Audio Engineering Society

Convention Paper

Presented at the 127th Convention
2009 October 9–12   New York, NY, USA

Dynamic Panner: An Adaptive Digital Audio Effect for Spatial Audio

Martin J. Morrell[1] and Joshua D. Reiss[1]

[1] Centre for digital Music, Queen Mary, University of London, London, E1 4NS, UK
Martin.morrell@elec.qmul.ac.uk

ABSTRACT

Digital audio effects usually have their parameters controlled by the user, whereas adaptive digital audio effects, or A-DAFx, have some parameters that are driven by the automatic processing and extraction of sound features in the signal content. In this paper we introduce a new A-DAFx, the Dynamic Panner. Based on RMS measurement of its incoming audio signal, the sound source is panned between two user defined points. The audio effect is described and discussed, detailing the technicalities of all the control parameters and the creative context in which the effect can be used. Objective results that can be obtained from the effect are also presented. The audio effect has been implemented as both stereo and 3-dimensional effects using Max/MSP.

Figure 1 Here is an image of the Stereo Dynamic Panner User Interface

1. INTRODUCTION

There is a huge marketplace for audio effects for musicians, producers, live sound engineers and recording engineers. Recently, effects have become more intuitive, and are sometimes built to be dynamically controlled and some re adaptive. The authors present a new type of dynamic effect, the Dynamic Panner. In a music mix, an instrument is usually panned to a certain position and left there in that position for the entire song or at least section. An extra dimension would be added to the music mix/instrument by making use of continual change of pan position. Previous pan based effects used a tempo based movement from a given Point A to given Point B. Other effects, however, are based on the dynamic input level of the audio signal to apply a certain effect when the signal passes a given threshold, for example a compressor reduces dynamic range over a threshold [1][2][3], and a dynamic equalizer[1] applies a certain equalization curve over a given threshold. The use of the Dynamic Panner has a more creative than technical use, although the technical workings of the device are discussed herein with some discussion of creative uses. The technical use of the effect would be to move the sound in and out of regions of spectral or spatial masking so that it is either easier or harder to be heard by the listener when the sound level increases or decreases.

To the best of our knowledge, an adaptive panning has only been presented in [2], however with some limitations. First, it was limited to a stereophonic version (implemented with constant-power panning) and a 2D-planar version (implemented with VBAP source positioning) [3]. Second, while the general framework proposed include a generalization of hysteretic functions in the mapping layer between sound descriptor and effect parameters; it has not been investigated in the specific case of adaptive panning. When first starting this work the authors were not aware of Verfaille's work and so this was originally developed independently.

To overcome those limitations, the Dynamic Panner we propose in this paper addresses those issues. Indeed, it uses 3D sound projection. It also makes use of hysteresis in the same way as the dynamic effect family does, while transforming sound descriptors into effect parameters.

2. HOW IT WORKS

The Dynamic Panner is a feed-forward auto-adaptive effect [2]. Indeed, it is built using a feed forward design, which is where the RMS measurement is done by taking a parallel signal from the input to the RMS detection. Its auto-adaptive denomination results from the fact that the sound processing technique is controlled by a sound descriptor, namely RMS, that is computed from the same sound.

The control parameter set of the Dynamic Panner is similar to that found on a dynamic processor; threshold, attack, hold, and release. Along with these the effect also has a master pan, dynamic pan amount, sensitivity and smoothness controls. There are a lot of controls for the effect, but as a new effect it is justifiable to give the user the most amount of useful control as possible before some standard ways of using the effect are established.

The effect uses these different parameters in order to move a sound towards a given location from a 'master' location over a user given threshold by a calculated amount based on the sensitivity of the effect.

$$T_1 = T_2 - \Delta dB \qquad (1)$$

Signal level in these equations is measured using RMS over 130ms [1] for both T_1 and T_2. ΔdB Is a set level that T_1 is below T_2. For ease T_1 is set using a dB measurement rather than a multiplier or other linear scale. This type of hysteresis is also known as a Schmitt Trigger [4].

[1] http://www.tcelectronic.com/dynamicEQ.asp
http://www.audyssey.com/technology/dynamicEQ.html
http://www.voxengo.com/product/glisseq/

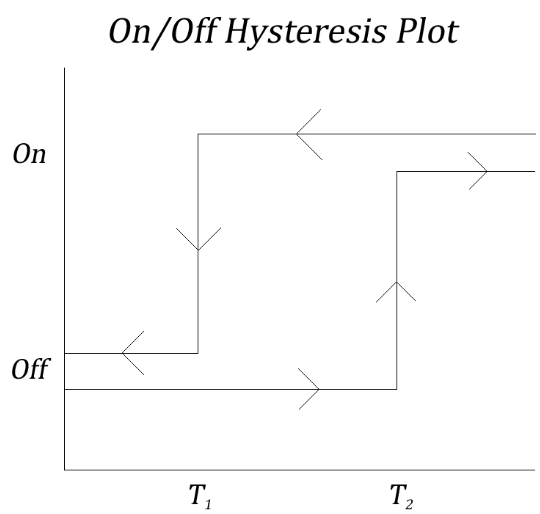

Figure 2 Hysteresis Plot of Threshold

2.1. Time Controls

There are four different parameters that control the effect over time once the thresholds have been passed; attack, hold, release and look ahead.

2.1.1. Attack

Attack is a measurement of time, t_{ATK} . In this instance it is the time taken for the sound to go from the *Master Pan* position to the *Dynamic Pan Location*, or in the 3D Dynamic Panner, *Master Azimuth* to *Dynamic Azimuth*, *Master Elevation* to *Dynamic Elevation* and *Master Inside* to *Dynamic Inside*. The time constraints may range from 0ms to 2000ms.

2.1.2. Hold

The hold adds an amount of time once attack time has passed before the release is applied. The hold is only applied once the level has gone below T_2 , the attack of the effect has passed t_{ATK} and the effect has been fully applied. The hold, t_{HLD} , can range from 0ms to 1500ms.

2.1.3. Release

The final time value in the Dynamic Panner, release, is the time for the effects on *Dynamic Pan Location* to return to the *Master Pan* position. The release, t_{RLS} , of the effect ranges between 0ms and 4000ms. The release can be used to add a slow movement to the sustain and release of a musical note further adding to the listener's perspective of the instrument moving out of the mix.

2.1.4. Look Ahead

A look ahead mode can be used so that the triggering RMS window happens at the same point in time as the full extent of the dynamic pan. To obtain this, the audio signal is delayed before having the effect applied, but the feed forward detection is not delayed. Therefore the look ahead uses a simple method where the look ahead delay time, t_{LAH} , is the same as the attack time, t_{ATK} .

2.2. Position Controls

The position controls determine the pan placement of the audio source when using the effect.

2.2.1. Master Pan

The Master Pan given by θ_{MASTER} is the pan position of the incoming mono sound source when the effect is not being applied, that is when $SignalLevel < T_1$. The master pan uses the sine/cosine law shown below:

$$G_L = cos\theta_{MASTER}$$
$$G_R = sin\theta_{MASTER} \quad (2)$$

2.2.2. Dynamic Pan

The Dynamic Pan Location, $\theta_{DYNAMIC}$, is the angular placement that the sound is trying to achieve by the use of this effect. The equation is the same as that for master pan (2) shown below in (3). In both cases θ is between 0 and 90 degrees in accordance with the standard stereo constant energy panning law.

$$G_L = cos\theta_{DYNAMIC}$$

$$G_R = sin\theta_{DYNAMIC} \quad (3)$$

It is most common that the stereo angle for loudspeaker reproduction be 60 degrees. Regardless of the stereo aperture it is still necessary to use the range from 0 to 90 degrees otherwise the panning law will not work. The user can use range from -1 to 1 where 0 is the centre, commonly found in commercial audio software to denote width as changeable and the panning as a ratio rather than an absolute angle. If users require, then we alter (2) and (3) to produce (4) using the user's stereo aperture.

$$G_L = cos\left(90\frac{\theta}{\theta_{APERTURE}}\right)$$
$$G_R = \sin\left(90\frac{\theta}{\theta_{APERTURE}}\right) \quad (4)$$

2.3. Smoothness

Smoothness controls how often a new pan position is calculated when moving between *Master Pan* and *Dynamic Pan Location* positions.

2.3.1. Millisecond Smoothness

In this millisecond based control the time that it takes for a new pan position to be processed ranges between 2ms and 1000ms.

