Harmonic Synchronous OverLap-and-Add Method

Department of Computer Science and Engineering
National Sun Yat-sen University

Spring 2012

Outline

- Engine Noise Model
- Input Signal
- Sample Dataset Selection
- Identifying Search Intervals
- Finding Pitch Marks
- Multiple Harmonic Phase Match
- Optimal Pitch Marks
- Synthesis
- Tentative Research Topics

Frequencies Related to a Running Engine

- Engine frequency

$$F_{en} = \frac{\text{rpm}}{60}$$

Frequencies Related to a Running Engine

- Engine frequency

$$F_{en} = \frac{\text{rpm}}{60}$$

- Fundamental frequency (4-stroke engine)

$$F_c = \frac{\text{rpm}}{120} = F_0$$

Frequencies Related to a Running Engine

- Engine frequency

$$F_{en} = \frac{\text{rpm}}{60}$$

- Fundamental frequency (4-stroke engine)

$$F_c = \frac{\text{rpm}}{120} = F_0$$

- Explosion (Ignition) frequency

$$F_{ex} = N_c F_c$$

Frequencies Related to a Running Engine

- Engine frequency

$$F_{en} = \frac{\text{rpm}}{60}$$

- Fundamental frequency (4-stroke engine)

$$F_c = \frac{\text{rpm}}{120} = F_0$$

- Explosion (Ignition) frequency

$$F_{ex} = N_c F_c$$

- Harmonics of fundamental frequency F_0

$$F_k = kF_0$$

Engine Noise Model

Engine noise model at constant speed (corresponding to fundamental frequency F_0)

$$x_{F_0}[n] = \sum_k A_{F_0,k} \sin\left(2\pi \frac{kF_0}{F_s} n + \phi_{F_0,k}\right) + N_{F_0}[n]$$

$A_{F_0,k}$: the amplitude of the k^{th} harmonic;
F_s: the sampling frequency;
$\phi_{F_0,k}$: the phase of the k^{th} harmonic;
$N_{F_0}[n]$: the stochastic engine noise.

Input Signal

HSOLA requires an input signal from which sound samples can be extracted.

Input Signal

HSOLA requires an input signal from which sound samples can be extracted.

- Record engine noise with no gear engaged over the available rpm range.

Input Signal

HSOLA requires an input signal from which sound samples can be extracted.

- Record engine noise with no gear engaged over the available rpm range.
- The variation of the engine speed must be slow enough such that the signal is short-time stationary at least over the duration of one period of the fundamental frequency.

Sample Dataset Selection

From the input signal, we want to iteratively collect the sound samples at equally spaced fundamental frequencies.

Sample Dataset Selection

From the input signal, we want to iteratively collect the sound samples at equally spaced fundamental frequencies.

Let the corresponding fundamental frequency range be $\left[F_0^{min}, F_0^{max}\right]$. The equally spaced fundamental frequencies for which we want to collect sound samples are

$$F_0^i = F_0^{min} + i \ \frac{F_0^{max} - F_0^{min}}{M-1}, \quad i = 0, \ldots M-1.$$

Sample Dataset Selection

From the input signal, we want to iteratively collect the sound samples at equally spaced fundamental frequencies.

Let the corresponding fundamental frequency range be $\left[F_0^{min}, F_0^{max}\right]$. The equally spaced fundamental frequencies for which we want to collect sound samples are

$$F_0^i = F_0^{min} + i \; \frac{F_0^{max} - F_0^{min}}{M-1}, \quad i = 0, \ldots M-1.$$

For common vehicles, the number of frequency samples M can be generally set to

$$M \approx 1000.$$

Identifying Search Intervals

Given the input signal, we need to decide the extraction instant, called the "pitch mark", for the sound sample of each fundamental frequency.

Identifying Search Intervals

Given the input signal, we need to decide the extraction instant, called the "pitch mark", for the sound sample of each fundamental frequency.

For each fundamental frequency, say F_0^i, we decide a search interval in the input signal, say $\mathcal{I}_i$, within which we find the pitch mark.

Identifying Search Intervals

Given the input signal, we need to decide the extraction instant, called the "pitch mark", for the sound sample of each fundamental frequency.

For each fundamental frequency, say F_0^i, we decide a search interval in the input signal, say $\mathcal{I}_i$, within which we find the pitch mark.

Specifically, $\mathcal{I}_i$ is the time segment during which the most energetic component is between

$$\left[k_{A_{max}} F_0^i - \Delta F, k_{A_{max}} F_0^i + \Delta F\right]$$

Finding Pitch Marks

A set of potential pitch marks for frequency F_0^i is defined as the discrete-time indexes in $\mathcal{I}_i$

$$\mathcal{P}_i = \{n_{i,l}\}_{l=1,2,\ldots},$$

where the initial phase of the harmonic component $k_{A_{max}} F_0^i$ is null (0) for the portion of input signal starting at instant $n_{i,l}$.

Finding Pitch Marks

A set of potential pitch marks for frequency F_0^i is defined as the discrete-time indexes in $\mathcal{I}_i$

$$\mathcal{P}_i = \{n_{i,l}\}_{l=1,2,\ldots},$$

where the initial phase of the harmonic component $k_{A_{max}} F_0^i$ is null (0) for the portion of input signal starting at instant $n_{i,l}$.

