Linear estimation based primary-ambient extraction for stereo audio signals

Jianjun HE[a], Ee-Leng Tan[a], Woon-Seng Gan[a]

[a] DSP Lab, School of Electrical and Electronic Engineering,

Nanyang Technological University, Singapore, 639798, Singapore

Emails(jhe007@e.ntu.edu.sg, etanel@ntu.edu.sg, ewsgan@ntu.edu.sg)

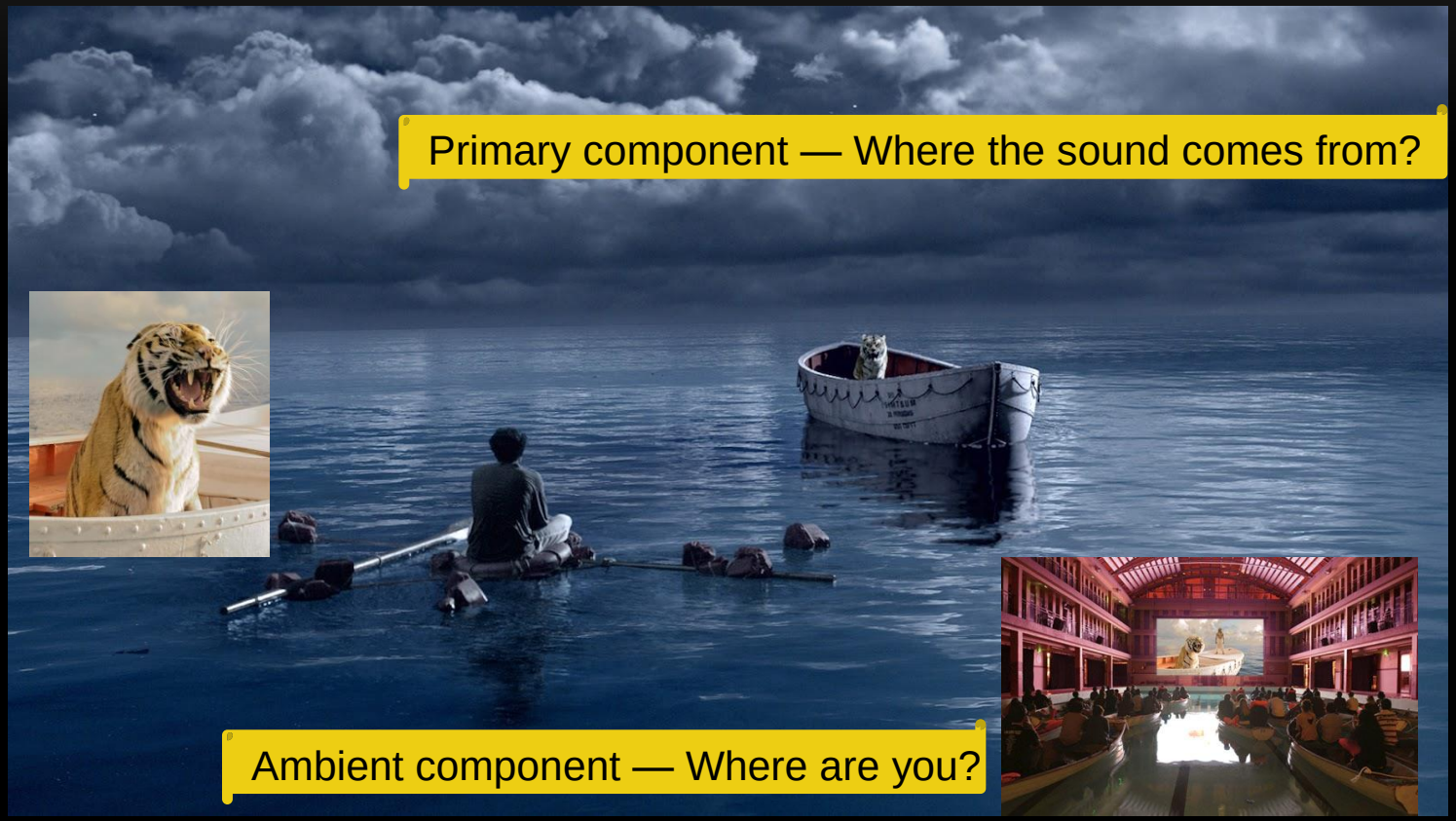
IEEE/ACM Transactions on Audio, Speech and Language Processing, vol. 22, no. 2, pp. 505-517, Feb. 2014

