SPACE WARPING BASED DIMENSIONALITY REDUCTION OF HIGHER ORDER AMBISONICS SIGNALS

Maximilian Kentgens, Peter Jax

Institute of Communication Systems (IKS), RWTH Aachen University, Germany
{*kentgens, jax*}@*iks.rwth-aachen.de*

ABSTRACT

A novel approach to non-uniform spatial resolution encoding and dimensionality reduction of Higher Order Ambisonics spatial audio content is proposed. Space warping is applied to modify the spatial arrangement of the audio scene such that the coefficient energy is packed in few low order coefficients. The underlying idea is to squeeze or stretch parts of the spherical coordinates. In contrast to conventional Mixed Order Ambisonics, our approach is based on a transformation. Unlike methods for dimensionality reduction such as PCA, the resulting signal remains a valid HOA signal and spatial relations are maintained. We present a mathematical framework for space warping, demonstrate the potential of the method using an example, and discuss prerequisites and applications.

***Index Terms*—** Higher Order Ambisonics, spherical harmonics, space warping, non-uniform spatial resolution, dimensionality reduction

1. INTRODUCTION

The Higher Order Ambisonics (HOA) approach of representing sound fields using a truncated spherical harmonics (SH) decomposition [1, 2] provides a powerful mathematical framework which facilitates recording [3, 4], representation [5], analysis [6, 7, 8, 9], modification [10, 11, 12], and reproduction with loudspeakers [13, 14] or headphones [15, 16] for various applications such as multimedia, telepresence [17], and virtual reality [18, 19]. Due to the rotational invariance of SH signals [20], the spatial resolution is independent of the direction of the incident sound waves which is a strength and a weakness at the same time. The human auditory system has better resolution in the horizontal plane than from elevated directions [21], and often sound waves mainly impinge from predominant directions. Direction dependent resolution can be achieved with *Mixed Order Ambisonics* (MOA) schemes [1, 22, 23, 24] by selecting a certain subset of SH coefficients.

The proposed method to obtain a reduced number of SH coefficients and enabling direction dependent spatial resolution differs from existing MOA approaches. It is based on the modification of the spatial arrangement of the audio scene by means of *space warping* [10, 11]. Its goal is to shift the energy towards lower order coefficients. Unlike conventional transform coding methods such as PCA [25], the transformed signal is still a valid SH signal but can be further truncated without significant loss of information. Thus, it can be processed with existing compression techniques for HOA. Spatial modification approaches are frequently based on the decomposition of the audio scene through statistical analysis, e.g., DirAC [26], eigenvalue decomposition-based COMPASS [12], independent component analysis [27], or compressed sensing [28]. Contrary to this, space warping is not necessarily signal adaptive. It consists of a linear operation that can be directly applied on the HOA signal vector. The underlying idea is to squeeze or stretch parts of the spherical coordinate system's axes. Space warping was independently proposed in [10] and [11]. Our new approach shows parallels to the *warped discrete cosine transform* which was introduced for the compression of images [29].

Sec. 2 summarizes the mathematical foundations of spherical harmonics based sound field representation. In Sec. 3, space warping is recapitulated and an equalization function is derived. Sec. 4 presents the proposed method for dimensionality reduction. Sec. 5 demonstrates the functionality of the proposed system using an example. Secs. 6 and 7 provide a discussion and conclusion, respectively.

2. SPHERICAL HARMONICS BASED DESCRIPTION OF SOUND FIELDS

Using the inclination angle θ, the azimuth angle ϕ, the Euclidian distance from the origin r, and the wavenumber $k = \frac{\omega}{c}$ with angular frequency ω and the speed of sound c, the pressure of a sound field $p(k, \theta, \phi, r)$ in the vicinity of the coordinate system's origin can be approximated with the series expansion [30]

$$p(k, \theta, \phi, r) \approx \sum_{n=0}^{N} \sum_{m=-n}^{n} 4\pi \mathrm{i}^n j_n(kr) a_{nm}(k) Y_n^m(\theta, \phi). \qquad (1)$$

