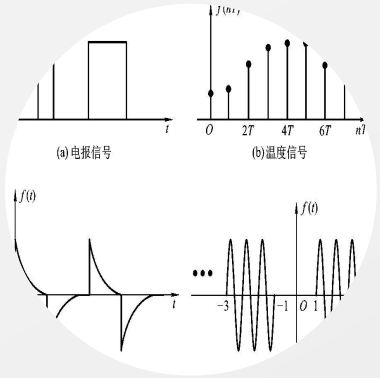
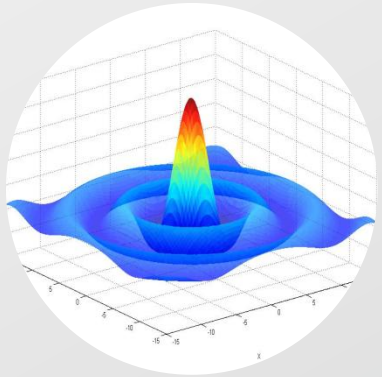
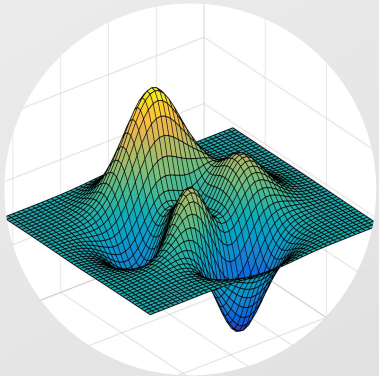
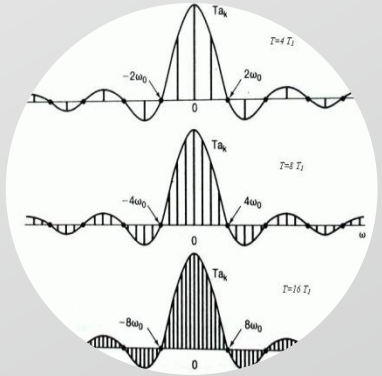




Project1：Speech Synthesis And Perception With Envelope Cue





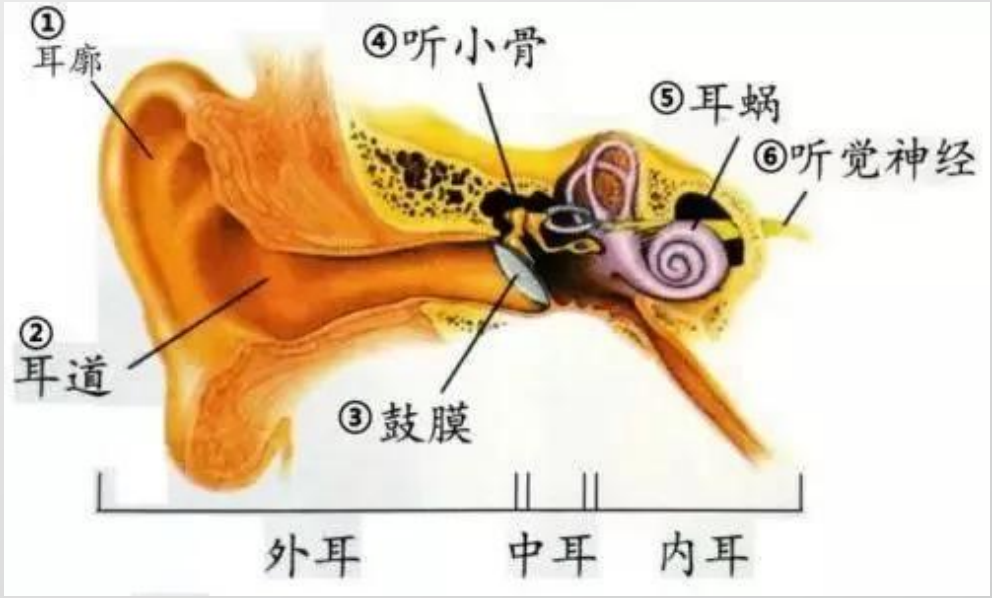
主讲老师：王小静
办公地点：慧园2栋411

Overview

- In this tutorial, you will learn to synthesize a speech signal based on ***multi-band envelope cues***.
- In lab 5, you've learned:
  - how to design a low-pass/band-pass filter
  - how to extract envelope
  - how to generate a speech-spectrum shaped noise (SSN)
  - how to do energy normalization
  - how to read/save a '*.wav' file

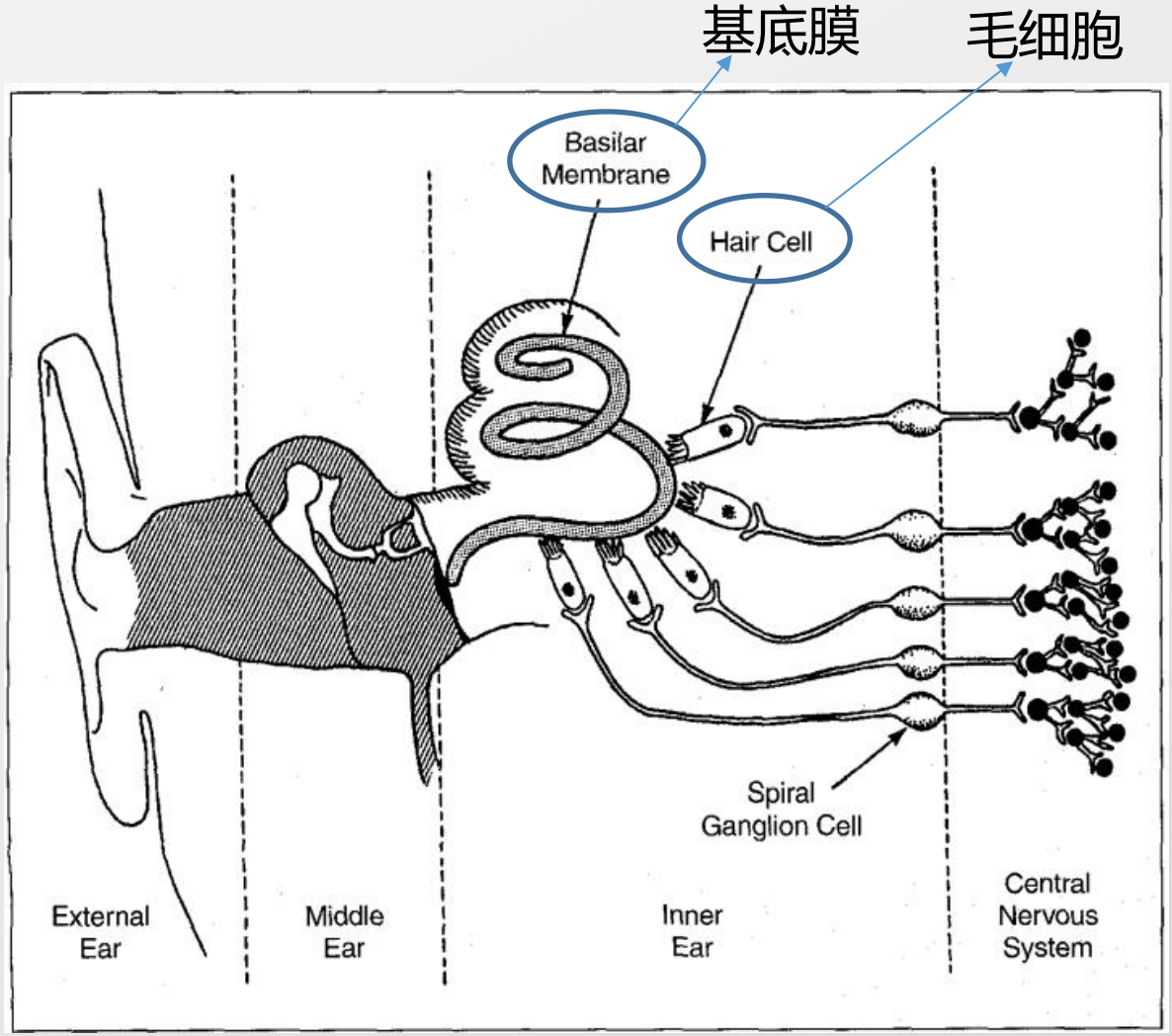
Background: The principle of hearing

acoustic pressure waves → outer ear → eardrum(鼓膜) → ossicles(听小骨) → cochlea(耳蜗) → auditory nerve → brain



Loizou (1999) “Introduction to cochlear implants,” IEEE Eng. in Med. and Bio. Mag., 18, 32-42.

The function of cochlea



mechanical vibrations

↓

vibrations in fluid

↓

pressure variations

↓

displacements of basilar membrane

↓

bending of hair cells

↓

electrochemical substance

↓

neurons fire

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Loizou (1999) “Introduction to cochlear implants,” IEEE Eng. in Med. and Bio. Mag., 18, 32-42.

Deafness & cochlear implants

The most common cause of deafness is the loss of hair cells, rather than the loss of auditory neurons.


