ℓ_1-minimization for Audio Inpainting

Václav MACH
with
Pavel RAJMIC, Zdeněk PRŮŠA, Nicki HOLIGHAUS

4th SPLab Workshop 2014

Motivation

Restoration of historic sound recordings:

- Phonograph Cylinders and their Copies *[Mach 2012]*
- Gramophone Records *[Rajmic, Klimek 2004]*
- Magnetic Tapes
- Oldest Compact Discs

Actual recordings:

- Amateurs using professional sound equipment
- Encoding errors
- VoIP packet drop-outs

Unsuitable for archiving but recommended for publishing!

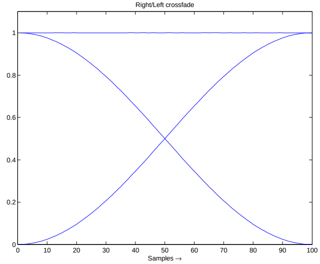

AR methods and related

State-of-the-art methods for audio interpolation

- Based on AR modelling of time samples (extrapolation)

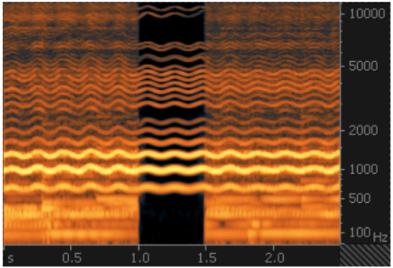
$$x[i] = \sum_{j=1}^{k} a_j x[i-j] + u[i],$$

- And crossfading of extrapolated parts
- *[Janssen 1986]*, *[Etter 1996]* etc.



Interpolation of partials from sinusoidal modelling

- *[Lagrange et al. 2005]*, *[Lukin, Todd 2008]*



Sparse representation of signals

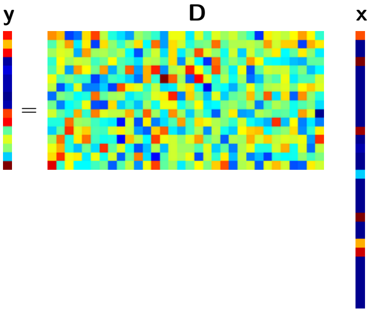
- Additive model of signal $\mathbf{y}$:

$$\mathbf{y} = \mathbf{D}\mathbf{x} \quad \text{or} \quad \mathbf{y} \approx \mathbf{D}\mathbf{x}$$

- known vector $\mathbf{y} \in \mathbb{C}^m$ (measurement, signal)
- matrix $\mathbf{D} \in \mathbb{C}^{m \times N}$, $m < N$, full column-rank (atoms, dictionary)
- infinitely many solutions of equations $\mathbf{y} = \mathbf{D}\mathbf{x}$
- We want the sparsest possible, i.e.

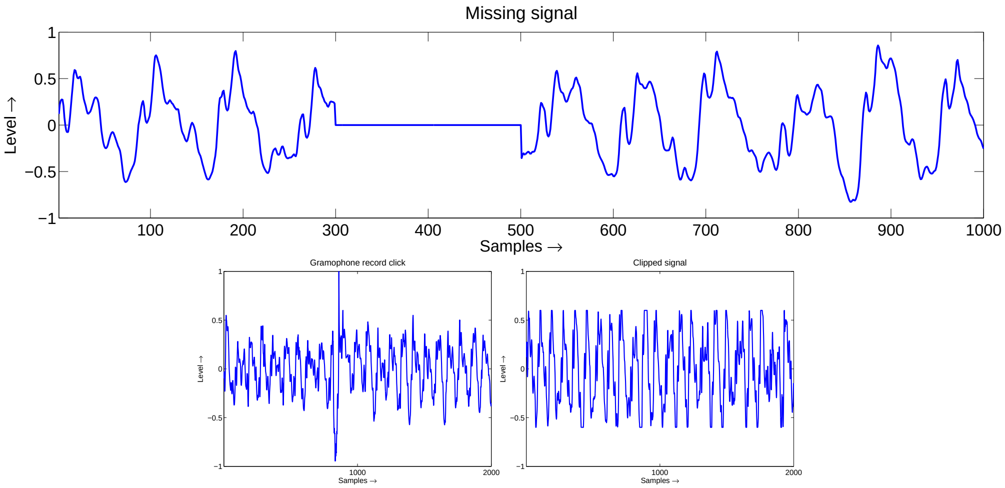
$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \mathbf{D}\mathbf{x} = \mathbf{y} \qquad \text{(P0)}$$

