Panorama Effect (Stereo Panning)

Using a "multi-channel sound reproduction system" we can change the apparent position of a source just by "feeding the channels with the same signal" and "adjusting the relative amplitude of the channels". By doing so these signals arrives at slightly different levels and times and we get virtual sound image at a different place from the real source locations.

θ=angle of virtual source
θ_L=angle formed by each loudspeaker with the frontal direction
g_L & g_R =gain applied to left & right stereo channels

In a standard loudspeaker setup it is assumed that listener stands in central position and form an angle $2\theta_L$ with the two loudspeakers. Two gains g_L & g_R are applied in order to set apparent azimuth at the desired value θ.

A unit magnitude "two channel signal corresponding to the central apparent source position ($\theta = 0$) can be represented by the column vector.

$$U = \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$$

So gains to be applied to the two channels in order to steer the sound source to the desired azimuth are obtained by the matrix vector multiplication.

$$A_\theta U = \begin{bmatrix} g_L \\ g_R \end{bmatrix}$$ Where A_θ=Rotation matrix

Amplitude panning by means of a rotation matrix preserves the loudness of the virtual sound source while moving its apparent azimuth.

It θ_L=45 the rotation matrix takes the form shown below, so that θ= +/- θ_L only one of the two channels is non zero. It is easily verified that the rotation by matrix given below corresponds to applying the tangent law to the configuration with θ_L=45^o.

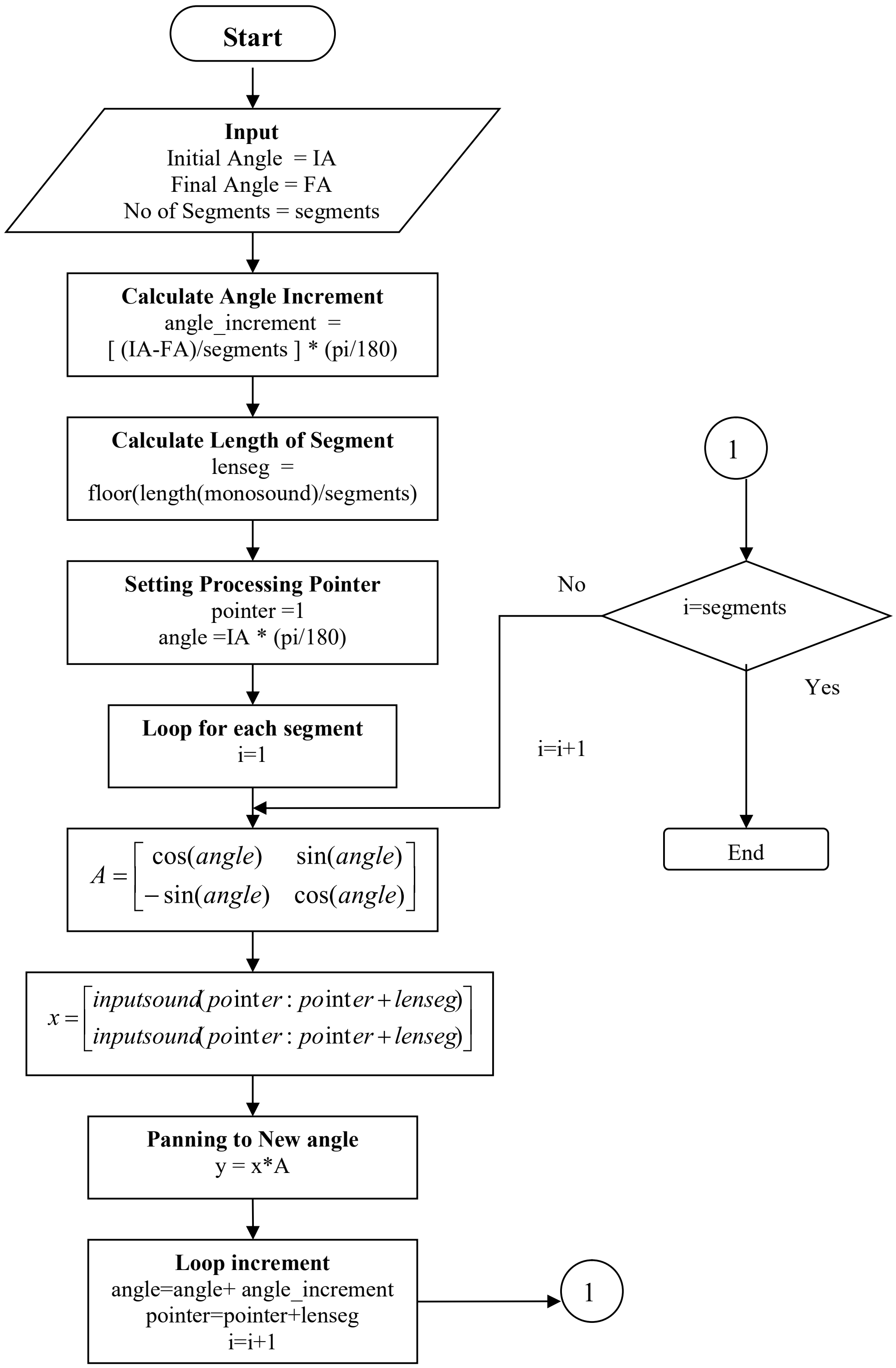
$$A_\theta = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

