

University of Southern California

Live recording of multi-track audio

Joao Luiz Carvalho

EE522 Project

04/24/2007

University of Southern California

Motivation

- Studio recording
 - Each sound track is recorded separately
- Live Recording
 - Concerts, band practice
 - All tracks recorded simultaneously
 - Each microphone perceive sound from multiple sources
 - Overall quality is lower

University of Southern California

Objective

- Resolve the dry sound coming from each source in a live recording

University of Southern California

System Model

University of Southern California

System Model

University of Southern California

$$y_A = x_1 * h_{1A} + x_2 * h_{2A} + x_3 * h_{3A} + x_4 * h_{4A}$$

System Equations

$$y_A = x_1 \cdot h_{1A} + x_2 \cdot h_{2A} + x_3 \cdot h_{3A} + x_4 \cdot h_{4A}$$

$$y_B = x_1 \cdot h_{1B} + x_2 \cdot h_{2B} + x_3 \cdot h_{3B} + x_4 \cdot h_{4B}$$

$$y_C = x_1 \cdot h_{1C} + x_2 \cdot h_{2C} + x_3 \cdot h_{3C} + x_4 \cdot h_{4C}$$

$$y_D = x_1 \cdot h_{1D} + x_2 \cdot h_{2D} + x_3 \cdot h_{3D} + x_4 \cdot h_{4D}$$

Frequency Domain – Linear System

$$Y_A = X_1 \cdot H_{1A} + X_2 \cdot H_{2A} + X_3 \cdot H_{3A} + X_4 \cdot H_{4A}$$

$$Y_B = X_1 \cdot H_{1B} + X_2 \cdot H_{2B} + X_3 \cdot H_{3B} + X_4 \cdot H_{4B}$$

$$Y_C = X_1 \cdot H_{1C} + X_2 \cdot H_{2C} + X_3 \cdot H_{3C} + X_4 \cdot H_{4C}$$

$$Y_D = X_1 \cdot H_{1D} + X_2 \cdot H_{2D} + X_3 \cdot H_{3D} + X_4 \cdot H_{4D}$$

Matrix representation

$$\begin{pmatrix} Y_A \\ Y_B \\ Y_C \\ Y_D \end{pmatrix} = \begin{pmatrix} H_{1A} & H_{2A} & H_{3A} & H_{4A} \\ H_{1B} & H_{2B} & H_{3B} & H_{4B} \\ H_{1C} & H_{2C} & H_{3C} & H_{4C} \\ H_{1D} & H_{2D} & H_{3D} & H_{4D} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{pmatrix}$$

microphone recordings logsweep measurements dry sounds (unkowns)

System solution

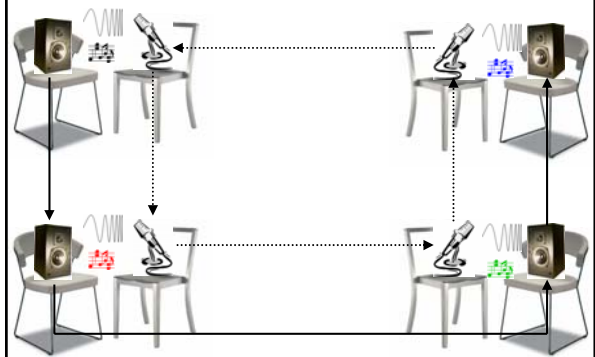
- Solve for each frequency component:

$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{pmatrix} = \begin{pmatrix} H_{1A} & H_{2A} & H_{3A} & H_{4A} \\ H_{1B} & H_{2B} & H_{3B} & H_{4B} \\ H_{1C} & H_{2C} & H_{3C} & H_{4C} \\ H_{1D} & H_{2D} & H_{3D} & H_{4D} \end{pmatrix}^{-1} \begin{pmatrix} Y_A \\ Y_B \\ Y_C \\ Y_D \end{pmatrix}$$

Experiments - Limitations

- Lots of noise
 - Repeated each recording 10 times
- Mic+Speaker System
 - Low and high frequencies very attenuated
 - Inverted response: very high gains (100dB +)
 - Noise amplification
 - Focused on 100 Hz to 8kHz band only
- Could only record one track at a time
 - Had to improvise

Experimental Setup




University of Southern California

Recordings

- Used Adobe Audition to obtain synchronized recordings
- 16 logsweeps (h_{ij})
- 16 audio clips (y_{ij})
 - $-y_A = y_{1A} + y_{2A} + y_{3A} + y_{4A}$
 - $-y_B = y_{1B} + y_{2B} + y_{3B} + y_{4B}$
 - $-y_C = y_{1C} + y_{2C} + y_{3C} + y_{4C}$
 - $-y_D = y_{1D} + y_{2D} + y_{3D} + y_{4D}$

University of Southern California

Results



SNR (dB)		
	Eq	Inv
Ch1	9.4	12.0
Ch2	6.4	11.2
Ch3	-0.4	11.6
Ch4	8.5	12.0

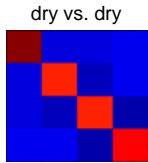
Simulations showed that this can be improved with more accurate room response measurements

University of Southern California

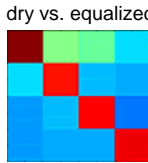
Auto-correlations

- Cross-correlation between the 4 audio clips (maximum value)

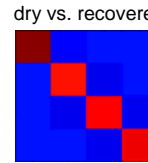
dry vs. dry

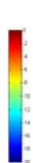


dry vs. equalized



dry vs. recovered





- Cross-talk successfully removed

University of Southern California

Conclusions

- Proposed method successfully removed cross-talk
- Consistently better than simply equalizing each channel
- SNR can be improved using more accurate room response measurements

Thank you!

Joao Carvalho 2007
jcarvalh@usc.edu