- But how to find it?



Audio inpainting

- Novel method for missing/distorted signal interpolation *[Adler et al., 2012]*
- Obtaining signal coefficients from "reliable" samples and reduced dictionary
- Restoration of missing samples with full dictionary



Obtaining coefficients by l1-minimization

LASSO - Least Absolute Shrinkage and Selection Operator *[Tibshirani 1996]*

$$\hat{\mathbf{y}}^{\mathrm{m}} = \mathbf{M}^{\mathrm{m}}\mathbf{D} \cdot \operatorname*{arg\,min}_{\mathbf{c}\in\mathbb{R}^N} \left(\frac{1}{2} \|\mathbf{M}^{\mathrm{r}}\mathbf{D}\mathbf{c} - \mathbf{M}^{\mathrm{r}}\mathbf{x}\|_2^2 + \lambda\|\mathbf{c}\|_1 \right)$$

- $\mathbf{M}^{\mathrm{r}}$... diagonal matrix, selection of 'reliable' dictionary atoms
- $\mathbf{M}^{\mathrm{m}}$... diagonal matrix, selection of 'missing' dictionary atoms
- $\mathbf{D}$... synthesis operator
- $\mathbf{x}$... measured signal (with gaps)
- $\hat{\mathbf{y}}^{\mathrm{m}}$... missing samples estimation
- $\mathbf{c}$... coefficient vector

Sparse Modeling

LASSO Solver: Proximal splitting algorithm

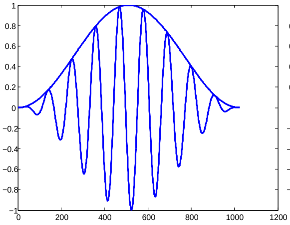
- Splits the optimization problem into two separate problems

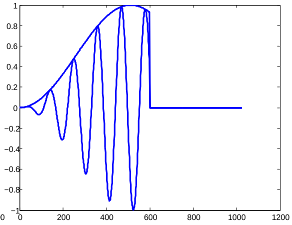
$$\hat{\mathbf{y}}^{\mathrm{m}} = \mathbf{M}^{\mathrm{m}}\mathbf{D} \cdot \arg\min_{\mathbf{x}\in\mathbb{C}^N} \left(\underbrace{\frac{1}{2}\|\mathbf{M}^{\mathrm{r}}\mathbf{D}\mathbf{c} - \mathbf{M}^{\mathrm{r}}\mathbf{x}\|_2^2}_{f_2} + \underbrace{\lambda\|\mathbf{c}\|_p}_{f_1} \right).$$