澳大利亚


奥地利

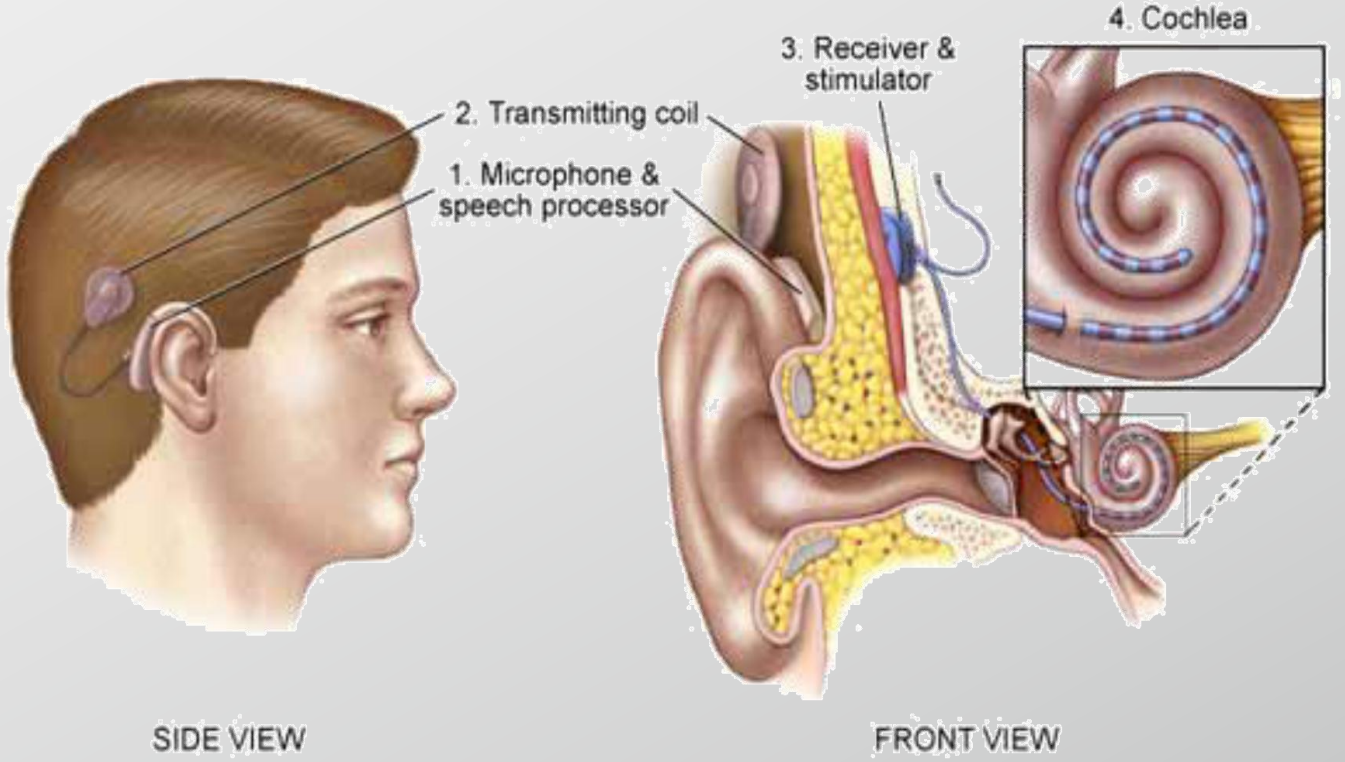
美国

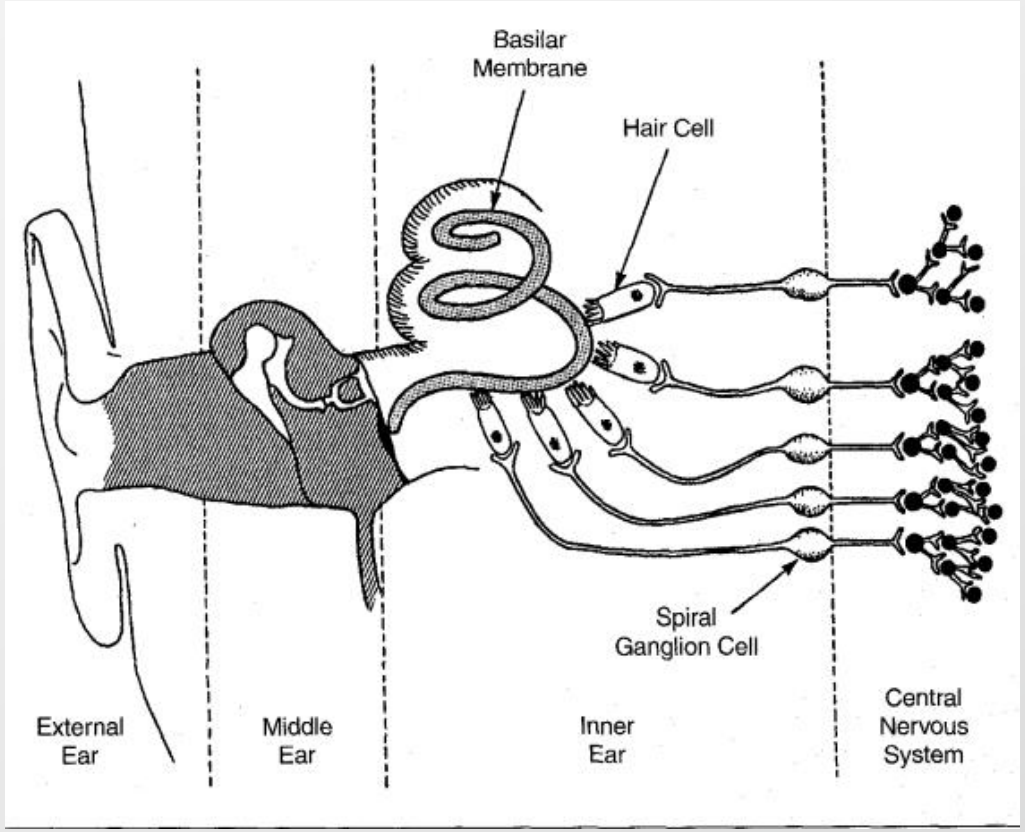
南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

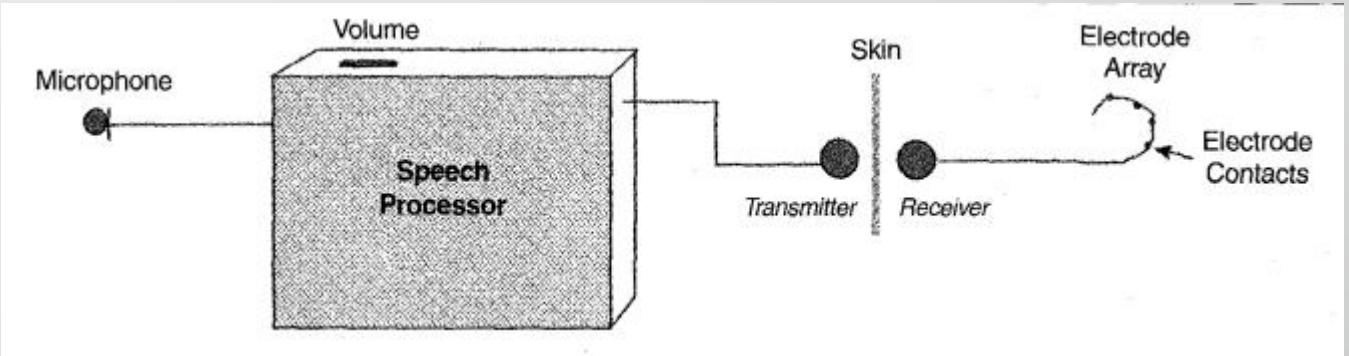
Belal, A., Maher, A., & Talaat, M. (1976). Histopathology of profound deafness. *Indian Journal of Otolaryngology*, 28(3), 108-111.

The principle of cochlear implants

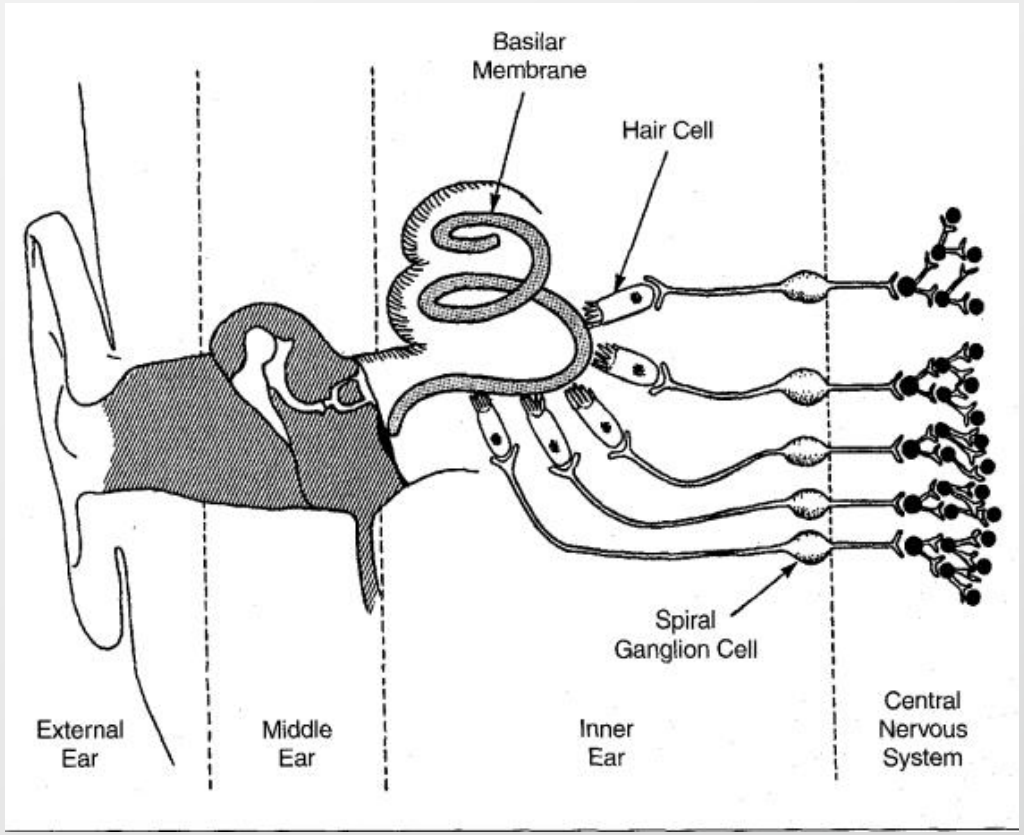
- The microphone collects signals, and the speech processor processes the collected signals.
- The wireless transmitter sends the processed signal to the inner side of the cochlea implants.
- Wireless receiver receives signals.
- Received signals stimulate the auditory nerve through electrodes to reconstruct signals.

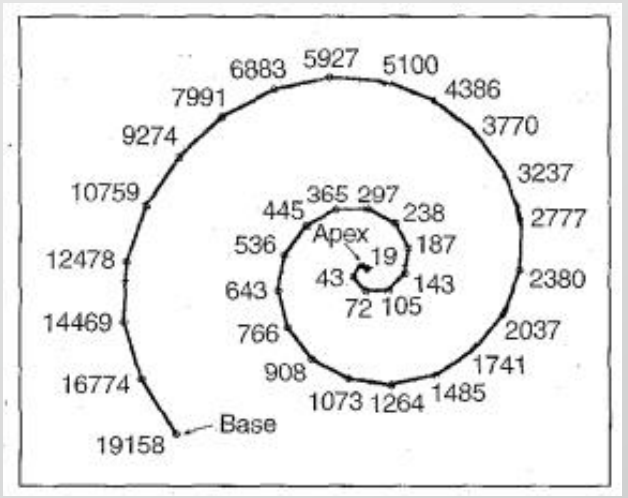






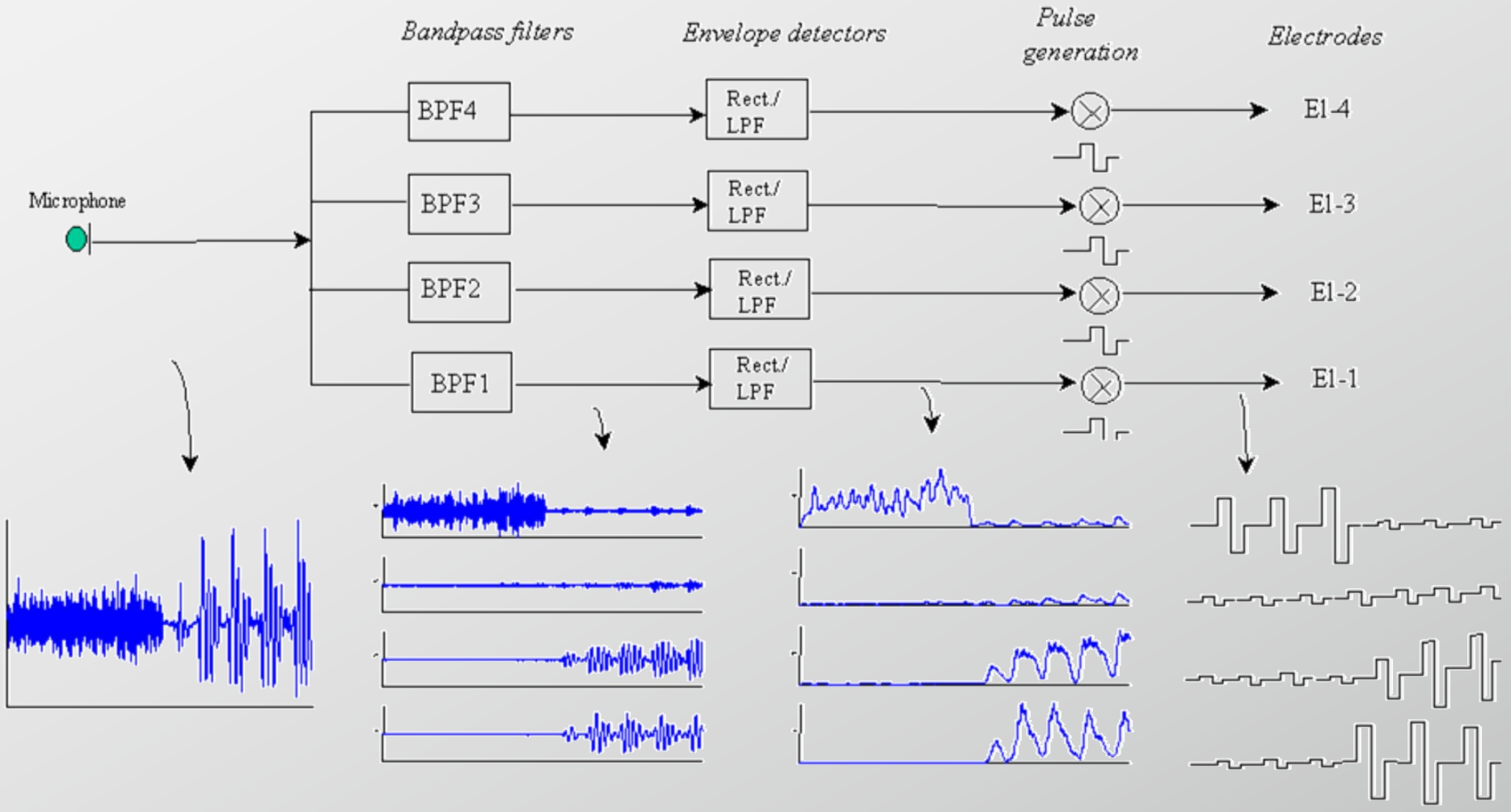
Loizou (1999) “Introduction to cochlear implants,” IEEE Eng. in Med. and Bio. Mag., 18, 32-42.





