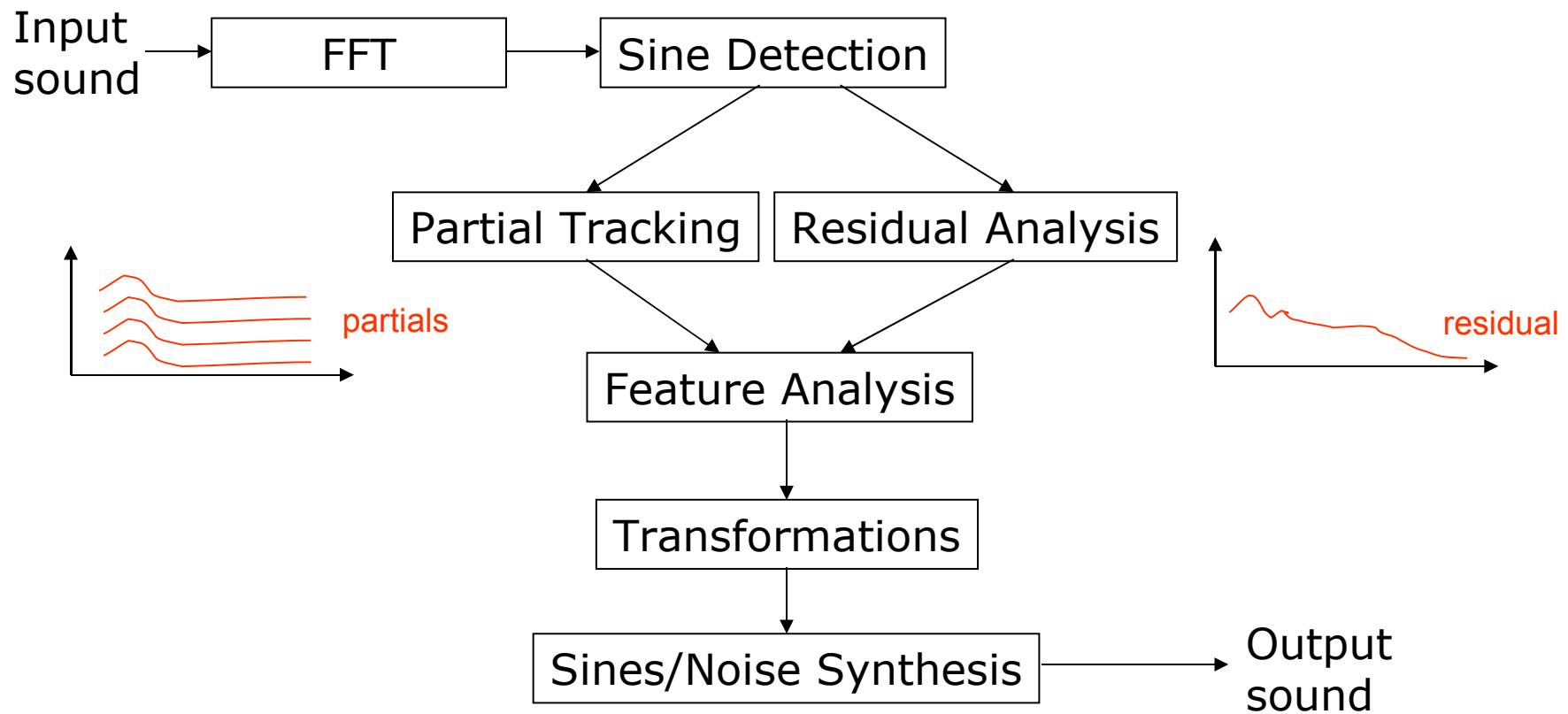


Approaches to Sinusoidal plus Residual Modeling

Xavier Serra
Music Technology Group
Universitat Pompeu Fabra, Barcelona

Sinusoidal plus residual framework



Relevant Research Topics (I)

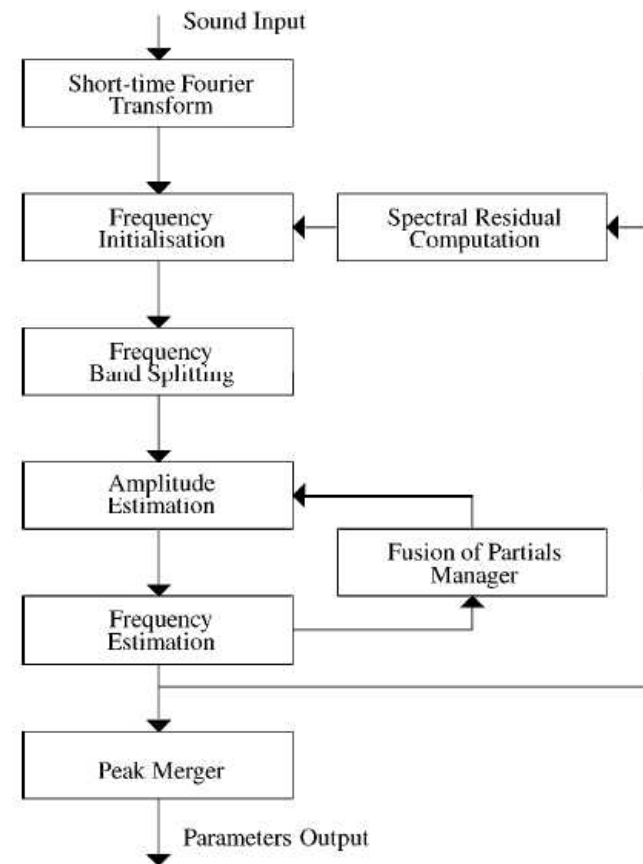
- Detection/Estimation of Sinusoids
- Partial Tracking
- Transient Modeling
- Multiresolution
- Residual Analysis/Modeling
- Feature-based Analysis/Synthesis
- Synthesis of Sinusoids/Noise