2.3.2. Tempo Smoothness

Smoothness can also be calculated by using a given BPM (Beats Per Minute) tempo and a subdivision (SUB) note value; semibreve, minim, crotchet, quaver, semiquaver and so on up to 128 subdivisions, with the final result displayed as a millisecond measurement. The BPM and subdivision are converted to milliseconds by using:

$$t_{TBM} = \left(\frac{1000*60}{BPM*SUB}\right) \quad (5)$$

2.4. Sensitivity

The sensitivity control is comparable to a ratio control on a compressor; it determines by how many dB over T_2 the signal level has to go to be at the full *Dynamic Pan Location*. Sensitivity has a range between 0dB and 70dB. The equation for this is:

SENS=(SignalLevel-T_2)/ Sensitivity
If SENS>1, then SENS=1
If SENS<0, then SENS=0 (6)

$$\theta_{TARGET} = \theta_{MASTER} + \left[SENS \cdot (\theta_{DYNAMIC} - \theta_{MASTER})\right] \quad (7)$$

This shows that the current location of the sound is dependent on multiplication factor SENS, which ranges between 0 and 1. When the signal level goes beyond $T_1 + Sensitivity$ the value of *Dynamic Pan Location* is used. When the signal level is between T_2 and $T_2 + Sensitivity$ the amount the sound is moved is linearly proportional to the difference between the stated dB measurements.

3. WORKING SYSTEM

The authors' implementation of this effect was built using Max/MSP software [5]. Max/MSP uses objects. Each object carries out a function such as multiplication or addition for simple objects or calculates VBAP pan position for more complex objects. No objects were created for this application by the authors; all objects used are part of the software or available as free objects[2] [6].

The function of the system happens over a given time period which is defined by the user controllable parameters as introduced in section 2.

$$t_{SYSTEM} = t_{ATK} + t_{HLD} + t_{RLS} \quad (8)$$

Since this is only valid for an instantaneous peak above the threshold we can add another coefficient to the equation to show a more realistic t_{SYSTEM} where the threshold is crossed for a period of time greater than a single RMS measurement (130ms). This produces:

$$t_{SYSTEM} = t_{ATK} + t_{HLD} + t_{RLS} + t_{OVER} \quad (9)$$

With the t_{OVER} shown in (9), calculating t_{SYSTEM} breaks down and is only possible to calculate if the entire audio signal is known and the RMS windows are calculated. This is due to the fact that the audio effect reacts to the incoming audio, and the audio in real mixing is always changing in gain value. t_{ATK} is only ever passed if the audio is at a constant level for the same value as t_{ATK}, which will not happen often. In reality a new θ_{TARGET} (stereo version given for ease

[2] http://www.acoustics.hut.fi/software/vbap/MAX_MSP/
http://www.icst.net/downloads/
http://www.jasch.ch/dl/default.htm

of example, but 3D version is simply three times the same equation) is created every RMS window and by doing so t_{ATK} starts again and t_{ATK} becomes a speed rather than an absolute time. It should be noted at this point that if the next RMS window is greater than $RMS_{CURRENT}$ then the t_{ATK} is used to move the sound to θ_{TARGET}. If the next RMS window is smaller then $t_{RELEASE}$ is used.

There is a look ahead mode in the authors' implementation that simply adds a delay to the audio signal going through the effect. The feed forward detection is left un-delayed. The delay is equal to t_{ATK}. It can give a more musical response as the sound finishes its movement on the triggering RMS window rather than starting the movement on the triggering RMS window, which would otherwise add a certain lagging feel to the audio effect.

Figure 3 is a simple block diagram of how different parts of the Dynamic Panner interact with one another. It can be seen, on the left hand, the flow of the audio that is outputted whilst on the right hand side there are the feed forward detection and user controls.

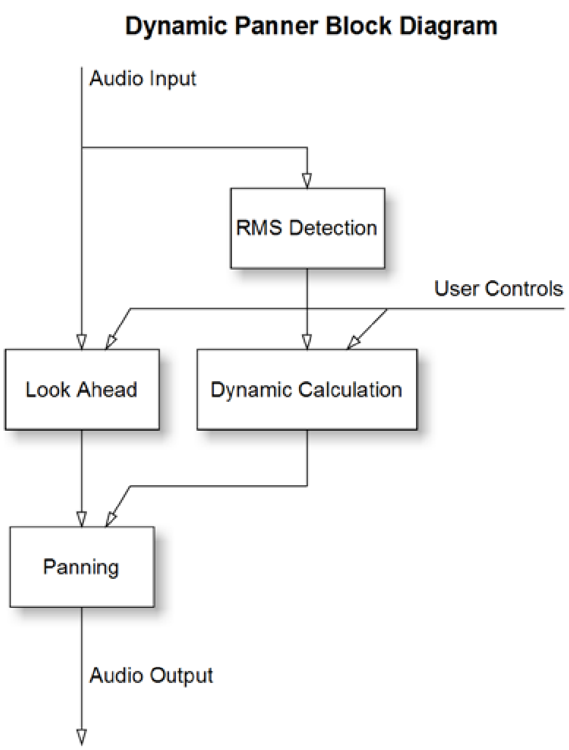

Figure 3 Dynamic Panner Block Diagram

4. GRAPHICAL DATA

Shown in this section are graphical representations of pan movement based on the dBFS input level of a sound. The representations are all of the stereo (1D) case using a 90° channel separation. The 3D graphs are trivial to reproduce and three graphs would be needed to show the movement of sound if azimuth, elevation and inside are used for the master and dynamic pan sections.

4.1. Case1: Master Pan 0°, Dynamic Pan 45°

In Graph 1 the θ_{MASTER} is 0° and the $\theta_{DYNAMIC}$ is 45°. The sensitivity is set to auto mode for the first 6 examples, which means the sensitivity is always the positive value of the threshold and always has the maximum dynamic pan at 0dBFS. The six thresholds used are labeled on the graph.

For the last 6 examples, the sensitivity is set at 30dB so the input signal has to rise 30dB over the threshold to be at the dynamic pan location. More than 30dB above threshold and the sound remains at the dynamic pan location and below the threshold the sound is always at the master pan. Six thresholds are shown and labelled on the graph. It is interesting when looking at Graph 1 to note that given certain settings the full dynamic pan may never be reached as can be seen for the -20dBFS threshold. This is true for all arbitrary threshold values greater than -30dBFS.

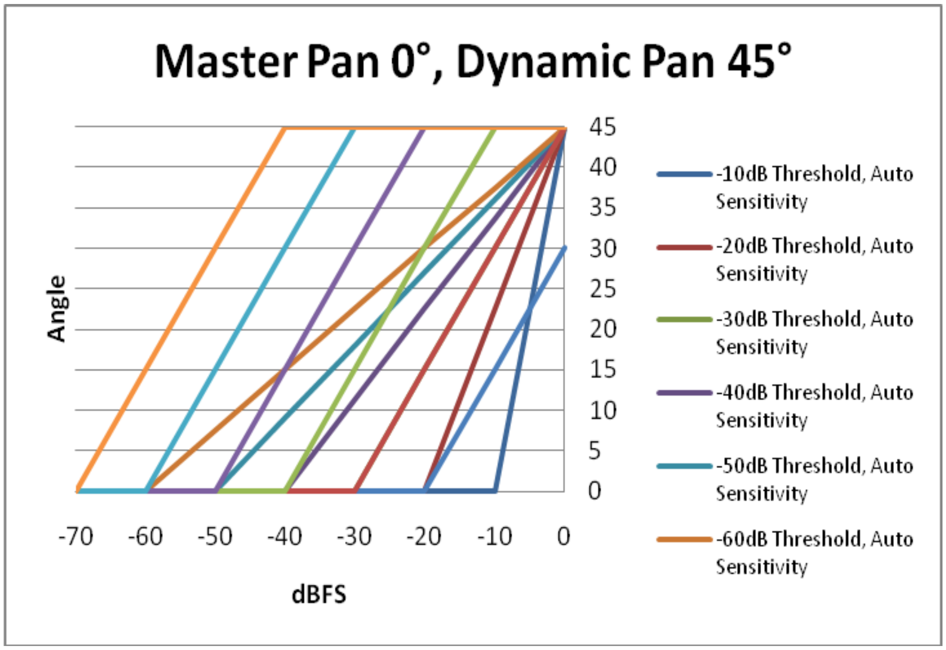

Figure 4 Case1: Pan Locations Produced by the Dynamic Panner Using Different Thresholds and Sensitivity Settings

4.2. Case2: Advanced Settings

Figure 4 shows an example where sensitivity of 28dB, a master pan of -30°, dynamic pan of 30° and a threshold of -50dBFS. The movement that will be applied to a sound using these settings can be clearly seen. This moves from the simplicity of the first two graphs to show that the master pan does not have to be 0° for the effect to work.