DOI: 10.1109/TASLP.2013.2297015

NANYANG TECHNOLOGICAL UNIVERSITY

Dsp Digital Signal Processing Lab

Introduction



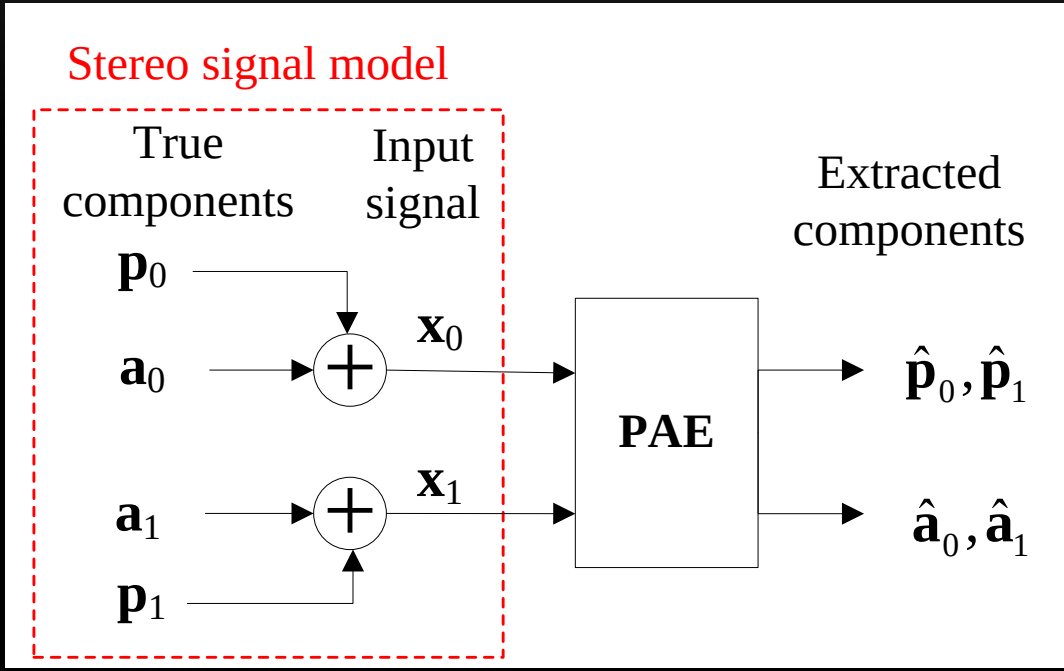
Applications of PAE

Various playback systems

PAE

- Spatial audio processing
- Spatial audio coding
- Audio re-mixing

Stereo Signal Model



Assumptions

Primary correlated	$\mathbf{p}_1 = k\mathbf{p}_0$
Ambient uncorrelated	$\mathbf{a}_0 \perp \mathbf{a}_1$
Primary ambient uncorrelated	$\mathbf{p}_i \perp \mathbf{a}_j$, $\forall i, j \in \{0,1\}$
Ambient power balanced	$P_{\mathbf{a}_0} = P_{\mathbf{a}_1}$

Primary panning factor PPF: $k = \dfrac{\mathbf{p}_1}{\mathbf{p}_0}$

Primary power ratio PPR: $\gamma = \dfrac{\text{Total primary power}}{\text{Total signal power}}$