Relevant Research Topics (II)

- Morphing
- Time Scaling
- Compression/Transmission
- Source Separation/Transcription
- Music/Audio Content Analysis
- Instrument/Voice Models
- Expanded Models
- Software Environments

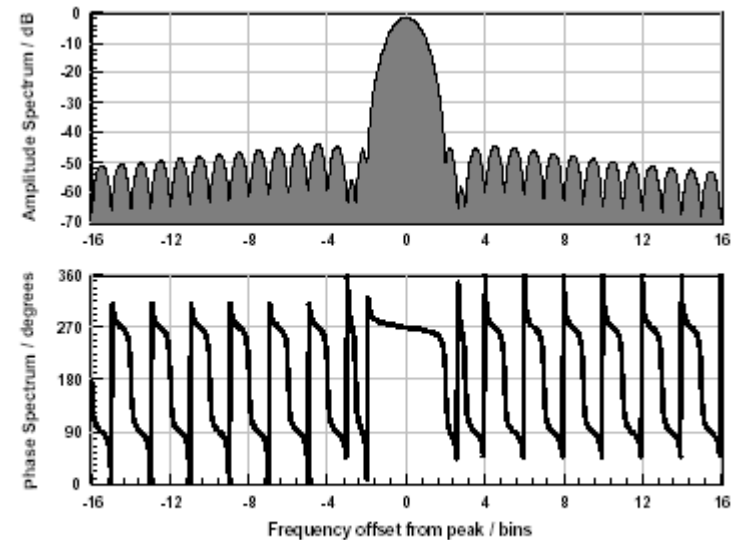
Detection/Estimation of Sinusoids (I)

- George, E. B. (thesis, 1991)
 - Analysis by synthesis where each sinusoid is subtracted one at a time.
- Depalle, P.; Hélie, T. (WASPAA, 1997)
 - parametric modeling of the STFT.
- Goodwin, M. (thesis, 1997)
 - Matching Pursuit.



Detection/Estimation of Sinusoids (II)

- Ding, Y.; Qian, X. (ICMC, 1997)
 - Global optimization of phase using B-spline quadratic polynomials.
- Masri, P. (DAFX, 1998)
 - Measurement of distortion.
- Marchand, S. (DAFX, 1998)
 - n signal derivatives.
- Hainsworth, S.; Macleod, M. (DAFX, 2003)
 - Frequency reassignment.



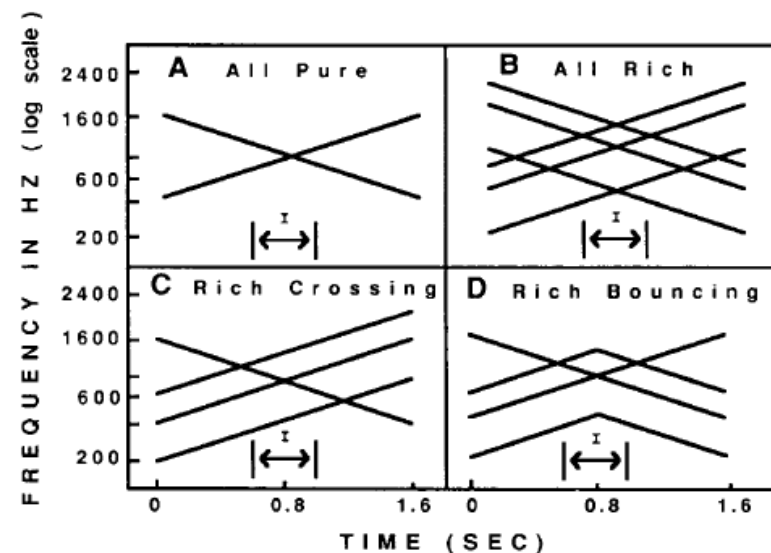
(a) - Rising amplitude: $\frac{d(\log A)}{dt} = +3$ dB per frame

Partial Tracking

- Depalle, Ph. et alt. (ICMC, 1993)
 - Hidden Markov Models.

