Audio Engineering Society

Convention Paper

Presented at the 117th Convention
2004 October 28–31 San Francisco, CA, USA

A Novel Multichannel Panning Method for Standard and Arbitrary Loudspeaker Configurations

Ramy Sadek[1] and Chris Kyriakakis[2]

[1] Institute for Creative Technologies / Univ. of Southern California, Marina Del Rey, California 90292, USA
sadek@ict.usc.edu

[2] Immersive Audio Laboratory, Integrated Media Systems Center, Univ. of Southern California, Los Angeles, California 90089, USA
ckyriak@imsc.usc.edu

ABSTRACT

This paper presents a novel panning algorithm called Speaker-Placement Correction Amplitude Panning (SPCAP) which guarantees conservation of loudspeaker power output. The method is appropriate for any speaker arrangement (e.g. ITU 5.1, 10.2, etc.), and scales with the number of speakers. SPCAP works by correcting initial pan values based on speaker placement to achieve constant power output. Because panning occurs over an arbitrary number of speakers (i.e. is not pair-wise), SPCAP provides two significant advantages over discrete panning schemes. First, pan values for current and future surround-sound formats (e.g. 5.1 and 10.2) are guaranteed to conserve power under any lower-resolution setup, making dynamic up/down mixing in non-standard setups feasible. Second, SPCAP provides a framework for producing *wide* (non point-source) sounds.

1. INTRODUCTION

Recent work on multichannel pan-pots has focused on particular loudspeaker configurations (e.g. ITU 5.1 and 10.2) [1] [2]. Diverse areas such as virtual reality and computer music require panning methods that allow reasonable performance on both standard and non-standard loudspeaker configurations. This work addresses the need for such flexible pan-pot generation. The VBAP method presented in [3] was the first to generate layout-independent pan-pots.

We have created a framework for designing pan-pots that are independent of loudspeaker configuration. The contributions of this work are a set of pan-pot rules guaranteed to conserve power under rotation and a framework for rendering *wide* (rather than point-source) sounds. In broad terms, the method works by determining the significance of each loudspeaker to the total output. By weighting each loudspeaker

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **OCT 2004**	2. REPORT TYPE	3. DATES COVERED **00-00-2004 to 00-00-2004**

4. TITLE AND SUBTITLE **A Novel Multichannel Panning Method for Standard and Arbitrary Loudspeaker Configurations**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Southern California,Institute for Creative Technologies,13274 Fiji Way,Marina del Ray,CA,90292**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**5**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

appropriately, SPCAP computes the number of *effective* speakers per channel. Any power loss is tracked; the lost power is then added back into the system and distributed according to each channel's contribution to the total power output. It can therefore be seen that power output is constant under rotation.

2. BACKGROUND

Researchers have sought optimal pan-pot laws since the beginning of stereophony [5]. Though the most popular and well-known pan-pot is pair-wise amplitude panning [1], there exist several panning methods that are not pair-wise, most notably [1] [2] [3] [4]. Pulkki's VBAP is the method most pertinent to the one presented here. Both VBAP and our method generate multi-speaker pan-pots dynamically, based on loudspeaker position information. However, the central aims of the two methods differ: while VBAP aims to produce maximally sharp images, our method is specifically targeted to producing wide sounds. The VBAP method aims to generalize pair-wise panning to three dimensions. While VBAP is the basic inspiration for our method, we endeavor to overcome three problems of VBAP:

(i) VBAP requires interpolation between source positions to avoid disturbing effects [3]

(ii) Because of *(i)*, power output is not constant during interpolation [3].

(iii) Implementing *spread* as described in [3] is costly, being linearly proportional to desired source width and speaker number.

The first two problems may be overcome within the VBAP framework by computing the proper gain values at each time-step. Problem *(iii)* is more fundamental than *(i)* or *(ii)*.