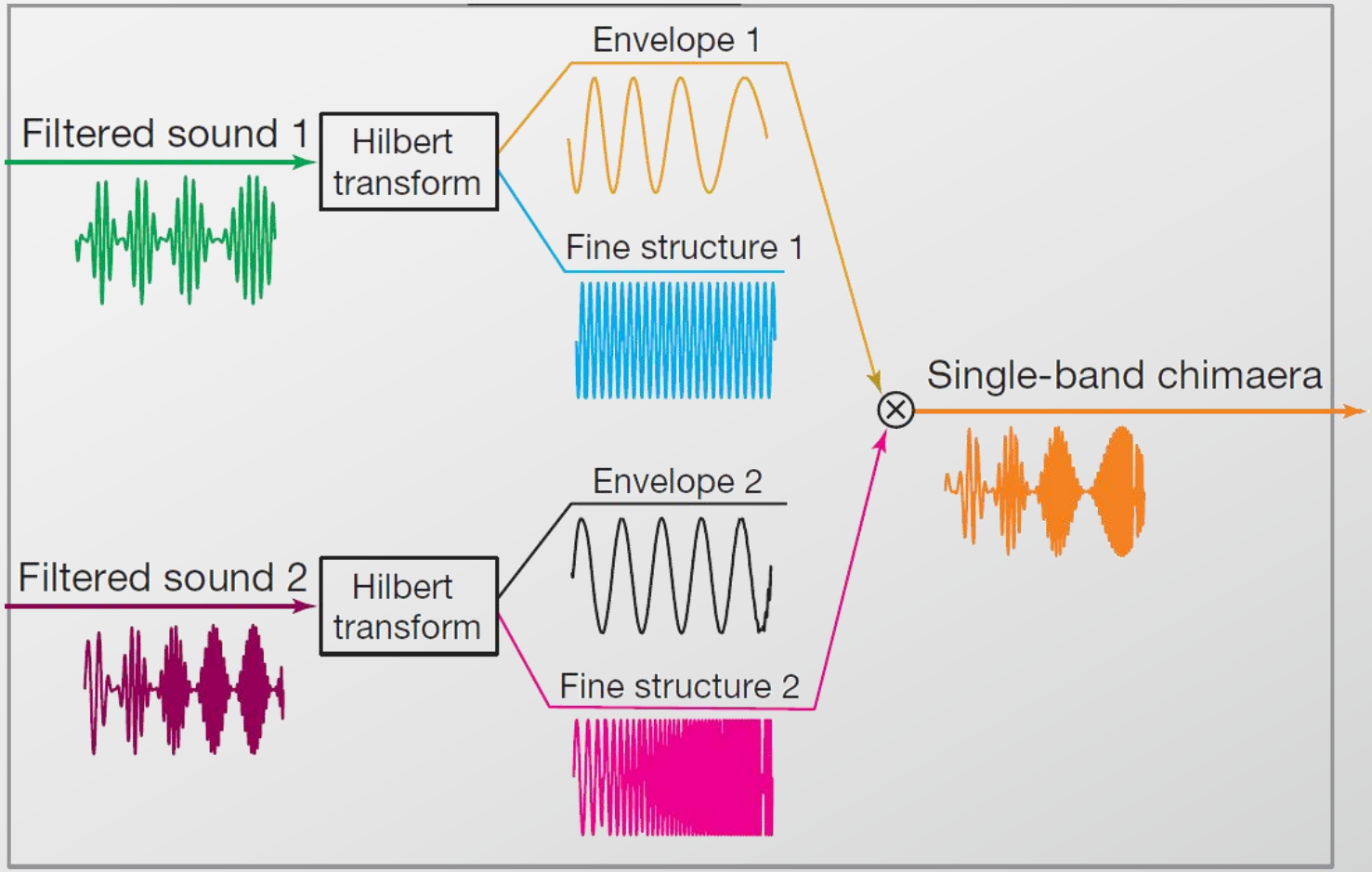
Loizou (1999) “Introduction to cochlear implants,” IEEE Eng. in Med. and Bio. Mag., 18, 32-42.

Speech processing in cochlear implants



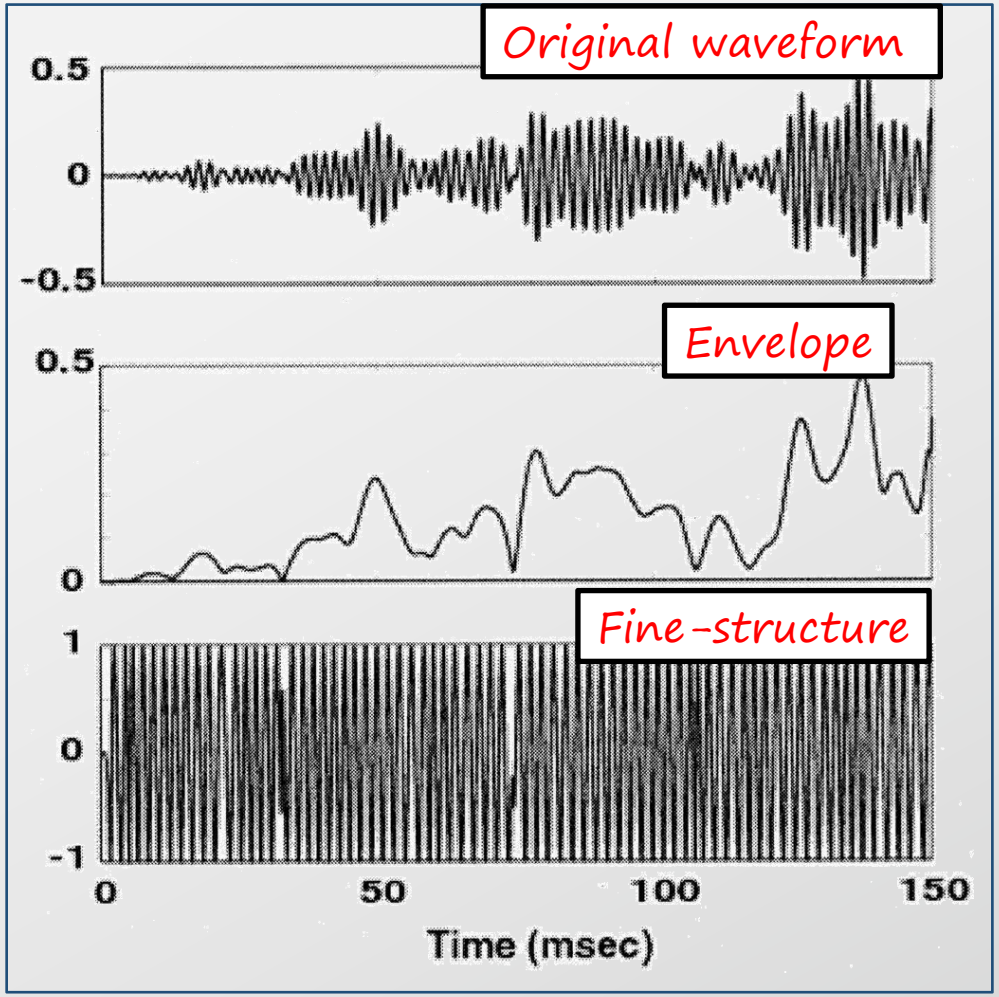
Loizou (1999) “Introduction to cochlear implants,” IEEE Eng. in Med. and Bio. Mag., 18, 32-42.

Acoustic cues of speech signal



- Envelope and fine- structure
  - **Envelope:** amplitude modulation, low- frequency
  - **Fine-structure:** frequency modulation, high- frequency

Z. M. Smith et al. (2002), "Chimaeric sounds reveal dichotomies in auditory perception ," Science.



- **Envelope:**

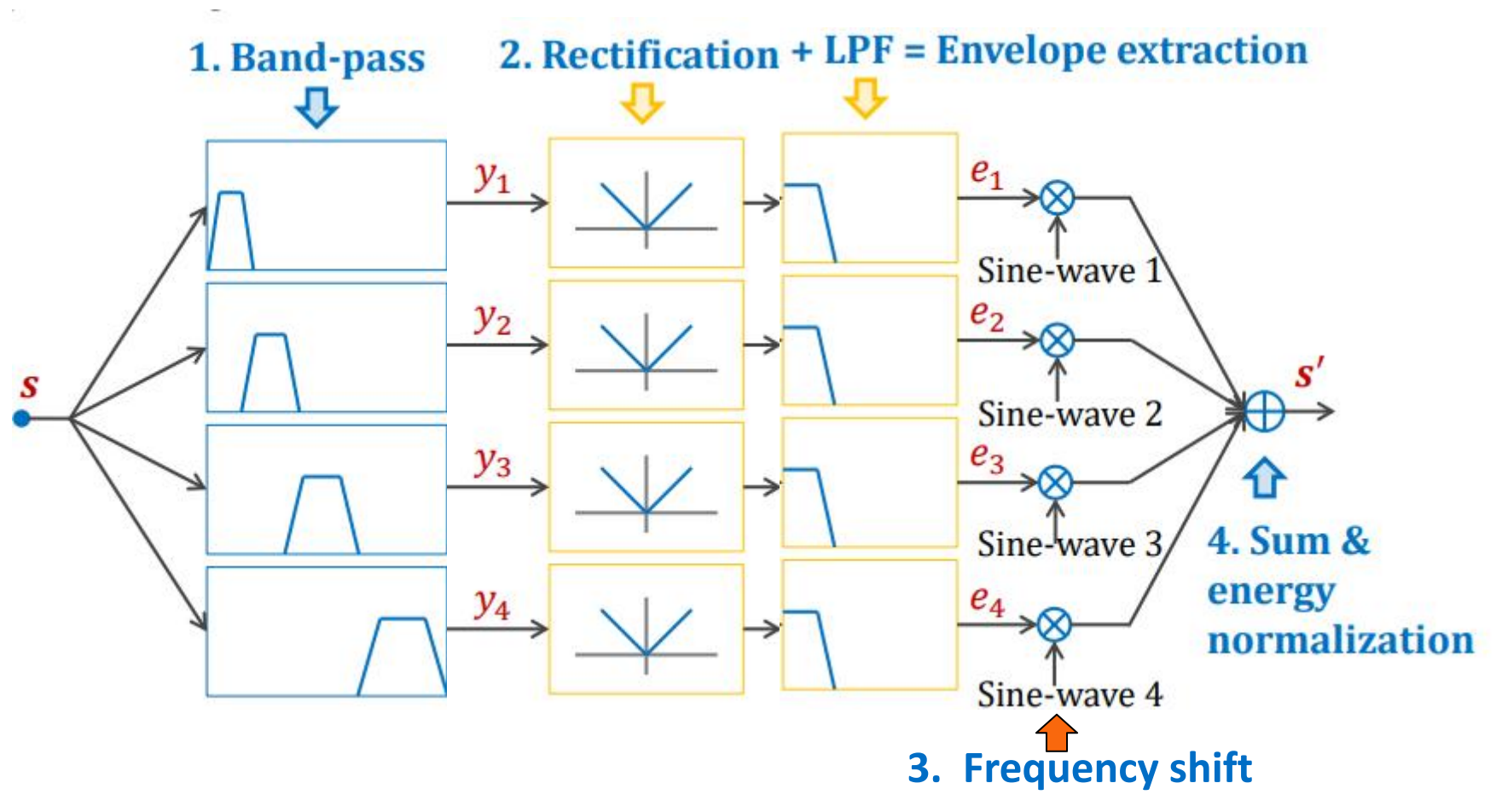
  The envelope is most important for speech reception