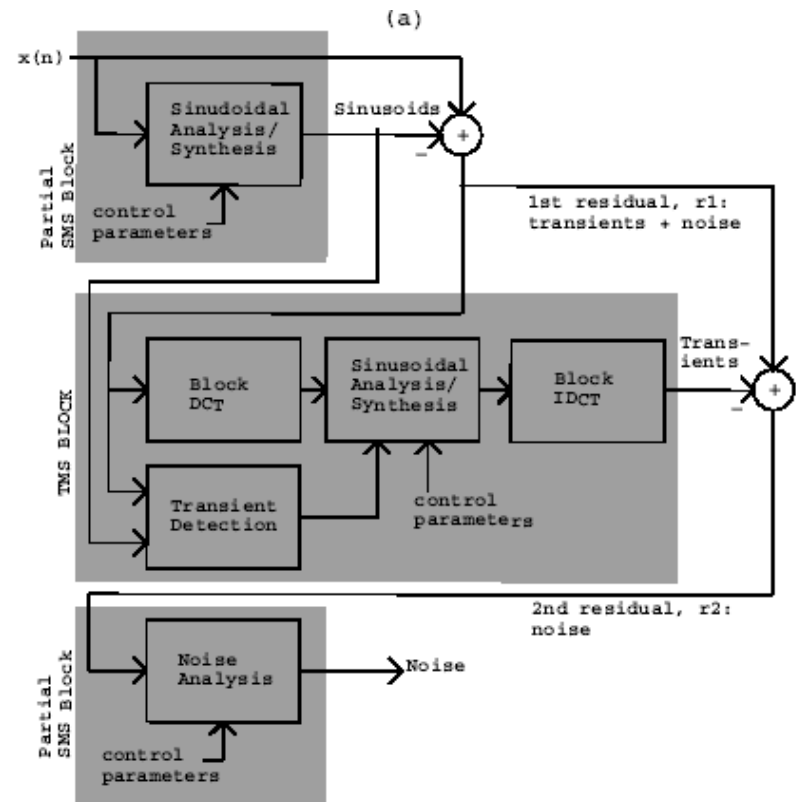
- Wang, A. (thesis, 1994)
 - Frequency-Lock Loop.

- Lagrange, M. et alt. (DAFX, 2003)
 - Linear prediction of the frequency evolutions.



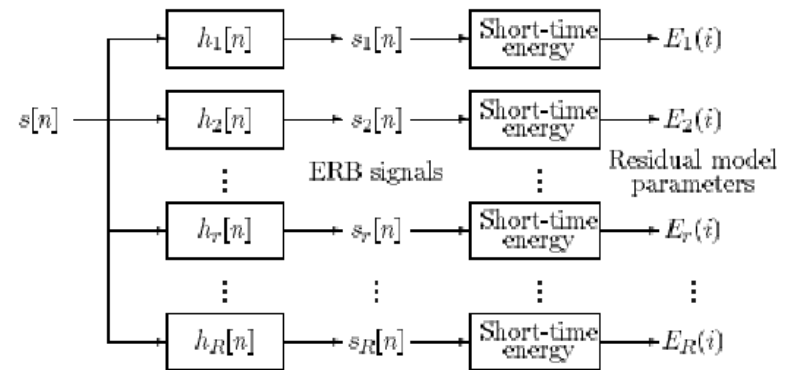
Transient Modeling

- Masri, P. (thesis 1996)
 - Analysis of transients to position analysis window.
- Ali, M. (thesis, 1996)
 - Wavelet Analysis for transients.
- Verma, T. et al. (ICMC, 1997)
 - Sinusoids+Transients+ Noise Model.



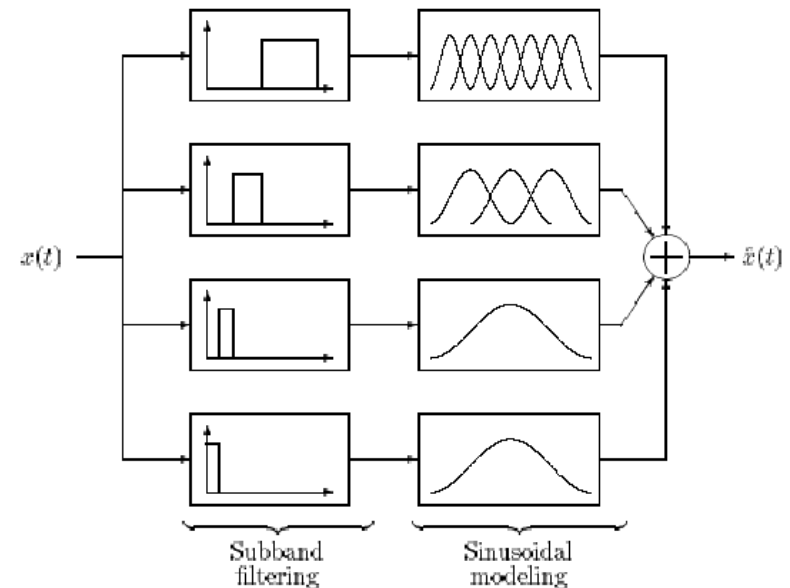
Residual Analysis/Modeling

- Hamdy, K. N. et al. (ICASSP, 1996)
 - Wavelet coding of residual.
- Goodwin, M. (thesis, 1996)
 - Filter-bank auditory model.
- Ding Y.; Qian, X. (ICMC, 1997)
 - LPC modeling.
- Desainte-Catherine, M; Hanna, P. (DAFX, 2000)
 - Parameterization of noise-like sounds.



