Enhanced sound field reproduction within prioritized control region

Hanchi Chen; Thushara D. Abhayapala; Wen Zhang

Australian National University, Australia

ABSTRACT

Higher-order ambisonics has been identified as a robust technique for synthesizing a desired sound field. However, the synthesis algorithm requires a large number of secondary sources to derive the optimal results for large reproduction regions and over high operating frequencies. This paper proposes an enhanced method for synthesizing the sound field using a relatively small number of secondary sources which allows improved synthesizing accuracy for certain subregions of the interested zone. This method introduces the spherical harmonic translation into the mode matching algorithm to acquire a uniform modal-domain representation of the sound fields within different sub-regions. Then by changing the weighing of each region, the least mean squares solution can be easily controlled to cater for certain prioritized reproduction requirements. Simulations show that this technique can effectively improve the matching accuracy of a given sub-region, while only slightly increasing the global reproduction error. This method is shown to be especially effective in the situations where the number of secondary sources is limited.

1. INTRODUCTION

The technique of spatial sound field synthesis is used to reproduce a certain acoustic scene, such as the sound field produced by a series of virtual sources, within a region. Listeners inside the reproduction region will thus have the same listening experience as in the original acoustic scene. In order to achieve this, a series of loudspeakers are placed on the boundary or outside the reproduction region, and a series of carefully derived signals are played by these loudspeakers to produce the desired sound field. Ambisonics, originally proposed by (1), uses zero and first order decomposition of the desired sound field to calculate the driving signals. Due to its low order, the system only works well at low frequencies. For high frequencies as well as large reproduction regions, higher order ambisonics (HOA) based on 2D cylindrical harmonics (2, 3, 4) as well as 3D spherical harmonics (5, 6) decomposition of the sound field were developed. These techniques are also referred to as the "mode matching" approach.

Another popular approach is the Wave Field Synthesis method first proposed by (7, 8). This technique was developed based on the Kirchhoff-Helmholtz integral (9), and uses a plurality of monopole and dipole loudspeakers placed on the boundary of the interested region to control the sound field (10). A further development of this technique (called "2.5D WFS") uses linear and planar arrays to reproduce the sound field in a planar region (11, 12). Due to the use of point sources rather than cylindrical sources, an 2.5D operator was proposed for the technique (10, 13).

However, when it comes to 3D sound field synthesis, a fundamental problem arises which makes implementation very difficult: the synthesis quality is strongly related to the number and position of the loudspeakers (14, 15, 16). The ideal placement of the loudspeakers for the mode-matching technique is to have the loudspeakers evenly distributed on a sphere surrounding the interested region (5), such structure is impractical in reality. To solve this problem, an array configuration for 3D sound field synthesis using multiple circular loudspeaker arrays was proposed by Zhang and Abhayapala (17), this method uses a functional analysis based algorithm to derive the driving signals.

Still, the trade off between the number of the loudspeakers and the size and frequency of the reproduction zone exists. The reproduction quality degrades rapidly as the number of loudspeakers becomes less than the minimal required number. In the case that the interested region can be separated and reduced into a few smaller regions, it is possible to control the sound field in these small regions through spatial multizone reproduction techniques (18). However, the calculation involves matrix inversion, and if done without proper regularization, the results may be highly unstable.

The goal of this paper is to introduce a spatial single zone sound field reproduction technique which allows for higher reproduction accuracy within certain sub-zones. This can be achieved by balancing between the single zone reproduction and the spatial multizone reproduction techniques. Through the use of spherical harmonic translation, the mode-matching method can be simultaneously applied to both the global interested zone and certain sub-region within it (refered to as high priority regions), and by adjusting the weighing factors in the LMS solution, one can easily control the reproduction quality of different regions. This technique is particularly useful when an insufficient number of loudspeakers are available, and / or in applications where a high reproduction accuracy is required for certain sub-zones, such as active noise cancellation.

This paper is organized as follows: in Section 2, the spherical harmonic expansion technique is briefly explained, also introduced is the spherical harmonics translation theorem, which gives the relationship between spherical harmonics in different coordinate systems. In Section 3, the combined least-mean-squares solution to find the optimal loudspeaker driving signals for both the global region and its sub-regions is formulated, it is shown that by adjusting the weighing of each region, the relative reproduction accuracy of each region can be adjusted with ease. Finally, in Section 4, a series of simulation results demonstrate the performance of this method, it is shown that this method offers superior performance compared to the normal LMS method, especially with a relatively small amount of loudspeakers. The proposed method also provides a much better regularized solution compared to the multi-zone method. The relationship between the weighing factors and the reproduction accuracy in the associated regions is also shown in this section.