- **Fine-structure:**

  The fine structure is most important for pitch perception and sound localization.

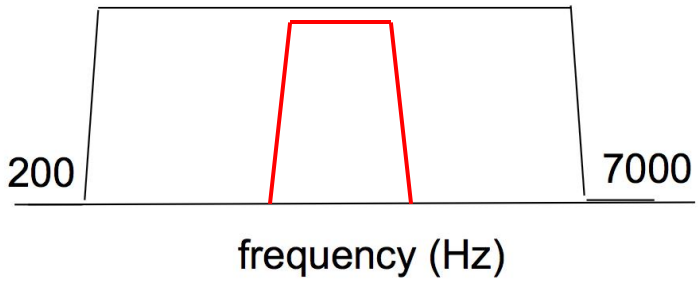
B.S. Wilson et al. (2005), "Two new directions in speech processor design for cochlear implants," Ear Hear.

Speech synthesis with envelope cue: Tone-vocoder



Example

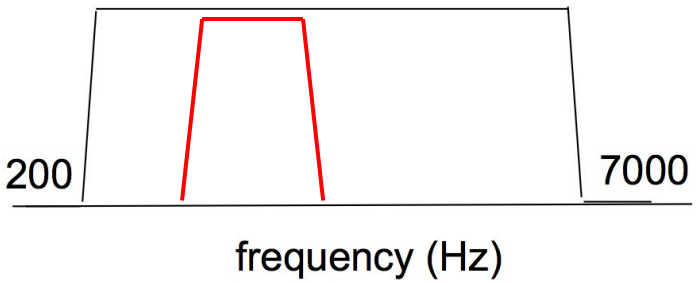
- Set the number of bands to N.
- For the i^{th} band
  1) Design band-pass (e.g., 400-1000 Hz) filter at band i
     >> fs=16000; %sampling frequency
     >> [b, a]=butter(4, [400 1000]/(fs/2)); %band-pass filter
  2) Do band-pass filtering at band i
     >> y= filter(b, a, s); % s is speech signal, and y is the band-passed signal at band i
  3) Do full-wave rectification, and low-pass filtering to get the envelope at band i
  4) Generate a sinewave, whose frequency equals to the center frequency of the i^{th} band-pass filter
  5) Multiply the envelope signal in 3) and sinewave in 4)



  6) Repeat for all N bands
  7) Sum up the outputs from all bands (denoting the summed outputs as s')
  8) Do energy normalization, i.e., letting the energy of s' equals to that of s
  9) Save the wavefile for signal s'

- Set the number of bands to N.
- For the i^{th} band
  1) Design band-pass (e.g., 400-1000 Hz) filter at band i
     >> fs=16000; %sampling frequency
     >> [b, a]=butter(4, [400 1000]/(fs/2)); %band-pass filter
  2) Do band-pass filtering at band i
     >> y= filter(b, a, s); % s is speech signal, and y is the band-passed signal at band i
  3) Do full-wave rectification, and low-pass filtering to get the envelope at band i
  4) Generate a sinewave, whose frequency equals to the center frequency of the i^{th} band-pass filter
  5) Multiply the envelope signal in 3) and sinewave in 4)

  6) Repeat for all N bands
  7) Sum up the outputs from all bands (denoting the summed outputs as s')
  8) Do energy normalization, i.e., letting the energy of s' equals to that of s
  9) Save the wavefile for signal s'



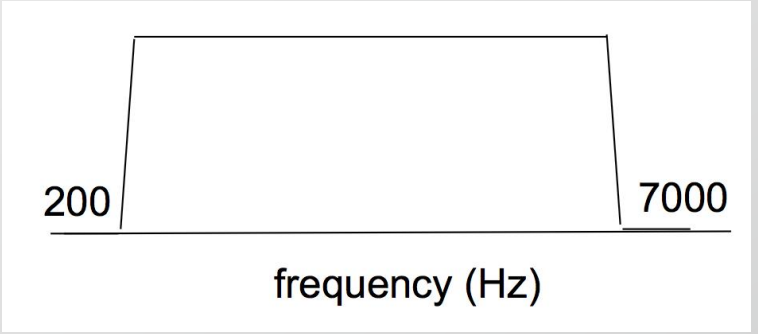
Question? How to divide pass-band from 200 Hz to 7000 Hz?

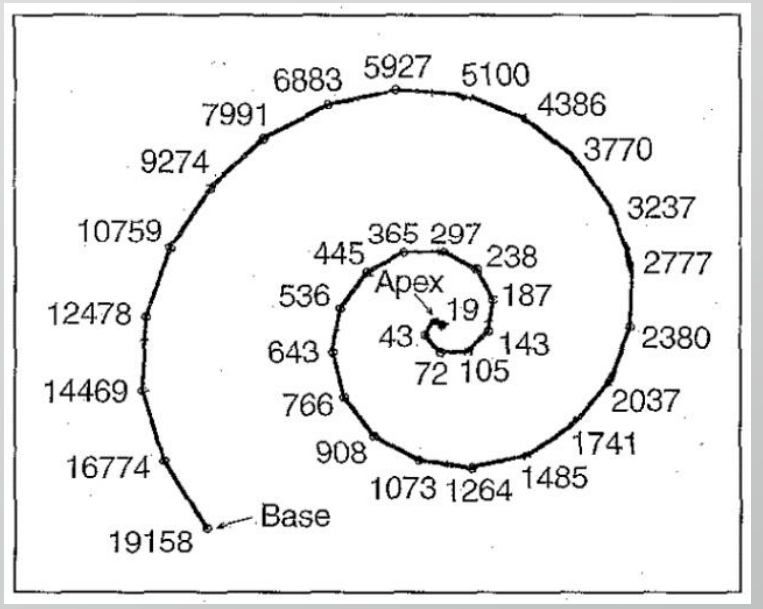
Determine the frequency range of the i^{th} band :

- Frequency-to-place mapping as:

$$f = 165.4 \times (10^{0.06 \cdot d} - 1)$$

where f is -3 dB cut-off frequency, and d is the distance (in millimeter) along the cochlea.





1) Equally divide the cochlea length.

2) Equal frequency interval division

Greenwood, D.D., 1990. A cochlear frequency-position function for several species—29 years later. J. Acoust. Soc. Am. 87, 2592–2605.

Project tasks -1

- **Sentences for pro 1: ‘C_01_01.wav’ & ‘C_01_02.wav’**
- **Task 1**
  - Set LPF cut-off frequency to 50 Hz.
  - Implement tone-vocoder by changing the number of bands to N=1, N=2, N=4, N=6, and N=8.
  - Save the wave files for these conditions, and describe how the number of bands affects the intelligibility (i.e., how many words can be understood) of synthesized sentence.

Project tasks -2

- **Task 2**
  - Set the number of bands N=4.
  - Implement tone-vocoder by changing the LPF cut-off frequency to 20 Hz, 50 Hz, 100 Hz, and 400 Hz.
  - Describe how the LPF cut-off frequency affects the intelligibility of synthesized sentence.

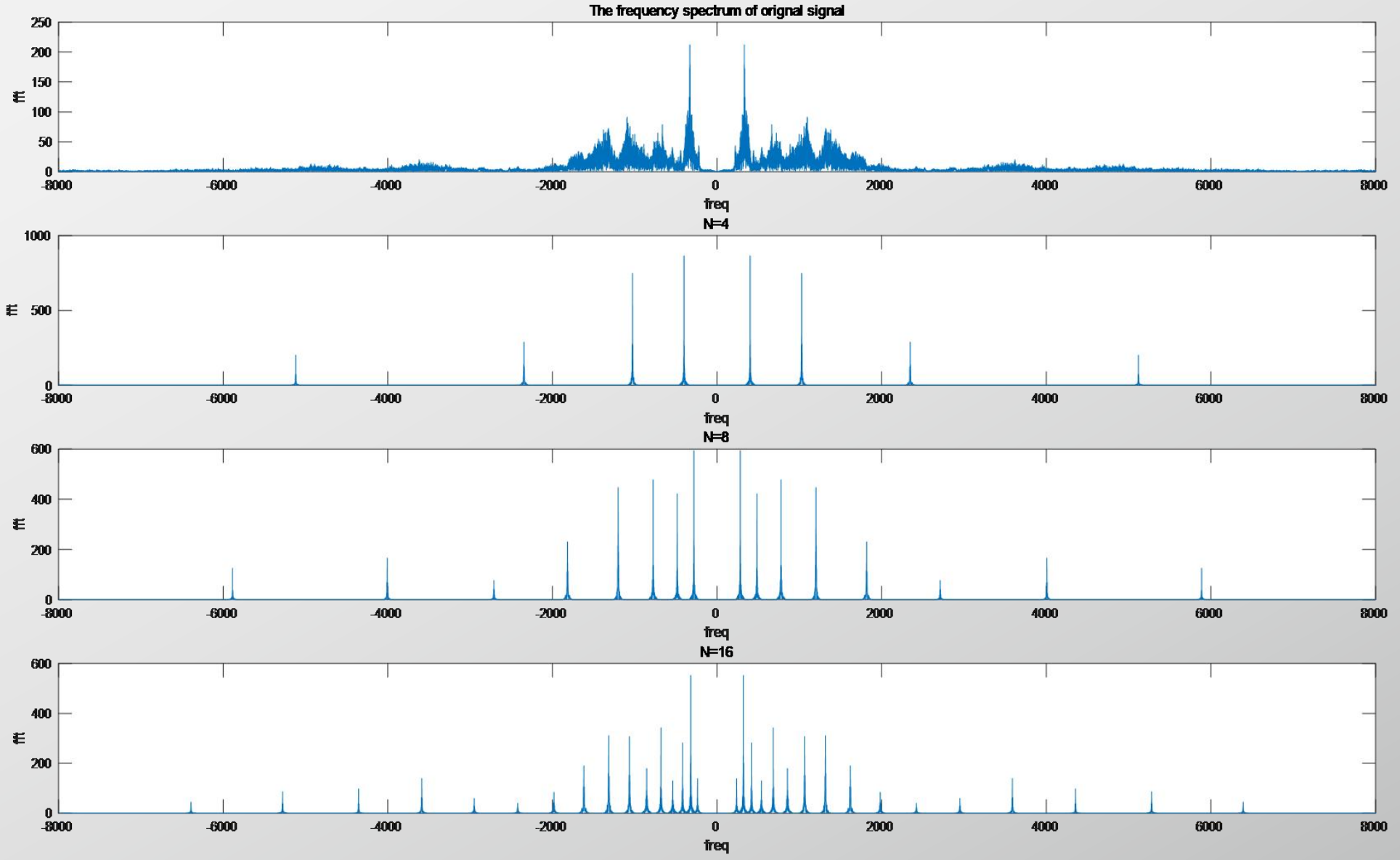
Project tasks -3

- **Task 3**
  - Generate a noisy signal (summing clean sentence and SSN) at SNR -5 dB.
  - Set LPF cut-off frequency to 50 Hz.
  - Implement tone-vocoder by changing the number of bands to N=2, N=4, N=6, N=8, and N=16.
  - Describe how the number of bands affects the intelligibility of synthesized sentence, and compare findings with those obtained in task 1.