The choice of truncation order N and wavenumber k (i.e., the frequency) determine the region where the approximation is accurate. Here, $j_n(\cdot)$ is the spherical Bessel function of the first kind, i is the imaginary unit, and $Y_n^m(\theta, \phi)$ are the complex-valued spherical harmonics functions[1]

$$Y_n^m(\theta, \phi) = \sqrt{\frac{2n+1}{4\pi} \frac{(n-m)!}{(n+m)!}} P_n^m(\cos\theta) \mathrm{e}^{\mathrm{i}m\phi} \qquad (2)$$

with the associated Legendre functions $P_n^m(\cdot)$ and factorial $(\cdot)!$. The actual frequency domain (or strictly speaking: wave domain) coefficients $a_{nm}(k)$ with order $n = 0, \ldots, N$ and degree $m = -n, \ldots, n$

[1] All considerations presented here are also valid for the real-valued N3D normalized SH definition commonly used in the Ambisonics community.

Supported by:

Federal Ministry for Economic Affairs and Energy

on the basis of a decision by the German Bundestag

This research was funded by the ZIM program of the German Federal Ministry for Economic Affairs and Energy under Grant No. ZF 4340601MS6.

do not depend on r but only on the wavenumber k. They can also be expressed in time domain, i.e., $a_{nm}(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} a_{nm}(k = \frac{\omega}{c})\mathrm{e}^{\mathrm{i}\omega t}\,\mathrm{d}\omega$, and are commonly known as *Higher Order Ambisonics* signals. In the following, a simplified notation is being used neglecting temporal and frequency dependencies, i.e., t and k are omitted.

The coefficients a_{nm} stand in a direct relation to the directional amplitude density $a(\theta,\phi)$ of the continuum of plane waves impinging into the coordinate system's origin [30],

$$a(\theta,\phi) \approx \sum_{n=0}^{N}\sum_{m=-n}^{n} a_{nm} Y_n^m(\theta,\phi). \tag{3}$$

The inverse relation is given by

$$a_{nm} = \int_{\mathbb{S}^2} a(\theta,\phi)\,[Y_n^m(\theta,\phi)]^*\,\mathrm{d}A, \tag{4}$$

where $\mathbb{S}^2$ denotes the surface of the unit sphere with the corresponding surface element $\mathrm{d}A = \sin\theta\,\mathrm{d}\theta\mathrm{d}\phi$. Eqs. (3) and (4) are sometimes referred to as the *inverse spherical Fourier transform* and the *spherical Fourier transform*, respectively.

3. SPACE WARPING

3.1. Mathematical Description

We aim at obtaining a transformed signal $\tilde{a}_{nm}$ which is a spatially modified version of a_{nm}. The warping operation is defined using the spatial domain signal $a(\theta,\phi)$ given in Eq. (3). In the scope of this paper and without loss of generality, we focus on rotation symmetric space warping operations. Due to the rotational invariance of a_{nm} (cf. [20]) we can then consider warping solely along the inclination angle θ.

The transformed signal $\tilde{a}_{nm}$ is obtained via

$$\tilde{a}_{nm} = \int_{\mathbb{S}^2} g(\theta) a(f(\theta),\phi)\,[Y_n^m(\theta,\phi)]^*\,\mathrm{d}A \tag{5}$$

with an arbitrary, strictly monotonous increasing space warping function $\tilde{\theta} = f(\theta)$ with $\theta,\tilde{\theta} \in [0,\pi]$, and a corresponding equalization function $g(\theta)$. The basic approach is borrowed from Pomberger and Zotter [10].

In contrast to [10], we derive a *generalized* equalization function which is valid for arbitrary warping functions $f(\theta)$. In order to preserve the energy of the warped signal, the equalization function is constrained accordingly, i.e.,

$$\int_{\mathbb{S}^2} |a(\theta,\phi)|^2\,\mathrm{d}A \overset{!}{=} \int_{\mathbb{S}^2} |g(\theta)a(f(\theta),\phi)|^2\,\mathrm{d}A. \tag{6}$$