2. BASIC THEORIES

2.1 Spherical harmonics

This paper uses spherical harmonic decomposition to describe and analyze the sound field within an interested region of spherical shape. It is assumed that the interested region is a free space with no scatterers or sound sources inside. The sound waves propagating inside the region are due to sources outside the region. Defining a spherical coordinate with its origin located at the center of the sphere, the sound pressure at any point $x(R,\theta,\phi)$ within the sphere can be represented as a weighted sum of spherical harmonics,

$$P(R,\theta,\phi) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} C_{nm} j_n(kr) Y_{n|m|}(\theta,\phi), \tag{1}$$

where $k = 2\pi f/c$ is the wave number, f and c are the frequency and the propagation speed of the wave, respectively. C_{nm} are the spherical harmonic coefficients, j_n is the spherical Bessel function of order n, and Y_{nm} denotes the spherical harmonic of order n and mode m, written as

$$Y_{nm}(\theta,\phi) = (-1)^m \sqrt{\frac{(2n+1)}{4\pi}\frac{(n-m)!}{(n+m)!}} P_{nm}(\cos\theta) e^{im\phi}, \tag{2}$$

P_{nm} denotes the normalized associated Legendre Polynomial of order n and mode m.
It is clear that the sound field at frequency k is completely described by the coefficients C_{nm} which act as weighing factors for each spherical harmonics. Once said coeffcients are found, the sound field information of the interested region is known. Various methods have been proposed to acquire the sound field coefficients, including using spherical and non-spherical microphone arrays, however, this paper focuses on manipulation of the acquired sound field coefficients for synthesizing a certain desired sound field.

2.2 Spherical harmonics translation

Similar to coordinate system translation, a spherical harmonic with respect to origin O_p $Y_{nm}(\theta_p,\phi_p)$ can be translated to another coordinate system O, represented as a linear combination of a series of spherical harmonics respect to O. The mathematical representation of the translation is described in this section.

Consider a global origin O, and a local origin O_q which is located at $\mathbf{R}_q = (R_q,\theta_q,\phi_q)$ respect to O; a point is located at $\mathbf{R} = (R,\theta,\phi)$ respect to O, its location respect to O_q is $\mathbf{r} = (r,\theta_r,\phi_r)$. It is easy to identify the relationship $\mathbf{R} = \mathbf{r} + \mathbf{R}_q$. The addition theorem that relates the spherical harmonics respect to O and O_q is given by (19)

$$j_n(kR)Y_{nm}(\theta,\phi) = \sum_{\nu=0}^{\infty}\sum_{\mu=-\nu}^{\nu} \hat{S}_{n\nu}^{m\mu}(\mathbf{R}_q) j_\nu(kr) Y_{\nu\mu}(\theta_r,\phi_r), \tag{3}$$

where

$$\hat{S}_{n\nu}^{m\mu}(\mathbf{R}_q) = 4\pi i^{\nu-n}\sum_{l=0}^{\infty} i^l(-1)^{2m-\mu} j_l(kR_q) Y^*_{l(\mu-m)}(\theta_q,\phi_q)\sqrt{\frac{(2n+1)(2\nu+1)(2l+1)}{4\pi}} W_1 W_2 \tag{4}$$

Here, W_1 and W_2 denote Wigner 3-j symbols, with

$$W_1 = \begin{pmatrix} n & \nu & l \\ 0 & 0 & 0 \end{pmatrix} \tag{5}$$

and

$$W_2 = \begin{pmatrix} n & \nu & l \\ m & -\mu & \mu-m \end{pmatrix} \tag{6}$$

Here, due to inherent properties of the Wigner 3-j symbols, the infinite summation of l that appeared in Eq.(4) can be truncated to $l \le n+\nu+1$ (19).

Combining equation Eq.(1) and equation Eq.(3) yields a solution to translate between sound field coefficients of different coordinate systems, which is given by (20)

$$C^q_{\nu\mu} = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} C_{nm}\hat{S}_{n\nu}^{m\mu}(\mathbf{R}_q), \tag{7}$$

where $C^q_{\nu\mu}$ represents the coefficients respect to O_q and C_{nm} represents the coefficients respect to q. It can be seen easily that by substituting Eq.(7) into a slightly modified Eq.(1), one can derive the sound pressure representation of a given point respect to O_q using the spherical harmonic coefficients respect to O.