The second example uses a sensitivity of 15dB, master pan 45°, dynamic pan -45° and threshold -40dBFS. This second example on Figure 4 shows how the sound would move from right to left between the two speakers.

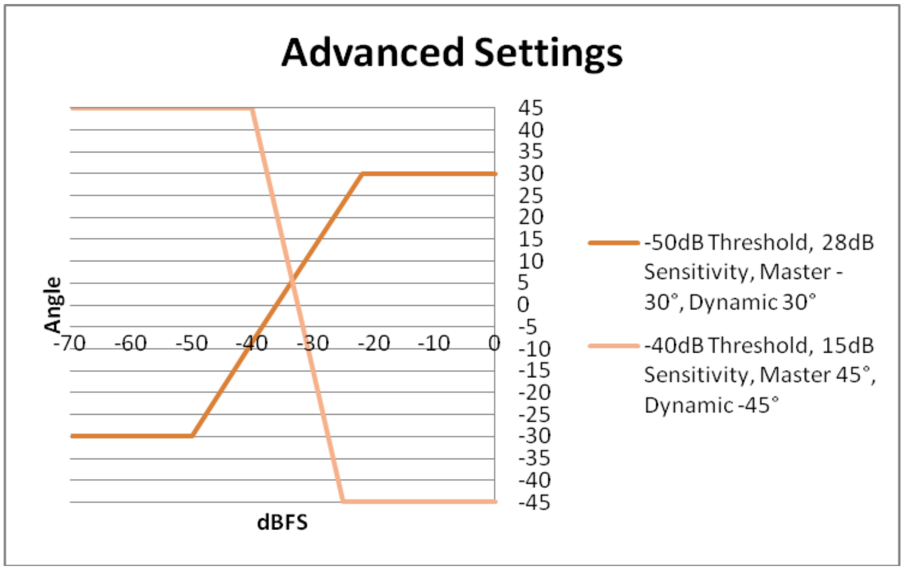

Figure 5 Case2: Pan Locations Produced by the Dynamic Panner for Two Examples

5. AUDIO EXAMPLES

Some audio examples were produced that show the effect under test, as well as being used in the context of a music mix. The time and pan position of the audio signal going through the effect was recorded as the mix was produced. For the first example the audio track was produced solely from the Max/MSP patch. For the second two examples, the multi-track music mix example, Max/MSP was used as a ReWire slave to a host Digital Audio Workstation, Cakewalk Sonar 8.

5.1. 1kHz Fade-In Sine Wave

A 10 second audio clip of a 1kHz sine wave had a fade-in applied to its entirety. The table of the dynamic panner settings used can be seen in table 1. Figure 6 shows the movement of the sound source as the effect was applied. The point of this test was to see if the Dynamic Panner would produce a smooth movement form left speaker to right speaker as the fade-in was applied.

Control Parameter	***1kHz Fade-In Sine Wave***	***"When We Weren't There"***	***"Le Disko"***
Threshold	-70.00dBFS	-50.75dBFS	-43.68dBFS
Attack	300ms	150ms	284ms
Release	300ms	400ms	632ms
Hold	0ms	0ms	0ms
Look Ahead	OFF	ON	OFF
θ_{MASTER}	-45°	-45°	-5°
$\theta_{DYNAMIC}$	45°	0°	-45°
Smoothness	2ms	2ms	170ms
Sensitivity	AUTO (70.00dB)	14.07dB	18.02dB

Table 1 Comparison of Settings Used for Dynamic Panner Stereo Audio Examples, This Shows The Various Control Parameters Available to the User to Obtain A Customised Effect

The test showed that there is a smooth movement between the left and right speakers as the RMS level is increased. During the authors' tests of this using an attack time below 300ms yielded less smooth results. It can be concluded that using the longer attack time will smooth out any inaccuracies produced from the RMS detection.

5.2. When We Weren't There

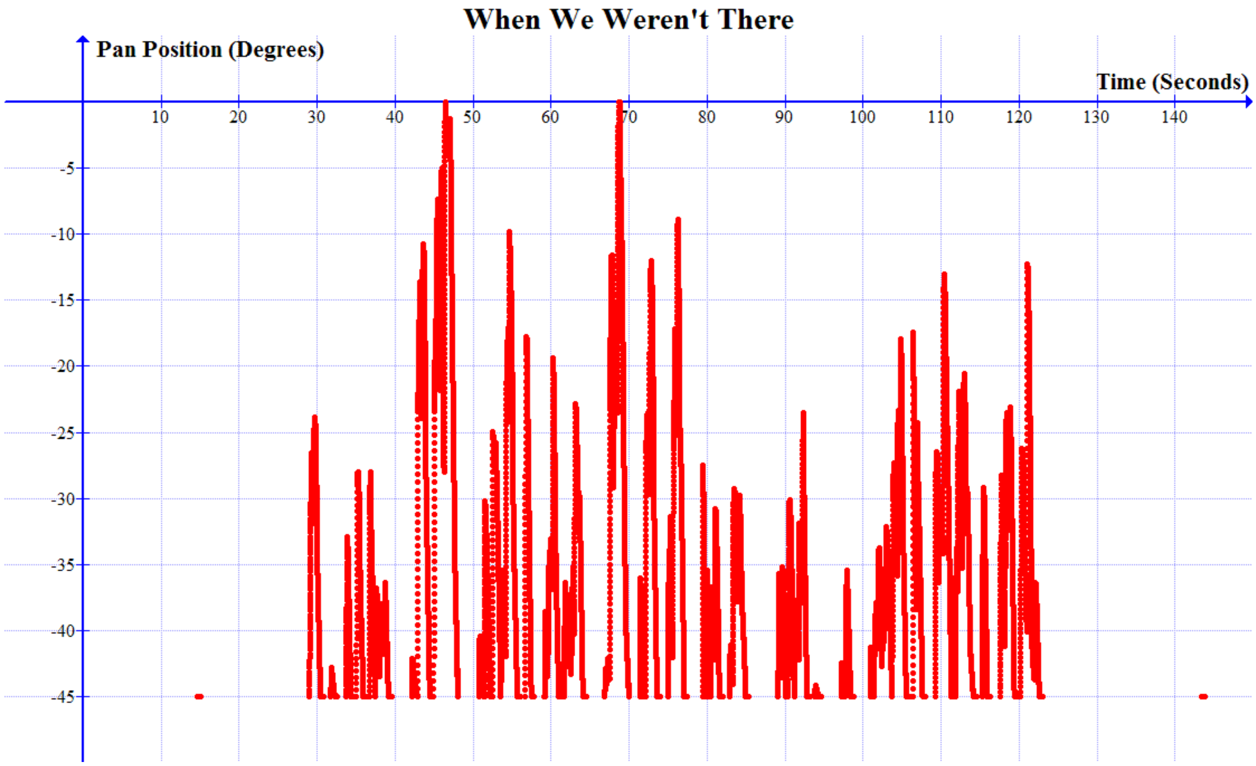

Figure 6 When We Weren't There – Plot of Sound Movement Produced by the 'Riff Guitar' in the Song. It Can Be Seen that Maximum Movement Only Occurs Twice Due to the Settings Used

The 'riff guitar' part was processed using the Dynamic Panner as part of a mix. The multi-track recording is part of the Centre for Digital Music's multi-track database. The settings used are shown below in Table 1.

Figure 6 shows the movement over time of the guitar riff part. As the guitar riff becomes louder so it moves towards the centre of the mix taking the place of where the vocalist is constantly panned.

Time and position values were recorded approximately every 2ms when the effect was being applied, in total over 32877 sets of values were obtained and plotted. The movement between the Master Pan and the Dynamic Pan position can be seen to be smooth with the release time longer than the attack as described in Table 1.

5.3. Le Disko

For this music mix the backing vocal was processed using the dynamic panner. The backing vocals are placed very close to the centre with the lead vocals when they are quiet but move outwards to the side when the effect is applied. The backing vocal becomes its own individual part rather than simply emphasising the lead vocal when it is panned further away. Table 1 shows the settings used on the Dynamic Panner for this audio example.

Figure 7 shows the movement of the backing vocals during the course of the song; Le Disko. The difference of using a higher value for the smoothness can be seen, the distance between points is far greater. This is also shown as only 556 sets of time and position values were recorded for this audio example.