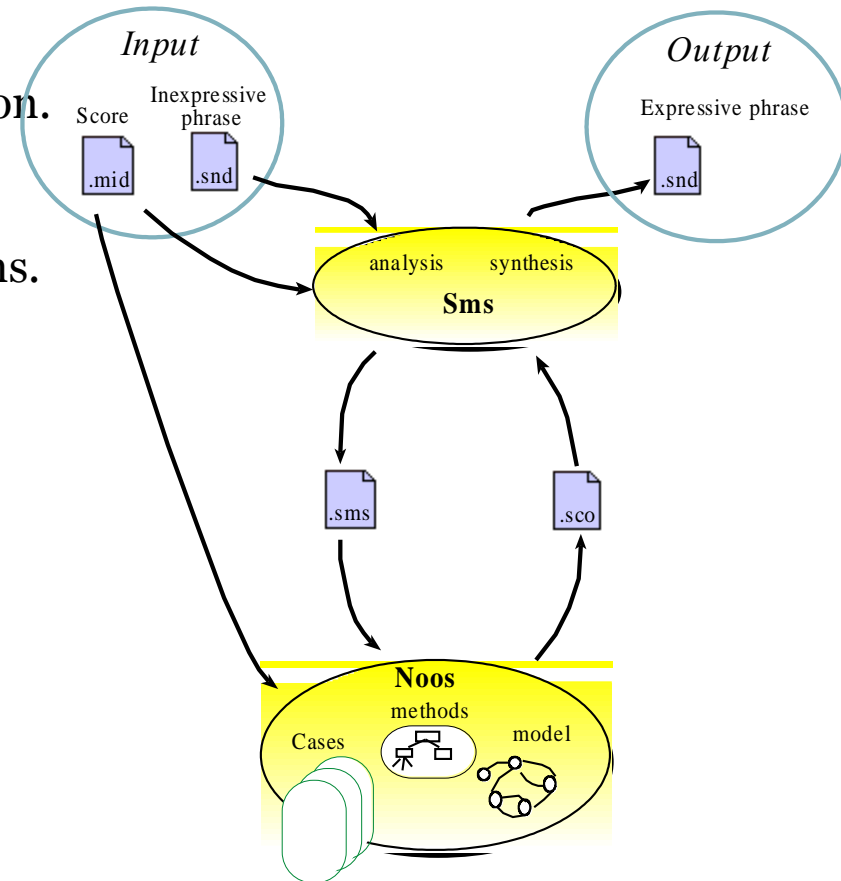
Multiresolution

- Ellis, D.; Vercoe, B. (ICMC, 1990)
 - constant-Q bandpass filters before sinusoidal modeling.
- Levine, S. et al. (WASPAA, 1997)
 - Bounded-Q transform.
- Polotti, P.; Evangelista G. (DAFX, 2001)
 - Harmonic-Band Wavelets.
- Dressler, K. (DAFX, 2006)
 - MR-FFT



