



A Hybrid Time and Frequency Domain Audio Pitch Shifting Algorithm



Nicolas Juillerat, *Pervasive & AI Research Group*, University of Fribourg, Switzerland
 Stefan Müller Arisona, *Media Arts & Technology*, University of Santa Barbara, California
 Simon Schubiger-Banz, *Computer Systems Institute*, ETH Zurich, Switzerland

One **Abstract Algorithm** working in **three steps**.

Three *concrete implementations* (two existing, one new):

1. The **Improved Phase Vocoder**, implementing all three steps entirely in the *frequency domain*
2. The **"Rollers"** algorithm, implementing all three steps entirely in the *time domain*
3. A **new hybrid implementation**, working in both the time and frequency domains
 Takes the best from the two existing approaches → higher quality

	Step 1	Step 2	Step 3
Abstract Algorithm	Choose a subband decomposition of the signal (can be time varying). Locate the dominant frequency f_i of each band i	Extract each band	Frequency shift each band, so that its dominant frequency f_i is scaled by the desired ratio k : $shift_amount = k f_i - f_i$ Shift amount is time varying
1. Improved PV (Laroche & Dolson, AES Journal (47)11)	 STFT + Spectral peak picking + define regions around the peaks. Interpolation for dominant frequencies. High quality ✓	Nothing to do: we stay in the frequency domain ✓	 Done in the frequency domain → Artifacts (due to overlapped STFT frames instead of interpolation) ✗
2. Rollers (ICALIP 2008)	 Fixed subband decomposition on a logarithmic scale (no signal analysis, low quality) ✗	One IIR band-pass filter per band. Butterworth, 4 th order → Low quality ✗	 Done in time domain. Sample by sample interpolation. → High quality ✓
3. New Hybrid Approach NEW!	 Logarithmic scale and peak picking ✓	<u>Many choices:</u> •Inverse STFTs • Multiresolution FIR filters • Adaptive tiling* •Etc Any filter bank with sufficient frequency resolution! No need to deal with phase at this step! ✓	Done in the time domain. (Same as the Rollers implementation above) Phase handled in a seamless way! ✓
The new hybrid approach produces higher quality. It was evaluated by: •A simple audio metric •A listening test			Music excerpts on-line: http://www.pitchtech.ch/Confs/AES125/index.html Other results (& VST plugins): http://www.pitchtech.ch

*not only theory, it really works