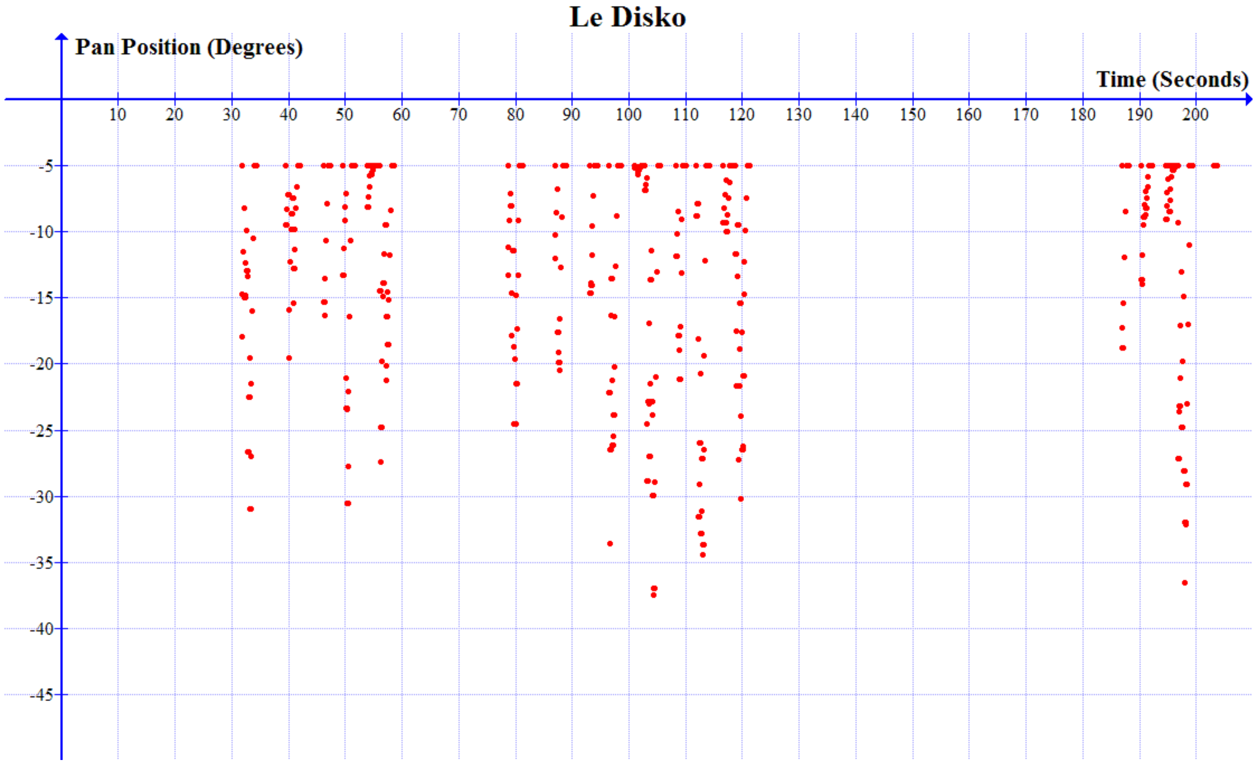

Figure 7 Le Disko – The Sound Movement Produced by The Dynamic Panner using the Settings as Shown in Table 1. Compared to Graph 3 There Are Significantly Fewer Positions Outputted Over Time

6. 3D DYNAMIC PANNER

There are very few audio effects that are designed to be specifically for multi-channel audio. The most notable effects designed for multichannel audio are compressors and reverberation. Some reverberation plug-ins are designed to accept a monophonic audio input and create a pseudo stereo output by means of diffusion and differences in early reflection and reverb tail in the different channels. With the 3D Dynamic Panner the sound is moved from one 3D coordinate to a Dynamic 3D coordinate based on the RMS level as previously discussed for the stereo version.

So as not to over extend this paper, only the different controls used for the 3D Dynamic Panner compared to the stereo version will be discussed. The different controls are for the 'Position Controls'. Figure 3 shows the user interface created for the 3D Dynamic Panner.

6.1. Master Position Controls

The 3D Dynamic panner has master controls for the Azimuth, the horizontal placement around the speaker array, Elevation, the height around the speaker array and Inside, an arbitrary scale to move the sound from on the speaker array to the listening position. These controls are all measured in degrees and the authors' application gives control for -360 to 360 for Azimuth (θ) and Elevation (ϕ) to give a wide range of movement control to the user. The Inside (r) control ranges from 1.0 at the speaker array to 0.0 for the listening position in the centre. This control is arbitrary as the end listener's 3D speaker array could have any radius.

6.2. Dynamic Position Controls

The Dynamic controls $\theta_{DYNAMIC}$, $\phi_{DYNAMIC}$ and $r_{DYNAMIC}$ determine where the sound will move to when the effect is fully applied ($T_2 + Sensitivity$). For RMS levels between the T_2 and $T_2 + Sensitivity$ values the $\theta_{CURRENT}$, $\phi_{CURRENT}$ and $r_{CURRENT}$ values are linearly and equally spaced.

6.3. 3D System

On the output section of the 3D Dynamic Panner the user has the option of Ambisonics [7][8] or VBAP (Vector Base Amplitude Panning) formats [9]. There is also an option for Ambisonics to use the authors opposite panning approach [10] for placing a sound in the listening area or using the Max Objects inbuilt distance function. For VBAP the opposite source is always used.

Visualisations are used in both the stereo and 3D versions of the Dynamic Panner. In the stereo version the user is shown a horizontal bar where the extreme left is the left speaker, extreme right the right speaker and the centre equally between the speakers. There is a ball that moves side to side showing $\theta_{CURRENT}$, as can be seen in Figure 8.

In the 3D version the user is presented with two circles that represent azimuth and elevation, pictured in Figure 2. The azimuth circle shows the azimuth position, $\theta_{CURRENT}$, combined with the inside value. The r value ranges from 1.0 (speakers) to 0.0 (listening position, LP). The same is true for the elevation circle where $\phi_{CURRENT}$ is combined with the inside value. This is so that the user is required to look at less visualization to realize where the sound should be at any given time. We can show the visualization as:

$$\theta_{VISUAL} = r\left(r\cos\theta, r\sin\theta\right)$$
$$\phi_{VISUAL} = \left(r\cos\phi, r\sin\phi\right) \qquad (10)$$