Linear estimation framework in PAE

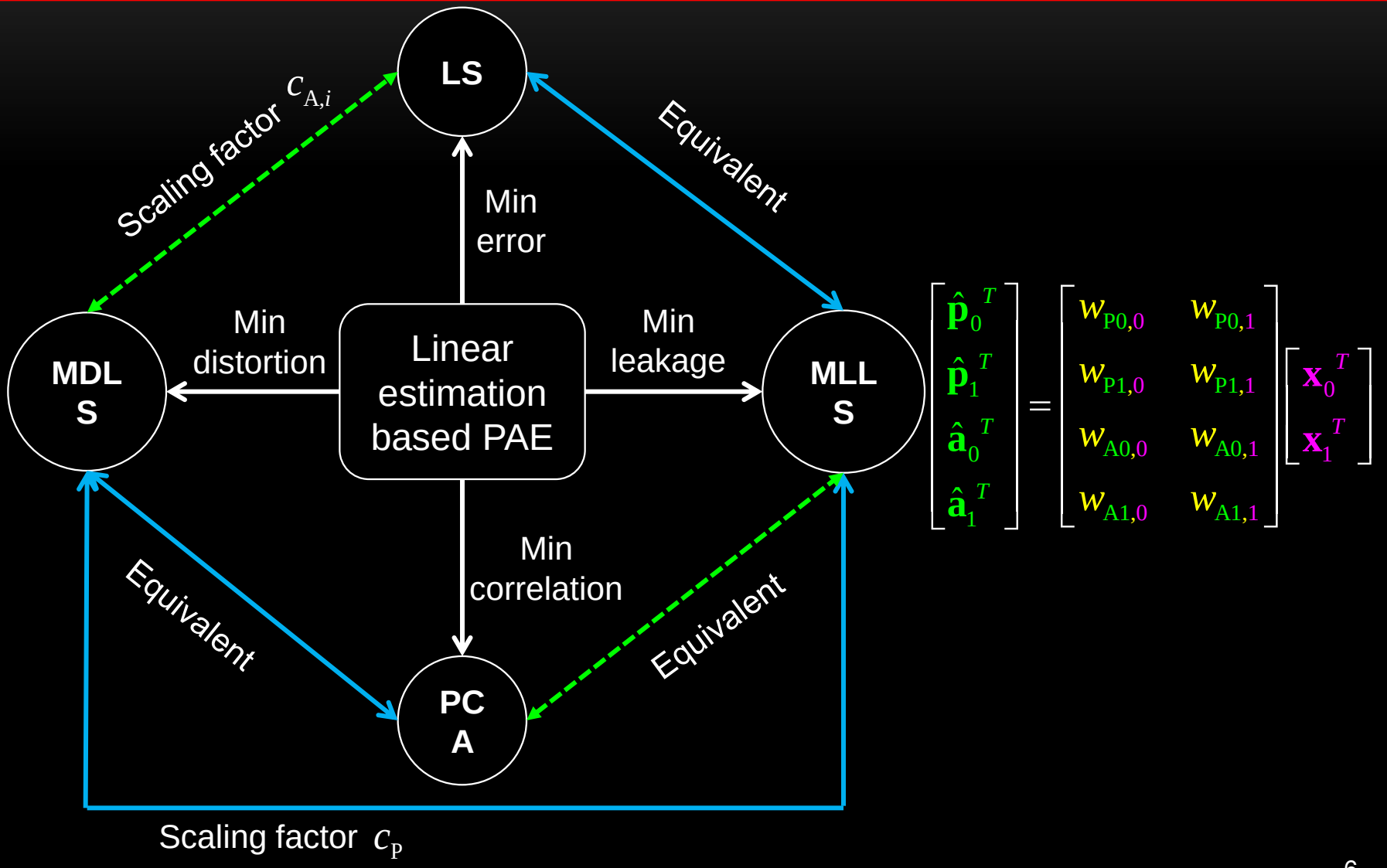
Signal = Primary + Ambient

$$\begin{bmatrix} \hat{\mathbf{p}}_0^{\,T} \\ \hat{\mathbf{p}}_1^{\,T} \\ \hat{\mathbf{a}}_0^{\,T} \\ \hat{\mathbf{a}}_1^{\,T} \end{bmatrix} = \begin{bmatrix} w_{\mathrm{P0},0} & w_{\mathrm{P0},1} \\ w_{\mathrm{P1},0} & w_{\mathrm{P1},1} \\ w_{\mathrm{A0},0} & w_{\mathrm{A0},1} \\ w_{\mathrm{A1},0} & w_{\mathrm{A1},1} \end{bmatrix} \begin{bmatrix} \mathbf{x}_0^{\,T} \\ \mathbf{x}_1^{\,T} \end{bmatrix} = \mathbf{W} \begin{bmatrix} \mathbf{x}_0^{\,T} \\ \mathbf{x}_1^{\,T} \end{bmatrix}$$

Error = **Dist**ortion + **Int**erf**e**rence + **Leak**age

Extraction Accuracy	ESR, DSR, ISR, LSR
Spatial Accuracy	ICC, ICTD, ICLD

PAE based on Linear estimation



Blue solid lines: primary component; Green dotted lines: ambient component.

