

PSI-5759 Codificação de Voz

Sugestões não-exclusivas nem exaustivas para os Trabalhos de Pesquisa

12 de setembro de 2017

1. Codificação Preditiva
 - Predição de curto prazo ou intraquadro
 - Predição de longo prazo ou interquadro
 - Predição em Codificadores sem e com Perdas
2. Codificação por Sub-bandas e por Transformadas
 - Sinais de Voz: SBC, MELP, MBE, STC, MLT, TDAC, bancos de filtros, wavelets, “pitch-synchronous transforms”, etc.
 - Sinais de Áudio: MPEG-1, MP3, MPEG-2, AAC, ATRAC, PAC, TwinVQ, TDAC, etc.
3. Codificação com Dicionários de Códigos
 - Em Codificadores sem Perdas
 - Em Codificadores Preditivos
 - Em Codificadores com Decisão Atrasada
 - Sinais de Voz: CELP, ACELP, MELP, AMR, AMR-WB, AMR-WB+, VMR, G.72x.x, etc.
4. Quantização
 - Escalar, Vetorial
 - Adaptativa
 - Algoritmos de treinamento: LBG, PCA, GMM, etc.
 - Algoritmos de busca: busca em árvore, etc.
5. Codificação e melhoria do sinal para transmissão
 - Variação da qualidade com a taxa de transmissão para um dado modelo de codificador

- Codificação otimizada em taxa-distorção
- Técnicas de processamento do sinal de voz para compensar perda de quadros em codificadores para telefonia móvel
- Telefonia pela internet ou intranet: compensação da perda de pacotes, compensação de variações no atraso de transmissão (“jitter”)
- Qualidade de serviço do sinal de voz telefônico
- Algoritmos para determinação da taxa de transmissão (“rate determination algorithms – RDAs”)

6. Temas avançados de pesquisa

- Amostragem não-uniforme dos parâmetros da voz
- Análise e localização de eventos da fala
- Modelos paramétricos da fala
- Codificação de voz em taxas baixas
- Codificação de voz em taxas muito baixas
- Diversidade espacial no processamento de voz
- Emprego da aprendizagem profunda – deep learning – na codificação, síntese, análise e realce da voz. Por exemplo, através de “autoencoders”, máquinas de Boltzmann restritas (RBMs) ou “deep belief networks” (DBNs)