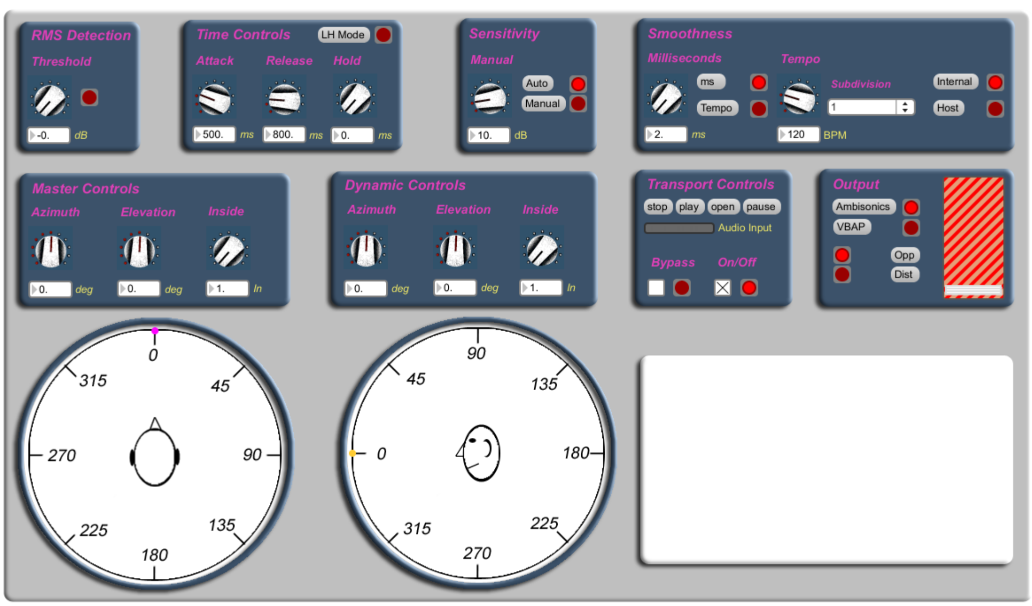

Figure 8 3D Dynamic Panner User Interface developed by the authors in Max/Msp

7. 3D GRAPHICAL DATA

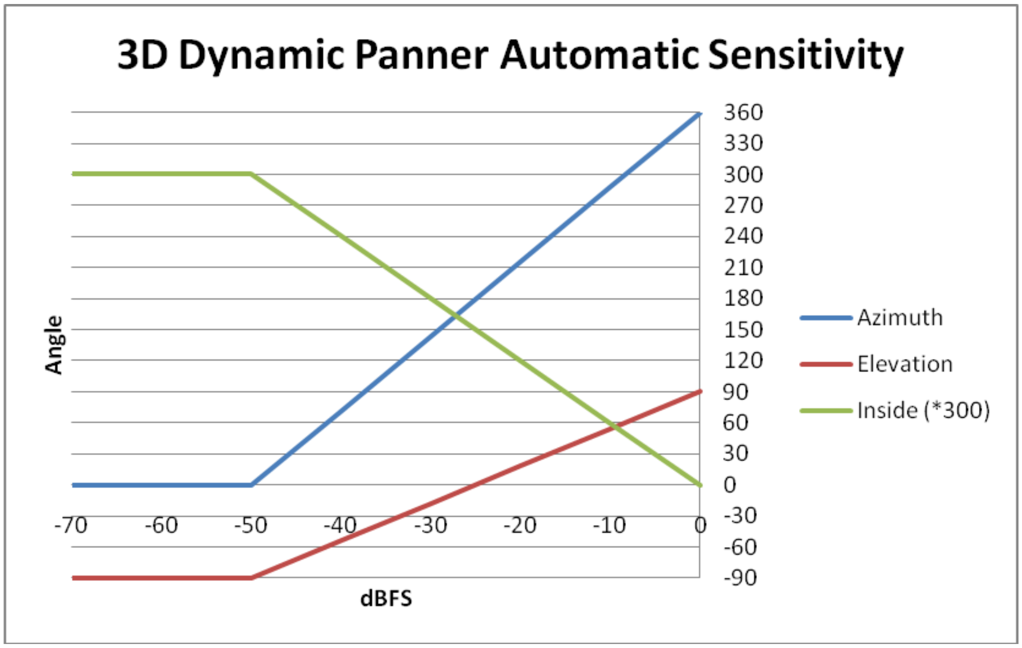

Figure 9 3D Dynamic Panner, Positions Produced for a θ, ϕ and $r*300$ Using a Threshold of -50dBFS and Automatic (50dB) Sensitivity

Figure 9 shows the movement plots for θ, ϕ and r when using a threshold of -50dBFS with automatic sensitivity. For ease of viewing, the inside measurement has been scaled so that 1.0 = 300. The master controls are $\theta_{MASTER} = 0$, $\phi_{MASTER} = -90$ and $r_{MASTER} = 1.0$, (300 on the graph). The dynamic controls are set to $\theta_{DYNAMIC} = 360$, $\phi_{DYNAMIC} = 90$ and $r_{DYNAMIC} = 0.0$.

8. 3D DYNAMIC PANNER AUDIO EXAMPLES

3D Audio examples were produced using Max/Msp software with Ambisonics objects. The produced recordings are 4 channel 1st order B Format audio files.

8.1. Radiohead – Reckoner

Control parameter	Radiohead – *"Reckoner"*	Plaid – *"Faster"*
Threshold	-29.12dBFS	-59.92dBFS
Attack	300ms	435ms
Release	800ms	435ms
Hold	0ms	0ms
Look Ahead	ON	ON
θ_{MASTER}	-31°	-120°
ϕ_{MASTER}	-40°	-0°
r_{MASTER}	1.0	1.0
$\theta_{DYNAMIC}$	0°	25°
$\phi_{DYNAMIC}$	0°	30°
$r_{DYNAMIC}$	0.344	1.0
Smoothness	2ms	150ms
Sensitivity	18.70dB	35.00dB

Table 2 Comparison of Settings Used for 3D Dynamic Panner Audio Examples. The Results of Using Different Settings Can Produce a Very Different Effect.

Figures 10 and 11 show the 3D movement of the sound source. Table 2 show the settings used and the graphs show that the full dynamic movement is only achieved at the end of the song. The graphs also show that the effects result depends on the RMS level and does not simple apply the same effect to the sound throughout the entire music mix.

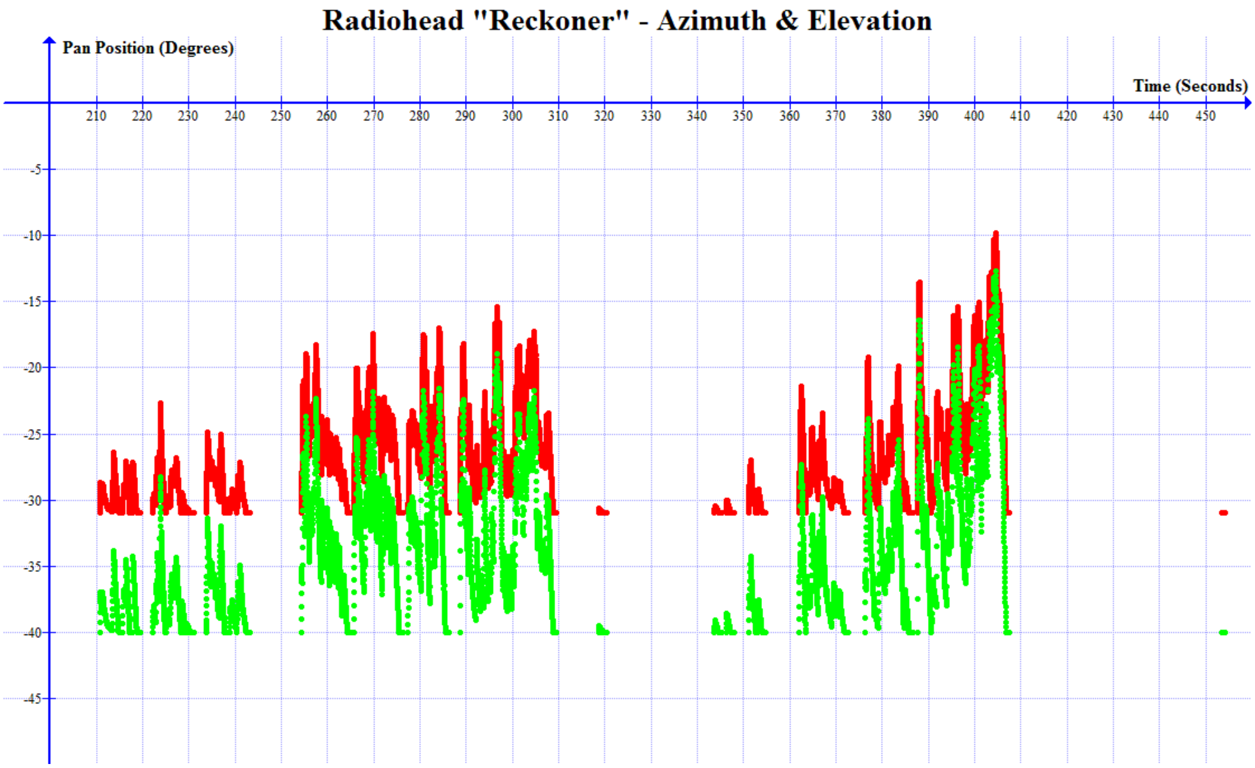

Figure 10 Radiohead “Reckoner”, θ and ϕ Movement Shown Over Time When The 3D Dynamic Panner is Applied to 1 Track in the Multi-Channel Recording

