Frequency Selective Pitch Transposition of Audio Signals

Sascha Disch, Bernd Edler

Bernd.Edler@audiolabs-erlangen.de

International Audio Laboratories Erlangen

A joint institution of Friedrich-Alexander-Universität Erlangen Nürnberg
and the Fraunhofer Institute for Integrated Circuits (IIS)

May 2011

FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG

Fraunhofer

IIS

Motivation

- Modern music production often relies on mixing of pre-recorded "samples"

- Need to adapt "samples" ex post to a different musical context

- Key mode conversion, e.g. major-to-minor or vice versa

- Modulation vocoder (MODVOC) has already been shown to be suitable for this task

- Special enhancements are proposed to address requirements for this application

FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG

AudioLabs
Disch, Edler, Pitch Transposition - 2

Fraunhofer
IIS

Modulation vocoder (MODVOC)

- **Single-pass** analysis, block-wise processing

- Signal adaptive BP filter aligned with spectral centers of gravity

- Additive synthesis with bonding: pairing of components between subsequent time blocks
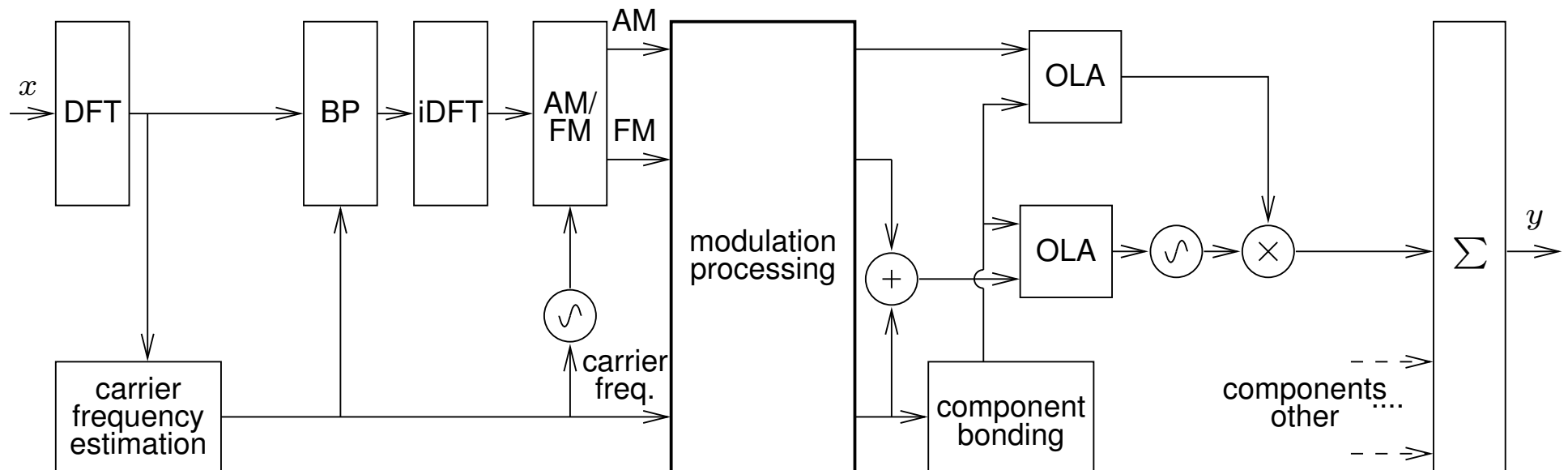
Enhancements: Envelope Shaping (ES)

- Global coherence is lost due to modifications on each individual MODVOC component

- Temporal smearing of transients (dispersion)