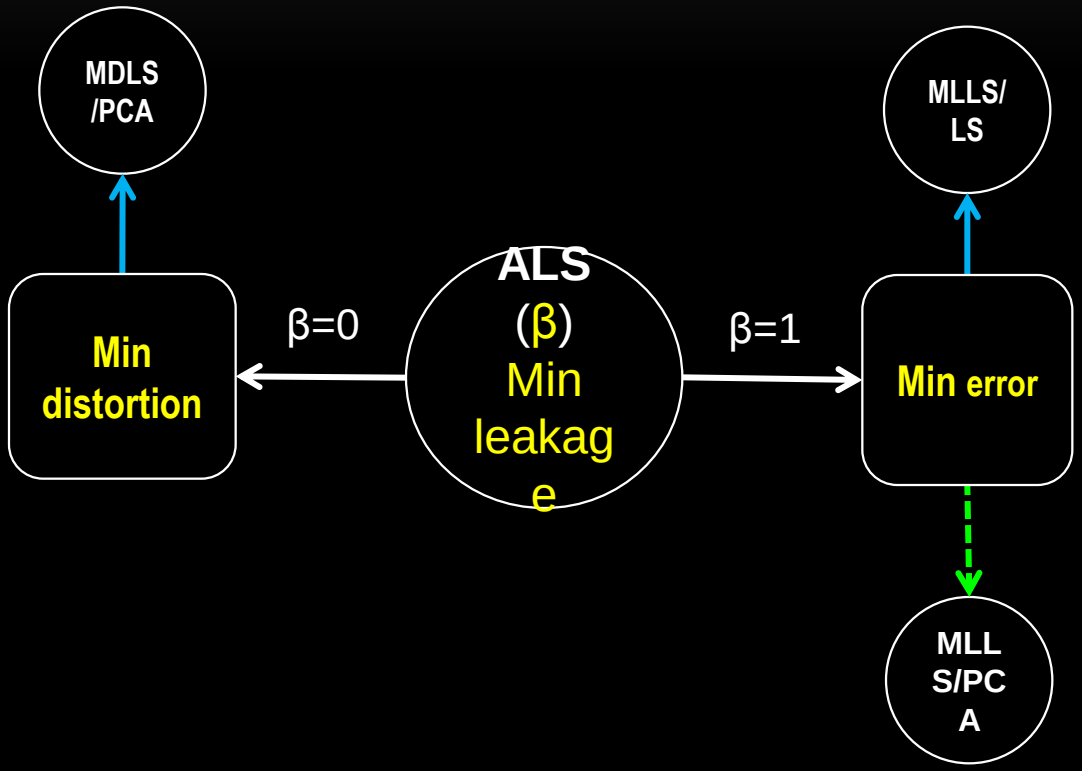
PAE based on Linear estimation
– Performance of the four approaches