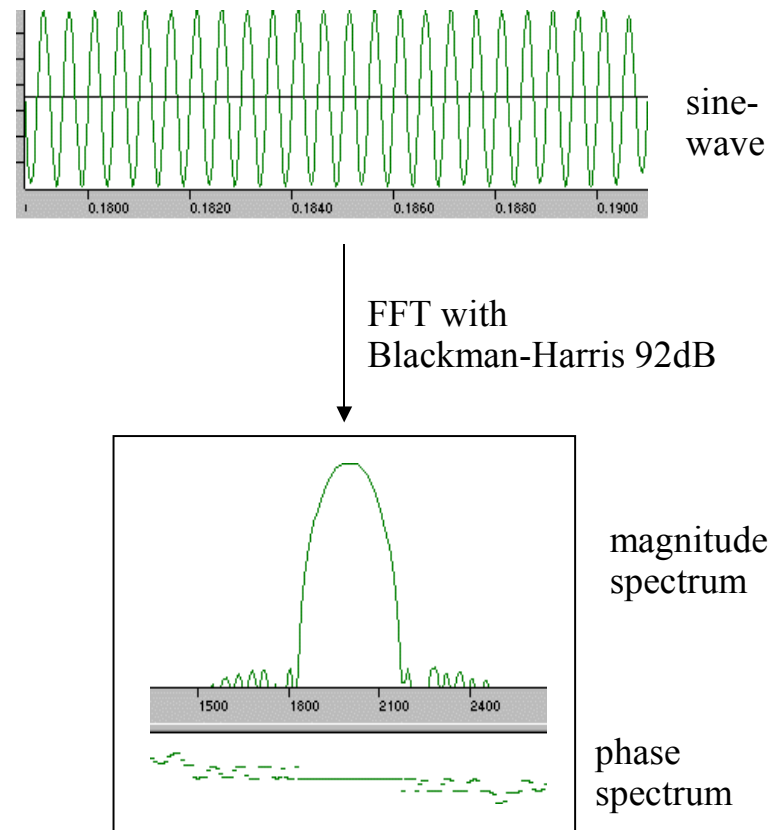
Feature-based Analysis/Synthesis

- Arcos, J. et al. (ICMC, 1997)
 - Expression analysis and transformation.
- Serra, X.; Bonada, J. (DAFX, 1998)
 - High-level features for transformations.
- Rossignol S. et al. (DAFX, 1999)
 - Vibrato processing.
- Gómez, E. et al. (DAFX, 2003)
 - Melodic transformations.
- Gouyon, F. et al. ((DAFX, 2003)
 - Rhythm transformations.



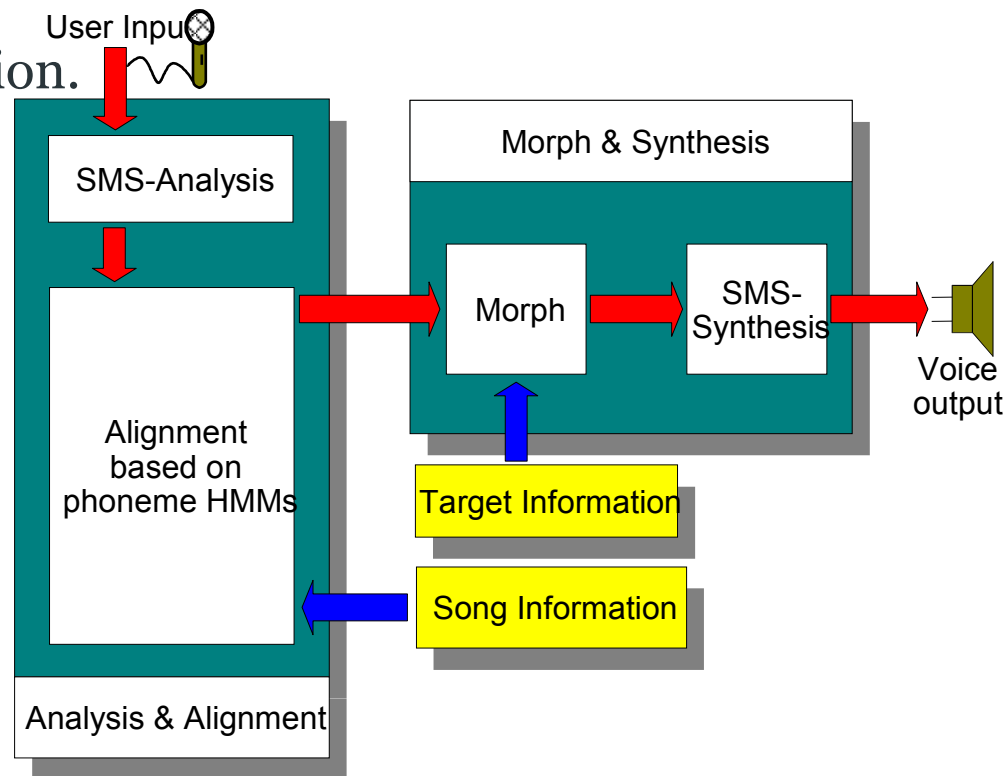
Synthesis of Sinusoids/Noise

- Rodet, X.; Depalle, Ph. (AES, 1992)
 - IFFT synthesis for sinusoids.
- Goodwin, M.; Rodet, X. (ICMC, 1994)
 - IFFT synthesis for nonstationary sines.
- Fitz, K.; Haken, L. (ICMC 1995)
 - Bandwidth Enhanced Sinusoidal Modeling.
- Goodwin, M. (thesis, 1996)
 - Noise synthesis using Equivalent Rectangular Bands.
- Hanna, P. (thesis, 2003)
 - Colored Noise by Sum of Sinusoids