Using the substitution method [31, Ch. 8.1.2], the right-hand side can be expressed as

$$\int_{\mathbb{S}^2} \frac{g^2(f^{-1}(\tilde{\theta}))}{f'(f^{-1}(\tilde{\theta}))}\frac{\sin f^{-1}(\tilde{\theta})}{\sin\tilde{\theta}}|a(\tilde{\theta},\phi)|^2\,\mathrm{d}\tilde{A}$$

with surface element $\mathrm{d}\tilde{A} = \sin\tilde{\theta}\,\mathrm{d}\tilde{\theta}\mathrm{d}\phi$. Here, $f^{-1}(\tilde{\theta})$ is the inverse function of $f(\theta)$, and $f'(\theta) = \frac{\mathrm{d}f(\theta)}{\mathrm{d}\theta}$ is its derivative. Since the original directional power shall be retained, the integrand is compared to the integrand of the left-hand side of Eq. (6) resulting in

$$g(\theta) = \sqrt{f'(\theta)\frac{\sin f(\theta)}{\sin\theta}}. \tag{7}$$

3.2. Space Warping for Order Limited Signals with Integral Approximation

Using a set of Q quasi-uniformly distributed sampling points on the sphere[2] $\Xi = \{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_Q\}$ where $\boldsymbol{\theta}_q = (\theta_q, \phi_q)$, Eq. (5) can be approximated by

$$\tilde{a}_{nm} \approx \frac{4\pi}{Q}\sum_{q=1}^{Q} g(\theta_q) a(f(\theta_q),\phi_q)\,[Y_n^m(\theta_q,\phi_q)]^*. \tag{8}$$

In the sequel, only coefficients $\tilde{a}_{nm}$ up to a truncation order $\tilde{N}$ are considered. These coefficients can be written in vector notation by introducing the index $\chi = 1,2,\ldots,(\tilde{N}+1)^2$ which is used to arrange the coefficients in a sequential order via $\chi = n^2+n+m+1$. Applying the short notation $\tilde{a}_\chi := \tilde{a}_{nm}$, we introduce

$$\tilde{\boldsymbol{a}}_{\boldsymbol{nm}} = \left(\tilde{a}_1, \tilde{a}_2 \ldots, \tilde{a}_{(\tilde{N}+1)^2}\right)^{\mathrm{T}} \tag{9}$$

where $(\cdot)^{\mathrm{T}}$ denotes the vector transpose. Similarly, the original signal coefficients a_{nm} are stacked in a vector $\boldsymbol{a}_{\boldsymbol{nm}}$ of length $(N+1)^2$. We can then express Eq. (8) as

$$\tilde{\boldsymbol{a}}_{\boldsymbol{nm}} = \frac{4\pi}{Q}\boldsymbol{Y}_{\Xi,\tilde{N}}^{\mathrm{H}}\,\mathrm{diag}\{\boldsymbol{g}\}\,\underbrace{\boldsymbol{Y}_{\tilde{\Xi},N}\boldsymbol{a}_{\boldsymbol{nm}}}_{=\boldsymbol{a}} \tag{10}$$

$$= \boldsymbol{W}\boldsymbol{a}_{\boldsymbol{nm}}, \tag{11}$$

where $(\cdot)^{\mathrm{H}}$ is the conjugate transpose of a matrix, $\mathrm{diag}(\cdot)$ denotes a diagonal matrix with the vector elements on its diagonal, and $\boldsymbol{a}$ contains the values of Eq. (3) evaluated at the warped sampling points $\tilde{\Xi} = \{\tilde{\boldsymbol{\theta}}_1, \tilde{\boldsymbol{\theta}}_2, \ldots, \tilde{\boldsymbol{\theta}}_Q\}$ with $\tilde{\boldsymbol{\theta}}_q = (f(\theta_q), \phi_q)$. The equalization weights are aggregated into vector $\boldsymbol{g} = (g(\theta_1), \ldots, g(\theta_Q))^{\mathrm{T}}$. Matrix $\boldsymbol{Y}_{\Xi,\tilde{N}}$ with dimension $Q \times (\tilde{N}+1)^2$ is defined as

