



پردازش صوت در تشخیص های پزشکی

Sound Processing in Medical Assessment

مقطع درس: تحصیلات تکمیلی

شماره درس: 1116079

تعداد واحد: 3 (نظری)

پیش‌نیاز: پردازش سیگنال های دیجیتال (DSP)

هدف: در این درس با پردازش های رایج بر روی سیگنال های صوت و گفتار آشنا می شوید. این پردازش ها در حوزه زمان ، فرکانس و کپستروم صورت می گیرند. نحوه استخراج انواع ویژگی های مفید از روی سیگنال های صوت و گفتار که در پروژه های تحقیقاتی شناسایی الگو کاربرد وسیع دارند مورد بررسی قرار گرفته و نحوه طراحی انواع طبقه بندهای رایج در این نوع پروژه ها ارائه خواهد شد. در نهایت کاربرد تکنیک های مطرح شده در چندین پروژه شناسایی بیماری ها از روی سیگنال های گفتار، اصوات ریوی و قلبی مطرح شده و جزئیات تکنیکی در این پروژه ها مورد بحث قرار می گیرند.

سرفصل مطالب درس: معرفی صوت و گفتار (امواج آکوستیک، مشخصات فیزیکی موج آکوستیک، امواج متناوب، غیرمتناوب و پیچیده، گفتار به عنوان یک موج آکوستیک)، سیستم تولید گفتار در انسان (فیزیولوژی و آناتومی تولید گفتار، پالس های چاکنایی و واکناری/بی واکی، فیلتر لوله صوتی در ارتباط با اندام تولید گفتار، سیگنال گفتار در حوزه زمان و فرکانس)، سیستم ادراک صوت و گفتار در انسان (عملکرد گوش انسان، کمیت های فیزیکی صوت در مقایسه با کمیت های ادراکی، شدت (Intensity) و مقیاس دسی بل، بلندی (Loudness) صدا، پیچ (Pitch) و فرکانس پایه گفتار، تیمبر (Timbre) پوش طیفی و زمانی گفتار)، واج شناسی گفتار (آواشناسی / واج شناسی، دسته بندی انواع واکه ها و همخوان های زبان، تعریف هجا و واج گونه های زبان، اثر هم تولیدی در گفتار، معرفی انواع واکه ها و همخوان های زبان فارسی)، پردازش های رایج سیگنال گفتار (مدلسازی سیگنال گفتار، ویژگی های مفید قابل استخراج از سیگنال زمانی گفتار، ویژگی های مفید قابل استخراج از سیگنال فرکانس پایه یا فرکانس پیچ یا فرکانس پایه از گفتار، استخراج پارامترهای LPC و مدل فیلتر لوله صوتی، استخراج ویژگی های دینامیکی جیتر و شیمر از روی گفتار، تبدیل کپستروم و کاربرد آن در پردازش صوت و گفتار (محاسبه تبدیل کپستروم از روی سیگنال در حوزه طیف و Z، استخراج ضرایب کپستروم در مقیاس مل (MFCC) ، کاربرد بردار بازنمایی MFCC و مشتقات آن)، طراحی مدل های بازشناسی مناسب در حوزه پردازش صوت و گفتار (معرفی طبقه بندهای NN ، SVM و نحوه استفاده در حوزه پردازش صوت، معرفی مدل های آماری GMM و HMM و نحوه استفاده در حوزه پردازش صوت، معرفی روش های پردازش گفتار جهت تشخیص برخی بیماری ها، تشخیص افسردگی با استفاده از پردازش گفتار، تشخیص پارکینسون با استفاده از پردازش گفتار، شناسایی انواع احساسات گوینده از روی گفتار، شناسایی گفتار پرخیشومی با استفاده از پردازش گفتار، شناسایی محل وقوع و نوع لکنت زبان با استفاده از پردازش گفتار، اصوات ریوی (ارتباط اصوات ریوی با آناتومی و فیزیولوژی آن، انواع ناهنجاری های ریوی و اثرات آنها بر روی صداهای ریوی، روش های ثبت چند کاناله صداهای ریوی، روش های مختلف پردازش چند کاناله اصوات ریوی)، اصوات قلبی (ارتباط اصوات قلبی با آناتومی و فیزیولوژی آن، انواع ناهنجاری های قلبی و اثرات آنها بر روی صداهای قلبی، پردازش صوت در شناسایی بیماری های قلبی).

منابع:

[1] Lawrence R. Rabiner, Ronald W. Schafer, "Introduction to Digital Speech Processing", 2007.

[2] Leontios J. Hadjileontiadis "Lung Sounds: An Advanced Signal Processing Perspective", Aristotle University of Thessaloniki, 2009.

- [3] Alain Marcha, "From Speech Physiology to Linguistic Phonetics", John Wiley, 2009,
- [4] Lawrence J. Raphael, Gloria J. Borden, Katherine S. Harris, "Speech Science Primer: Physiology, Acoustics, and Perception of Speech", Lippincott Williams & Wilkins. 2011.
- [5] Ben Gold, Nelson Morgan, Dan Ellis, "Speech and Audio Signal Processing: Processing and Perception of Speech and Music", second edition, John Wiley, 2011.
- [6] Jinyu Li, Li Deng, Reinhold Haeb-Umbach, Yifan Gong, "Robust Automatic Speech Recognition: A Bridge to Practical Applications", Elsevier Inc., 2015.
- [7] Nilanjan Dey, "Intelligent Speech Signal Processing", 1st Edition, Academic Press, 2019.
- [8] Uday Kamath, John Liu, James Whitaker, "Deep Learning for NLP and Speech Recognition", Springer, 2019.
- [9] Paul Hill, "Audio and Speech Processing with MATLAB", CRC Press .2020.