

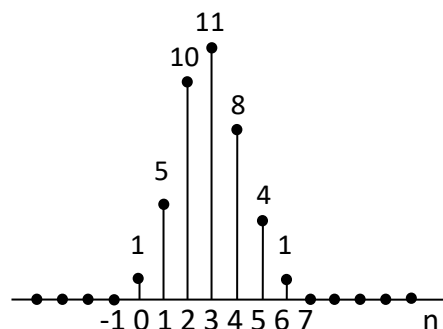
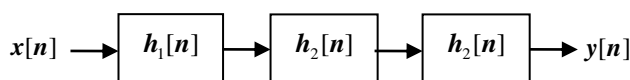


۱. (۱۰٪) [تحقیق] همان گونه که در کلاس بیان شد، پردازش گفتار در کاربردهای مختلف کاربرد دارد. هدف این سوال مروری بر برخی از این کاربردها یا روش های آنها در حوزه مورد علاقه شما (جهت تعیین موضوع پروژه پایانی درس) است. برای این کار، هر دانشجو حداقل دو موضوع (کاربرد یا روش) را بررسی کرده و پس از مطالعه منابع مرتبط (مقاله، پایان نامه، گزارش و ...) از آنها یک گزارش مختصر تهیه و به عنوان پاسخ این سوال ارائه می کند. توجه شود که برای هر موضوع، حداقل یک منبع معتبر را که در تهیه گزارش از آن استفاده کرده اید، همراه پاسخ ارسال کنید. انتظار می رود، پس از این تمرین، شما قادر باشید یکی از عناوین مطالعه شده را به موضوع پروژه درسی خود انتخاب کنید.

۲. (۲۰٪) [کانولوشن] همان گونه که در شکل زیر (سمت چپ) راست نشان داده شده است، رابطه ورودی و خروجی یک سیستم علی از اتصال سری سه زیرسیستم دیگر ایجاد شده است. پاسخ ضربه کل سیستم $h[n]$ در شکل (سمت راست) نشان داده شده است. اگر $h_2[n] = u[n] - u[n-2]$ باشد:

الف) پاسخ ضربه زیرسیستم اول $h_1[n]$ را بدست آورید.

ب) پاسخ کل سیستم را به ورودی $x_1[n] = u[n] - 2u[n-1] + u[n-2]$ محاسبه کنید.





۳. (۳۰٪) [تبدیل فوریه]

الف) تبدیل فوریه پاسخ ضربه (پاسخ فرکانسی) یک سیستم به صورت $H(e^{j\omega}) = 1 + \frac{1}{2}e^{-j\omega}$ است. خروجی این سیستم در حوزه زمان، $y[n]$ را برای سیگنالی با تبدیل فوریه $X(e^{j\omega}) = \frac{3e^{-4j\omega}}{(1 - \frac{1}{4}e^{-4j\omega})^2}$ بدست آورید. (راهنمایی: چون تبدیل فوریه کانولوشن دو سیگنال

(در حوزه زمان)، معادل با ضرب تبدیل فوریه این دو سیگنال (در حوزه فرکانس) است، برای پاسخ این سوال دو تبدیل فوریه را در هم ضرب کرده و سپس معکوس تبدیل فوریه بگیرید).

ب) پاسخ فرکانسی $H(e^{j\omega})$ سیستم میانگین‌گیری با ۱۱ نقطه، $y[n] = \frac{1}{11} \sum_{k=0}^{10} x[n-k]$ را

بدست آورید (برای این کار از طرفین رابطه تبدیل فوریه بگیرید و $H(e^{j\omega}) = \frac{Y(e^{j\omega})}{X(e^{j\omega})}$ محاسبه کنید).

۴. (۴۰٪) [پایاده‌سازی: کار با سیگنال دیجیتال] هدف از این تمرین آشنایی با سیگنال‌های دیجیتال و پردازش آنهاست. برای این کار، موارد خواسته شده را در زبان برنامه‌نویسی دلخواه پایاده‌سازی کرده و در هر بخش از تمرین، خروجی برنامه را (در قالب یک گزارش، شامل نتیجه‌گیری و نمودارها) به همراه کد برنامه نوشته شده، به تفکیک هر بخش از سوال، ارسال کنید.

در یک برنامه ضبط صدا مانند Cool Edit (یا نسخه جدیدتر آن Adobe Audition) حرف "آ"، حرف "س" و کلمه "آسمان" را ضبط کنید. برای ضبط کردن سیگنال‌ها از نرخ نمونه‌برداری 16 KHz، چندی‌سازی 16 Bits، به صورت مونو و فرمت Wav استفاده کنید. سعی کنید سکوت‌های ابتدا و انتهای فایل‌ها را به صورت دستی حذف کنید. این فایل‌ها را همراه پاسخ سوال ارسال کنید.