- f_2 is differentiable → Forward-Backward algorithm

Incorporating of weights for depleted atoms/coefficients

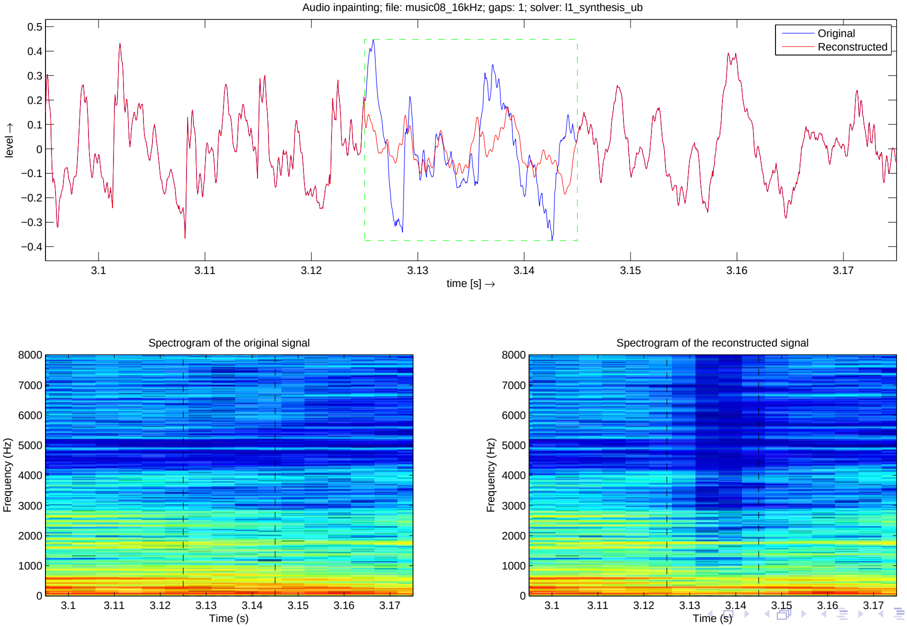
$$\mathbf{w}_i = \|\mathbf{g}\|_2^2 - \frac{|\mathbf{m}^{\mathrm{m}} * \mathbf{g}|_i}{\|\mathbf{g}\|_2}$$





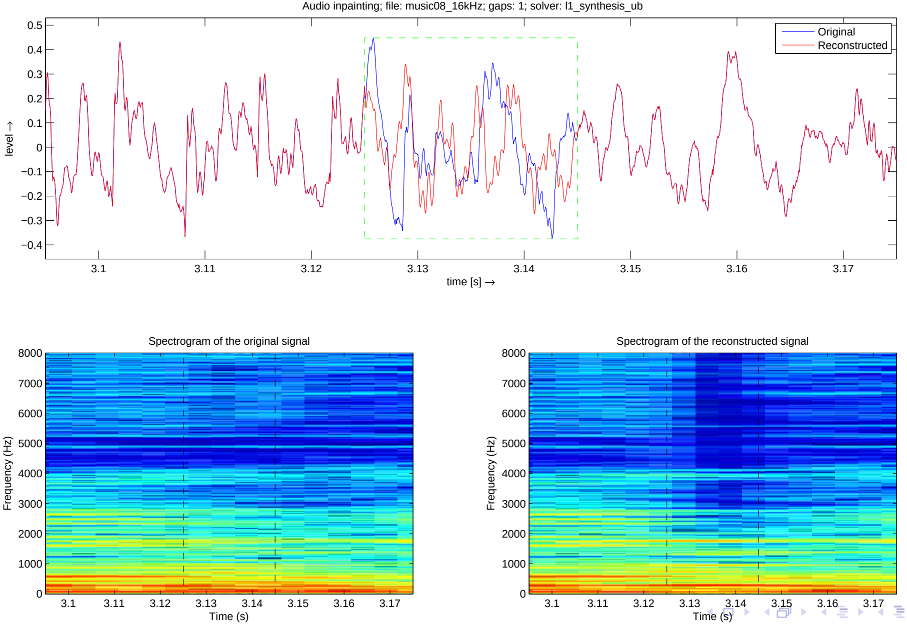
Experiments

NOT Using weights: SNR= 0.9342 dB



Experiments

Using weights: SNR$= -0.8793\,$dB



Conclusion, Future work

- Sparse representations in audio inpainting problem are beneficial
- Toolbox for Matlab developement

- **Structured sparsity** - persistency in time and/or frequency
- Analysis model - currently experimenting with several algoritms
- Complex testing on musical recordings of various genres
- Subjective evaluation

Thank you for your attention!