3. METHOD

Speaker-Placement Correction Amplitude Panning (SPCAP) is a novel panning algorithm for 3D audio. This method extends the notion of cosine-weighted panning by correcting gain values according to source and loudspeaker locations. By weighting each loudspeaker channel appropriately, SPCAP computes the number of *effective* speakers per channel. It then computes appropriate gain values in two passes over the speaker array. The first pass achieves gain normalization (Section 3.1) and the second pass ensures power-conservation (Section 3.2). See Figure 1. For clarity all equations assume that a listener is located at the origin, surrounded by speakers on a unit sphere. [1]

3.1. Gain Normalization

We compute the initial pan values using a normalized cosine-weighting based on the angle between the virtual source and each loudspeaker. In an ITU setup, for example, a source at zero degrees would result in a pan value of 1.0 in the center channel. So we have simply

$$P_i(\theta_s) = \frac{1}{2}\left[1+\cos\left(\theta_i - \theta_s\right)\right] \qquad (1),$$

where P_i is the initial pan-value assigned to the i^{th} loudspeaker, θ_i its placement angle and θ_s is the direction of the virtual source.

The pan values $\{P_i\}$ vary with speaker placement. Normalizing by the number of *effective* speakers per loudspeaker removes this dependency. Consider two co-located loudspeakers. If each loudspeaker receives a pan value as given by(1), the power from their direction is twice as high as appropriate. By viewing this pair as the number of effective speakers it represents, we can reduce their power by a factor of two, achieving the correct gain value from their direction.

We compute the effective number of speakers β_i for a given speaker *i* as a sum of the cosines of the angles between *i* and each loudspeaker and *j*. Since this includes the zero-angle between each speaker and itself, there is always at least one effective speaker.

So we have,

$$\beta_i = \sum_{j=0}^{n} \frac{1}{2}\left[1+\cos\left(\theta_i - \theta_j\right)\right] \qquad (2)$$

[1] Allowing speakers beyond the unit sphere and allowing the listener to move away from the origin are goals for our future work.

where θ_i, θ_j are the placement angles of the i^{th} and j^{th} speakers, respectively.

Once the effective speaker weights have been computed, we determine an initial gain value G_i via

$$G_i(\theta_v) = \frac{P_i(\theta_s)}{\beta_i} \qquad (3).$$

Dividing each pan value P_i by the corresponding β_i, we normalize each G_i against the speaker distribution. However, this is not sufficient to guarantee that power is conserved under rotation.

3.2. Power Conservation

Since power varies inversely with distance-squared [2] [7], we can determine the amount of power the loudspeaker array should emit. The emitted power P_e is given by the sum of the squared gains.

$$P_e(\theta) = \sum_{i=1}^{n} \left(G_i(\theta)\right)^2 \qquad (4)$$

Yet the correct power L is inversely proportional to distance-squared, yielding

$$L = \frac{\mathrm{k}}{d^2} \qquad (5),$$

where k is a constant [3]. Since amplitude-panning alone is not able to represent sounds within the boundary formed by the loudspeakers, there is a minimum allowed distance between the virtual source and the listener. We suppose that the minimum distance is the radius of the unit-sphere upon which the loudspeakers are located. Therefore, when a source is located on the unit-sphere, its power output should be maximal. (i.e. $d = 1$ and $L = \mathrm{k}$) .

Since we know the contribution of each i^{th} loudspeaker we can compute the ratio of each loudspeaker's output to the total power as

$$R_i = \frac{G_i^2}{P_e} \qquad (6).$$

Dividing L into portions $\{\gamma_i\}$ dictated by $\{R_i\}$ shows how much power should go to each i^{th} loudspeaker. Thus we define $\{\gamma_i\}$ in Equation (7).

$$\gamma_i = R_i L \qquad (7)$$

By definition, $\sum R_i = 1$. Therefore $\sum \gamma_i = L$, which means that using powers $\{\gamma_i\}$ ensures constant power output from the speakers. The final gain values $\{A_i\}$ are thus

$$A_i = \sqrt{LR_i} \qquad (8),$$

meaning that power is constant for all source angles.

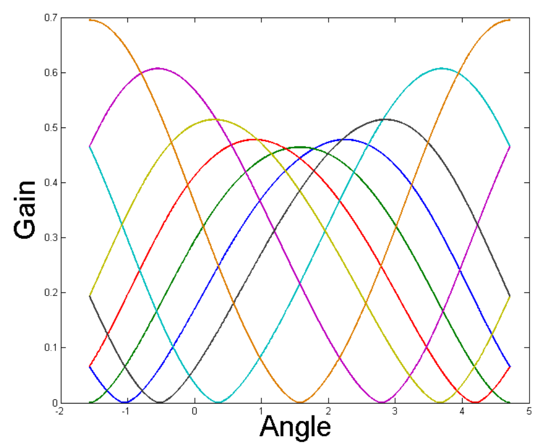

Figure 1: 8-channel gains yielding constant power.

