریتم GLA ۱۱
۲-۳-۳- الگوریتم Stochastic Relaxation ۱۵
۲-۴- ساختار های چندی سازی برداری ۱۶
۲-۴-۱- چندی سازی برداری چند مرحله ای (MSVQ) ۱۶
۲-۴-۲- Split VQ ۱۹
۲-۵- ضرائب پیش بینی خطی (LPC) ۲۰
۲-۶- معیار اعوجاج طیفی (SD) ۲۰
۲-۷- نمایش ضرائب LPC در حوزه فرکانس طیف خطی (LSF) ۲۲
۲-۸- کدر های سیگنال صحبت ۲۴
۲-۸-۱- کدر LPC (FS1015) ۲۶
۲-۷- کدر CELP (FS1016) ۳۰
فصل سوم: مروری بر کارهای انجام شده ۳۴
۳-۱- مقدمه ۳۵
۳-۲- مرور کلی ۳۵
فصل چهارم: روش چندی سازی برداری وفقی پیشنهادی ۴۴
۴-۱- مقدمه ۴۵

۴-۲- چندی سازی برداری حافظه دار و روش های پیش بینی ۴۵
۴-۲-۱- چندی سازی برداری پیشگویانه (PVQ) ۴۶
۴-۲-۲- روش پیش بینی یک بعدی ۴۹
۴-۲- ۳- روش پیش بینی دو بعدی ۵۱
۴-۳- Safety Net VQ ۵۶
۴-۴- روش وفقی سازی پیشنهادی اول ۶۰
۴-۵- شبکه های عصبی مصنوعی ۶۵
۴-۶- روش وفقی سازی پیشنهادی دوم ۶۷
۴-۷- بهره برداری از شاخص کتاب کد با استفاده از مرتب سازی کتاب کد بر اساس اعوجاج ۶۸
۴-۸- روش وفقی سازی پیشنهادی سوم ۷۰
۴-۹- کاهش پیچیدگی های جستجو ۷۱
فصل پنجم: ارزیابی ها ونتایج شبیه سازی روش های پیشنهادی ۷۲
۵-۱- مقدمه ۷۳
۵-۲- طراحی چندی سازی برداری با استفاده از ابزار طراحی چندی سازی برداری(VQDTool) ۷۳
۵-۳- استخراج بردارهای LSF ۷۷
۵-۴- معیار مقایسه وارزیابی های انجام شده ۷۹
۵-۴-۱- ارزیابی شماره یک ۷۹
۵-۴-۲- ارزیابی شماره دو ۸۲
۵-۴-۳- ارزیابی شماره سه ۸۴
۵-۴-۴- ارزیابی شماره چهار ۸۶
۵-۴-۵- ارزیابی شماره پنج ۸۷
۵-۴-۶- ارزیابی شماره شش ۸۸
فصل ششم: نتیجه گیری و پیشنهادات ۹۰
مراجع ۹۲

فصل اول: پیش درآمد

۱-۱- مقدمه

پارامتر های رمزگذاری پیش بینی خطی[۲] (LPC) بطور گسترده در اکثر کدر های سیگنال صحبت به منظور نمایش اطلاعات پوش طیف سیگنال صحبت مورد استفاده قرار می گیرند. پارامتر های LPC عموماً از یک سیگنال صحبت با نرخ ۵۰ فریم در هر ثانیه و با استفاده از یک آنالیز LPC مرتبه دهم بدست می آیند. سپس ضرائب LPC بدست آمده چندی سازی[۳] شده و به سمت گیرنده فرستاده می شوند.