Tangent Law (results from vector formulation of amplitude panning):

$$\tan\theta = \frac{g_L - g_R}{g_L - g_R}\tan\theta_L$$

In practice the steering angle θ does not necessarily correspond to the perceived localization azimuth. The perceived location is influenced by the frequency content of the sound. Some theories of directional psychoacoustic have been developed in the past in order to drive the rotation matrix with appropriate coefficients .Accurate implementation use frequency dependent rotation matrix, at least discriminating between low (less than about 500Hz) and high (between 500Hz and 3500 Hz) frequency components.

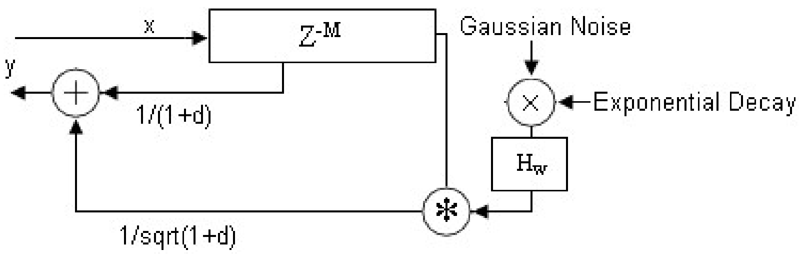
In implementation input sound is segmented in a number "segment" of blocks and each block is rotated of a matrix by vector multiplication. Flow chart on next page represents implements the amplitude panning between an initial and a final angle.