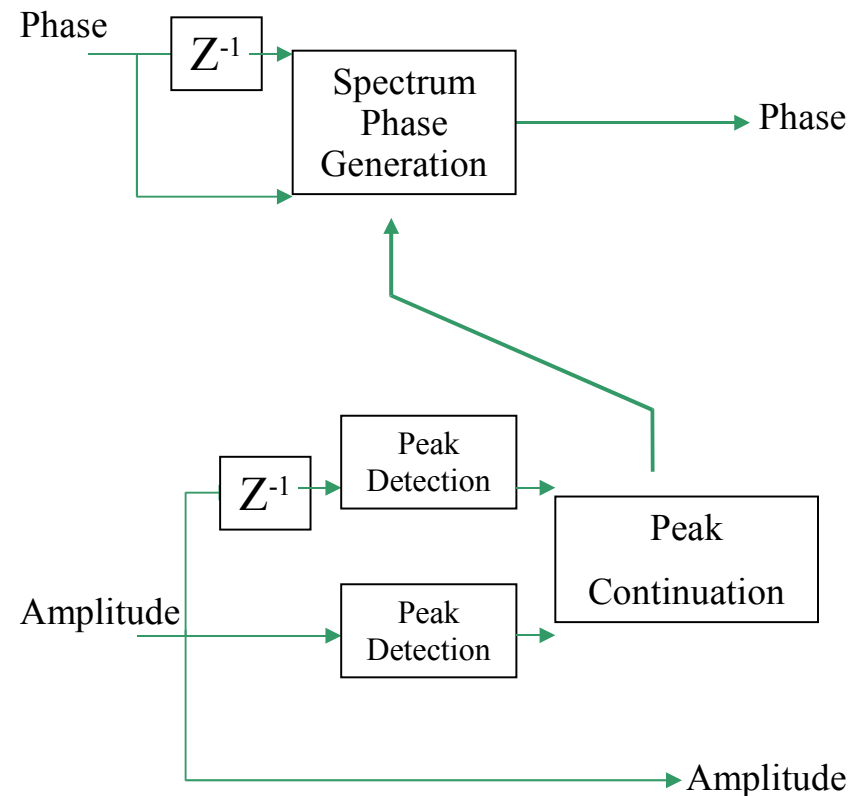
Morphing

- Serra, X. (ICMC, 1994)
 - Featured-based interpolation.
- Tellman, E. et al. (ICMC, 1994)
 - Sinusoidal interpolation.
- Cano, P. et al. (ICMC, 2000)
 - Voice morphing.



Time Scaling

- Verma, T.; Meng, H. (DAFX, 1998)
 - Using Sines + Transients + Noise Signal Model.
- Laroche, J.; Dolson, M. (IEEE, 1999)
 - Spectral Peaks.
- Bonada, J. (ICMC, 2000)
 - Spectral Peak Processing.



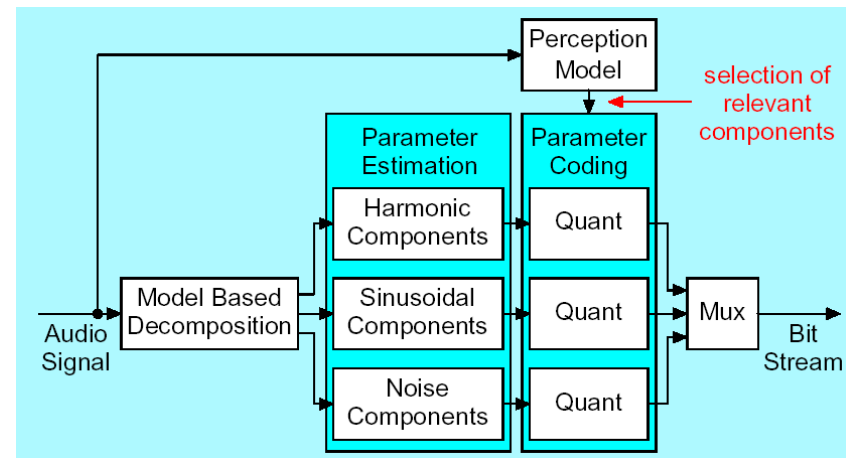
Coding/Transmission

- Hamdy, K. N. et al. (ICASSP, 1996)
 - Harmonics plus wavelets.

- Levine, S. N. (thesis, 1998)
 - Transformations in transform domain.

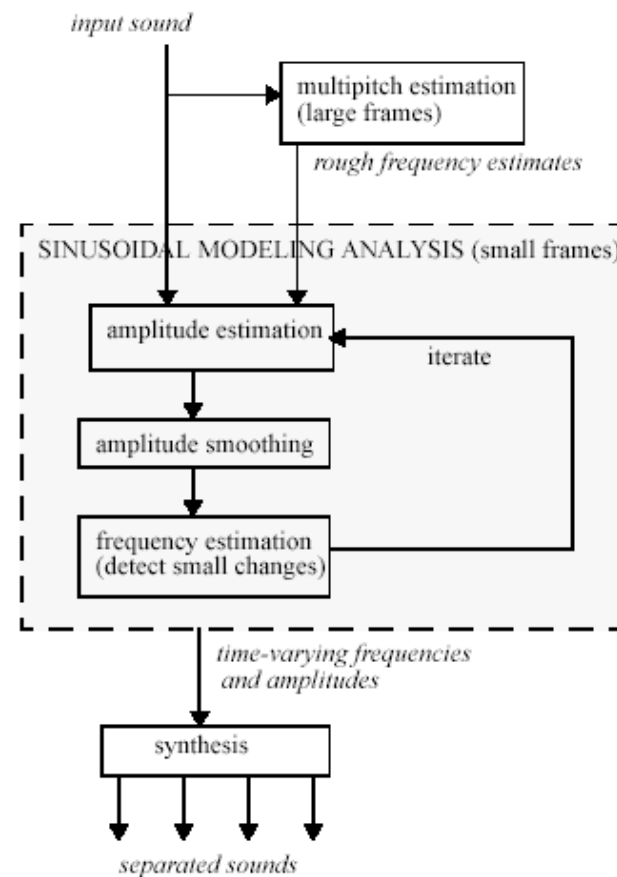
- Purnhagen, H.; Meine N. (ISCAS, 2000)
 - MPEG4: Harmonic and Individual Lines plus Noise (HILN).