Project tasks -4

- **Task 4**
  - Generate a noisy signal (summing clean sentence and SSN) at SNR -5 dB.
  - Set the number of bands to N=6.
  - Implement tone-vocoder by changing the LPF cut-off frequency to 20 Hz, 50 Hz, 100 Hz, and 400 Hz.
  - Describe how the LPF cut-off frequency affects the intelligibility of synthesized sentence.

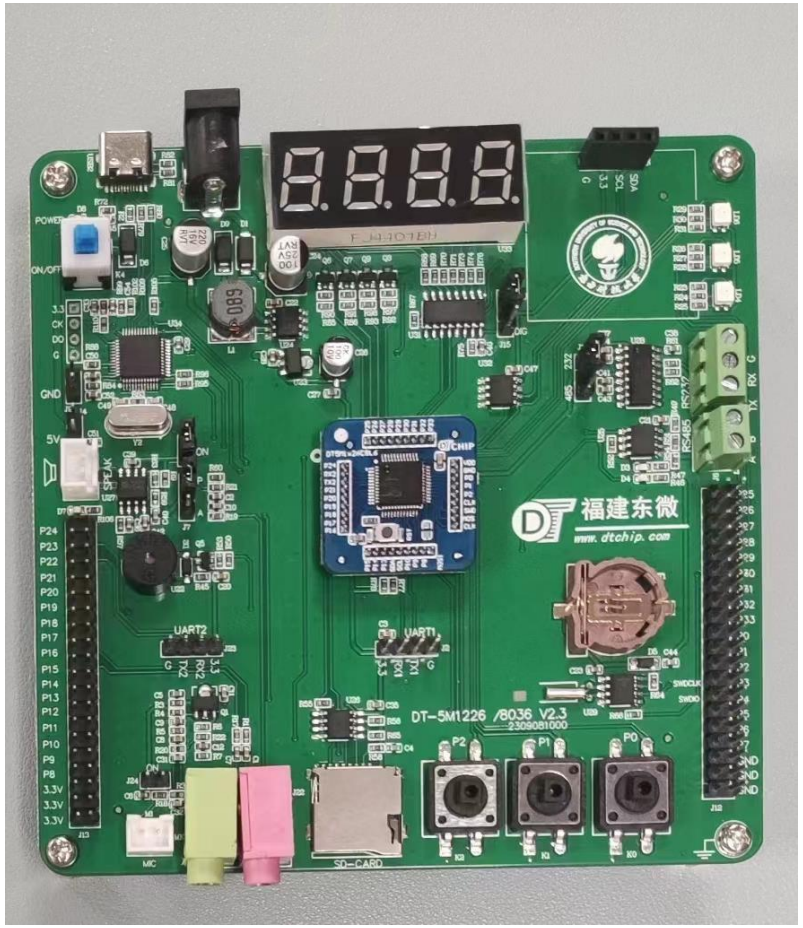
Results



硬件. DT5M1x26学习板简介—12.7培训会

实验目标：基于DT5M1x26芯片实现合成后的音频播放。

硬件介绍：DT5M1x26学习板是基于ARM Cortex-M0的架构的低功耗MCU板。板载Daplink通过Type-C口连接学习板和电脑，实现程序下载，将耳机或音响的音频线插入PCM OUT口，可以实现语音播放。