[2] We make the simplifying assumption that the virtual source represents an omnidirectional (monopole) emitter.

[3] The exact value of k is not needed, since we aim only to produce constant power, not replicate the power of a real source. Therefore, we take $\mathrm{k} = 1$.

4. DISCUSSION: BENEFITS OF SPCAP

The SPCAP method as formulated above determines n gain values - one per loudspeaker - that yield constant power during source rotations. Development of this method was in part motivated by the need to overcome three shortcomings of VBAP as described in Section 2. Problems *(i)* and *(ii)* are solved by the constant power guarantees of SPCAP.

By raising the cosine term in (1) to a power α, SPCAP can control *tightness*, or how steeply the signal dies off relative to angle from the source as in Figure 2. This control allows SPCAP to achieve the same effect as VBAP to approximate non-point-source sounds. In [3], Pulkki refers to this effect as *spread*.

$$P_i(\theta_s)=\frac{1}{2}\left[1+\cos\left(\theta_i-\theta_s\right)\right]^{\alpha} \qquad (9)$$

Because the runtime complexity of SPCAP is determined by the number of speakers (i.e. complexity O(n)), it is not a function of desired source width. Therefore, SPCAP also overcomes problem *(iii)*, as listed in Section 2.

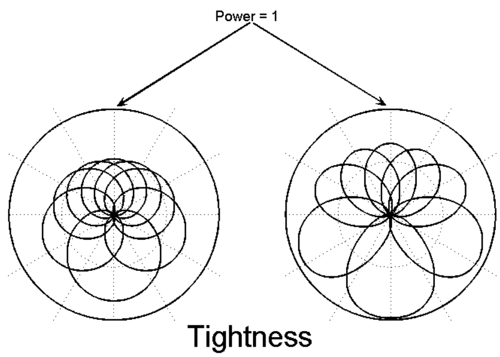

Figure 2 : Gains with moderate and high values for α

Using α as a function of distance has the additional property that as sounds move away from the loudspeakers, they become closer to point-sources [6]. Figures 2 and 3 illustrate the effect of raising α .

The flexibility of the SPCAP method provides a framework for our ongoing research into rendering truly *wide* sounds. Like VBAP, our method splits an input signal into an arbitrary number of output signals. While this approach creates a vague sense of *spread* or *tightness*, it does not capture *width* because the signals are correlated.

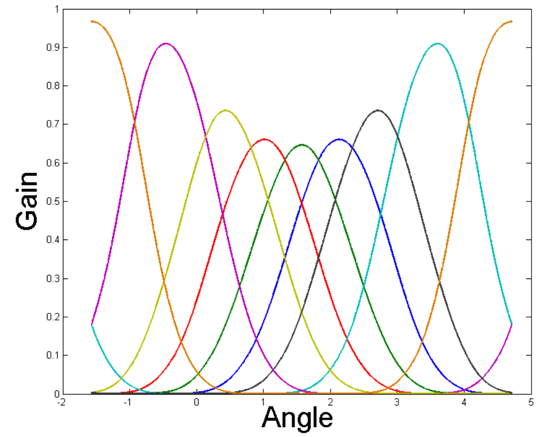

Figure 3: 8-channel Gains with moderate α value.

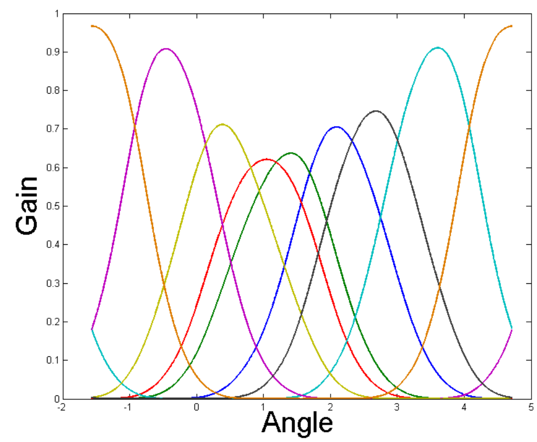

Figure 4 : 8-channel gains with center channel displaced 15 degrees.