Flow Chart: Panning Effect algorithm

Distance Rendering

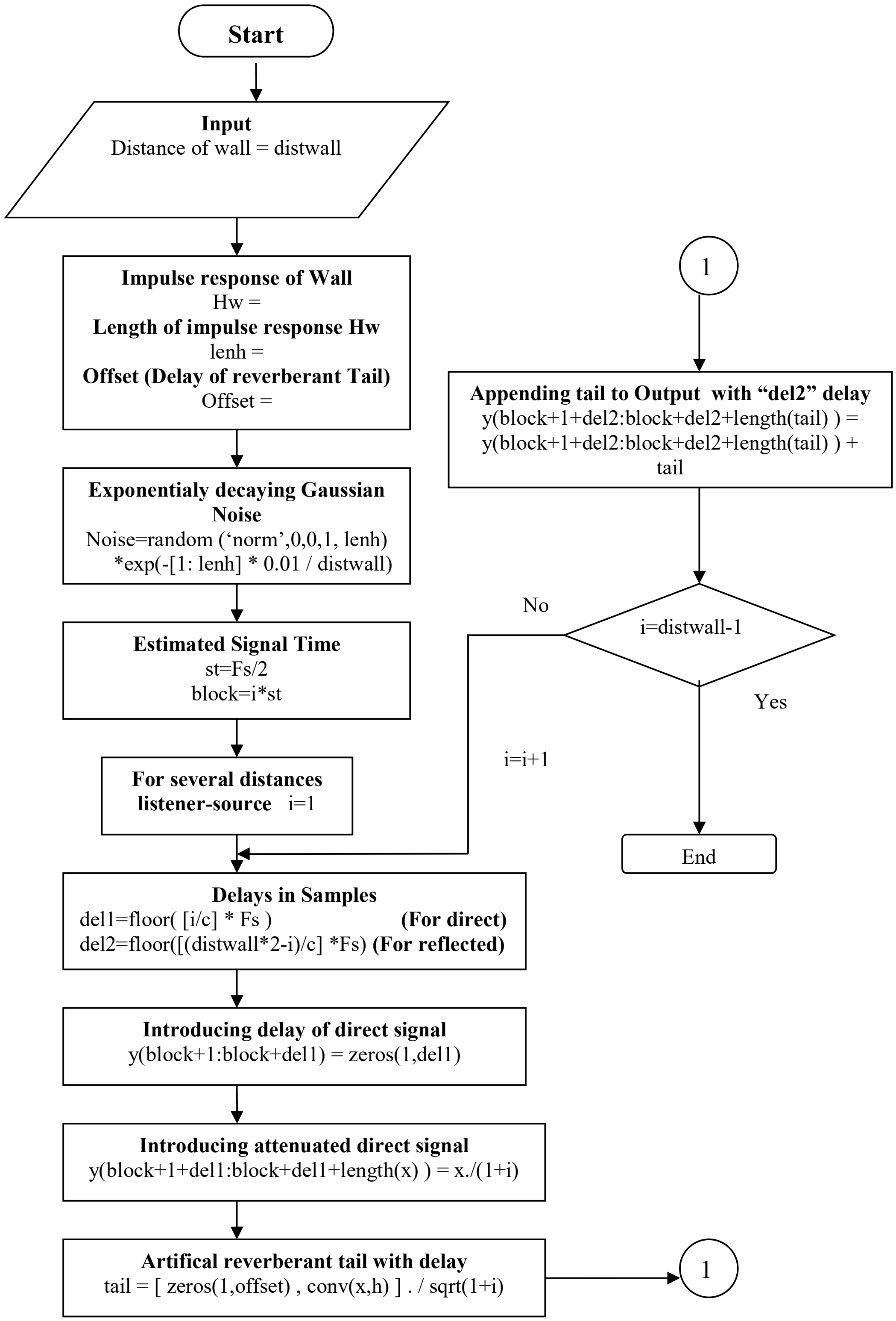
In digital audio effects the control of apparent distance can be effectively introduced even in monophonic audio systems. In fact the impression of distance of a sound source is largely controllable by insertion of artificial wall reflections or reverberant room response.



There are no reliable cues for distance in open space but in enclosure the ratio of reverberant to direct acoustic energy has proven to be robust distance cue.

It is assumed that a small space the amplitude of the reverberant signal changes little with distance and that in a large space it is roughly proportional to $1/\sqrt{Distance}$ [Cho71]. The direct sound attenuates as 1/distance if spherical waves are propagated.

- A single reflection from a wall can be enough to provide some distance cues in many cases.
- A single delay line with two taps is enough to reproduce this basic effect.
- If the virtual sound source is close enough to the listening point the first tap can be taken directly from the source thus reducing the signal processing circuitry to single non-recursive comb filter.
- To be physical consistent the direct sound and its reflection should be attenuated as much as the distance they travel.
- The wall reflection should also introduce some additional attenuation and filtering in the reflected sound (represented by H_w).
- The distance attenuation coefficient is set in such a way that they become one when the distance goes to zero just to avoid the divergence to infinity that would come from the physical laws of a point source.
- From this simple situation it is easy to see how the direct sound attenuates faster than the reflected sound as long as the source approaches the wall. This idea can be generalized to closed environments adding a full reverberant tail to the direct sound.
- An artificial sound yet realistic reverberant tail can be obtained just by taking an exponentially decayed Gaussian noise and convolving it with the direct sound.
- The reverberant tail should be added to the direct sound after some delay (proportional to the size of the room) and should be attenuated with distance in a lesser extent than the direct sound.