$$\boldsymbol{Y}_{\Xi,\tilde{N}} = \begin{pmatrix} Y_0^0(\boldsymbol{\theta_1}) & Y_1^{-1}(\boldsymbol{\theta_1}) & Y_1^0(\boldsymbol{\theta_1}) & \cdots & Y_N^N(\boldsymbol{\theta_1}) \\ Y_0^0(\boldsymbol{\theta_2}) & Y_1^{-1}(\boldsymbol{\theta_2}) & Y_1^0(\boldsymbol{\theta_2}) & \cdots & Y_N^N(\boldsymbol{\theta_2}) \\ \vdots & \vdots & \vdots & & \vdots \\ Y_0^0(\boldsymbol{\theta_Q}) & Y_1^{-1}(\boldsymbol{\theta_Q}) & Y_1^0(\boldsymbol{\theta_Q}) & \cdots & Y_N^N(\boldsymbol{\theta_Q}) \end{pmatrix},$$

and $\boldsymbol{Y}_{\tilde{\Xi},N}$ with dimension $Q \times (N+1)^2$ is defined accordingly. By combining the matrix factors of Eq. (10), the space warping operation can be expressed as a single matrix $\boldsymbol{W}$. Matrix $\boldsymbol{W}$ is referred to as the space warping matrix.

4. SPACE WARPING BASED DIMENSIONALITY REDUCTION

The effect of the space warping operation is comparable to an acoustic zoom where certain regions of the $\mathbb{S}^2$ signal are enlarged at the expense of other regions which shrink in size. As a result, the energy of signal components in the enlarged regions moves towards lower SH coefficients while higher order components are needed to represent the fine spatial structure in the squeezed regions. Given a finite $\tilde{N}$, the warped signal $\tilde{\boldsymbol{a}}_{\boldsymbol{nm}}$ may be seen as a non-uniform angular resolution representation of $\boldsymbol{a}_{\boldsymbol{nm}}$ which compared to the previously proposed methods, cf. [1, 22, 23, 24], is a fundamentally different approach to Mixed Order Ambisonics.

Given that directions with dominant signal components in a HOA signal can be identified—either by online adaptation or with

[2] These points may be designed using spherical design methods such as the ones presented in [32], [33], or [34].

a priori assumptions on the signal—space warping can be used to concentrate energy in few low order coefficients. This can be seen as a dimensionality reduction approach as typically used in transform coding, cf. [25]. The choice of $\tilde{N} < N$ is a trade-off between distorting the signal and achieving appropriate dimensionality reduction. The tremendous advantage of the proposed method is that the transform domain signal is still a valid HOA signal which can be further treated with existing methods and tools.

Classical transform coding mechanisms are often based on orthonormal transformations. This means that they are energy preserving, and their inverses are given by their matrix Hermitians [25]. Space warping is not order-preserving. In general, the operation excites SH modes higher than the input truncation order N. Due to the limitation to a finite output order $\tilde{N}$, we sacrifice semi-orthogonality. As a result, space warping is a leaky transform. The better the warping function is adapted to the signal, i.e., few energy impinges from regions to be squeezed, the fewer higher modes are excited.

The "unwarping" of $\tilde{\boldsymbol{a}}_{\boldsymbol{nm}}$ for the case $\tilde{N} < N$, i.e., the reverse operation to Eq. (10), can be realized with the Moore–Penrose pseudoinverse,

$$\boldsymbol{W}_{\text{inv}} = \boldsymbol{W}^{\text{H}}(\boldsymbol{W}\boldsymbol{W}^{\text{H}})^{-1}. \tag{12}$$

Thus, the reconstructed signal

$$\hat{\boldsymbol{a}}_{\boldsymbol{nm}} = \boldsymbol{W}_{\text{inv}}\tilde{\boldsymbol{a}}_{\boldsymbol{nm}} \tag{13}$$

is an approximation of $\boldsymbol{a}_{\boldsymbol{nm}}$. Here, $(\cdot)^{-1}$ denotes the matrix inverse.

