

Title: Inpainting methods for 2D data imputation

ID: L13

- Goals
 - Overview of image inpainting methods and their application to audio
 - Implementation of an audio inpainting method based on 2D solutions
 - Report
- Methodology
 - Become familiar with one inpainting method for each of the following classes:
 - Simple model-based method (e.g., https://scikit-image.org/docs/stable/auto_examples/filters/plot_inpaint.html, https://docs.opencv.org/master/df/d3d/tutorial_py_inpainting.html, etc.)
 - CNN-based method (e.g., <https://paperswithcode.com/paper/image-to-image-translation-with-conditional>, etc.)
 - Deep Image Prior (<https://github.com/DmitryUlyanov/deep-image-prior>)
 - Build a small dataset for testing considering different scenarios of interest (i.e., inpainting of regular or irregular regions of different size)
 - Select a short audio track
 - Compute a time-frequency representation of it
 - Provide an inpainting mask that shows the regions to be inpainted
 - The main experiment on an audio track is
 - Run an inpainting method on the time-frequency representation magnitude
 - Go back in the time domain using: (i) the original phase; (ii) Griffin-Lim method (<https://librosa.org/doc/main/generated/librosa.griffinlim.html>)
 - Run the main experiment for
 - Different time-frequency representation (e.g., STFT, Log-Mel Spectrogram, etc.)
 - The three selected inpainting methods
 - Compare the achieved results by means of quality metrics (e.g., SNR of the reconstructed audio, PSNR on the time-frequency magnitude, listening quality, etc.) considering the different tested scenarios (e.g., how “large” can the missing part be in order to be correctly inpainted?).
- Report
 - Motivation behind audio inpainting for audio applications (e.g., restoration, real-time communication error concealment, etc.)
 - Explain that it is possible to use image inpainting methods also in the audio realm
 - Overview of 2D data inpainting methods
 - Description of the used methods
 - Description of the performed experiments
 - Results and comments to compare the different techniques. Which one is the best and why?