The double summation in Eq.(7) can be conveniently represented using a matrix equation, in the form of (20)

$$\mathbf{C}_q = \mathbf{T}_q\mathbf{C} \tag{8}$$

Here, $\mathbf{C}_q$ and $\mathbf{C}$ are the local and global coefficient vectors, arranged as $\mathbf{C} = \begin{bmatrix} C_{00} & C_{1-1} & C_{10} & \dots & C_{NN}\end{bmatrix}^T$ and $\mathbf{C}_q = \begin{bmatrix} C_{00} & C_{1-1} & C_{10} & \dots & C_{VV}\end{bmatrix}^T$. $\mathbf{T}_q$ is the translation matrix that maps the global coefficients to the local coordinate system q. $\mathbf{T}$ consists of all the $\hat{S}_{n\nu}^{m\mu}(\mathbf{R}_q)$ needed to translate $\mathbf{C}$ into $\mathbf{C}_q$, the orders of $\hat{S}_{n\nu}^{m\mu}(\mathbf{R}_q)$ are arranged in correspondence with $\mathbf{C}_q$ and $\mathbf{C}$, thus $\mathbf{T}$ can be written as (20)

$$\mathbf{T}_q = \begin{bmatrix} \hat{S}_{00}^{00} & \hat{S}_{10}^{(-1)0} & \hat{S}_{10}^{00} & \dots & \hat{S}_{N0}^{N0} \\ \hat{S}_{01}^{0(-1)} & \hat{S}_{11}^{(-1)(-1)} & \hat{S}_{11}^{0(-1)} & \dots & \hat{S}_{N1}^{N(-1)} \\ \hat{S}_{01}^{00} & \hat{S}_{11}^{(-1)0} & \hat{S}_{11}^{00} & \dots & \hat{S}_{N1}^{N0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{S}_{0V}^{0V} & \hat{S}_{1V}^{(-1)V} & \hat{S}_{1V}^{0V} & \dots & \hat{S}_{NV}^{NV} \end{bmatrix} \tag{9}$$

For a given maximum order N, there are a total number of $(N+1)^2$ spherical harmonics available, thus the size of $\mathbf{T}_q$ should be $(V+1)^2$ by $(N+1)^2$.

3. COMBINED LEAST MEAN SQUARE SOLUTION

A common way of deriving loudspeaker driving signals to produce a certain desired sound field is by pressure matching in the modal domain. Given N_l loudspeakers and N_y sound field coefficients, the channel information between the l^{th} loudspeaker and the y^{th} mode can be denoted H_{ly}, the channel matrix is then expressed as $\mathbf{H}$, where

$$\mathbf{H} = \begin{bmatrix} H_{11} & H_{21} & \dots & H_{N_l 1} \\ H_{12} & H_{22} & \dots & H_{N_l 2} \\ \vdots & \vdots & \ddots & \vdots \\ H_{1N_y} & H_{2N_y} & \dots & H_{N_l N_y} \end{bmatrix} \tag{10}$$

The sound field coefficients can be lined up in any order, for a 3D sound field with a maximum order of N, the total number of modes is given by $N_y = (N+1)^2$. for convenience, it is assumed that the sound field coefficients are arranged for $n = 0,1,2...N$ and $m = -n, -n+1...n$, thus the $y^t h$ coefficient satisfies $y = n^2 + n + m + 1$, for example H_{L1} represents the channel from the L^{th} speaker to coefficient C_{00}, and H_{L2} corresponds to the channel to coefficient $C_{1(-1)}$, etc.

A desired sound field of the same order can be expressed as a column vector of sound field coefficients

$$\mathbf{P} = \begin{bmatrix} P_{00} & P_{1(-1)} & P_{10} & \dots & P_{NN} \end{bmatrix}^T \tag{11}$$

The least mean square solution to find a series of driving signals $\mathbf{D}$ for the loudspeaker array to produce a sound field that resembles the desired sound field $\mathbf{P}$ can be found by solving for $\mathbf{D}$ from

$$\mathbf{P} = \mathbf{HD}, \tag{12}$$

which has the solution

$$\mathbf{D} = (\mathbf{H}^*\mathbf{H} + \lambda \mathbf{I})^{-1}\mathbf{H}^*\mathbf{P}, \tag{13}$$

where λ is a regularization parameter.
Since the sound field at a sub-region q of the interested zone, can be expressed using the global sound field coefficients $\mathbf{C}_G$ and a translation matrix $\mathbf{T}_Q$, as $\mathbf{C}_Q = \mathbf{T}_Q\mathbf{C}_G$. A LMS solution for synthesizing the desired sound field only within the sub-region Q can be found by solving

$$\mathbf{T}_Q\mathbf{P} = \mathbf{T}_Q\mathbf{HD} \tag{14}$$

Thus $\mathbf{D}$ can be expressed as

$$\mathbf{D} = [\mathbf{T}_Q\mathbf{H}]^{-1}\mathbf{T}_Q\mathbf{P} \tag{15}$$

Here, $[\mathbf{T}_Q\mathbf{H}]^{-1}$ denotes the Moore-Penrose pseudo inverse, which takes a similar form as in Eq.(13). Regularization may be required in the calculation.
It should be noted that the solution provided by Eq.(15) normally requires regularization, since the matrix $\mathbf{T}_Q\mathbf{H}$ is usually not well conditioned, which may result in very large driving signals for the loudspeakers. This is especially true when the sub-region Q is small.