- Amatriain, X. et al. (AES, 2002)
 - Content-based transmission.



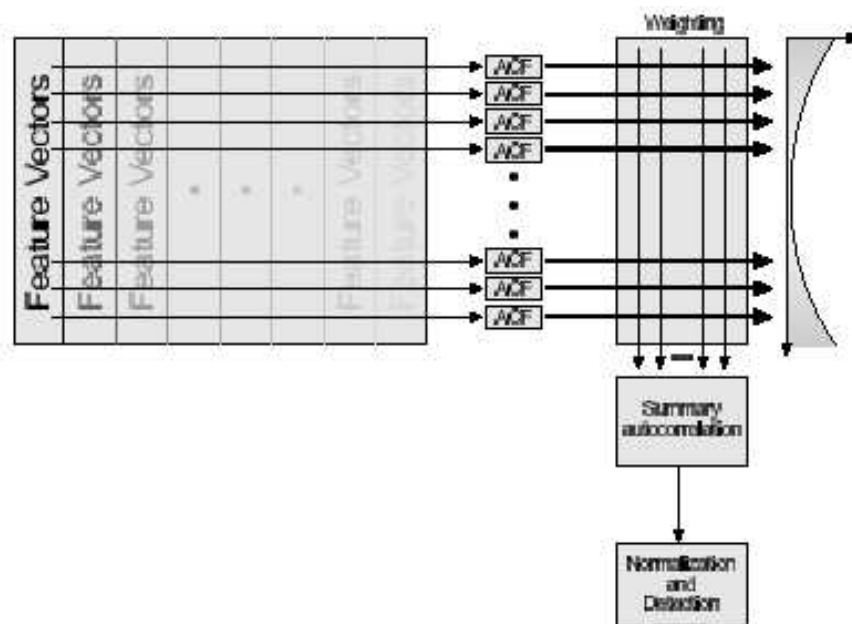
Source Separation / Transcription

- Maher, R. (thesis, 1989)
 - Partial collision and Two Way Mismatch algorithm for Fo detection.
- Virtanen, T. et al. (ICASSP, 2000)
 - Multipitch analysis and iterative parameter estimation.



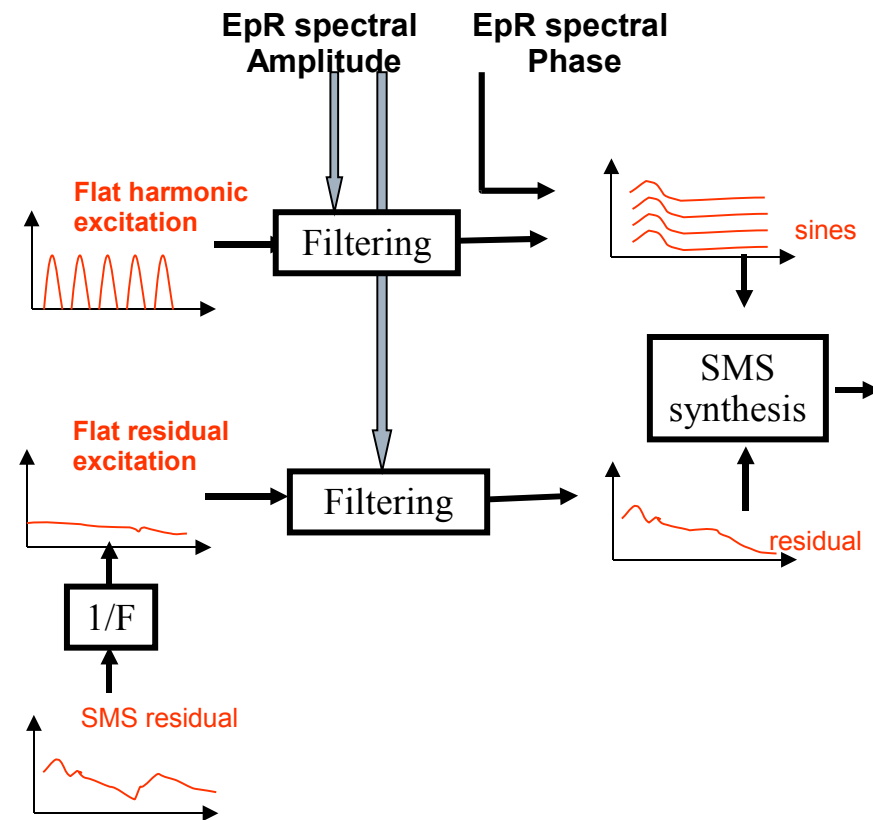
Music/Audio Content Analysis

- Herrera, P. et al. (CBMI, 1998)
 - Descriptors for MPEG-7.
- Heittola, T.; Klapuri, A. (ISMIR, 2002)
 - Identification of drums.
- Gómez, E. et al. (JNMR, 2003)
 - Melodic description.
- Wang, A. (Shazam, 2003)
 - Audio identification.
- Misra, A. et al. (DAFX, 2006)
 - Audio Scene processing (TAPESTREA).