برای کاربرد های رمزگذاری سیگنال صحبت با نرخ بیت پایین، مهم می باشد که ضرائب LPC با تعداد بیت کمتر و بدون کاهش کیفیت سیگنال صحبت چندی سازی شوند. کار های قابل ملاحظه ای برای توسعه روش های چندی سازی برداری[۴] (VQ) و اسکالر به منظور نمایش اطلاعات پوش طیف سیگنال صحبت با کمترین تعداد بیت انجام شده است. در عین حال این روش ها به نحوی ارائه شده اند که هزینه محاسبات چندان افزایش نمی یابد. بلکه در برخی مواقع منجر به کاهش هزینه محاسباتی شده اند. نظیر استفاده از ساختار های چندی سازی Split VQ و چندی سازی برداری چند مرحله ای. یکی از روش های مؤثر در چندی سازی ضرائب LPC استفاده از نمایش معادل آنها در فرکانس طیف خط[۵] (LSF) و چندی سازی آنها به جای خود ضرائب LPC می باشد. این کار به خاطر ویژگی های مطلوب پارامترهای LSF انجام می شود. چندی سازی پارامتر های LSF با استفاده از ساختار های چندی سازی Split VQ و چندی سازی چند مرحله ای بطور گسترده توسط آقای Paliwal و Atal بررسی شده است [۱] .

گفته می شود که سیگنال صحبت یک سیگنال متغیر با زمان کند و غیرایستان می باشد. در نتیجه عملکرد چندی سازی برداری بهبود خواهد یافت اگر بتوان کتاب کد یا چندی سازی برداری را همراه با تغییر خواص آماری بردار ورودی وفقی کرد. بنابراین یک چندی سازی برداری وفقی می باشد اگر کتاب کد برای تطبیق با خواص آماری محلی دنباله ورودی تغییر کند. در نتیجه برای یک چندی سازی برداری وفقی می توان دستیابی به عملکرد بهتری را انتظار داشت. روش های متفاوتی برای ساخت یک چندی سازی برداری وفقی

---

[۲] Linear Prediction Coding.
[۳] Quantization
[۴] Vector Quantization
[۵] Line Spectrum Frequency

وجود دارد. بعنوان مثال کتاب کد می تواند خود را مطابق با مشخصه های آماری محلی دنباله ورودی نظیر میانگین، واریانس، تابع چگالی احتمال و یا حتی تابع چگالی طیف توان وفقی کند.

۱-۲- رئوس مطالب پایان نامه

فصل دوم این پایان نامه به توصیف مفاهیم اولیه می پردازد. بدلیل استفاده از ابزار وفقی سازی برای یک چندی سازی برداری، ابتدا به توصیف مفهوم چندی سازی برداری و مفاهیم مرتبط با آن پرداخته می شود. از آنجا که اصلی ترین جزء هر کدگذار سیگنال صحبت ضرائب کدگذار پیش بینی خطی می باشند به توصیف اجمالی آنها پرداخته می شود. از آنجا که هدف این تحقیق بر پایه چندی سازی پارامترهای LSF می باشد به توصیف آن نیز پرداخته می شود. در ادامه بدلیل سروکار داشتن با کدرهای سیگنال صحبت به تشریح دو کدر پایه گذاری شده سیگنال صحبت خواهیم پرداخت.

فصل سوم، مروری بر روش های چندی سازی برداری وفقی موفق برای ضرائب کدگذاری پیش بینی خطی خواهد بود.

در فصل چهارم نیز روش پیشنهادسیبل (dB)ان خواهد شد. این روش در حضور پیش بینی، شبکه عصبی و ایده مرتب سازی محتویات کتاب کد بر اساس اعوجاج صورت می پذیرد. در عین حال با توجه به اینکه اینها از همبستگی بین فریمی در بین فریم های متوالی سیگنال صحبت استفاده می کنند و در برخی مواقع بدلیل همبستگی پایین در بین فریم های متوالی نتیجه تخریبی خواهند داشت، لذا روش های مذبور را در داخل ایده Safety Net پیاده سازی می کنیم. به همین جهت در این فصل به توصیف تک تک آنها و روش پیشنهادی خواهیم پرداخت.

در فصل پنجم نتایج شبیه سازی ها انجام شده و با یک چندی سازی برداری Split ثابت مقایسه خواهد شد. در عین حال روش پیشنهادی در انتهای این فصل نیز با روش پیشنهادی در مرجع [۲۴] مقایسه خواهد شد. در انتها و در فصل ششم نتیجه گیری کلی و پیشنهادات بیان خواهد شد.