- Envelope shaping (ES) uses duality:
  autocorrelation of spectrum $\leftrightarrow$ square of Hilbert envelope
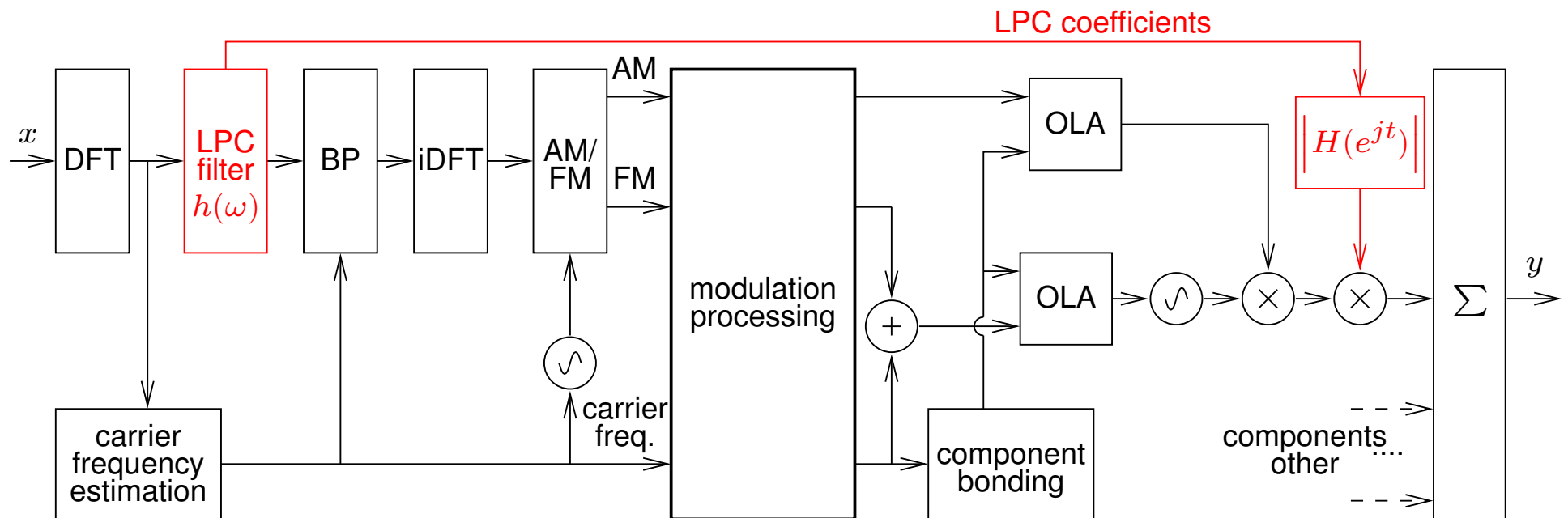
Enhancements: Harmonic Locking (HL)

- An instrument sound consists of fundamental and overtones

- Overtone series are quasi-harmonic on a linear frequency scale (integer multiples of f_0)

- Musical intervals are based on logarithmic scale

- Meaning of MODVOC component in terms of musicology dependant on predominant origin

- This must be considered in selective transposition for preservation of timbre $\rightarrow$ ambiguity

- What is the musical interpretation of each component?
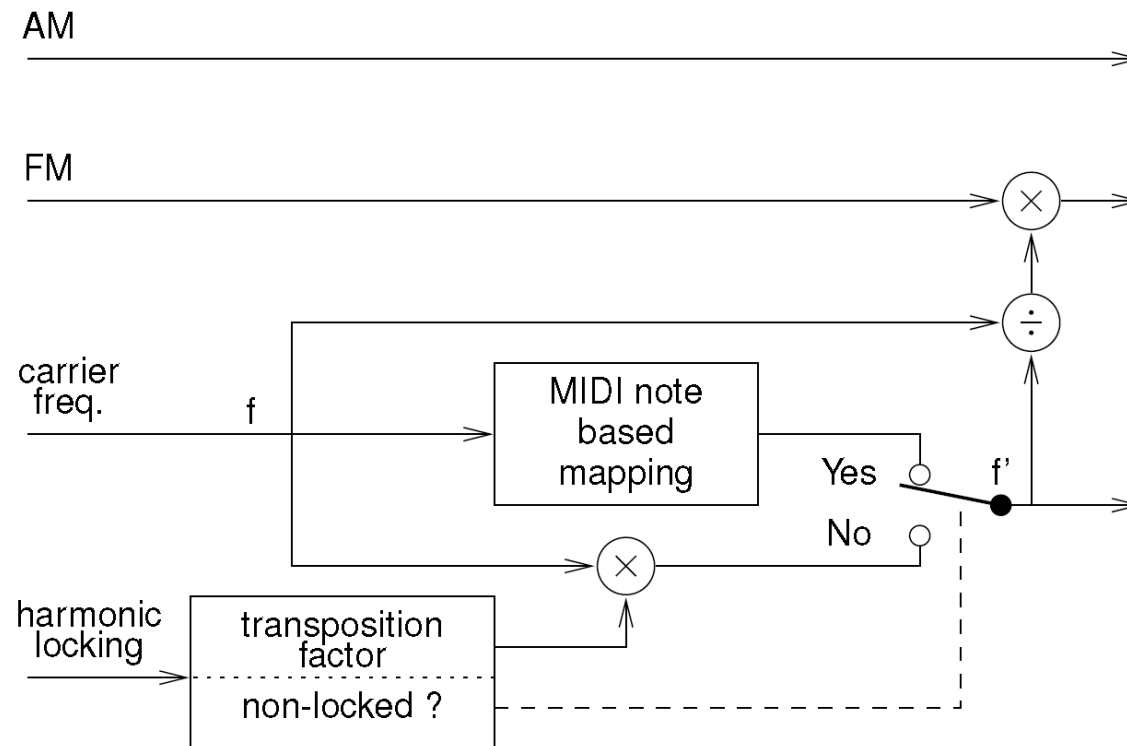 - Fundamental or
  - Overtone/Harmonic?

