26.9.2024

Unplugged Plug-And-Play audio declipper based on consensus equilibrium of DNN and sparse optimization <https://ieeexplore.ieee.org/document/10095928>

* Řeší se consensus ekvilibrium jako optimalizace pro audio declipping
* Výhodou CE je tady to, že kombinuje silné stránky obou přístupů k declipingu
  + První jsou metody založené na řídkosti, které:
    - Přímo používají vzorky původního signálu
    - Mají slabou performance v případě, že je signál silně poškozen clippingem
    - Předpokládají řídkost signálu ve spektru
  + Druhé jsou metody založené na Neuronových sítích, které
    - používají původní vzorky pouze nepřímo
    - mají dobrou úspěšnost i u silně poškozených signálů
    - DNN extrahují informace z datasetů, naučí se jak vypadá clipnutý signál a pak to použije pro declipping (somehow…)
    - Nemohou však přímo použít neclipnuté vzorky původního signálu, takže nemají tak dobrou performance na signály, které jsou jen jemně poškozeny clippingem. Neuronka se totiž obtížně učí co je to konzistentní clipping na signálech, které jsou jen málo poškozené
      * Nedaří se ji najít jakousi identity mapu?
  + **Gaborova transformace = STFT!**
  + **ILL-posed** = problém, který nemá jednoznačné řešení
* **Uhhh…**
* **Obsah obrázku text, snímek obrazovky, Písmo, dokument

  Popis byl vytvořen automaticky**
  + **Co definuje tato rovnice?**



* Proximity operator = operátor, který z vektoru dělá jiný vektor a závisí na funkci „malé f“
* Nemam šajn co se děje:
* Obsah obrázku text, snímek obrazovky, Písmo, číslo

  Popis byl vytvořen automaticky
* Obsah obrázku text, snímek obrazovky, Písmo, číslo

  Popis byl vytvořen automaticky
* Netušim vůbec jak číst kapitolu results
  + Co je to SDR a PESQ? Jak byly tedy porovnány výsledky jednotlivých metod a CE?

Plug and Play unplugged <https://arxiv.org/pdf/1705.08983>

* Plug & Play prior method – je existující metoda řešení restaurace signálu? (declipping, denoising..)
* Regularized inversion – je tedy ta úloha, pod kterou spadá declipping? **Najít původní reprezentaci neúplných dat**? To nemá jednoznačné řešení a proto jsou použity regularizéry, které to řešení mají stabilizovat.
* Cost function je nějaká funkce, která určuje, **jak moc signál splňuje kritéria pro to (měří kvalitu**), aby byl považován za declipnutý?
  + Jeho komponenty forward a prior jsou pak ta existující řešení na declipping?
  + V CE tedy nepotřebuji k rozhodování cost function?
  + S(G(x))
* MAP (něco jako ML cos dělal v robotě) se používá k nalezení původního signálu
  + V CE je ale nahrazeno novými rovnicemi, tzn CE zobecňuje MAP
* Už tady moc netušim:

Obsah obrázku text, snímek obrazovky, Písmo, číslo

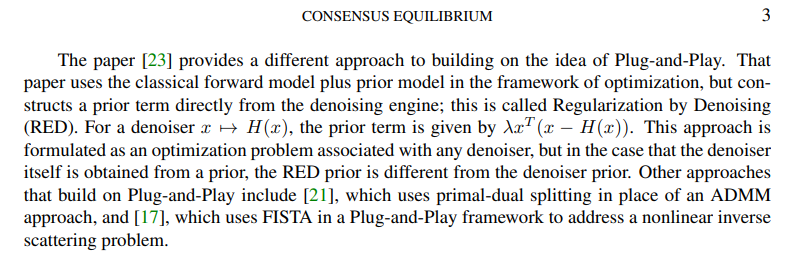
Popis byl vytvořen automaticky

* f 🡪 jak moc dobře signál odpovídá rekonstruovaným datům
* h 🡪 nějaké další požadavky, které na něj mám (třeba ta řídkost spektra)
* Přechod od MAP k CE

Obsah obrázku text, snímek obrazovky, Písmo

Popis byl vytvořen automaticky

* Tohle chápu tak, že minimalizuju součet výstupů, který dostanu z jednotlivých řešení (neuronek/algoritmů)
* **x\_i 🡪 je jeden dílčí restaurovaný signál jedné mé komponenty/agenta?**
* **f\_i 🡪 je nějaká funkce, která nějak analyzuje každý z těchto signálů?**
  + A já se tedy snažím všechna tyto f\_i opakovaně počítat, dokud se mi jejich hodnoty nepodaří minimalizovat pod nějaký threshold?
* Cílem tohoto přístupu, tzv. Plug-and-Play prior method**, je nahradit apriorní model v bayesovské formulaci** (ta rovnice s pravděpodobností) denoisujícím (declipping) operátorem
* Na stranách 2 a 3 jsou různý odkazy na články, kde se někdo pokusil podobný model pro denoising použít
* Tady už se úplně ztrácim:



* **Co jsou to proximal maps?**
  + Hladká funkce = taková, která má spojitou první a druhou derivaci, tzn není tam nikde skok?
    - h(x), neboli „prior“ je často nehladká funkce, proto se to tam řeší
  + je to nějaká mapa funkcí, které mi analizují výstup každého z mých agentů a někam si ukládají nějaké hodnoty?
* **Co je to forward model a prior model?** Lze to chápat jedno jako klasickou iterační metodu pro řešení problému a to druhé jako metodu založenou na strojovém učení?
* **Spare = řídký,** tzn problémy sparse-optimization-based znamenají, že optimalizuji řešení problému za předpokladu „řídkosti“? (jako třeba ve spektru?)

Obsah obrázku text, snímek obrazovky, Písmo, řada/pruh

Popis byl vytvořen automaticky

* **Co představuje u?** je to nějaký aproximovaný signál – třeba průměr ze všech jednotlivých agentů?

## Matlab a Python

* oboje jde použít…
* knihovny v Pythonu
  + Pyroomacoustics
  + SigPY
  + Librosa
  + Spleeter
* Knihovny v Matlabu
  + Audio Toolbox
  + LTFAT (Large Time-Frequency Analysis Toolbox)
  + SPGL1 – řešení řídkých optimalizačních problémů, lze tam použít rekonstrukci signálů s penalizací (L1 norma) – netušim o co jde ale bylo to zmíněno v článku
* Pyton vs. Matlab:
  + Python má hodně knihoven a nástrojů, opensource
    - Je to standardní jazyk pro AI a machine learning, takže kdybych chtěl používat deep learning pro decliping, bude vhodnější
  + Matlab
    - Dělal sem v něm víc
    - Má hodně hotových nástrojů
    - Dobře se v něm dělají prototypy, protože umožňuje rychle vykreslovat průběhy
* Matlab a Python je možný propojit:
  + Matlab Engine for Python – z pythonu spouštim přímo kod z matlabu
  + Python-Matlab Bridge – volam Matlab funkce v pythonu
  + MATLAB’s py module – volam funkce z Pythonu v Matlabu – muzu tam tak pouzit pythoní knihovny

Z konzultace 26.9

Problematika kterou budeme řešit (Audio Interpolation):

* Inpainting/PLC (Packet loss conceilment)/Declicking(gramo)
* Podobnou situací je clipping v signálu
* Dekvantizace, opet podobna uloha, moc sem zaokrouhlil, chci to zvrátit
* Denoising taky spada sem
* Poškození signálu simulujeme abychom mohli porovnávat

GACELA – síť, která umí audio impainting

MAYAVOZ

Z konzultace středa 2.10 12:30