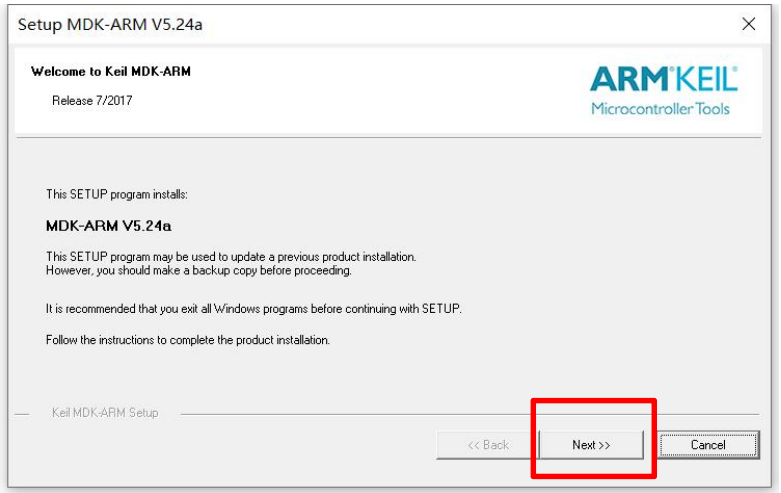


实验过程：

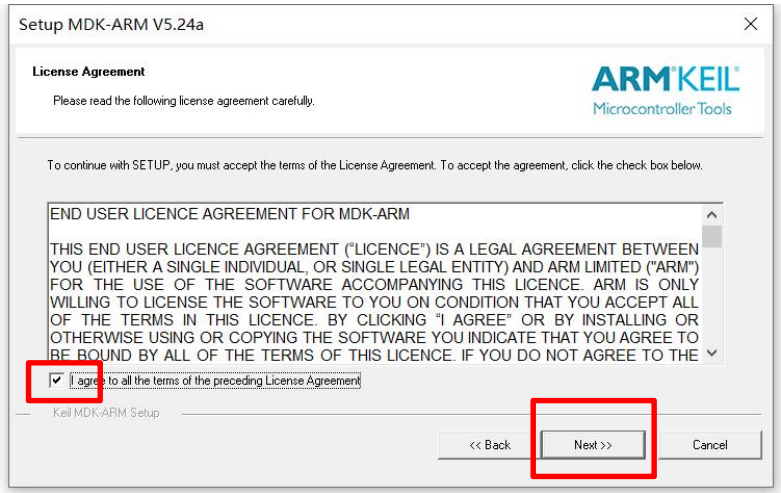
1、MATLAB进行语音合成信号处理

2、将合成后的音频数据导入嵌入式开发板DT5M1x26中，并利用其音频输出端口将语音信号播放出来。

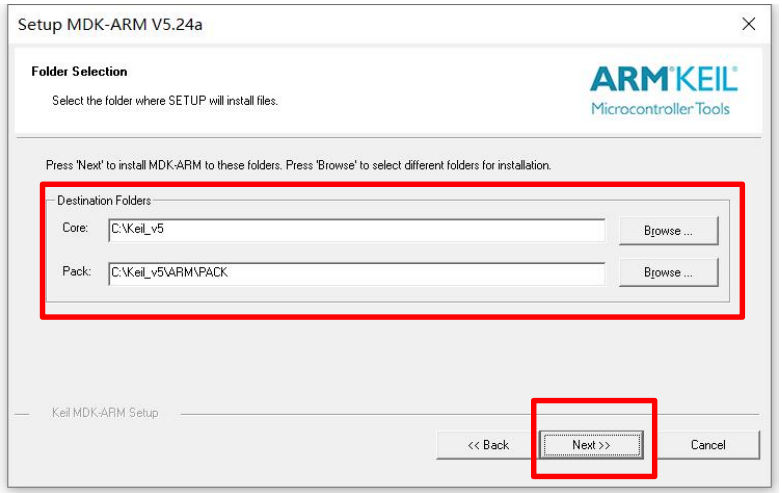
Preparation1： ARM KEIL 安装步骤-1



确认MDK-ARM

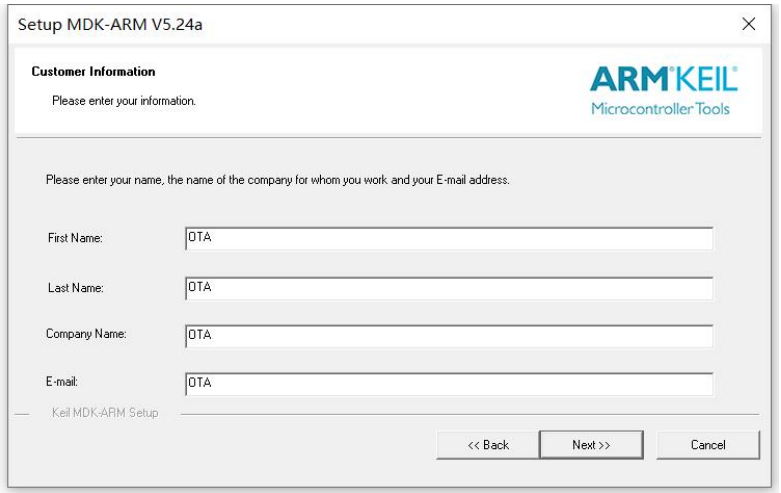


确认License

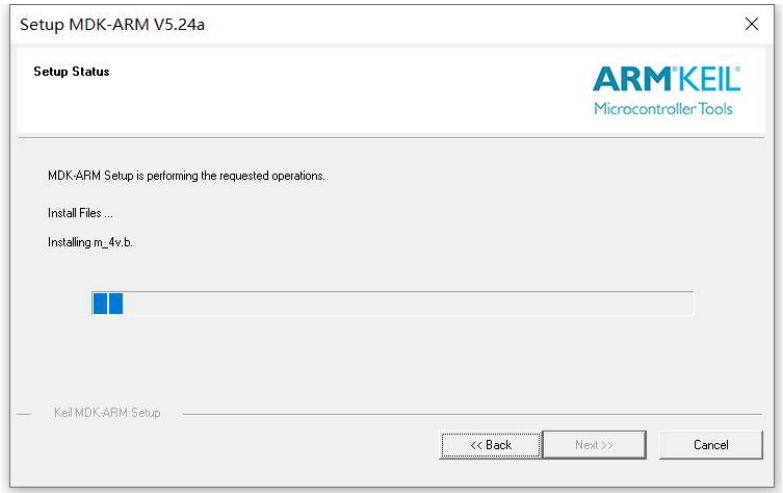


确认安装路径

ARM KEIL 安装步骤-2



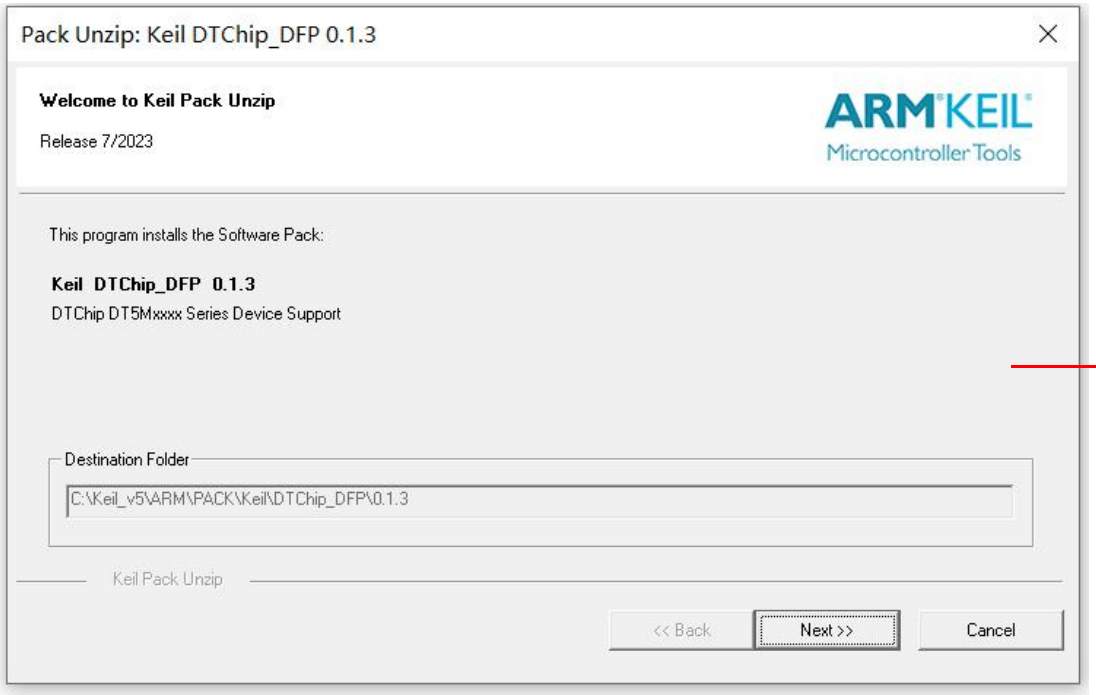
确认用户名

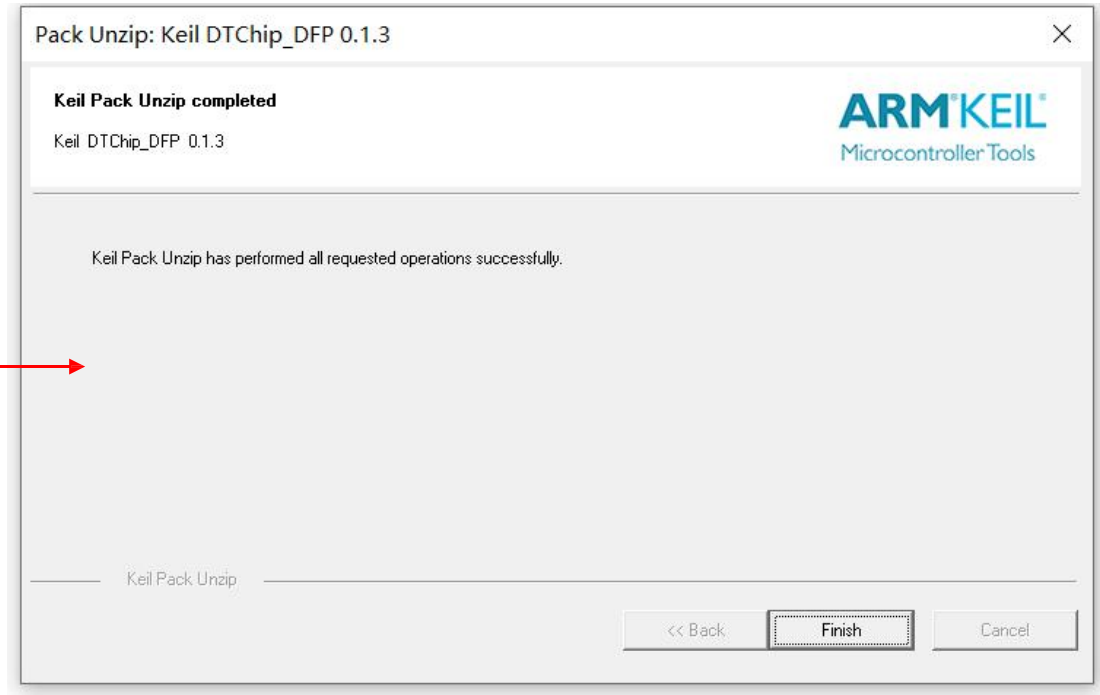


安装进程

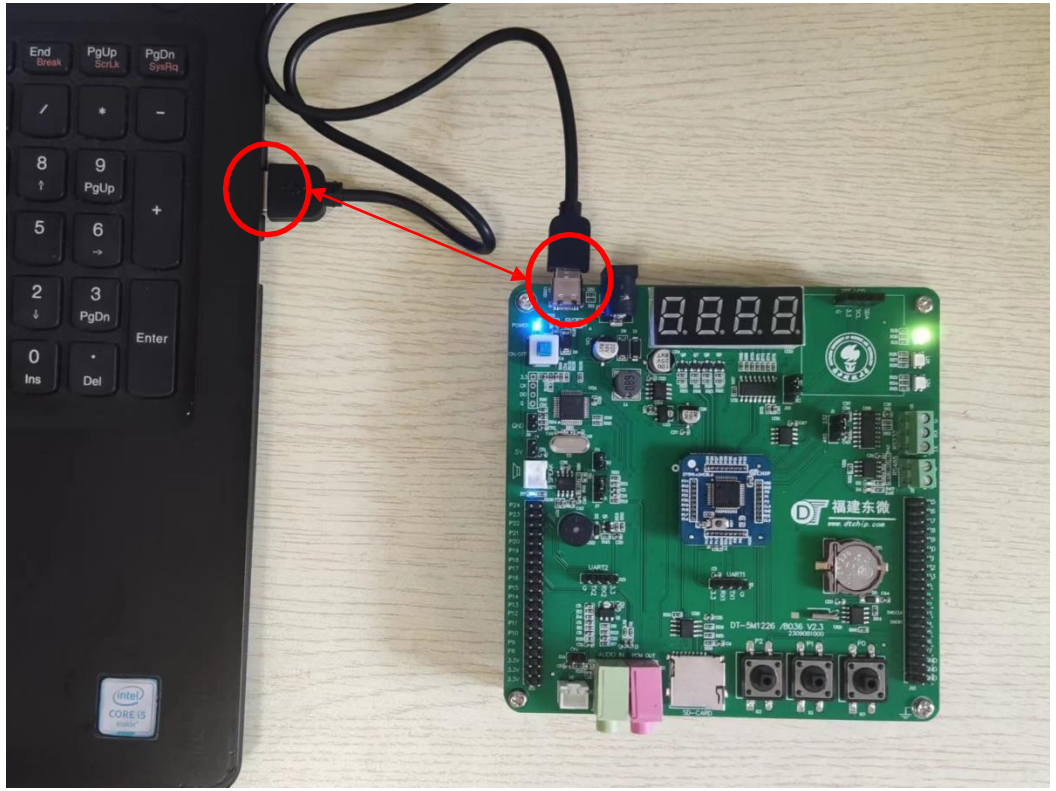
确认完成

安装keil.DTChip_DFP.0.1.3.pack芯片支持包





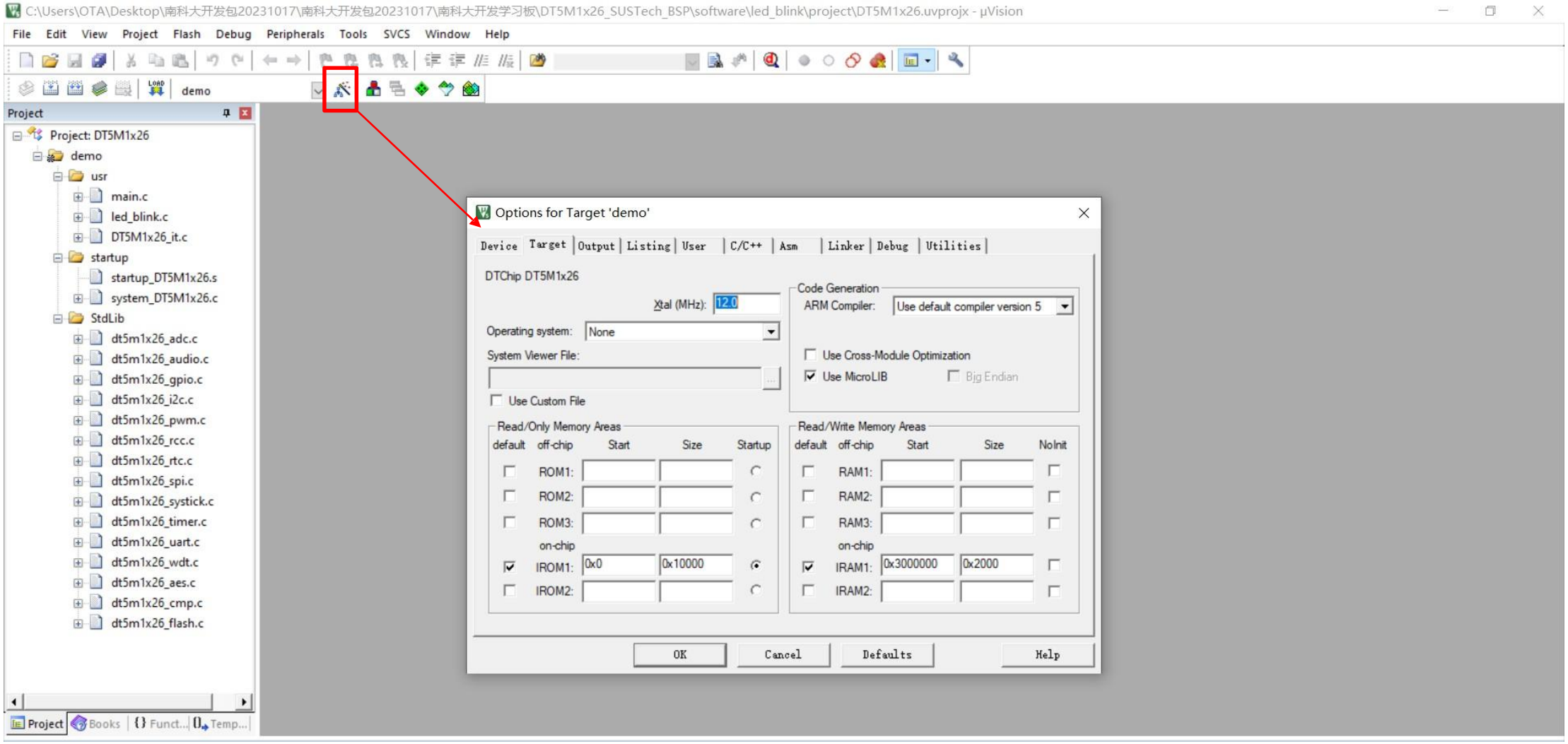
连接DT开发板和电脑，运行led_blink程序



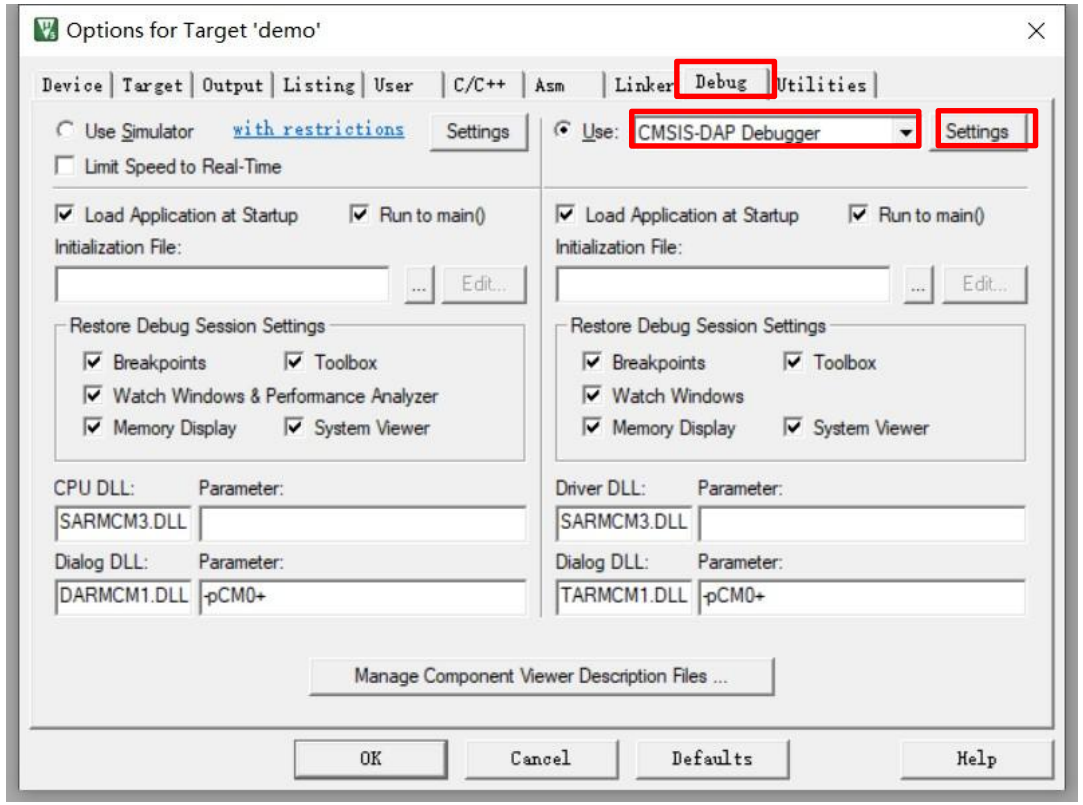
南科大开发包20231017 › 南科大开发学习板 › DT5M1x26_SUSTech_BSP › software › led_blink › project

名称	修改日期	类型	大小
Listings	2023/10/16 16:45	文件夹	
Objects	2023/10/17 15:53	文件夹	