The solution in Eq.(13) minimizes the cost function

$$L = [\mathbf{P} - \mathbf{HD}]^H[\mathbf{P} - \mathbf{HD}], \tag{16}$$

which is the sum of the squared errors in all spherical harmonic coefficients. Similarly, Eq.(15) minimizes the cost function

$$L = [\mathbf{T}_Q\mathbf{P} - \mathbf{T}_Q\mathbf{HD}]^H[\mathbf{T}_Q\mathbf{P} - \mathbf{T}_Q\mathbf{HD}], \tag{17}$$

which corresponds to the sum of the squared errors in all local spherical harmonic coefficients in the subregion Q.

In order to take both global and regional coefficient error into consideration, it is necessary to have a cost function that contains both Eq.(16) and Eq.(17). It is intuitive to write a new cost function that is simply the weighed sum of Eq.(16) and Eq.(17), as

$$L = \mathbf{E}_G^H\mathbf{E}_G + \alpha\mathbf{E}_Q^H\mathbf{E}_Q, \tag{18}$$

where $\mathbf{E}_G = [\mathbf{P} - \mathbf{HD}]^H[\mathbf{P} - \mathbf{HD}]$ is the global error vector, and $\mathbf{E}_Q = [\mathbf{T}_Q\mathbf{P} - \mathbf{T}_Q\mathbf{HD}]^H[\mathbf{T}_Q\mathbf{P} - \mathbf{T}_Q\mathbf{HD}]$ is the local error vector. α is the scaling factor which controls the relative importance of the reproduction accuracy at region q.

By combining Eq.(12) and Eq.(14), a LMS solution that minimizes Eq.(18) can be implemented

$$\begin{bmatrix} \alpha \mathbf{T}_Q \mathbf{P} \\ \mathbf{P} \end{bmatrix} = \begin{bmatrix} \alpha \mathbf{T}_Q \mathbf{H} \\ \mathbf{H} \end{bmatrix} \mathbf{D} \tag{19}$$

, whose solution for $\mathbf{D}$ is

$$\mathbf{D} = \begin{bmatrix} \alpha \mathbf{T}_Q \mathbf{H} \\ \mathbf{H} \end{bmatrix}^{-1} \begin{bmatrix} \alpha \mathbf{T}_Q \mathbf{P} \\ \mathbf{P} \end{bmatrix} \tag{20}$$

Intuitively, the solution Eq.(20) considers not only the set of global coefficients, but also a linear mapping of these coefficients which correspond to the sound field in a sub-region Q within the interested zone. By adding a weighing factor α, the priority of the sub-region can be controlled. When $\alpha = 0$, the sub-region is ignored and the solution becomes identical to Eq.(13); if $\alpha = 10$, the local reproduction accuracy becomes 10 times more significant than the global accuracy, and as a result the driving signals $\mathbf{D}$ would construct a sound field where the reproduction error within Q is 10 times less than the global average.

Obviously, Eq.(20) can be easily extended to the multiple sub-region case, with each region controlled by a separate weighing factor

$$\mathbf{D} = \begin{bmatrix} \alpha_1 \mathbf{T}_{Q1} \mathbf{H} \\ \alpha_2 \mathbf{T}_{Q2} \mathbf{H} \\ \vdots \\ \alpha_n \mathbf{T}_{Qn} \mathbf{H} \\ \beta \mathbf{H} \end{bmatrix}^{-1} \begin{bmatrix} \alpha_1 \mathbf{T}_{Q1} \mathbf{P} \\ \alpha_2 \mathbf{T}_{Q2} \mathbf{P} \\ \vdots \\ \alpha_n \mathbf{T}_{Qn} \mathbf{P} \\ \beta \mathbf{P} \end{bmatrix} \tag{21}$$

It can be seen that by setting $\alpha_1, \alpha_2 ... \alpha_n = 1$ and $\beta = 0$, Eq.(21) becomes a solution for spatial multizone sound field reproduction.