Enhancements: Harmonic Locking (HL)

- Intervals of harmonics with respect to their fundamental

- Ambiguity example: scale degree: tonic, 5th harmonic → a major third or the fifth harmonic of the fundamental?

Harmonic number			Interval name
1	2	4	perfect unison (P1)
			minor second (m2)
		9	major second (M2)
			minor third (m3)
	5		major third (M3)
			perfect fourth (P4)
			tritone
	3	6	perfect fifth (P5)
			minor sixth (m6)
			major sixth (M6)
		7	minor seventh (m7)
			major seventh (M7)

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
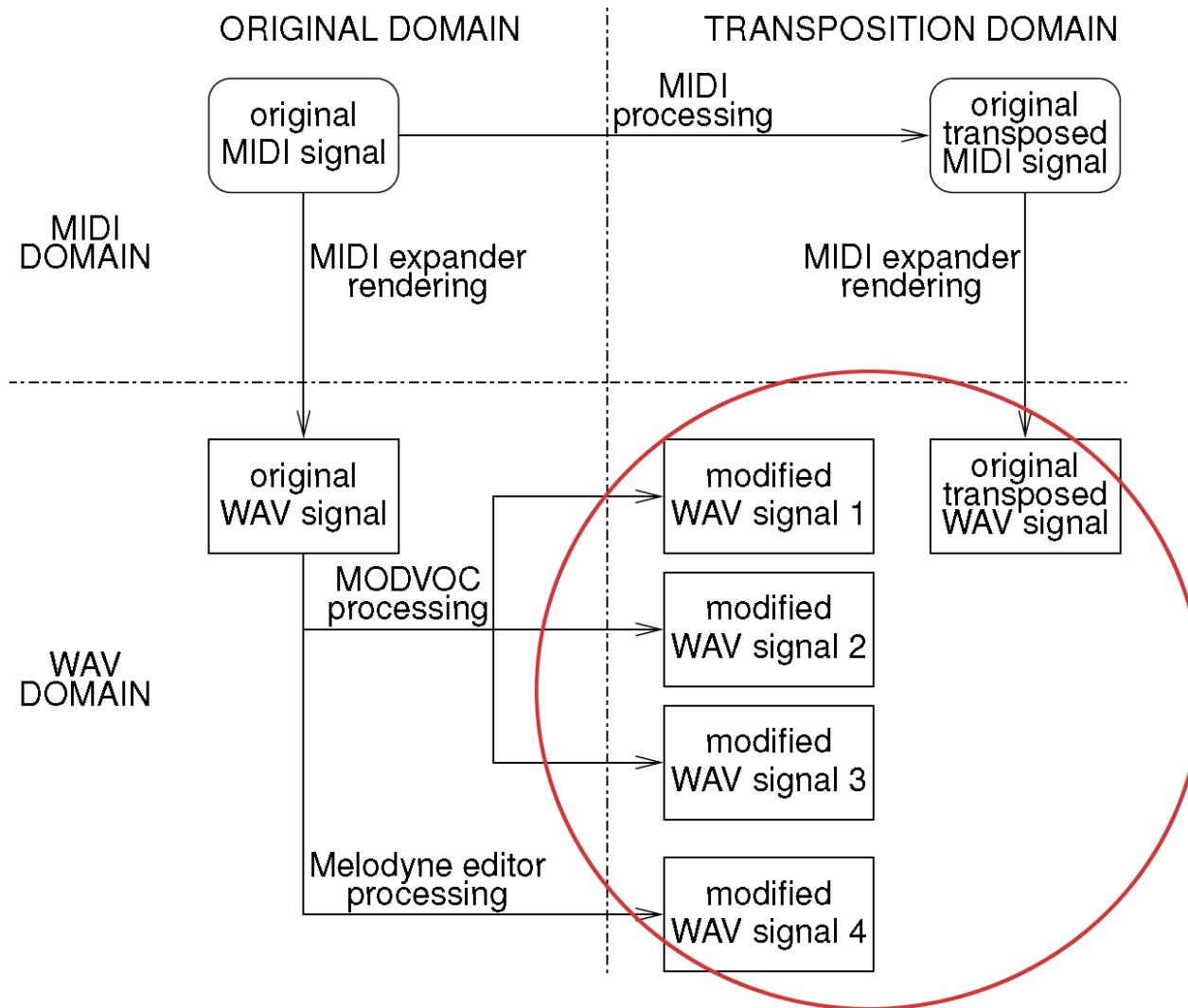
Fraunhofer IIS