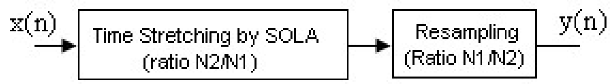


Flow Chart : Distance Rendering algorithm

Pitch Shifting by Time Stretching and Re-sampling

This effect shifts the frequency spectrum of the input signal. It can be achieved by "variable speed replay" which also leads to a compression or expansion of the duration of a sound. This is accomplished by re-sampling in the time domain.

The spectrum of the sound is compressed or expanded over the frequency axis. The harmonic relations $f_i = i.f_{fundamental}$ of the sound are not altered but are scaled according to $f_i^{new} = \alpha f_i^{old}$. The amplitudes of harmonics remain the same $\alpha_i^{new} = \alpha_i^{old}$.



In order to rescale the pitch shifted sound towards the original length a further time stretching algorithm can be applied to the sound. The order of pitch shifting and time scaling can be changed. First a time scaling algorithm expands the input signal from length N_1 to length N_2. Then a re-sampling operation with the inverse ratio N_1/N_2 performs pitch shifting and a reduction of length N_2 back to length N_1.

Variable speed replay

The phrase "variable speed replay" is used to mean that what has happened initially during the time period nT_{in} is now happening during nT_{replay} at relative speed v. where nT_{in} and nT_{replay} are the initial and replay sampling periods.

$$nT_{repaly} = nT_{in} / v$$

It means a straight forward method of implementing the variable speed replay is hence to modify the sampling frequency while playing back the sound according to

$$f_{replay} = f_{in} .* v$$

Note if sound signal is slowed down pitch decreases and duration increases and if fasten pitch increases and duration decreases.

At output sampling conversion has to be performed between the desired replay frequency and output frequency which is equal to input frequency.

If v<1 (time expansion) then $f_{in} > f_{replay} < f_{out}$ and more output samples are needed than available from input signal. The output signal is an interpolated version (over sampled) by the factor 1/v of the input signal.

If v>1 (time compression) then $f_{in} < f_{replay} > f_{out}$ and less output samples are needed than available from input signal. The input signal is decimated by the factor v. Before decimation the bandwidth of the input signal has to be reduced to $f_{replay}/2$ by a digital low pass filter.

Time Stretching

With the "time stretching" we mean the contraction or expansion of the duration of an audio signal.

We can alter duration of sound by using "variable speed replay" technique. But it has draw back of simultaneously transposing the sound. Harmonizer could be used to transpose the sound in the opposite direction and the combination of both methods leads to a time stretching algorithm.

The traditional technique for time stretching based on the variable speed replay introduces a pitch shift so we have to use **time segment processing** which divides the input sound into segments.

The main task of time stretching algorithm is to shorten or lengthen a sound of M samples to a new particular length $M^{'}=\alpha.M$, where α is the scaling factor. If the sound is to be lengthened some segments are repeated while if the sound is to be shortened some segments are discarded.

The problem of time segment processing is amplitude and phase discontinuity at the boundaries of the segments.

- Amplitude discontinuity are avoided by partial overlapping the blocks
- Phase discontinuity are avoided by a proper time alignment of the blocks

Two techniques of time stretching based on time segment are.