Compared to the spatial single zone sound field reproduction technique, this method allows more accurate reconstruction within some more critical areas; compared to the multizone reproduction method, which provides optimal reconstruction result within a few smaller regions but has no guarantee on the sound field in between these regions, this approach offers a balanced solution, where the the whole zone of interest is reproduced while a few sub-regions are given higher priorities. This method is also more stable and more predicable than the multi-zone method, since the global channel matrix $\mathbf{H}$ makes the matrix to be inversed in Eq.(21) well-conditioned, as a result, the derived driving signal $\mathbf{D}$ normally has limited power, and further regularization is normally unnecessary.

Another merit of this technique is when only a relatively small number of loudspeakers are accessible. Specifically, when the total number of loudspeakers is insufficient for producing the desired sound field for the whole interested zone, or the placement of the loudspeaker array disallows synthesizing sound waves impinging from certain directions, the normal LMS algorithm finds the best solution that gives the minimal average error across the whole zone. However, if the application has very high requirements on reproduction accuracy, for example active noise cancellation, this technique may fail to meet the demand, especially when insufficient loudspeakers are available. Using this new technique, it is possible to utilize the limited amount of loudspeakers to reproduce the sound field accurately in certain critical sub-zones, such as the area where the listener is most likely positioned, while only slightly degrading the performance in the rest of the zone.

4. SIMULATION RESULTS

The simulations are set up to synthesis a certain sound field based on microphone array recordings of the sound field. First, one or more point sources are set up to be primary sources, the sound fields in the interested region due to these primary sources are captured by a microphone array with less than 1% error, this captured sound field becomes the desired field $\mathbf{P}$. Then, the sound fields due to each secondary loudspeaker is recorded separately, which forms the channel matrix $\mathbf{H}$. Next, different algorithms, including the LMS mode matching, multizone and the proposed prioritized LMS mode matching are used to derive loudspeaker driving signals $\mathbf{D}$ to best synthesize the sound field. For the latter two algorithms, one or more sub regions within the interested zone are selected as high priority area. Finally, the performance of the three methods are compared by plotting the synthesized sound fields on multiple elevations.