Enhancements: Harmonic Locking (HL)

- **Harmonic locking** (HL)

- Estimated fundamentals are mapped directly

- Frequency shift of overtone is locked to shift of its estimated fundamental

AudioLabs

Disch, Edler, Pitch Transposition - 7

FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG

Fraunhofer

IIS

Listening Tests: Methodology

- Generation of synthetic test conditions

Listening Tests: Commercial System

- Additional comparison of MODVOC to commercial system

- **Direct note access** (DNA) in "Melodyne editor" by Celemony

- Available since autumn 2009

- Selective editing of polyphonic music

- Principle

  ○ **Multi-pass analysis**
  ○ Automatic decomposition into "notes"
  ○ Heuristic classification rules

- Can be operated to perform key mode conversion

FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG

Fraunhofer

IIS

Listening Tests: Conditions, Items

name	description	instruments	key mode
A	Violin Concerto J. S. Bach, BWV1041	Orchestra	Amin
B	Eine kleine Nachtmusik W. A. Mozart, KV525 Mv1	String Quartet	Gmaj
C	Berceuse G. Fauré, Op56	Flute and Guitar	Emaj
D	Nocturno F. Strauss, Op7	Horn and Piano	Dbmaj
E	Waltz F. Carulli, Op241 No1	Guitar	Cmaj
F	Ein Musikalischer Spass W.A. Mozart, KV522 Mv1	Horns, Violin, Viola, Cello	Fmaj
G	Ode an die Freude L. v. Beethoven	Piano	Gmaj
H	Piano Trio L. v. Beethoven, Op11 Mv3	Clarinet, Cello, and Piano	Bbmaj