5. SIMULATIONS & RESULTS

An exemplary warping function (cf. [10])

$$f(\theta) = \cos^{-1}\left(\frac{\alpha + \cos\theta}{1 + \alpha\cos\theta}\right) \tag{14}$$

with warping factor $\alpha \in (-1, 1)$ is illustrated in Fig. 1. In this case, Eq. (7) is simplified to

$$g(\theta) = \frac{\sqrt{1-\alpha^2}}{1+\alpha\cos\theta}, \tag{15}$$

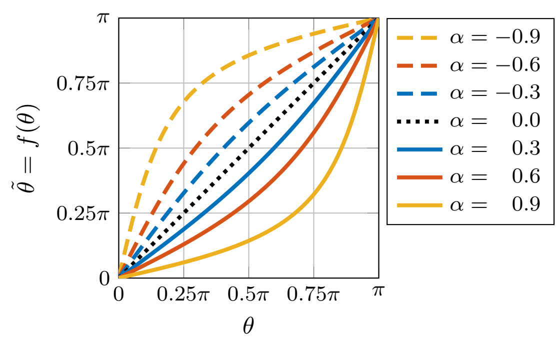


Fig. 1: Exemplary space warping function from Eq. (14) with variants of the warping factor α.

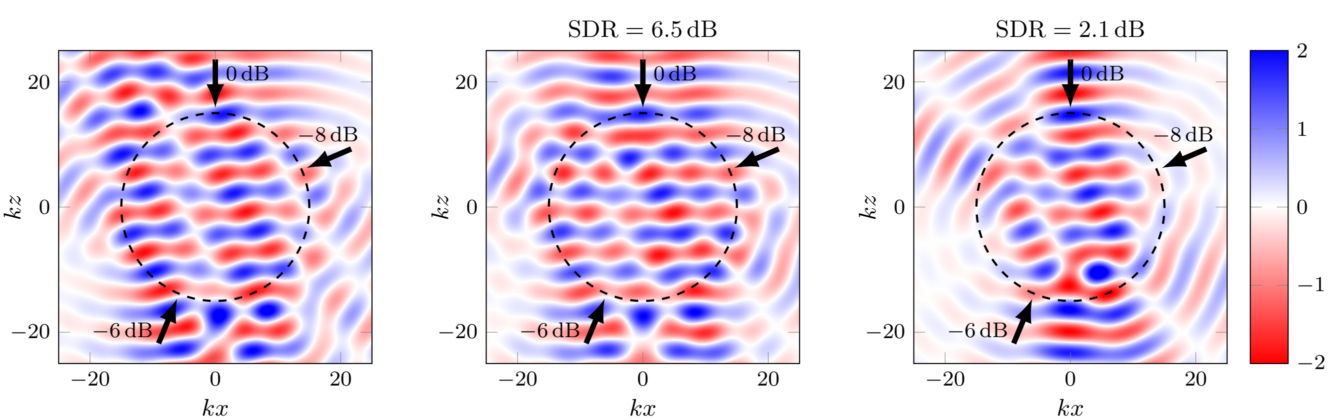
which is exactly the derivative of $f(\theta)$, i.e., $g(\theta) = f'(\theta)$ (again, cf. [10]). The effect of this operation is that for $\alpha > 0$ the region around $\theta = 0$ is enlarged while the region around $\theta = \pi$ is squeezed, and vice versa for $\alpha < 0$.

A simulation with a dominant plane wave impinging from $(\theta, \phi) = (0, 0)$ and two attenuated waves from other directions, is used to show the effectiveness of the proposed method. The three waves are sinusoidal with arbitrary phases[3]. They are encoded in a HOA representation $\boldsymbol{a}_{\boldsymbol{nm}}$ with $N = 15$. The corresponding sound field is illustrated in Fig. 2a. In region $kr < N$ the order $N = 15$ truncated HOA representation of the generated sound field introduces a small error [30]. This region is shown with a dashed circle. The warping function from Eq. (14) is applied. Warping factor α and output order $\tilde{N}$ are varied. A quasi-uniform sampling grid with $Q = 5200$ sampling points was chosen.

The performance is assessed with two measures. The first one is the energy ratio

$$\xi = \frac{\|\tilde{\boldsymbol{a}}_{\boldsymbol{nm}}\|^2}{\|\boldsymbol{a}_{\boldsymbol{nm}}\|^2} \tag{16}$$

[3]Here, the wave number k is irrelevant since the following results are given relative to k.