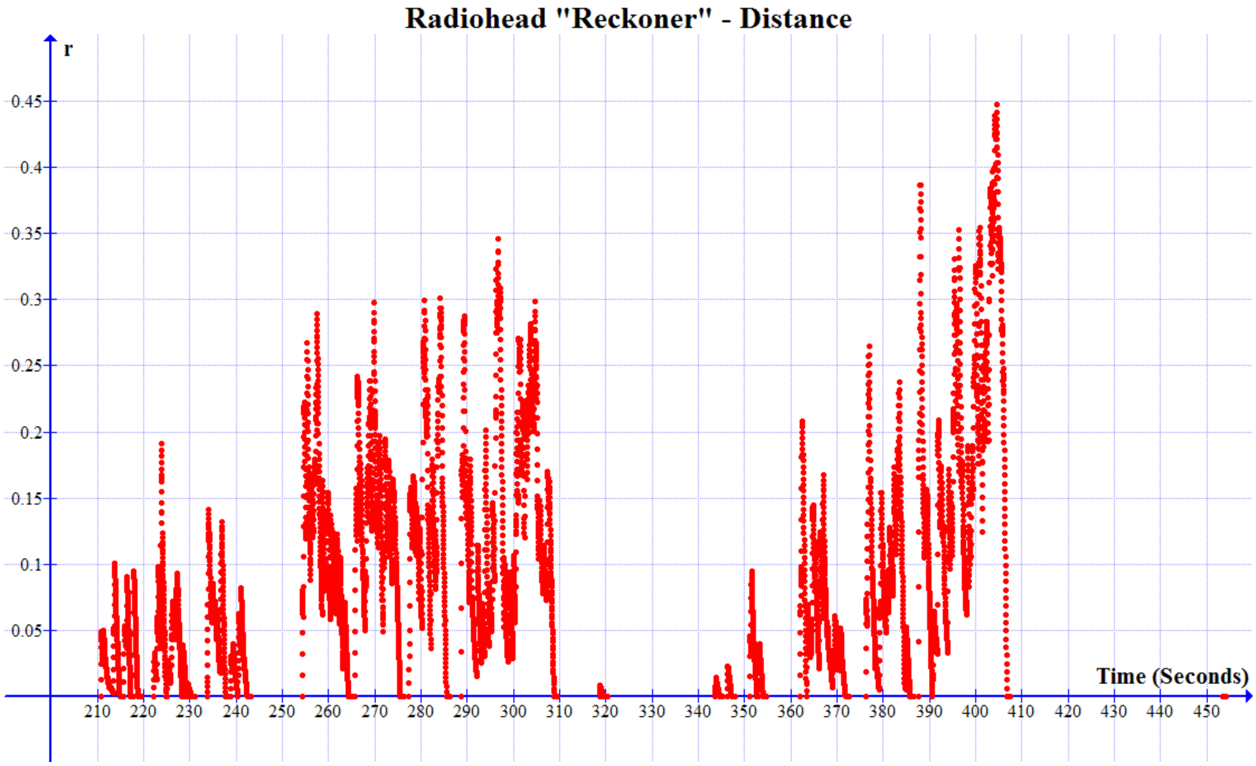

Figure 11 Radiohead “Reckoner”, *r* Movement Shown Over Time When The 3D Dynamic Panner is Applied to 1 Track in the Multi-Channel Recording

8.2. Plaid - Faster

Table 2 shows the settings used for the second 3D audio example.

For this audio example a greater azimuth smaller movement in the elevation and no movement in the inside were used. Figure 12 shows the azimuth and elevation movements, a graph for inside is omitted due to there being no movement in that direction.

We can see from Figure 12 the affect of using 150ms smoothness on the movement of the sound. Because of the smoothness it is much harder to tell the movement of the sound as it jumps around and so misses out all the interpolating points. It is difficult to see the elevation movement on the graph due to the small distance that it uses. It is again clear to see that the effect is adaptive and the movement varies largely throughout the song. We could, however, set the dynamic to have 0dB sensitivity so it would jump between the master positions and the dynamic positions.

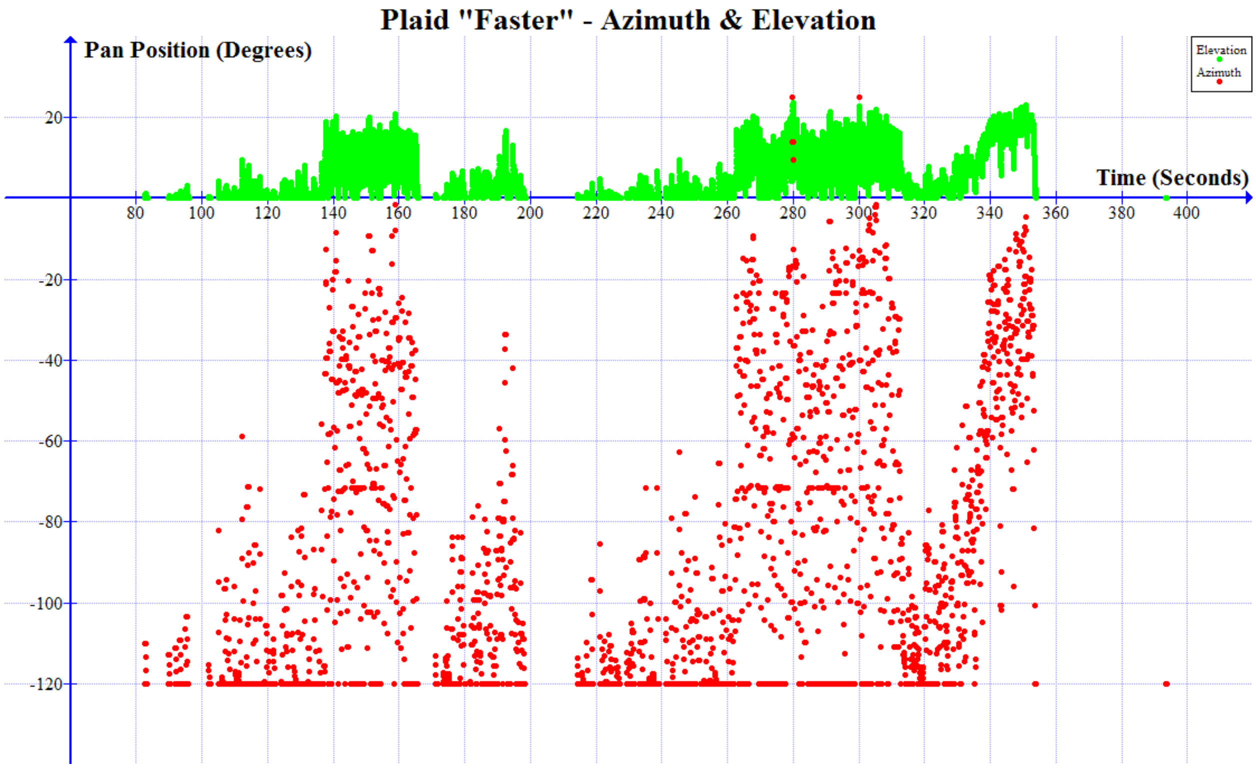

Figure 12 Plaid "Faster", The Azimuth and Elevation Plots Produced for This Audio Example Clearly Show The Effect of a Higher Value of Smoothness, This Gives The Audio Perception of The Sound Jumping Around

9. SPECTRAL CENTROID PANNER

The Spectral Centroid Panner is also a pan position movement audio effect, although instead of using RMS level detection to determine the sound sources output position, frequency detection is used to position a source. The use of this effect is also mainly creative for music making but can be used to solve the problem of making a monophonic input into a pseudo stereo audio mix. The best example of this would be to take a mono recording of a piano and create a stereo mix that sounds similar to what the pianist would hear. For a 3D example of the effect, one could use a monophonic recording of a wind instrument and give pseudo height information as to where the key holes on the instrument are placed. Another use for the Spectral Centroid Panner is to place low frequencies central and keep the energy equal between speakers. At higher frequencies the sound can be moved to a position that is not masking another sound. This is also useful as low frequencies use more energy than higher frequencies and so the overall gain of the sound could be higher without distorting the output of the musical mix.

To avoid repetition in the paper only the controls that are different to the Dynamic Panner will be presented in this section. The differences for the stereo and 3D versions is the Threshold detection which uses a frequency detection and instead of using a Sensitivity control a Max Frequency control is used instead.

9.1. Threshold Frequency

The authors tested 5 different methods for detecting the dominating frequency in real time of the incoming audio signal. The methods tested were Spectral Centroid (object made by Tristan Jehan) [11], Pitch Detection (by Tristan Jehan) [11], Resonance Frequency Filter Bank, 3rd Order Crossover Filter available in Max/Msp and a 4th Order Linkwitz-Riley Crossover Filter. The last 3 methods used a hierarchical approach using 4 banks of 5 filters to hone in on the 'hot' frequency. This hierarchical approach means it is equivalent to having 5^4 (625) filters although because the frequency is averaged over 130ms, it takes 520ms to calculate through the hierarchical filter bank.

The choice of frequency detection was Spectral Centroid, because it was the method that was most accurate when testing with a pure sinusoidal signal. The spectral centroid is calculated by [12]:

$$Centroid = \frac{\sum_{n=0}^{N-1} f(n)\, x(n)}{\sum_{n=0}^{N-1} x(n)} \qquad (11)$$

Where $f(n)$ is the centre frequency of a frequency bin, $x(n)$ is the magnitude of the frequency bin.

The threshold frequency is used on a logarithmic scale because human hearing is logarithmic. This means there is a higher degree of control at low frequency where more notes are in the same. The controls range between 1Hz and 20,000Hz to cover the range of human hearing.

9.2. Max Frequency

The second control specific to the Spectral Centroid Panner is the replacement of sensitivity with Max Frequency. This sets the frequency at which the full Dynamic Pan settings will be achieved. The control also works logarithmically as with the Threshold Frequency. The angular distances between the Master and Dynamic Pan controls is then linearly spaced across the range of Threshold Frequency to Max Frequency.