فصل دوم: مفاهیم اولیه

۲-۱- مقدمه

در ارتباطات بیسیم فشرده سازی سیگنال صحبت نقش مهمی در افزایش ظرفیت کلی شبکه ارتباطی و استفاده همزمان کاربران دارد. فشرده سازی سیگنال صحبت اساساً با نمایش سیستم تولید سیگنال صحبت انسان توسط مدل های پارامتری قابل دستیابی است. هدف اصلی این مدل دستیابی به سیگنال صحبت ساختگی با کیفیت مناسب ونرخ بیت پایین می باشد. دو نوع افزونگی در سیگنال صحبت قابل مشاهده است.

۱) افزونگی کوتاه مدت: به علت همبستگی بین نمونه های همسایه سیگنال صحبت می باشد که یک نمونه از سیگنال صحبت بصورت ترکیب خطی از نمونه های گذشته تقریب زده می شود. ضرائب حاصل از ترکیب خطی ضرائب LPC نامیده میشوند [۸].

۲) افزونگی بلند مدت: به علت همبستگی بین نمونه های سیگنال صحبت در فاصله یک دوره تناوب Pitch در طول قسمت هایی از سیگنال صحبت صدا دار می باشد [۸].

یکی دیگر از تکنیک های فشرده سازی در سیگنال صحبت استفاده از چندی سازی برداری (VQ) می باشد که در اکثر کدرهای سیگنال صحبت استفاده می شود. بویژه از سال ۱۹۹۴ به بعد برخی از کدر های استاندارد از VQ برای LPC استفاده کرده اند [۸]. در این فصل به توصیف مفاهیم کلیدی مرتبط با چندی سازی برداری می پردازیم. بعد از این به توصیف روش های کلی طراحی چندی سازی های برداری و ساختار های آن می پردازیم. به علت اهمیت تحقیق در خصوص چندی سازی ضرائب LPC در ادامه به توصیف کلی این ضرائب پرداخته می شود. با توجه به تحقیقات گسترده در زمینه نمایش معادل ضرائب LPC و چندی سازی برداری ضرائب LPC در اینگونه نمایش ها، به توصیف کاربردی ترین آنها در حوزه فرکانس طیف خطی (LSF) پرداخته خواهد شد. در انتها به توصیف برخی کدر های پایه گذاری شده برای سیگنال صحبت پرداخته می شود.

۲-۲- مفهوم چندی سازی برداری

چندی سازی برداری تعمیم چندی سازی اسکالر از یک بعد به چند بعد می باشد. یک چندی کننده برداری با بعد M وسایز N نگاشتی از یک بردار x در فضای اقلیدسی M بعدی (R^M)، به یک مجموعه محدود

Y می باشد [۴]. این مجموعه شامل N بردار M بعدی می باشد که بردار کد یا کلمه رمز نامیده می شوند.

بنابراین:

$$Q: \mathrm{R}^M \to Y, \qquad (۱-۲)$$

که در اینجا داریم:

(۲-۲)

$$\begin{cases} x = [x_1, x_2, \ldots, x_M]^T, \\ (y_1, y_2, \ldots, y_N) \in Y, \\ y_i = [y_{i1}, y_{i2}, \ldots y_{iM}]^T;\ i = 1, 2, \ldots N. \end{cases}$$

مجموعه Y بعنوان کتاب کد (Codebook) چندی سازی برداری شناخته شده است.

عملکرد نگاشت برای چندی سازی برداری بصورت زیر تعریف می شود.

(۳-۲)

$$Q(x) = y_i;\ i = 1, 2, \ldots, N,$$

در ادامه به توصیف مفاهیم مربوط به چندی سازی برداری می پردازیم.

۲-۲-۱- وضوح

وضوح برای یک چندی سازی برداری بصورت $r = \log_2 N$ تعریف می شود. این تعریف تعداد بیت مورد نیاز برای آدرس دهی یک کلمه رمز معین، بطور واحد را مشخص می کند. بعنوان مثال یک کتاب کد با وضوح ۱۲ بیتی، تعداد $N = 2^{12} = 4096$ کلمه کد دارد که به هر بردار کد درون کتاب کد یک کلمه کد تخصیص داده می شود [۸].