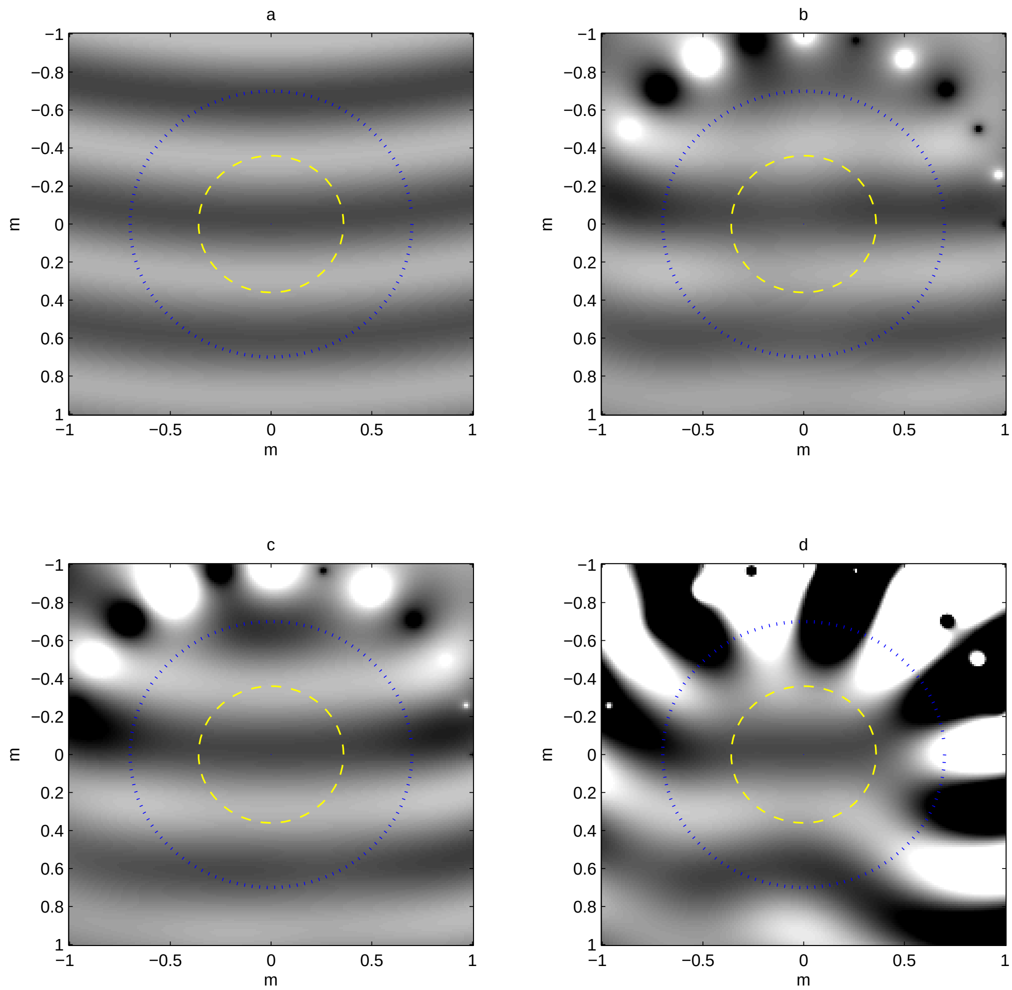

Figure 1 – Comparison of three methods for sound field reproduction on $\theta = \pi/2$ plane

Figure 1 plots the case where 36 loudspeakers are arranged into three semi-circles, placed 1 meter away from the origin, each semi-circular array consists of 12 loudspeakers, spanning from $\phi = 0$ to $\phi = \pi$. The three arrays have elevation angles $\theta = 3\pi/8, \pi/2$,and $5\pi/8$, respectively. A single point source is placed at $(R, \theta, \phi) = (6\text{m}, 2\pi/5, \pi/2)$, which generates a 540Hz sine wave.The region of interest is a 0.7m radius sphere centered at the origin, in addition, a priority sub-region is chosen to be a sphere of 0.4m radius, centered at $(0.2\text{m}, \pi, 0)$.

Figure 1 (a) plots the real sound field produced by the primary source; (b) shows the wave field synthesis result using the LMS method; (c) plots the result using prioritized LMS method; and finally (d) plots the LMS result for the high priority region only. It can be seen that both the conventional LMS method and the proposed method yield acceptable reproduction result. Closer observation would show that the proposed method gives a more accurate reproduction of the sound field within the central circle, which represents the high priority zone. On the other hand, in the case of (d), although the number of loudspeakers provided is sufficient for an approximate reproduction in the high priority zone, due to the relatively long distance between the secondary sources and the reproduction area, the algorithm resulted in very large signal power, even with regularization parameters inserted.

The proposed prioritized LMS algorithm, however, managed to synthesis the sound field with decent accuracy within the high priority zone P, while also giving a reasonable reconstruction result for the entire reproduction region. Furthermore, the driving signals for the secondary sources are also limited, due to the global coefficients also acting as regularization parameters for the matrix solution. Clearly, in this case, the proposed method yields the best result among the three approaches.