b2f.exe	2023/8/31 16:29	应用程序	10 KB
bios.hex	2023/8/31 16:29	HEX 文件	9 KB
build_hex.bat	2023/8/31 16:29	Windows 批处理...	1 KB
code_rom	2023/8/31 16:29	文件	5 KB
code_rom.hex	2023/8/31 16:29	HEX 文件	9 KB
DT5M1x26.bin	2023/10/17 15:53	binimage	1 KB
DT5M1x26.uvguix.Administrator	2023/10/17 16:01	ADMINISTRATO...	91 KB
DT5M1x26.uvguix.HIAPAD	2023/8/31 16:29	HIAPAD 文件	101 KB
DT5M1x26.uvoptx	2023/8/31 16:37	UVOPTX 文件	14 KB
DT5M1x26.uvprojx	2023/10/17 11:45	礦ision5 Project	19 KB
EventRecorderStub.scvd	2023/8/31 16:29	SCVD 文件	1 KB
flash.ini	2023/8/31 16:29	配置设置	1 KB
JLinkLog.txt	2023/8/31 16:29	文本文档	4 KB
JLinkSettings.ini	2023/8/31 16:29	配置设置	1 KB

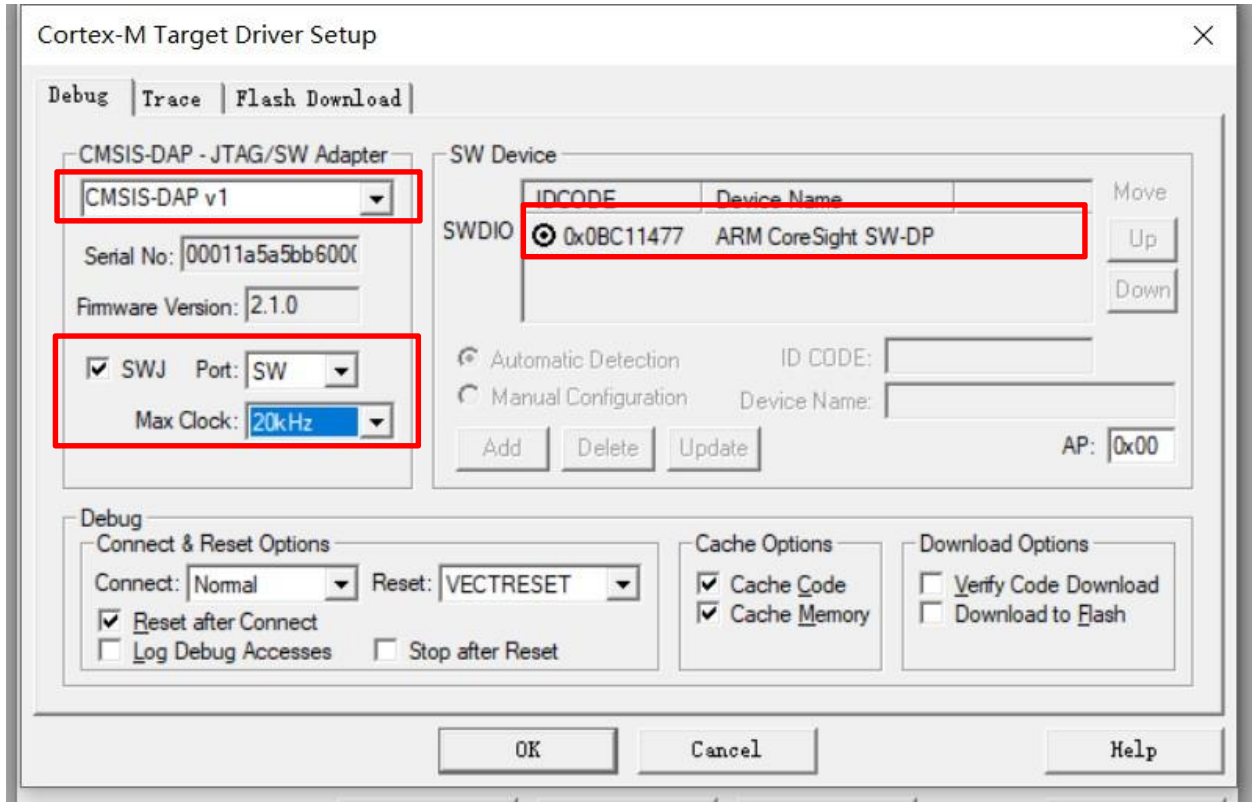
Options for Target



安装keil.DTChip_DFP.0.1.3.pack芯片支持包



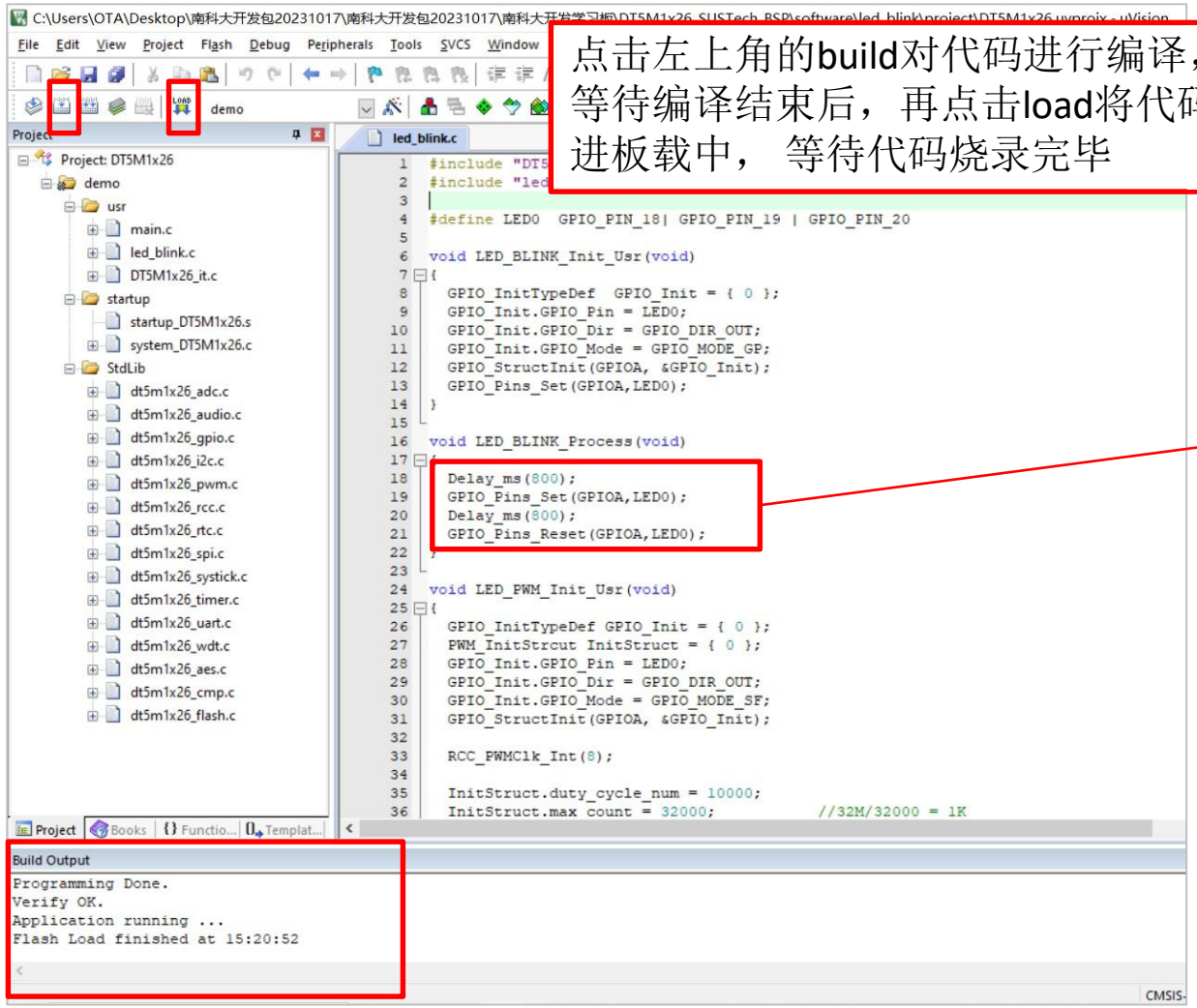
选择CMSIS-DAP Debugger选项，再点击setting进入下一步设置