۲-۲-۲- سلول

متناظر با هر N نقطه از چندی سازی برداری M بعدی، فضای R^M به N ناحیه یا سلول تقسیم می شود، $R_i, i = 1, 2, \ldots N$. سلول iام بصورت زیر تعریف می شود [۸].

(۴-۲)

$$R_i = \left\{ x \in R^M : Q(x) = y_i \right\} = Q^{-1}(y_i)$$

۲-۲-۳- انکدر[۶] و دیکدر[۷]

یک چندی سازی برداری می تواند به دو قسمت کدگذار (انکدر) و کدبردار (دیکدر) تجزیه شود. کدگذار E وظیفه نگاشت از فضای R^M به مجموعه شاخص $I=\{1,2,\ldots,N\}$, را به عهده دارد. کد بردار D مجموعه شاخص I را به مجموعه بازسازی شده بردار Y تبدیل می کند. در اینجا داریم :

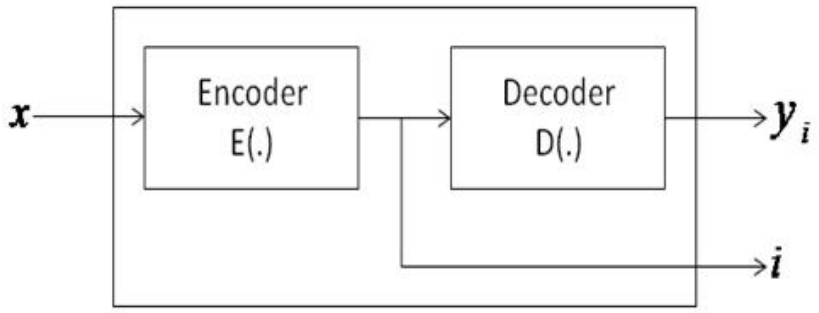
$$E:R^M \to I \;\&\; D:I \to R^M, \qquad (۵-۲)$$

دیکدر در واقع تعیین می کند که بردار ورودی در کدام یک از N ناحیه یا سلول قرار دارد. در عین حال انکدر بعنوان یک جدول مراجعه (Lookup table) عمل می کند که با توصیف کتاب کد تعیین می شود. عملکرد کلی یک چندی سازی برداری با اتصال سری انکدر و دیکدر در رابطه زیر مشخص شده است:

(۶-۲)

$$\hat{x}=Q(x)=D(E(x))=y_i$$

مفهوم انکدر و دیکدر در شکل ۲-۱ نشان داده شده است.



شکل ۲-۱: مفهوم چندی سازی برداری با اتصال سری انکدر و دیکدر.

۲-۲-۴- فاصله یا معیار اعوجاج

فاصله یا معیار اعوجاج یک تخصیص از یک هزینه غیر منفی $d(x,Q(x))$ متناظر با چندی سازی هر بردار ورودی نسبت به بردار باز سازی شده $Q(x)$ می باشد.

$$d(x,Q(x))=\begin{cases} 0, & Q(x)=x \\ >0, & otherwise \end{cases} \qquad (۷-۲)$$

[6] Encoder
[7] Decoder

با چنین معیاری می توانیم عملکرد یک سیستم را با داشتن مقدار معلوم d مشخص کنیم. معیار اعوجاجی که بطور گسترده استفاده می شود معیار مربع خطا می باشد که توسط رابطه زیر داده شده است.

(۸-۲)

$$d(x,\hat{x}) = \left\| x - \hat{x} \right\|^2 = \sum_{j=1}^{M} (x_j - \hat{x}_j)^2$$

به این رابطه مربع فاصله اقلیدسی گفته می شود. نوع دیگری از این معیار مربع فاصله اقلیدسی وزن دار شده می باشد [۱]. این رابطه بصورت کلی در زیر نشان داده شده است.

(۹-۲)

$$d(x,\hat{x}) = \sum_{j=1}^{M} w_j * (x_j - \hat{x}_j)^2,$$

در این رابطه w_j یک ضریب محاسبه شده برای هر مؤلفه از بردار می باشد.