Listening Test 1: MUSHRA Results

Average and 95% Confidence Intervals

1. hidden_reference   2. 3k5Hz_reference   3. MODVOC   4. MODVOC_hl
5. MODVOC_hl_es   6. DNA
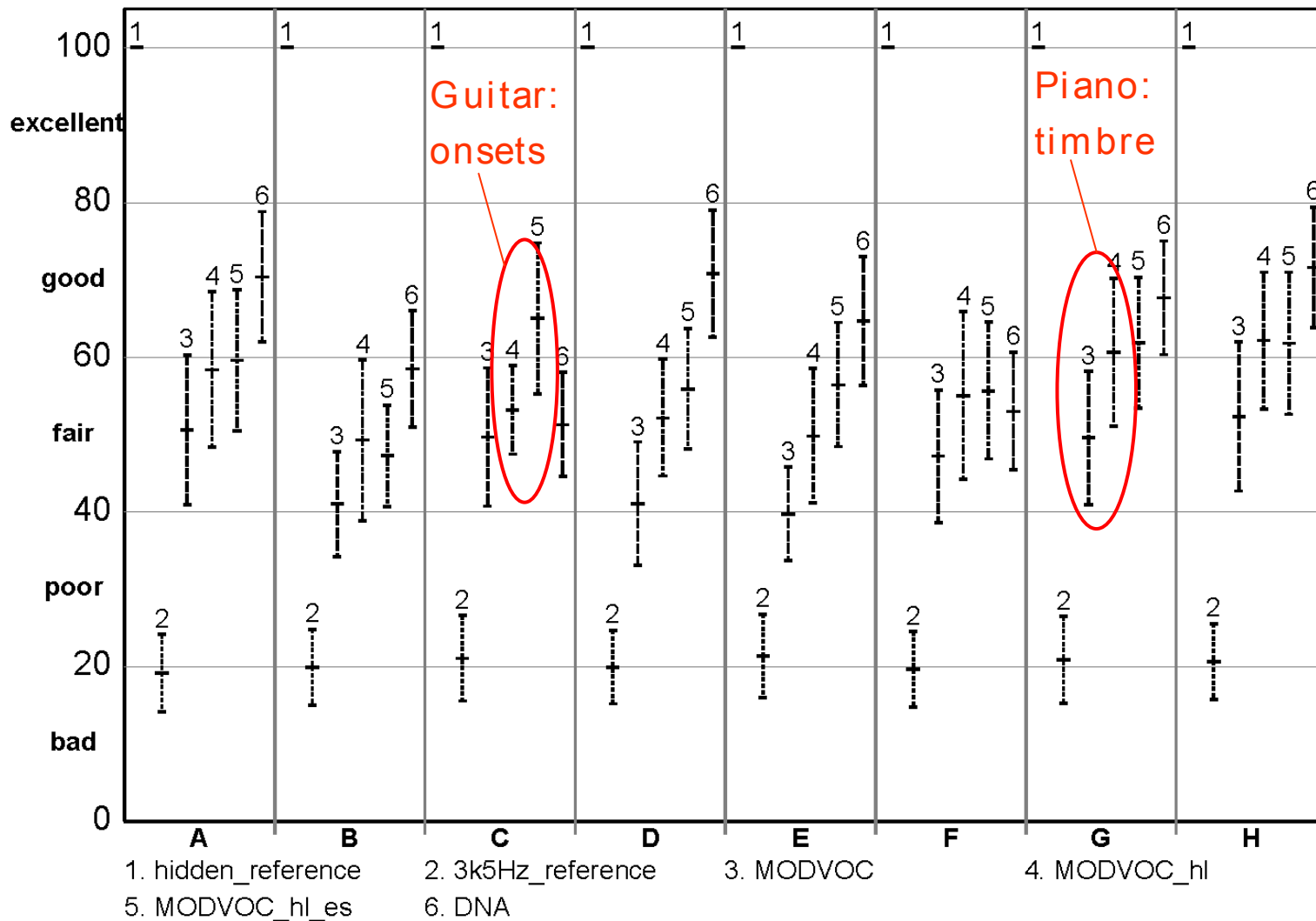
Listening Test 1: Results

- HL indeed improves timbre of all items

- ES indeed improves transient parts of items (e.g. onsets of guitar, piano)

- DNA is rated better than MODVOC in 5 out of 7 items

- Overall rating covers many aspects
  - Unnatural sounding artifacts (clicks, pre- and post-echoes, etc.)
  - Melody and chord transposition
  - Timbre preservation

- Listeners informally reported a trend for transposition errors in DNA and timbre problems for MODVOC

→ formal preference tests on main quality aspects

FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG

Fraunhofer

IIS

Listening Test 2: Methodology, Items

- Preference tests on these aspects, 12 expert listeners having both technical and musical background

- Conditions: MODVOC with ES and HL, DNA

- Unknown melody is hardly remembered by subjects $\rightarrow$ now key mode changed in MIDI; items converted back to original key mode in waveform domain

name	description	instruments	key mode
1	Violin Concerto J. S. Bach, BWV1041	Orchestra	Amin
2	Berceuse G. Fauré, Op56	Flute and Guitar	Emaj
3	Ode an die Freude L. v. Beethoven	Piano	Gmaj
4	Concerto for Violin and Orchestra L. v. Beethoven, Op61	Violin and Orchestra	Dmaj