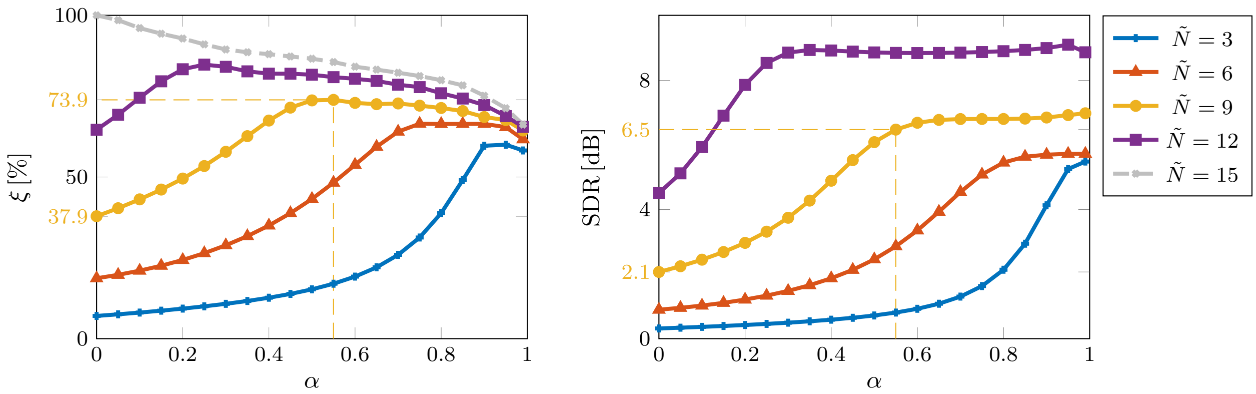


(a) Original signal $\boldsymbol{a}_{\boldsymbol{nm}}$ with $N = 15$

(b) From order $\tilde{N} = 9$ reconstructed signal $\hat{\boldsymbol{a}}_{\boldsymbol{nm}}$ with proposed method

(c) Signal $\boldsymbol{a}_{\boldsymbol{nm}}$ directly truncated to order $\tilde{N} = 9$

Fig. 2: xz-planes of sound field snap-shots corresponding to different HOA signals. All HOA signals represent three plane waves impinging from the indicated directions with amplitudes as labeled.



(a) Energy ratio ξ

(b) Signal-to-distortion ratio SDR

Fig. 3: Simulation results of proposed method for an input signal of order $N = 15$ consisting of three plane waves as indicated in Fig. 2.

which is a measure of how much energy is maintained after warping and order reduction. The second one is the signal-to-distortion ratio

$$\mathrm{SDR} = 10\log_{10}\left(\frac{\int_{\mathbb{S}^2} |a(\theta,\phi)|^2 \,\mathrm{d}A}{\int_{\mathbb{S}^2} |\hat{a}(\theta,\phi) - a(\theta,\phi)|^2 \,\mathrm{d}A}\right) \mathrm{dB} \qquad (17)$$

after reconstruction. Using Parseval's relation [30], the SDR can be expressed as

$$\mathrm{SDR} = 10\log_{10}\left(\frac{\|\boldsymbol{a}_{nm}\|^2}{\|\hat{\boldsymbol{a}}_{nm} - \boldsymbol{a}_{nm}\|^2}\right) \mathrm{dB}. \qquad (18)$$

The simulation results are illustrated in Fig. 3. The warping factor $\alpha = 0$ is equivalent to a direct order truncation of $\boldsymbol{a}_{nm}$. When reducing $\boldsymbol{a}_{nm}$ to order $\tilde{N} = 9$, i.e., retaining $(\tilde{N}+1)^2 = 100$ of the $(N+1)^2 = 256$ coefficients, only $37.9\,\%$ of the original energy is maintained. The distortion is only $2.1\,\mathrm{dB}$ higher than the signal energy. The effect is a significant decrease of the sweet spot size as evident from Fig. 2c. With the proposed method and a warping factor $\alpha = 0.55$, the energy is shifted towards lower coefficients so that $73.9\,\%$ of the original energy is maintained. The SDR of the reconstructed signal $\hat{\boldsymbol{a}}_{nm}$ increases to $6.5\,\mathrm{dB}$. After reconstruction, the sweet spot size is well preserved, see Fig. 2b. The dominant signal component impinging from $\theta = 0$ is well reconstructed at the expense of signal components from $\theta \approx \pi$, which are significantly distorted.