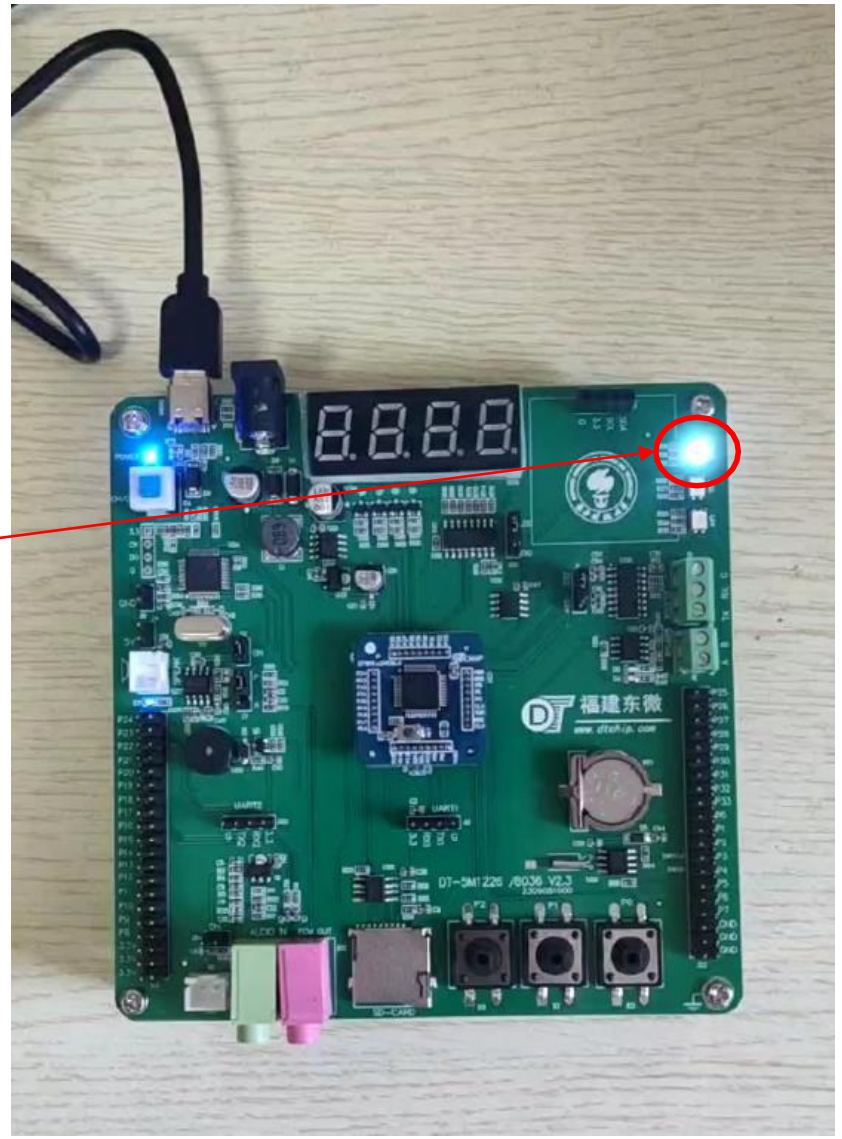


在SW Device中，若出现设备则说明开发板与电脑连接成功

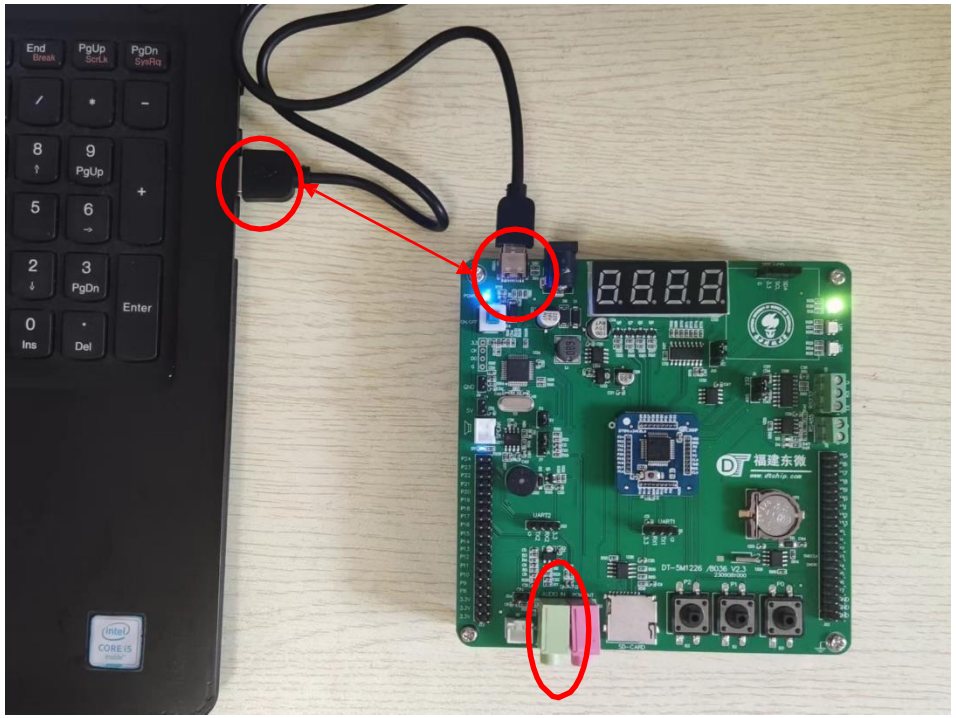
编译—>下载

点击左上角的build对代码进行编译，
等待编译结束后，再点击load将代码烧录进板载中， 等待代码烧录完毕





实验内容：使用DT5M1x26播放音频



Listings	2023/9/24 23:53	文件夹	
Objects	2023/9/24 23:53	文件夹	
b2f.exe	2023/8/8 14:01	应用程序	10 KB
bios.hex	2023/8/8 14:01	HEX 文件	10 KB
build_hex.bat	2023/8/8 14:01	Windows 批处理...	1 KB
code_rom	2023/8/8 14:01	文件	5 KB
code_rom.hex	2023/8/8 14:01	HEX 文件	10 KB
DT5M1x26.bin	2023/9/24 23:49	BIN 文件	62 KB
DT5M1x26.uvguix.Administrator	2023/8/9 15:46	ADMINISTRATO...	180 KB
DT5M1x26.uvguix.HIAPAD	2023/8/8 14:01	HIAPAD 文件	100 KB
DT5M1x26.uvguix.toLanzen	2023/9/24 23:53	TOLANZEN 文件	96 KB
DT5M1x26.uvoptx	2023/9/24 23:53	UVOPTX 文件	14 KB
DT5M1x26.uvprojx	2023/9/24 23:53	礦ision5 Project	18 KB
EventRecorderStub.scvd	2023/8/8 14:01	SCVD 文件	1 KB
flash.ini	2023/8/8 14:01	配置设置	1 KB
fliter.c	2023/9/8 20:59	C 文件	2 KB
JLinkLog.txt	2023/8/9 11:50	文本文档	286 KB
JLinkSettings.ini	2023/8/8 14:01	配置设置	1 KB

注意事项：

1. MATLAB中音频数据为浮点数，样例的c语言代码中音频数据需要int16整数，需要进行转换。
2. 由于板载Flash的内存有限，每次在代码中只能储存一定量的数据，请勿因存储过多而内存满。
3. 样例中使用的wav音频在长度和采样率上与需要播放的音频不同，需要修改部分数据。

step1：数据转换

1. MATLAB中音频数据为浮点数，样例的c语言代码中音频数据需要int16整数，需要进行转换。

step2：main.c函数中修改数据

testaudio.m *

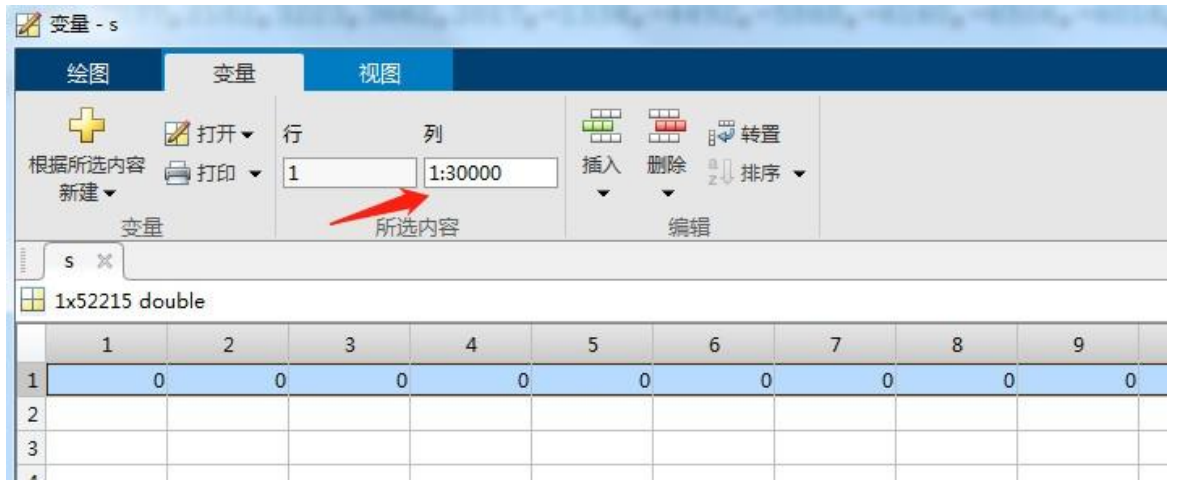
```
1  [s,fs]=audioread("C_01_01.wav");
2  s=s';
3  s=round(s*32768);
4  %round是四舍五入取整数，在取滤波后值时有效
```

注意事项：

1.设置行为1，列为1：30000（C_01_02为1:29636）

2.Ctrl+c复制数组文件，再Ctrl+v粘贴到命令行窗口

3.ctrl+a全选数组文件，并粘贴至main.c对应位置并进行修改



step3 设置采样率—>编译—>运行

查看采样率

在工作区中找到采样率fs，并在main.c代码中修改Sound_Init.sample_rate的赋值。

```
Sound_Init.sound_format = SOUND_FORMAT_PCM;//wav音频使用PCM采样
Sound_Init.sample_rate  = SOUND_SAMPLE_RATE_8K;//wav音频的采样率，对应fs
//对应SOUND_SAMPLE_RATE_8K,SOUND_SAMPLE_RATE_16K,SOUND_SAMPLE_RATE_44.1K
Sound_Init.pcm_frist_value = 0;
Sound_StructInit(&Sound_Init);
```

Organization

- **Each group consists of three or four students.**
- **Each group need present one Lab project (submit reports for both projects):**
- The presentation date is Dec. 14th or 15th for project 1， and Jan. 4th or 5th for Project 2.
- **Each presentation is 15 minutes (including Q & A)**
  - All team members need to contribute to the presentation.
  - Presentation in English or Chinese.

The presentation should...

- Introduce
  - team
  - objective of the project
  - background review (search more additional information)
  - methodology
- Present
  - relevant data, figure, etc.
  - the results for project tasks (e.g., with demo, Figure, etc.)
  - interpretation of project findings
- Discuss
  - what you have learned from this study?
  - problems during this project and your solution
  - investigation beyond project tasks
  - critical thinking
- Appendix (if any)
- Team effort (e.g., individual contribution)
- Reference
- Q & A (answer questions raised from audience)

Questions

saveas--Save figure to specific file format

Syntax

```
saveas(fig,filename)
saveas(fig,filename,formattype)
```

Description

`saveas(fig,filename)` saves the figure or Simulink® block diagram specified by `fig` to file `filename`. Specify the file name as a character vector that includes a file extension, for example, `'myplot.jpg'`. The file extension defines the file format. If you do not specify an extension, then `saveas` saves the figure to a FIG-file.To save the current figure, specify `fig` as `gcf`.

`saveas(fig,filename,formattype)` creates the file using the specified file format, `formattype`. If you do not specify a file extension in the file name, for example, `'myplot'`, then the standard extension corresponding to the specified format automatically appends to the file name. If you specify a file extension, it does not have to match the format. `saveas` uses `formattype` for the format, but saves the file with the specified extension. Thus, the file extension might not match the actual format used.

print--Print figure or save to specific file format
savefig--Save figure and contents to FIG-file