However, the values $\{R_i\}$ in SPCAP provide insight into how much each channel contributes to the output. This information could be used to inform a de-correlation process which will properly widen the sound. Determining these processes is the aim of our current research.

The SPCAP approach has several additional advantages over discrete panning schemes. First, this method provides a framework in which content distribution is abstracted from delivery mechanism. That is, pan values in SPCAP are extracted from 3D virtual source positions, which do not depend on the configuration of the playback system. Therefore, sessions created using one setup map nicely to another – possibly very different – setup. Content mixed on an ITU setup, for example, will mix-down to 4.1 or two-channel system in a reasonable way. Conversely, SPCAP will up-mix a session arranged on a 4.1 system in order to exploit the additional channels of a high-resolution setup (e.g. 10.2). In particular, SPCAP gracefully handles non-standard setups because it performs the gain determination and power-conservation steps on a per-speaker basis. For example, if logistical considerations demand a setup with no center channel, SPCAP will compensate, diverting power to the nearest speakers. Figure 4 shows the gains for an 8-channel setup where the center channel has been displaced by 15 degrees.

Note that the SPCAP method does not place restrictions on the criteria for assigning initial gains or effective weights. The only requirements are as follows:

a) $0 \leq \frac{P_i}{\beta_i} \leq 1$,

b) this ratio varies smoothly and continuously as the source rotates.

In this formulation, we use cosines in equations (1) and (2) because we desire global support to facilitate our research on rendering *wide* sounds. Applications that do not desire global support could replace (1) with a limited-support function. Similarly, applications that seek to use panning criteria other than angle (taking into account psycho-acoustic phenomena, for example) should replace the weighting function in (2) with a more suitable function.

5. CONCLUSION

We have presented a method called SPCAP for dynamic pan-pot generation. While SPCAP takes inspiration from VBAP [3] it offers a number of significant contributions. First, it guarantees conservation of power, independent of the playback setup. This feature allows for dynamic re-mixing of source material. Second, it offers a framework for our research in rendering *wide* sounds. Finally, SPCAP is computationally inexpensive. For these reasons SPCAP has many applications in areas as diverse as content distribution, long-distance collaborative music, or computer games. The method is particularly well suited to tasks in which numerous mono-sounds are to be mixed dynamically (e.g. videogames and computer music) or to create and deliver content to venues with non-standard audio setups.

6. ACKNOWLEDGEMENTS

This paper was developed with funds of the Department of the Army under contract number DAAD 19-99-D-0046. Any opinions, findings and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the Department of the Army.

Thanks are due to Tomlinson Holman for his assistance in this project and to Johnny Li for the software [6] which inspired several ideas in this work. Aaron Isaksen and Tim Hawkins have been vital in ensuring correctness of the work presented here. Additional thanks are due to John DeWeese, Kumar Iyer and Farah Kidwai for extensive proof-reading, editing and support.

7. REFERENCES

[1] Craven, Peter G., "Continuous Surround Panning for 5-Speaker Reproduction", AES 24th International Conference on Multichannel Audio, Banff, June 2003.

[2] Gerzon, M.A., "Panpot Laws for Multispeaker Stereo", AES 92nd convention, March 1992, Vienna: AES preprint no. 3309.

[3] V. Pulkki, "Virtual Source Positioning Using Vector Base Amplitude Panning", J. Audio Eng. Soc., vol. 45, no. 6, pp. 456-466, 1997 June.

[4] West, J.R., "Five-Channel Panning Laws: An Analytical and Experimental Comparison", Master's Thesis, Music Engineering, University of Miami, 1998.

[5] Torick, E., "Highlights in the History of Multichannel Sound", J. Audio Eng. Soc., vol. 46, no.1/2, 1998 January/February.

[6] Li, J., "JLAPan", ProTools Plug-in software source code, 2001.

[7] Crocker, M. J., Ed. *Handbook of Acoustics*, Wiley-Interscience, 1998.