$k : PPF$

$\gamma : PPR$

Performance Measure	Primary component		Ambient component		
	MDLS/PCA	MLLS/LS	MLLS/PCA	LS	MDLS
ESR	$\frac{1-\gamma}{2\gamma}$	$\frac{1-\gamma}{1+\gamma}$	$\frac{1}{1+k^2}$	$\frac{1}{1+k^2}\frac{2\gamma}{1+\gamma}$	$\frac{2\gamma}{1+k^2+\left(k^2-1\right)\gamma}$
LSR	$\frac{1-\gamma}{2\gamma}$	$\frac{1-\gamma}{2\gamma}\left(\frac{2\gamma}{1+\gamma}\right)^2$	0	$\frac{1}{1+k^2}\frac{2\gamma(1-\gamma)}{(1+\gamma)^2}$	$\frac{\left(1+k^2\right)(1-\gamma)2\gamma}{\left[\left(1+k^2\right)(1+\gamma)-2\gamma\right]^2}$
DSR	0	$\left(\frac{1-\gamma}{1+\gamma}\right)^2$	$\left(\frac{1}{1+k^2}\right)^2$	$\left(\frac{1}{1+k^2}\frac{2\gamma}{1+\gamma}\right)^2$	0
ISR	0		$\left(\frac{k}{1+k^2}\right)^2$	$\left(\frac{k}{1+k^2}\frac{2\gamma}{1+\gamma}\right)^2$	$\left[\frac{2k\gamma}{\left(1+k^2\right)(1+\gamma)-2\gamma}\right]^2$
ICC(ICTD)	1(0)		1	$\frac{2k\gamma}{\sqrt{\left(1+k^2\right)^2-\left(1-k^2\right)^2\gamma^2}}$	
ICLD	k^2		$\frac{1}{k^2}$	$\frac{1}{k^2}\frac{1+\gamma+k^2(1-\gamma)}{1+\gamma+\frac{1}{k^2}(1-\gamma)}$	$\frac{1}{k^2}\frac{1-\gamma+k^2(1+\gamma)}{1-\gamma+\frac{1}{k^2}(1+\gamma)}$

PAE using Adjustable Least squares (ALS)



Adjustable factor $\beta \in [0,1]$

$$\mathbf{W}_{\mathrm{ALS}} = \begin{bmatrix} \frac{1}{1+k^2}\left(1-\beta\frac{1-\gamma}{1+\gamma}\right) & \frac{k}{1+k^2}\left(1-\beta\frac{1-\gamma}{1+\gamma}\right) \\ \frac{k}{1+k^2}\left(1-\beta\frac{1-\gamma}{1+\gamma}\right) & \frac{k^2}{1+k^2}\left(1-\beta\frac{1-\gamma}{1+\gamma}\right) \\ 1-\beta\frac{1}{1+k^2} & -\frac{1}{k}\left(1-\beta\frac{1}{1+k^2}\right) \\ -k\left(1-\beta\frac{k^2}{1+k^2}\right) & 1-\beta\frac{k^2}{1+k^2} \end{bmatrix}$$

Blue solid lines: primary component; Green dotted lines: ambient component.

PAE based on Linear estimation
– Recommendations

Approach	Strengths	Weaknesses	Recommendations
PCA	• No distortion in the extracted primary component; • No primary leakage in the extracted ambient component; • Primary and ambient components are uncorrelated;	• Ambient component severely panned;	Spatial audio coding and interactive audio in gaming, where the primary component is more important than the ambient component.
LS	• Minimum MSE in the extracted primary and ambient components;	• Severe primary leakage in the extracted ambient component;	Applications in which both the primary and ambient components are extracted, processed, and finally mixed together.
MLLS	• Minimum leakage in the extracted primary and ambient components; • Primary and ambient components are uncorrelated;	• Ambient component severely panned;	Spatial audio enhancement systems, and applications in which different rendering or playback techniques are employed on the extracted primary and ambient components.
MDLS	• No distortion in the extracted primary and ambient components;	• Severe interference and primary leakage in the extracted ambient component;	High-fidelity applications in which timbre is of high importance.
ALS	• Performance adjustable;	• Need to adjust the value of the adjustable factor;	For applications without explicit requirements.

Conclusions

Formulated the linear estimation framework for PAE.

Introduced two groups of performance measures.

- Extraction accuracy: ESR, DSR, ISR, LSR
- Spatial accuracy : ICC, ICTD, ICLD

Proposed MLLS, MDLS, ALS and compared them with PCA and LS in PAE.

❑ **Primary extraction**

- PCA=MDLS: minimum distortion

 : a scaling factor difference

- LS=MLLS: minimum leakage & error

❑ **Ambient extraction**

- MLLS (=PCA), LS, MDLS minimize the leakage, error, and distortion, respectively

❑ ALS can achieve an adjustable performance for both primary and ambient components.

Different approaches are preferred in different spatial audio applications.