At a certain point, the SDR starts to saturate. An increase of α does not result in an increase of the SDR anymore. The component at $\theta = 0$ gets less disorted and the other components more. The saturation depends on the directional energy distribution of the original signal $\boldsymbol{a}_{nm}$.

The case $N = \tilde{N} = 15$ in Fig. 3a (grey line) shows that a certain energy leakage is inevitable with the proposed method. This is due to the non-orthogonality of the space warping transform as explained in Sec. 4. However,in this case, the SDR is high enough to exceed the limits in plot Fig. 3b.

6. DISCUSSION

The performance of the proposed method highly depends on the signals' characteristics. This cannot be shown further due to the limited space of this paper. When sound impinges from a limited predominant angular region and squeezed regions are sparse in energy or perceptually less important, the proposed method seems promising. This applies to many practical scenarios since high energy direct sound often impinges from $\theta \approx \pi/2$, or even from few frontal directions only. Other warping characteristics and the abolition of the rotational symmetry constraint might be necessary, cf. [35].

Applications for the proposed method are twofold. Firstly, non-uniform resolution encoding might be necessary in cases where non-uniformly distributed content is acquired, e.g., [24]. Secondly, it may be applied in transform coding-based compression of HOA. In the latter case, space warping may be used as a preprocessing step. Unlike the statistically optimal Karhunen-Loève transform (also known as principal component analysis, PCA) [25], side-information only consists of the parametric warping characteristics and the resulting signal maintains spatial relations. Thus, further compression that exploits both temporal and spatial redundancy is possible. Future research is needed for formal perceptual validation, investigations on the impact of quantization, and information theoretic considerations from a signal compression perspective.

7. CONCLUSION

A novel approach to non-uniform spatial resolution encoding and dimensionality reduction of HOA based on space warping was proposed. Contrary to conventional Mixed Order Ambisonics, our approach is based on a transformation. Unlike other methods for dimensionality reduction, the resulting signal remains a valid HOA signal and spatial relations are maintained. We presented a mathematical framework for space warping and derived a novel equalization function. Using an exemplary warping function, we performed simulations. Based on this simple example it is obivous that the signal to distortion ratio of HOA reconstruction can be substantially optimized by choosing a signal-dependent warping function.

8. REFERENCES

[1] Jérôme Daniel, *Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia*, Ph.D. thesis, Université Paris 6, 2001.

[2] Darren B. Ward and Thushara D. Abhayapala, "Reproduction of a plane-wave sound field using an array of loudspeakers,"

IEEE Transactions on Speech and Audio Processing, vol. 9, no. 6, pp. 697–707, 2001.

[3] Jens Meyer and Gary Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2002.

[4] Thushara D. Abhayapala and Darren B. Ward, "Theory and design of high order sound field microphones using spherical microphone array," in *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, may 2002.

[5] Jérôme Daniel, Rozenn Nicol, and Sébastien Moreau, "Further investigations of high order ambisonic and wavefield synthesis for holophonic sound imaging," in *AES 114th Convention*, 2003.

[6] H. Teutsch and W. Kellermann, "Detection and localization of multiple wideband acoustic sources based on wavefield decomposition using spherical apertures," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2008.

[7] Tahereh Noohi, Nicolas Epain, and Craig T. Jin, "Direction of arrival estimation for spherical microphone arrays by combination of independent component analysis and sparse recovery," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2013.

[8] Joonas Nikunen and Tuomas Virtanen, "Direction of arrival based spatial covariance model for blind sound source separation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 3, 2014.

[9] Daniel P. Jarrett, Emanuel A. P. Habets, and Patrick A. Naylor, *Theory and Applications of Spherical Microphone Array Processing*, Springer, 2017.

[10] Hannes Pomberger and Franz Zotter, "Warping of 3D Ambisonic recordings," in *3rd International Symposium on Ambisonics and Spherical Acoustics*, 2011.

[11] Peter Jax and Johann-Markus Batke, "Method and apparatus for changing the relative positions of sound objects contained within a Higher-order-Amibsonics representation," 2012, US Patent 9338574 B2.

[12] Archontis Politis, Sakari Tervo, and Ville Pulkki, "COMPASS: Coding and multidirectional parameterization of ambisonic sound scenes," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018.