الف) برنامه‌ای بنویسید که چهار فایل ضبط شده را بخواند و پارامترهای طول سیگنال (برحسب ثانیه)، انرژی و توان سیگنال را در خروجی چاپ کند. همچنین نمودار حوزه زمان سیگنال‌ها را رسم کنید. با توجه به مقادیر بدست آمده و شکل موج، در مورد تفاوت‌های قابل درک در دو سیگنال "آ" و "س" نظر خود را بیان کنید.

ب) برنامه‌ای بنویسید که سه سیگنال ضبط شده را از یک فیلتر هموارسازی میانه (Median Smoother) عبور دهید و خروجی هر کدام را به ازای مقادیر $N=100, 1000, 3000$ بدست آورده و شکل موج‌های هر کدام را رسم کنید. از نظر ظاهری و شنیداری، سیگنال‌های حاصل چه تفاوتی با سیگنال اصلی دارند؟ در مورد مقایسه این سه سیگنال با هم چه می‌توان گفت؟
 ب) از دو سیگنال "آ" و "س" تبدیل فوریه بگیرد. برای هر سیگنال مقدار تبدیل را برای چهار نمونه اول و چهار نمونه آخر به عنوان خروجی چاپ کنید. مشاهده خود را از این مقادیر بیان کنید. نمودار اندازه طیف و فاز را برای سیگنال‌ها رسم کنید. در مورد تفاوت‌های طیف دو سیگنال "آ" و "س" چه نتیجه‌ای می‌توان گرفت؟

د) در تکمیل برنامه بخش ج، انرژی طیف را برای دو سیگنال ضبط شده محاسبه کنید و با مقادیر انرژی محاسبه شده در قسمت الف مقایسه کرده و نتیجه مشاهده خود را بیان کنید.

راهنمایی: در متلب، می‌توانید برای خواندن فایل صوتی از Wavread (یا Audioread) و برای تبدیل فوریه از fft استفاده کنید.

۵. (۴۰٪) [پایاده‌سازی: سنتز گفتار فارسی] برای ساخت یک سیستم ساده سنتز گفتار فارسی همه واج‌های فارسی را که در جدول جدا آورده شده است، هر کدام در یک فایل ضبط کنید (نرخ نمونه‌برداری 16 KHz، چندی‌سازی 16 Bits، به صورت مونو و فرمت Wav). قسمت‌های سکوت ابتدا و انتهای هر کدام از فایل‌ها را به صورت دستی حذف کنید. با پشت سر هم قرار دادن این واج‌ها، سیگنال صوتی کلمات "ستایش"، "فلسطین" و "بروید"؛ و



عبارت "دانشکده علوم و فنون نوین" را تولید کنید. برای افزایش قابل فهم بودن صداهای تولیدی چه روشی را پیشنهاد می کنید.

همه واج های ضبط شده و صداهای تولیدی کلمات را به همراه پاسخ این تمرین ارسال کنید.

#	IPA	Char	Code	Farsi Letter	Phonetic Description
1	I	i	105	ای	high front unrounded
2	e	e	101	!	mid front unrounded
3	a	a	97	آ	low front unrounded
4	u	u	117	او	high back rounded
5	o	o	111	أ	mid back rounded
6	/	/	47	آ	low back rounded
7	p	p	112	پ	unvoiced bilabial plosive
8	b	b	98	ب	voiced bilabial plosive
9	t	t	116	ت، ط	unvoiced dental plosive
10	d	d	100	د	voiced dental plosive
11	k	k	107	ک	unvoiced velar plosive
12	g	g	103	گ	voiced velar plosive
13	q	q	113	ق، غ	voiced uvular plosive
14]]	93	أ، و، ع	glottal stop
15	\$	\$	36	چ	unvoiced alveopalatal affricate closure
16	,	,	44	ج	voiced alveopalatal affricate
17	f	f	102	ف	unvoiced labiodental fricative
18	v	v	118	و	voiced labiodental fricative
19	s	s	115	س، ث، ص	unvoiced alveolar fricative
20	z	z	122	ز، ذ، ظ، ض	voiced alveolar fricative
21	.	.	46	ش	unvoiced alveopalatal fricative
22	[[91	ژ	voiced alveopalatal fricative
23	x	x	120	خ	unvoiced uvular fricative
24	h	h	104	ه، ح	unvoiced glottal fricative
25	l	l	108	ل	lateral alveolar
26	r	r	114	ر	trill alveolar
27	m	m	109	م	nasal bilabial
28	n	n	110	ن	nasal alveolar
29	y	y	121	ی	approximant palatal