$\mathcal{P}_i$ corresponds to the set of local maxima of the cross correlation function between a sinusoid of frequency $k_{A_{max}} F_0^i$ and a portion of the input signal over one period of F_0^i.

Finding Pitch Marks

A set of potential pitch marks for frequency F_0^i is defined as the discrete-time indexes in $\mathcal{I}_i$

$$\mathcal{P}_i = \{n_{i,l}\}_{l=1,2,\ldots},$$

where the initial phase of the harmonic component $k_{A_{max}} F_0^i$ is null (0) for the portion of input signal starting at instant $n_{i,l}$.

$\mathcal{P}_i$ corresponds to the set of local maxima of the cross correlation function between a sinusoid of frequency $k_{A_{max}} F_0^i$ and a portion of the input signal over one period of F_0^i.

The baseline approach, called the *single harmonic phase match*, n_{i,l_i^*} corresponds to the portion in $\mathcal{I}_i$ with the global maximum of the cross correlation function in $\mathcal{P}_i$.

Multiple Harmonic Phase Match

In the second approach, called the *multiple harmonic phase match*, the phases of higher-order harmonics are taken into account.

Multiple Harmonic Phase Match

In the second approach, called the *multiple harmonic phase match*, the phases of higher-order harmonics are taken into account.

The phase for the k^{th} harmonic at potential pitch mark with index $n_{i,l}$ is denoted by $\phi_{F_0^i,k,l}$.

Multiple Harmonic Phase Match

In the second approach, called the *multiple harmonic phase match*, the phases of higher-order harmonics are taken into account.

The phase for the k^{th} harmonic at potential pitch mark with index $n_{i,l}$ is denoted by $\phi_{F_0^i,k,l}$. It is estimated via the cross correlation function (CCF) between a sinusoid of frequency kF_0^i and the input signal at $n_{i,l}$.

Multiple Harmonic Phase Match

In the second approach, called the *multiple harmonic phase match*, the phases of higher-order harmonics are taken into account.

The phase for the k^{th} harmonic at potential pitch mark with index $n_{i,l}$ is denoted by $\phi_{F_0^i,k,l}$. It is estimated via the cross correlation function (CCF) between a sinusoid of frequency kF_0^i and the input signal at $n_{i,l}$.

Specifically,

$$\phi_{F_0^i,k,l} = 2\pi \; \Delta N_{l,k} \; \frac{kF_0^i}{F_s},$$

where $\Delta N_{l,k}$ is the difference in samples between potential pitch marks and the local optimum of the above CCF, and F_s is the sampling frequency.

Optimal Pitch Marks

The optimal pitch mark, n_{i,l_i^*}, is chosen as follows.

Optimal Pitch Marks

The optimal pitch mark, n_{i,l_i^*}, is chosen as follows.

- We compute the average phase of each harmonic over N_s previously extracted samples. That is

$$\overline{\phi_{F_0^i,k}} = \frac{\sum_{j=1}^{N_s} \phi_{F_0^{i-j},k,l_{i-j}^*}}{N_s}, \quad (N_s \leftarrow 2\alpha).$$

Optimal Pitch Marks

The optimal pitch mark, n_{i,l_i^*}, is chosen as follows.

- We compute the average phase of each harmonic over N_s previously extracted samples. That is

$$\overline{\phi_{F_0^i,k}} = \frac{\sum_{j=1}^{N_s} \phi_{F_0^{i-j},k,l_{i-j}^*}}{N_s}, \quad (N_s \leftarrow 2\alpha).$$

- Define a phase-matching criterion which is minimized for l_i^*

$$l_i^* = \arg\min_{l \in \mathcal{P}_i} \left\{ C(l) = \sum_k \left| \overline{\phi_{F_0^i,k}} - \phi_{F_0^i,k,l} \right| A_{i,l,k} \right\},$$

  where $A_{i,l,k}$ is the magnitude in dB of the local maximum of the CCF used to calculate $\phi_{F_0^i,k,l}$.

Synthesis

With the decided pitch marks, $\{n_{i,l_i^*}\}_{i=0}^{M-1}$, any arbitrary evolution of engine speed can be synthesized by OLA (overlap-and-add) algorithm on properply selected samples.

Synthesis

With the decided pitch marks, $\{n_{i,l_i^*}\}_{i=0}^{M-1}$, any arbitrary evolution of engine speed can be synthesized by OLA (overlap-and-add) algorithm on properply selected samples.

Use the Tukey window for extraction. It contains one period of the fundamental frequency plus Δ, with

$$\Delta = 128 \text{ samples.}$$

Here Δ is the overlap between adjacent samples. (Figure 2)

Synthesis

With the decided pitch marks, $\{n_{i,l_i^*}\}_{i=0}^{M-1}$, any arbitrary evolution of engine speed can be synthesized by OLA (overlap-and-add) algorithm on properply selected samples.

Use the Tukey window for extraction. It contains one period of the fundamental frequency plus Δ, with

$$\Delta = 128 \text{ samples.}$$

Here Δ is the overlap between adjacent samples. (Figure 2)

In the case of slow-varying or constant engine speed, a random oscillation in the range $[-\alpha, \alpha]$ ($\alpha \leftarrow 5$) is added to the selected sample index.

Tentative Research Topics

- number of cylinders
- 2-stroke vs. 4-stroke
- number of pitch marks to be extracted from the input signal
- rate of frequency change during recording
- number of periods in a sample
- shape and overlap ratio