۲-۲-۵- شرط نزدیکترین همسایگی

فرض کنید که $d(x,y)$ معیار اعوجاج بر روی فضای بردار ورودی نسبت به خروجی باشد. شرط نزدیک ترین همسایگی کل فضای ورودی را به سلول یا نواحی معین تقسیم بندی می کند.

(۱۰-۲)

$R_i = \{x : d(x,y_i) \le d(x,y_j); \forall j \in I\}, i \in I.$ $I = \{1,2,\ldots,N\}$ همان مجموعه شاخص ها می باشد. بنابراین با توجه به شرط نزدیک ترین همسایگی، هر سلول R_i شامل آن نقاطی از x می باشد که اعوجاج کمتری را با استفاده از بردار کد بازسازی شده y_i نسبت به هر بردار کد دیگر دارند.

برای یک کتاب کد Y با اندازه N، تقسیم بندی بهینه سلول ها رابطه زیر را ارضاع می کند.

$$R_i = \{x : d(x,y_i) \le d(x,y_j)\},\ i \ne j\ \&\ i,j = 1,2,\ldots,N \qquad (۱۱-۲)$$

این بدین معنی می باشد که $Q(x) = y_i$ اگر و فقط اگر $d(x,y_i) \le d(x,y_j)$. بنابراین با استفاده از رابطه زیر کتاب کد معین Y ، شامل مینیمم اعوجاج یا نزدیکترین نگاشت همسایگی است.

$$d(x,Q(x)) = \min_i (d(x,y_i)) \qquad (۱۲-۲)$$

۲-۲-۶- شرط مرکز ثقل (Centriod)

مرکز ثقل $cent(R_0)$ بدین گونه تعریف می شود که هر مجموعه غیر تهی $R_0 \in R^M$ بطوریکه بردار y_0 (اگر وجود داشته باشد)، اعوجاج بین هر نقطه $X \in R_0$ و y_0 را مینیمم می کند ، بر روی توزیع احتمال معین X که $X \in R_0$ می باشد میانگین گرفته می شود. این تعریف در رابطه زیر نشان داده شده است.

$$cent(R_0) = \left\{ y_0 : \mathrm{E}\left\{ d(X, y_0) \mid X \in R_0 \right\} \le \mathrm{E}\left\{ d(X, y) \mid X \in R_0 \right\} \right\}, \forall y \in R^M \qquad (۱۳-۲)$$

برای یک افراز معین $\{R_i ; i = 1,2,\ldots,N\}$، بردارهای رمز بهینه رابطه زیر را ارضاع می کنند.

$$y_i = cent(R_i) \qquad (۱۴-۲)$$

با توجه به رابطه فاصله اقلیدسی، مرکز هر سلول طبق رابطه زیر مشخص می شود.

$$y_i = \frac{1}{N_i} \sum_{x_k \in R_i} X_k \qquad (۱۵-۲)$$

۲-۲-۷- مجموعه مرزی

برای یک افراز سلول معین $R_i, i = 1,2,\ldots,N$ مجموعه مرزی بصورت زیر تعریف می شود.

$$B = \left\{ x : d(x, y_j) = d(x, y_i),\ i \ne j \right\}, i, j = 1,2,\ldots,N \qquad (۲۶-۱)$$

بنابراین مجموعه مرزی شامل نقاطی است که بطور یکسان به y_i و برخی از y_jها نزدیک باشد و از این رو نزدیکترین همسایگی یکتایی ندارند. شرط لازم برای اینکه یک کتاب کد بهینه باشد این است که :

$$B = \varnothing \qquad (۲۷-۱)$$

این بدان معنی است که مجموعه مرزی بایستی تهی باشد.

۲-۳- الگوریتم های طراحی چندی سازی برداری

کتاب کد فرضی Y با سایز N موجود می باشد. مطلوب است که سلول های R_i و کلمات رمز را به نحوی پیدا کنیم که میانگین اعوجاج $D = E\left\{ d(X, Q(X)) \right\}$ مینیمم شود. X بردار تصادفی ورودی با یک PDF معین می باشد. در عمل به جای استفاده از مجموعه نامتناهی X از یک مجموعه بردار آموزشی شامل N_t بردار، برای بهینه سازی چندی سازی برداری استفاده می کنیم.