10. SPECTRAL CENTROID PANNER AUDIO EXAMPLES

Audio examples were created to test and demonstrate the Spectral Centroid Panner in use. The first demonstration tests the effect when applied to a sweeping sine wave. The second example shows the Spectral Centroid Panner being used in the context of a music mix.

10.1. Sweeping Sine Wave

The results with a sweeping sine wave should show a constant movement between the Threshold Frequency and the Max Frequency, 1Hz and 20,000Hz respectively. The settings used on the Spectral Centroid Panner as shown below in Table 3, Figure 13 shows the movement of the sound source from this test.

Control parameter	**Sweeping Sine Wave**	**Roots – *“Marathon Man”***
Threshold Frequency	1Hz	219.20Hz
Attack	200ms	100ms
Release	200ms	200ms
Hold	0ms	0ms
Look Ahead	ON	OFF
θ_{MASTER}	-45°	-40°
$\theta_{DYNAMIC}$	45°	40°
Smoothness	2ms	300ms
Max Frequency	20000Hz	796.32Hz

Table 3 Comparison of Settings Used for Dynamic Panner Stereo Audio Examples

The graph indicates that the intention of the Spectral Centroid Panner is fulfilled apart from a slight bump around 22.5seconds, but otherwise a smooth curve of movement is produced. The error produced in the movement can be due to the window size difference used in Max/Msp and the Spectral Centroid object.

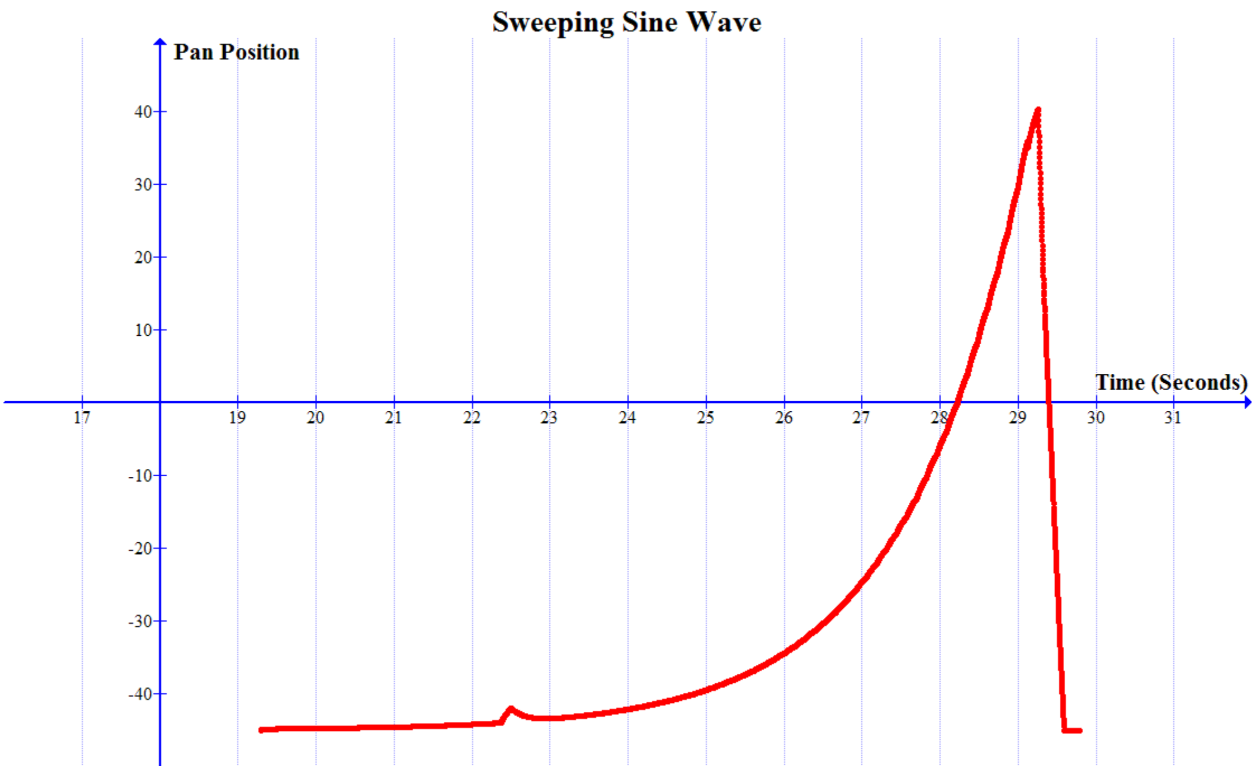

Figure 13 Sound Movement Produced Between -45° and 45° By The Spectral Centroid Panner to a Sweeping Sine Wave

10.2. Marathon Man

In this audio example the arpeggiating synthesizer that constantly runs throughout the music mix has the Spectral Centroid Panner applied to it. By using a smoothness value that is three times that of the attack time and one and a half that of the release time, the sound is moved abruptly between pan positions. Much like if frequencies were allowed above the nyquist frequency, we cannot join the values to find the original movement between points. Table 3 shows the settings used and Figure 14 shows the movement of the sound.

No discernable movement lines can be obtained from Figure 14. This shows that the effect can be used in such a way as to not be able to see the full movement points on a graph or to calculate the parameters used by looking at a graph of the outputted position values, other than the pan positions which may be obtained by the outer limits on the y axis.

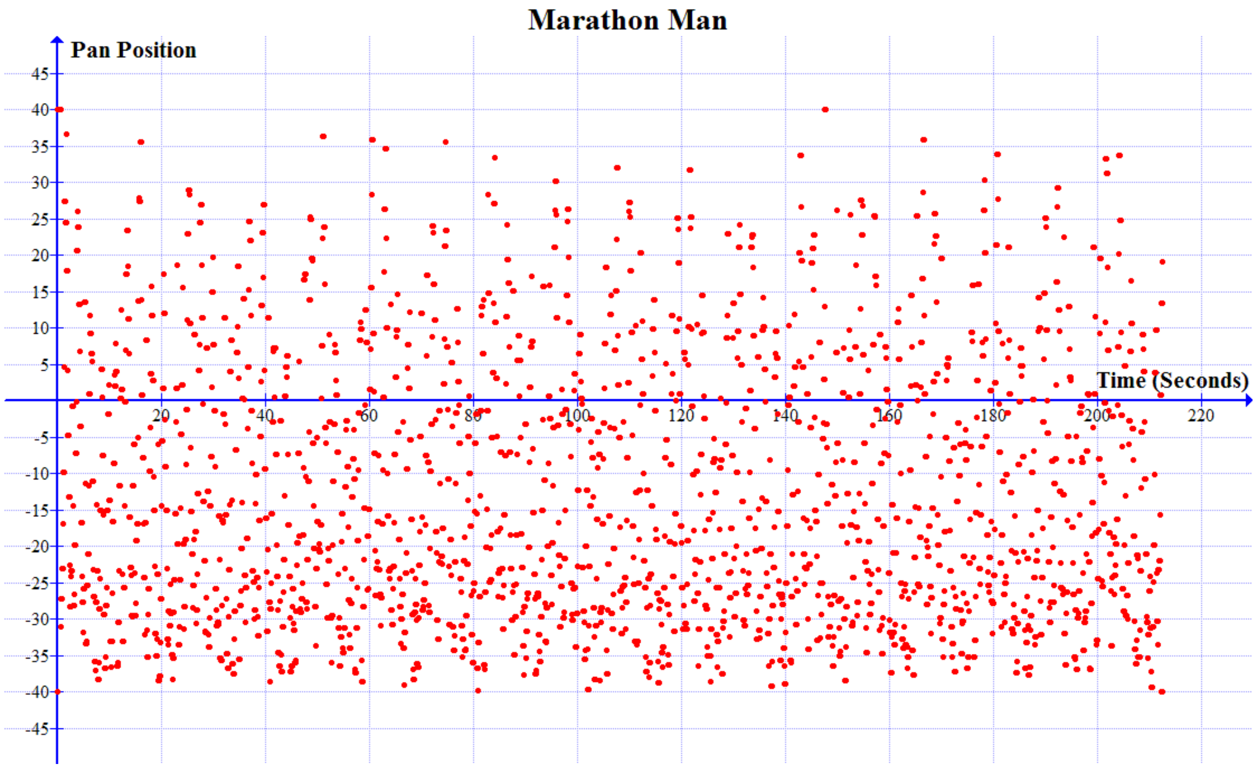

Figure 14 The movement produced by the Spectral Centroid for Marathon Man uses smoothness time three times that of the attack time and twice that of the release time. Using settings in this was produces a graph where the movement is difficult to see.