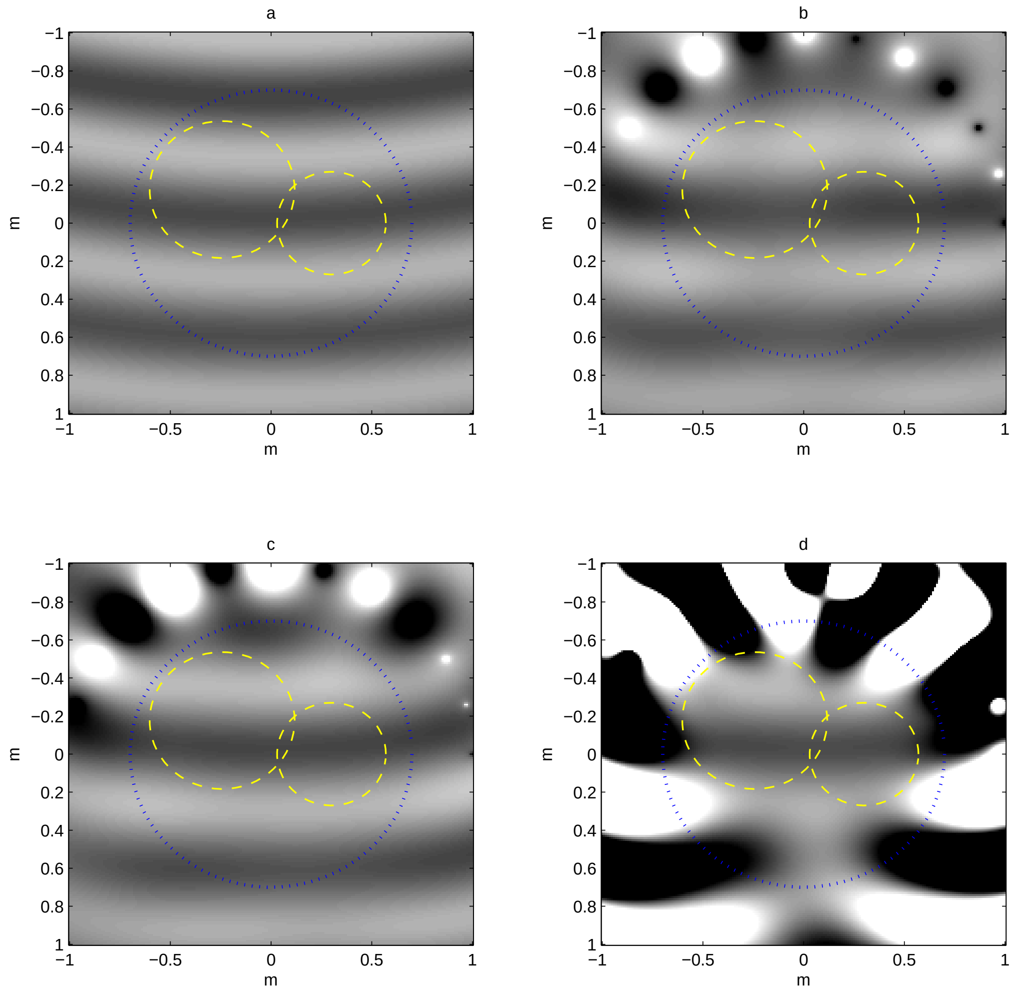

Figure 2 – Comparison of three methods for sound field reproduction with 2 high priority zones

When multiple priority regions are defined, the proposed algorithm will optimize the result for all of these regions according to the weighing factor given to these regions. A simulation result using the same setup as Figure 1 is shown in Figure 2. The only difference is that in Figure 2, two priority sub-regions were defined, their locations were set to be $(0.3\text{m}, \pi/2, 0)$, and $(0.4\text{m}, \pi/2, 0.8\pi)$, and the radius are $0.4m$ and $0.3m$, respectively. The weighing factor for both regions were set to be $\alpha = 10$, and the weighing for the global coefficients was $\beta = 1$. It can be seen that, once again, only the proposed method was able to reconstruct the sound field. Although there are inevitable errors, the synthesized sound field was a good approximation to the desired one, especially within the two prioritized sub-regions.

It has been mentioned in Section 3 that the weighting of each priority zone can be adjusted independently to change the solution of the LMS algorithm, so as to further improve the reproduction accuracy of certain priority area. This is demonstrated in Figure 3, where two priority zones of 0.3m radius were chosen. Again plot (a) shows the desired sound field, which is reconstructed using the secondary sources and the reconstruction errors are shown in plots (b) and (c). In (b), both regions were set to have the same weighing $\alpha = 10$ while the global sound field has a weighing $\beta = 1$. In (c), however, the upper right priority zone has its weighting increased to

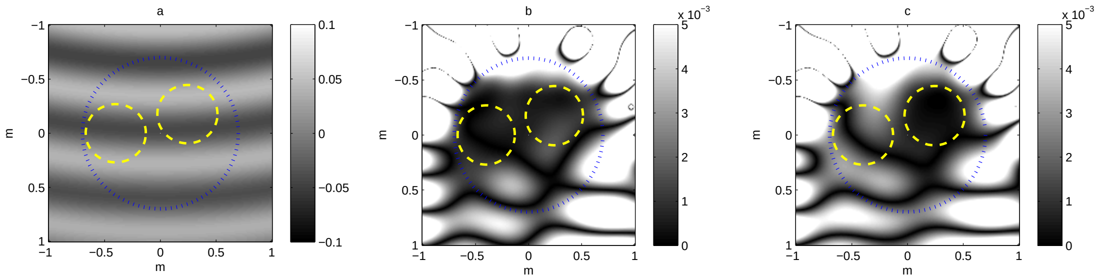

Figure 3 – Reproduction error plots for different high priority zone weight settings

$\alpha_1 = 30$, while the other priority zone's weighing is reduced to $\alpha_2 = 5$. It can be observed that compared to (b), (c) gives a more accurate reproduction in the upper right region, while the region on the left has a slightly worse accuracy. The global reproduction accuracy also degraded slightly, due to the increased weighting for priority zone Q_1.