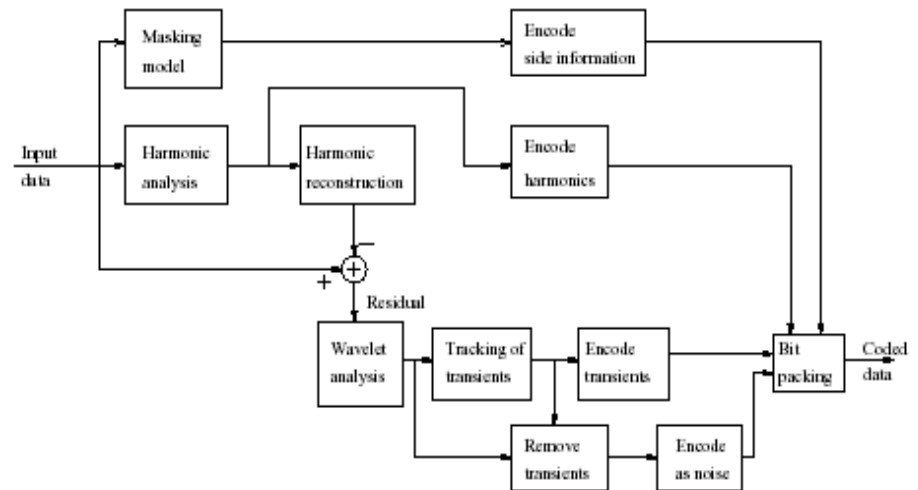
Instrument/Voice Models

- Macon, M. W. et al. (AES, 1997)
 - Singing voice (Lyricos)
- Haas, J. (MOSART, 2001)
 - Sax synthesizer
- Bonada et. al. (ICMC, 2001)
 - Singing voice (Vocaloid)
- Laroche, J. (DAFX, 2003)
 - Spectral processing plus TD-PSOLA



Expanded Models

- Laroche, J. et al. (ICASSP, 1993)
- Fitz, K. et al. (ICMC, 1995)
- Ding Y. et al. (ICMC, 1997)
- Goodwin, M. (thesis, 1997)
- Verma, et al. (ICMC, 1997)
- Peeters, G. et al. (ICMC, 1999)



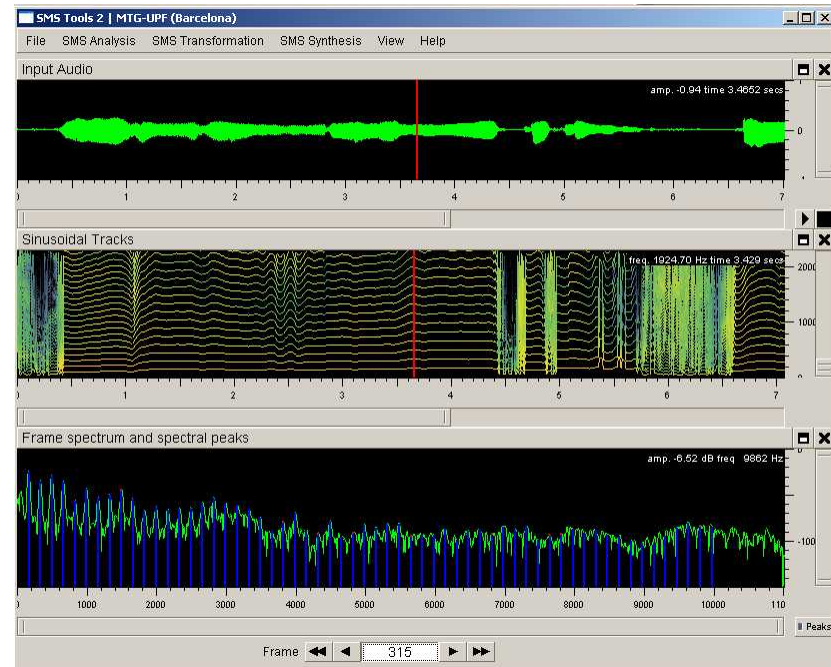
Software Environments

- Serra, X. (LMJ, 1991)
 - SANSY: Lisp environment based on SPIRE

- Fitz, K. et al. (ICMC, 1995)
 - Lemur

- Loscos, A. et al. (DAFX, 1998)
 - SMSPerformer

- Amatriain, X. et al. (ACM, 2002)
 - CLAM



Conclusions

- From speech to audio to music.
- From analysis/synthesis to content processing.
- Beyond signal processing techniques.
- Techniques are ready for many practical applications.
- Need to combine bottom-up with top-down approaches.