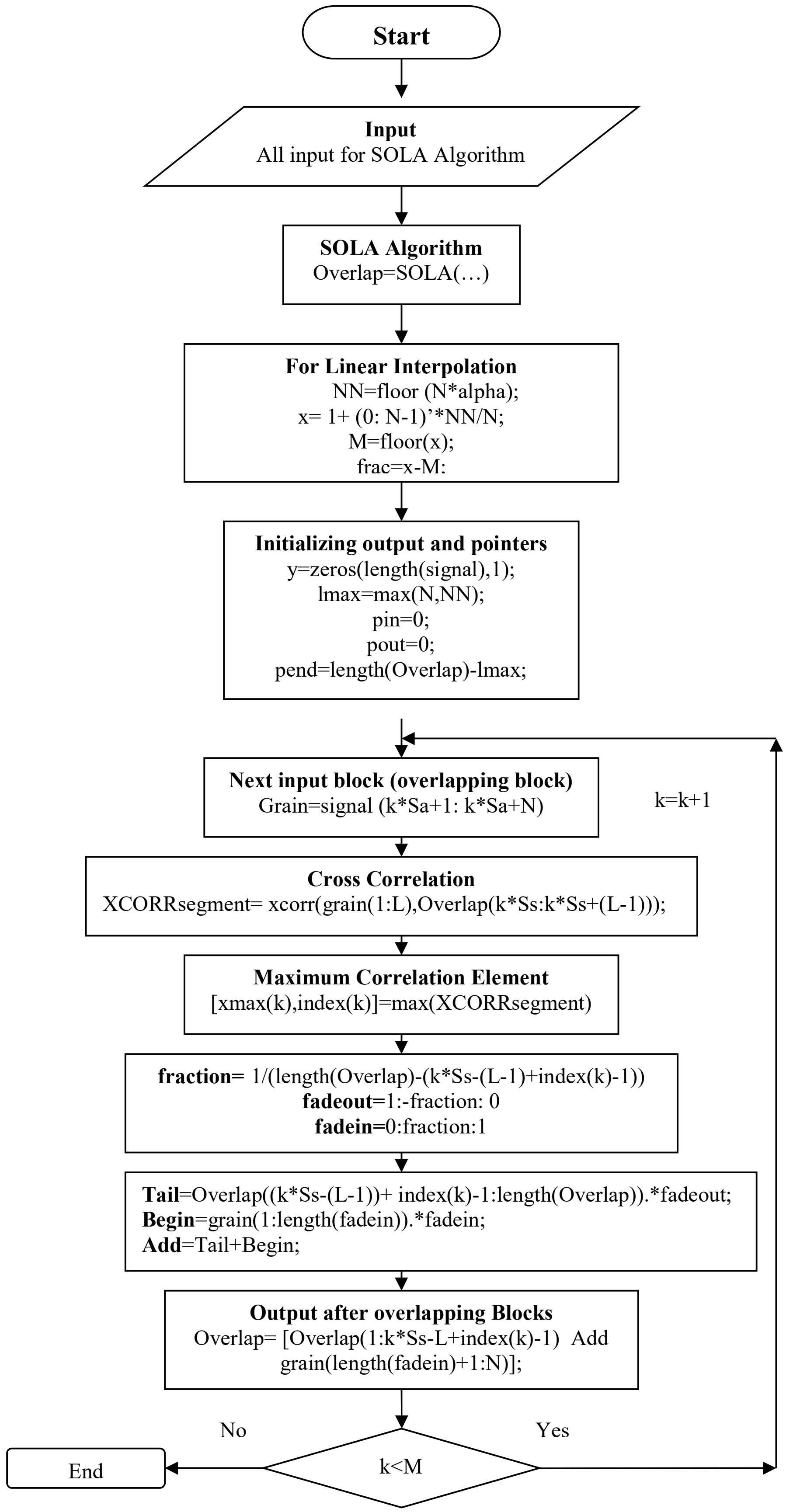
- Synchronous overlap and add (SOLA)
- Pitch Synchronous overlap and add (PSOLA)

Synchronous overlap and add (SOLA)

