

# ICASSP 2022 Paper Reading List

---

SPE-55 - Speech Synthesis: Prosody

---

## 石颖

- SPE-3.6 序列对齐
- MLSP-12.1 NAS for KWS
- SPE-22.4 一种对关键词增强的搜索方法
- MLSP-18.2 curriculum-based data augmentation
- SPE-34.5 来给普及下生物学知识
- SPE-49.1 研究研究, 高端的样子 generative models for SE
- SPE-51.1/SPE-51.2 neural HMM
- SPE-58.3/SPE-58.5 acoustic modeling
- MLSP-36.3 feature-imitating network 带我们学习一下
- SPE-73.1 deliberation network 带我们学习一下
- SPE-76.6 CNN interpretation
- MLSP-54.1 universal audio representations

## 孙浩然

- SPE-11 - Speech Synthesis: Style & Expressiveness
- SPE-19 - Voice Conversion: Representation 几篇文章都挺不错
  - SPE-19.2 VC toolkit
  - SPE-19.3 有点意思, 是 cycle-consistency loss + random-sampling SSL
  - SPE-19.4 结论挺有意思, 离散->说话人信息, 连续->发音信息。
  - SPE-19.6 好文章, VAE + DA 实现 spk 和 content 解耦
- SPE-23.6 与你的工作相关, speech disentanglement: AIC + GAN
- SPE-27 - Voice Conversion I
  - SPE-27.2 Flow for text-free VC
  - SPE-27.3 Noisy factor
- SPE-35.3 Cyclic Training
- SPE-35.6 BNFs and disentanglement
- SPE-61.5 long-short speech coding
- SPE-66.1 6666666666
- MLSP-43.3 zero-shot in TTS
- MLSP-46.3 remix-cycle-consistent loss for separation.
- AUD-30.1 neural vocoder benchmark

## 陈琛、陈仁苗、江昊宇

- CHAL-5 - Audio Deepfake Detection 有些热度, 给我们简单分享一下 (苗)
- CHAL-6 - Multimodal Information Based Speech Processing
- SPE-2.1/SPE-2.2/SPE-2.3 多模态语音识别 (琛)
- MMSP-1.1 音视频事件检测

- AUD-6.1 深度好文
- AUD-6.4 多模态预训练
- AUD-6.5 说话人因子辅助音视频语音识别
- SPE-21.2 对抗样本 (苗)
- SPE-21.3 大规模录音重放数据集 (苗)
- SPE-45.5/SPE-45.6 AV for SE [VAE/GAN] (琛)
- SPE-47.3 Confidence estimation (苗)
- SPE-47.4/MMSP-6.2 Modality missing
- SPE-54.6 AV for WWS

## 陈琛

- IVMSP-28.4/SPE-85.4/SPE-85.5/ lipreading model
- SPE-60.5 data aug for AV learning
- SPE-70.4 lip-speech synchronization

## 陈仁苗、江昊宇

- IVMSP-30.4 将图片隐藏到音频中
- IFS-4.6 Open source for image generation
- MMSP-8.2 Text2Poster 娱乐一下

## 严子曦、李思瑞

- SPE-2.5 预训练模型用于 noisy ASR
- SPE-3.4 预训练模型用于 TTS
- MLSP-3 - Self-supervised Learning for Speech and Audio Processing I
- MLSP-6 - Self-supervised Learning for Speech and Audio Processing II
- SPE-14.2 Joint unsupervised, supervised and self-supervised training
- SPE-22.5 Wav2Vec + Cross-lingual adaptation (与你们之前的实验现象不太一致)
- SPE-47.5/SPE-47.6 Wav2Vec + ASR -> SRE
- AUD-11.6 W2V as SE prior
- SPE-30.3 W2V for LID
- SPE-31.3 W2V for SER
- MMSP-7.3 MAML for low-resource ASR

## 李鹏琦

- SPE-5 - Speaker Recognition I: Self Supervision
- SPE-13.2 Confidence estimator
- SPE-13.5 Explaining DNN for anti-spoofing
- SPE-21.4 Mix-up SSL 可能会对你下一步工作有所启发
- SPE-25 - Speaker Recognition IV: Attention Mechanism
- SPE-37.2 Household speaker identification
- SPE-45.1 简单一看 DOA for target speaker extraction
- SPE-48.4 简单一看 interpretability
- SPE-57.4 Mixup data aug
- SPE-57.6 MI for disentanglement

- MMSP-7.6 saliency masking
- SPE-68.4 graph-based attention
- SPE-76.6 CNN interpretation

## 候瑞海

- MLSP-4.2 哈希搜索
- MLSP-10.2 AE-based deep clustering
- SPE-21.5 GCN-based speaker clustering
- MMSP-6.4/MMSP-6.6 Deep hashing
- SPE-60.2 speech-image retrieval
- IVMSPP-31.6 can be applied to speaker diarization?
- SPE-72.1 speaker turn detection
- SPE-82.3 speaker diarization

## 文强

- SPE-6 - Speech and Spoken Language Corpora (重点是 SPE-6.4, 10种语言的辱骂声)
- AUD-2.1 婴儿哭声
- AUD-2.5 哭声、笑声、咳嗽、擤鼻涕
- SPE-9.1 抑郁症数据
- SPE-10.2 ASR + NLP -> 分音塔“敏感词”
- SPE-34.3 TorchAudio program
- SPE-48.5 CQCC 婴儿哭声
- AUD-18.3 COVID-19 数据, 下来听听
- MMSP-7.2 pre-trained audio model
- MLSP-43.4 GAN for noisy data simulation
- SPE-76.3 low-resource ASR
- AUD-35.1 engineering on audio event detection

---

## 下载地址:

---

- 链接: [https://pan.baidu.com/s/1JHDGzmqdnsseVER\\_CZByYA](https://pan.baidu.com/s/1JHDGzmqdnsseVER_CZByYA)
- 提取码: nyhf