Listening Test 2: Preference Results

- MODVOC is indeed preferred for melody transposition

- DNA is preferred for the most part wrt. timbre preservation

Preference for melody transposition in synthetic signals

Preference for timbre preservation in synthetic signals
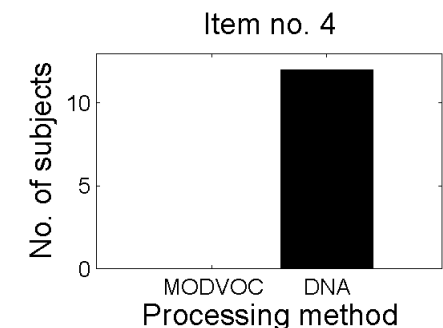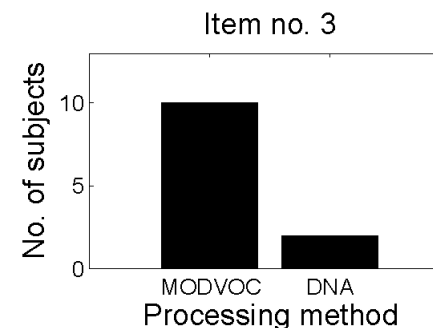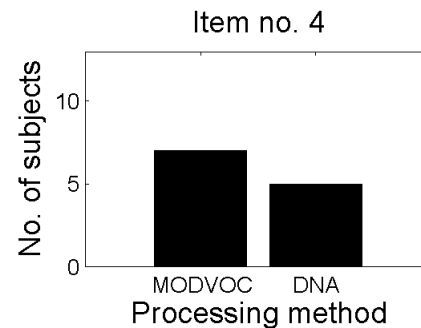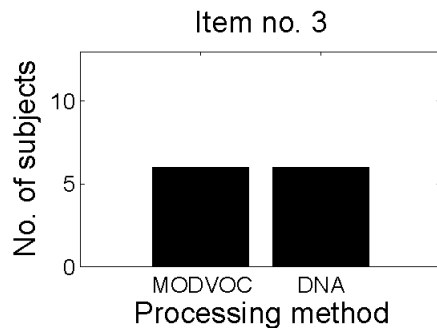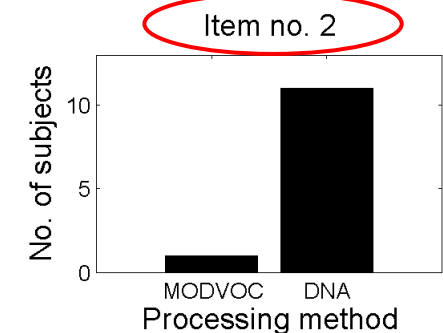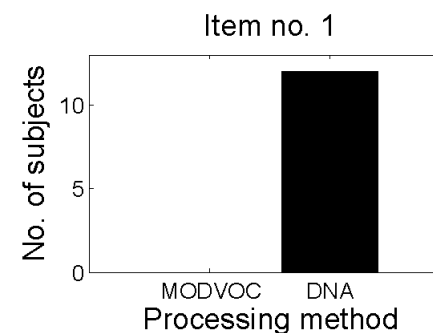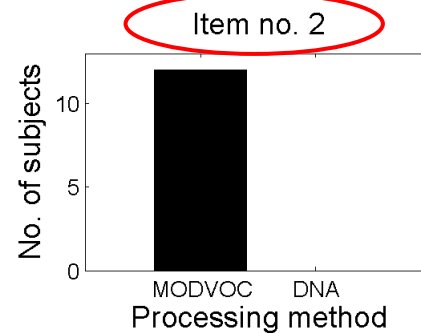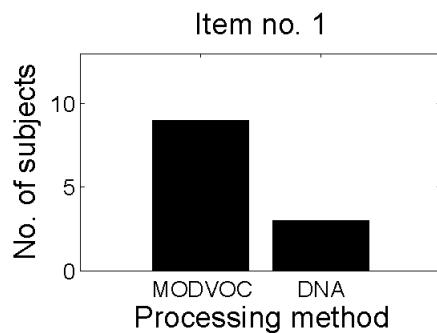
Original   Orig. tr.   MODVOC   DNA   All (short)

Listening Test 2: Solo Piano

- Both methods perform comparably in melody transposition for synthetic solo piano (item no.3)

- MODVOC is preferred in terms of melody and timbre
  - → HL performs well for this test signal

- DNA has instabilities and segmentation induced artifacts

Original   Original transposed

MODVOC

DNA   All

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

Fraunhofer

IIS

Summary

- Enhanced modulation vocoder (MODVOC) for selective transposition of pitch

- MODVOC capable of real-time processing (single-pass)

- Reproduction of transients improved by LPC based envelope shaping (ES)

- Timbre improved by harmonic locking (HL)

- MODVOC is preferred over commercial system (DNA) in transposition of melody and chords

- DNA is preferred over MODVOC in timbre preservation

- Absolute scores for both systems range from "fair" to "good"