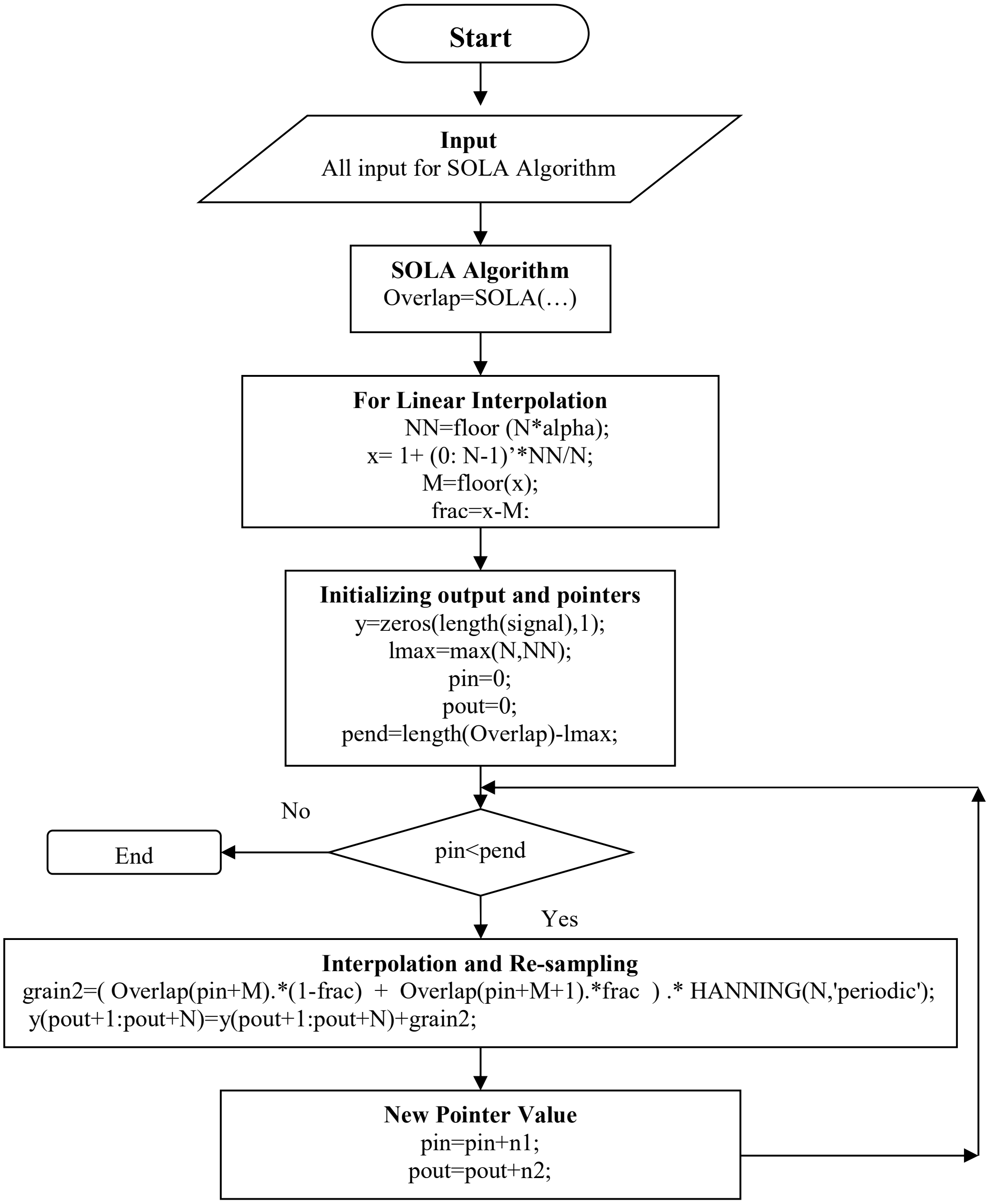
- The input signal is divided into overlapping blocks of a fixed length.
- The overlapping blocks are shifted according to the time scaling factor α.
- Then the similarities in the area of the overlap intervals are searched for a discrete time lag of maximum similarity.
- At this time point of maximum similarity the overlapping blocks are weighted by a fade in and fade out function and summed sample by sample.

Algorithm description

1. Segmentation of the input signal into blocks of length N with shift of S_a samples.
2. Repositioning of blocks with time shift $S_s=\alpha*S_a$ with scaling factor α.
3. Computation of the cross correlation $r_{xL1xL2}(m)$ between x_{L1} and x_{L2} which are segments of $x_1(n)$ and $x_2(n)$ in the overlap interval of length L.
$$r_{xL1xL2}(m) = \frac{1}{L}\sum_{n=0}^{L-m-1} x_{L1}(n).x_{L2}(n+m) \quad 0 \le m \le L$$
4. Extracting the discrete time lag tag k_m where the cross correlation $r_{xL1xL2}(k_m)$ $=r_{max}$ has its maximum value.
5. Using this discrete time lag km fadeout $x_1(n)$ and fade-in $x_2(n)$.
6. Overlap-add of $x_1(n)$ and $x_2(n)$ for new output signal.



Flow Chart : Working of SOLA algorithm



Flow Chart : Pitch shifting with SOLA RE-sampling