11. FURTHER DISCUSSION

During the testing, there were some recommended settings and artifacts noticed worth of note.

11.1. Settings Recommendation

Good results were achieved using attack times above 100ms and sensitivity value between 10 and 40dB. Attack times of above 300ms were shown from the test to give the most fluid movement and below 100ms artifacts may be apparent to the listener. The sensitivity range recommended is similar to that used by instrumentalists in modern pop songs. It would also be recommended to place the Dynamic Panner before any compression devices so that the greatest movement is achieved, although this would require multi-channel compressors rather than a single channel, which would be more computationally heavy.

11.2. Artifacts

Artifacts became apparent when the attack and/or release time was too quick. With a small t_{ATK} the sound can jump between locations rather than move smoothly. Take for example a snare drum. When the short t_{ATK} and t_{RLS} were used there was an initial hit in the master pan location, a quick hit in the dynamic pan location and then a final hit back in the master pan location. This broke the single sound into three distinct sounds losing musicality and definition of the sound source.

11.3. Doppler Effect

As with any effect that involves movement, be that of the source or the listener, the Doppler Effect can impact on the frequency of the sound source. The effect on frequency due to a moving source can be calculated using (12) given in [1].

$$f_d = f_s\left(1 + \frac{c_s}{c}\right) \quad (12)$$

Where c is the speed of sound, c_s is the speed of the sound source or listener's movement, f_s is the intended frequency and f_d is the frequency heard by the listener.

Although Doppler will undoubtedly affect the sound, when the Dynamic Panner is being used it is the author's opinion that it will not be moved fast enough to hear a large deviation in frequency. However, the effect is to be used creatively so it is the choice of the user to have as many or as little artifacts or distortions to the original sound source as they choose. If we take a scenario where the listener sets the speaker on a line 1.5m away from the listening position and places the speakers at +/-30°. The vectors produced are [1.5, -0.8666] for the left speaker and [1.5, 0.8666] for the right speaker, both having a Euclidian distance of 1.7321m. The distance between the speakers is also 1.732m as shown in Figure 15. By first using the recommended attack time of 500ms, using the full aperture for travel and taking c to be 340m/s we obtain f_d=1.010f_s. This will make the sound source increase in pitch by 1%. If now we look at the recommended release time of 800ms then f_d=1.006f_s, a change in frequency of 0.6% making the 1kHz tone 1.006kHz.

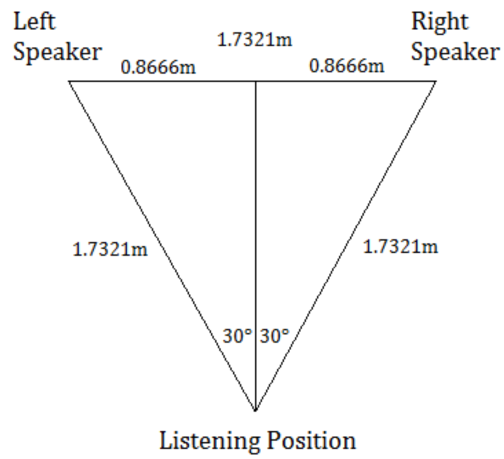

Figure 15 Speaker and Listener Placement for Calculating The Doppler Effect

11.4. Speed vs. Time

In the 3D situation, the t_{ATK} and t_{RLS} become dependent on the radius of the speaker array. In this case a more appropriate control to be used would be speed rather than attack time; based on the user's radius and speed wanted, the t_{ATK} and t_{RLS} can be calculated. By using:

$$c = \frac{d}{t} \quad (13)$$

$$d = r\Delta\sigma \quad (14)$$

d is the distance, r the radius and σ the angular distance, we calculate $\Delta\sigma$ from the Vincenty formula [13]:

$$\Delta\sigma = \left(\frac{\sqrt{\left(\cos\phi_f \sin\Delta\lambda\right)^2 + \left(\cos\phi_s \sin\phi_f - \sin\phi_s \cos\phi_f \cos\Delta\lambda\right)^2}}{\sin\phi_s \sin\phi_f + \cos\phi_s \cos\phi_f \cos\Delta\lambda} \right) \quad (15)$$

Where the azimuth angles, λ, and the elevation angles, ϕ, are given in radians. The λ_s and ϕ_s in the equation denotes the start position and λ_f and ϕ_f denotes the finish point.

12 Conclusion

In this paper we present the Dynamic Panner, that offers substantial improvements to the adaptive panner. Implemented both in stereo and 3D as Max/MSP patches, its mapping of RMS to the control parameters is similar to that of the specific control of dynamic effect family. Among the possible mapping strategies, the use of spectral centroid information to drive the effect provides interesting auditory effects. Although the effects can have artifacts depending on the t_{ATK} and t_{RLS} settings, this does not differ from adverse artifacts produced by other audio effects such as compressors. Moving a sound during a song can add extra interest to a musical mix.

In its stereo version, the Dynamic Panner adds interest to different parts in a music mix, enabling one to make a particular instrument stand out of the mix, a task which is usually assigned to equalizers, compressors, chorus or flanger. From the fade-in test it can be concluded that attack and release times of 300ms or above are needed for smooth movement and also the authors have noted that a sound can become disjointed if settings below 100ms are used. This is also explained by perceptual knowledge that speeds under approximately 20Hz are not heard as movement. Doppler could be a potential problem when using this effect, but has not been audibly noticeable by the author in his audio examples. With the use of higher smoothness settings a gate like movement effect is created which is a further creative control for this effect as it allows a sound to move at the same tempo as the song.

The authors have presented a true multi-channel effect for use within multi-track music recordings. The effect in a 3D environment can be much more noticeable than in conventional stereo as there is a larger area for the sound to travel. Doppler can have a greater effect in these situations if quick attack and release times are used in combination with large distances. This could inevitably detract from the musical mix as the instrument would become out of tune. There is no way to counteract the Doppler Effect when mixing unless a particular venue is mixed for, as the distance of speakers in an individual's playback system will be different.

The Spectral Panner has also been shown to be a useful and usable effect for music making. The choice of using spectral centroid as the frequency detection is emphasized by it showing the perceived brightness of a sound which will make the results match to human perception better. The optimistic ranges of frequency controls between 1Hz and 20,000Hz in reality are set so that there is only a gap of a few hundred Hz between Threshold Frequency and Max Frequency. This is due to the limited range of musical instruments and more so by range used by an instrument in a particular song.

13 References

[1] U. Zölzer, *DAFX: Digital Audio Effects,* Wiley, 2002.

[2] V. Verfaille, U. Zölzer and D. Arfib, "Adaptive Digital Audio Effects (A-DAFx): A New Class Of Sound Transformations", *IEEE Transactions on Audio, Speech and Language Processing*, 14(5), pp. 1817-1831, 2006.

[3] V. Verfaille, M. M. Wanderley and Ph. Depalle, "Mapping Strategies for Gestural and Adaptive Control of Digital Audio Effects", Journal of New Music Research, 35(1), pp. 71-93, 2006.

[4] M. Filanovsky, H. Baltes, "CMOS Schmitt Trigger Briefs", *IEEE Transactions on Circuits and Systems – 1:Fundamental Theory and Applications,* Vol 41, No. 1, January 1994.

[5] Max/Msp, Cycling74, http://www.cycling74.com/products/max5

[6] V. Pulkki, "Generic panning tools for MAX/MSP", http://lib.tkk.fi/Diss/2001/isbn9512255324/article4.pdf

[7] M.A. Gerzon, "Periphony: With-Height Sound Reproduction," *Journal. Audio Eng. Society*, vol. 21, pp. 2-10, February 1973.

[8] D. Malham, *Higher order Ambisonics systems*, http://www.york.ac.uk/inst/mustech/3d_audio/higher_order_ambisonics.pdf, 2003.

[9] V. Pulkki, "Virtual Sound Source Positioning Using Vector Base Amplitude Panning," *Journal Audio Engineering Society*, Volume 45, Issue6, pp. 456–466, June 1997.

[10] M. Morrell, J. Reiss, "A Comparative Approach to Sound Localisation within a 3D Sound Field", *Journal Audio Engineering Society*, Paper 7663, 126th Convention, May 2009.

[11] http://web.media.mit.edu/~tristan/

[12] G. Peeters, "A Large Set of Audio Features for Sound Description (Similarity and Classification) in the CUIDADO Project" Ircam, 2004.

[13] T. Vincenty, *Survey Review*, Directorate of Overseas Survey of the Ministry of Overseas Development, Surrey, 1975.