[13] Franz Zotter and Matthias Frank, "All-round ambisonic panning and decoding," *Journal of the Audio Engineering Society*, vol. 60, no. 10, 2012.

[14] W. Bastiaan Kleijn, Andrew Allen, Jan Skoglund, and Felicia Lim, "Incoherent idempotent ambisonics rendering," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, 2017.

[15] Markus Noisternig, Alois Sontacchi, Thomas Musil, and Robert Holdrich, "A 3d ambisonic based binaural sound reproduction system," in *AES 24th International Conference on Multichannel Audio*, 2003.

[16] Gerald Enzner, Michael Weinert, Stefan Abeling, Jan-Mark Batke, and Peter Jax, "Advanced system options for binaural rendering of ambisonic format," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2013.

[17] Jeremy R. Cooperstock, "Multimodal telepresence systems," *IEEE Signal Processing Magazine*, vol. 28, no. 1, 2011.

[18] Shankar Shivappa, Martin Morrell, Deep Sen, Nils Peters, and S. M. Akramus Salehin, "Efficient, compelling, and immersive vr audio experience using scene based audio/higher order ambisonics," in *AES International Conference on Audio for Virtual and Augmented Reality (AVAR)*, 2016.

[19] Maximilian Kentgens, Stefan Kühl, Christiane Antweiler, and Peter Jax, "From spatial recording to immersive reproduction - design & implementation of a 3DOF audio-visual VR system," in *AES 145th Convention*, 2018.

[20] Peter J. Kostelec and Daniel N. Rockmore, "FFTs on the rotation group," *Journal of Fourier Analysis and Applications*, vol. 14, no. 2, pp. 145–179, feb 2008.

[21] Jens Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization*, MIT Press, 1997.

[22] Chris Travis, "A new mixed-order scheme for ambisonic signals," in *Ambisonics Symposium Graz*, 2009.

[23] Sylvain Favrot, Marton Marschall, Johannes Käsbach, Jörg Buchholz, and Tobias Weller, "Mixed-order ambisonics recording and playback for improving horizontal directionality," in *AES 131st Convention*, 2011.

[24] Jiho Chang and Marton Marschall, "Periphony-lattice Mixed-order Ambisonic scheme for spherical microphone arrays," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, , no. 99, 2018.

[25] Jens-Rainer Ohm, *Multimedia Signal Coding and Transmission*, Springer, Heidelberg/Berlin, 2015.

[26] Ville Pulkki, "Spatial sound reproduction with directional audio coding," *Journal of the Audio Engineering Society*, vol. 55, no. 6, 2007.

[27] Nicolas Epain and Craig T. Jin, "Independent component analysis using spherical microphone arrays," *Acta Acustica united with Acustica*, vol. 98, no. 1, 2012.

[28] Andrew Wabnitz, Nicolas Epain, André van Schaik, and Craig Jin, "Time domain reconstruction of spatial sound fields using compressed sensing," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011.

[29] Nam Ik Cho and S.K. Mitra, "Warped discrete cosine transform and its application in image compression," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 8, pp. 1364–1373, 2000.

[30] Boaz Rafaely, *Fundamentals of Spherical Array Processing*, Springer, 2015.

[31] I. N. Bronshtein, K. A. Semendyayev, Gerhard Musiol, and Heiner Mühlig, *Handbook of Mathematics*, Springer, 5th edition, 2007.

[32] R. H. Hardin and Neil J. A. Sloane, "McLaren's improved snub cube and other new spherical designs in three dimensions," *Discrete and Computational Geometry*, vol. 15, no. 4, 1996.

[33] Jörg Fliege and Ulrike Maier, "The distribution of points on the sphere and corresponding cubature formulae," *IMA Journal of Numerical Analysis*, vol. 19, no. 2, 1999.

[34] Manuel Gräf and Daniel Potts, "On the computation of spherical designs by a new optimization approach based on fast spherical fourier transforms," *Journal Numerische Mathematik*, vol. 119, no. 4, 2011.

[35] Matthias Kronlacher and Franz Zotter, "Spatial transformations for the enhancement of ambisonic recordings," in *2nd International Conference of Spatial Audio (ICSA)*, 2014.