	Max Order	Coefficient Count	Weight	Mean Square Error	Error Percentage
Global	10	121	1	100%	7.24%
Global	10	121	1	138.53%	7.75%
Zone 1	3	16	10	0.26%	0.48%
Global	10	121	1	131.49%	7.41%
Zone 1	4	20	5	5.43%	1.93%
Global	10	121	1	185.24%	8.53%
Zone 1	4	20	4	7.13%	2.22%
Zone 2	3	16	12	0.59%	0.70%
Global	10	121	1	135.38%	8.91%
Zone 1	3	16	3	12.75%	3.85%
Zone 2	2	9	10	0.43%	0.83%
Zone 3	4	20	2	37.21%	7.67%

Table 1 – Reproduction accuracy for different priority zone settings

Table 1 lists a series of simulated data. The normalized mean square error and average error percentage are calculated for different scenarios, covering LMS based sound field synthesis with $0, 1, 2$ and 3 high priority sub-regions. In order to acquire more accurate synthesis results, a total of 60 loudspeakers are place in a semi-circle one meter away from the origin. Despite the increased number of secondary sources, there is still insufficient secondary sources to completely synthesis the sound field. This is to show the advantage of the proposed algorithm in the cases where insufficient secondary loudspeakers can be used.

The first row in Table 1 shows the performance of the normal LMS algorithm as a reference. The mean square error of the LMS method without prioritized control is used as a reference for comparison of MSE of different setups. It can be seen that an average error of 7.24% is observed from a total of 121 sound field coefficients. The next four rows show the simulation results for 1 priority zone, with its maximum order set to 3 and 4, respectively. In both situations, an increase in the global MSE is observed; the global error percentage also saw a slight increase. Most importantly, the error percentages in the priority zones are much smaller than the global error percentage (0.48% and 1.93%), the error in row 5, Table 1 is greater because of the larger size of the priority region, as well as the lower weight applied to the priority zone.

The rest of Table 1 shows the simulation results for 2 and 3 high priority regions, in both cases each sub-region is given a different weighting factor. The effect of the weighting factors can be easily seen, as the sub region with the largest weighting assigned always result in the lowest error percentage, while the regions with low weights and large radii only see a slight improvement over the global zone. Another observations is that the global synthesis precision degrades more greatly when a large weight is given to the high priority zone. Therefore, in practice, it is advised to choose the weighting and the radii of the sub-regions according to the needs, rather than simply using overly large values.

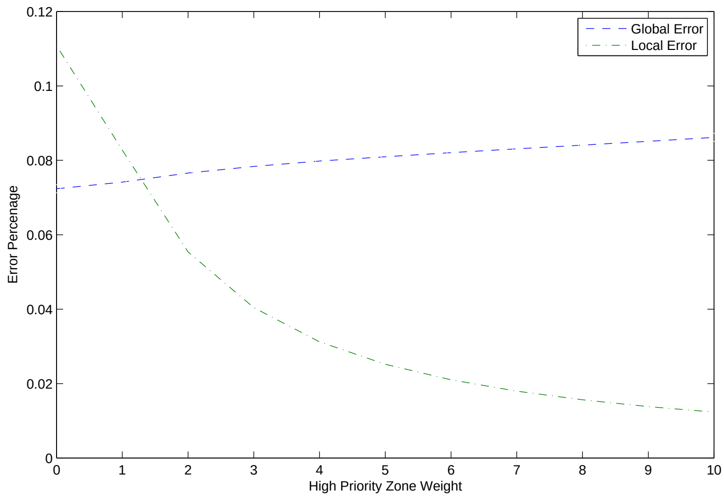

Figure 4 – Effect of the weighting factor on the local and global error percentage

In order to investigate the impact of the weighting factor on the prioritized sub-region and the global sound field, a series of simulations are carried out. The simulations evaluate the reproduction error of a sub-region, located at $(R,\theta,\phi) = (0.1\text{m}, \pi/2, \pi)$ with a maximum order of 3, within the global region of interest whose maximum order is 10. The global region is given the weight $\beta = 1$ while the weighting factor for the sub-region varied from 0 to 10. The resulting error percentages are plotted in Figure 4.

It is easy to identify the exponentially decaying line in Figure 4 as the sub-region error percentage. As the weighting factor changed from 0 to 10, the average reproduction error within this sub-region went from 11% down to a very low 1.2%. Meanwhile, the global error grew slowly from 7.2% to 8.6%. Further investigation would show that the rate at which the global error increases tend to reduce as the local weighing goes larger. Clearly, when using the proposed method to synthesis a sound field, the global reproduction accuracy only suffers slightly, in exchange for the dramatically improved reproduction quality in the prioritized sub-region. In practice, Figure 4 can be used as a trade-off guidance when choosing the appropriate weighting factors for each region.

5. CONCLUSION

A method to prioritize the the reconstruction quality of a sound field within a region is proposed. This method is based on spherical harmonics translation. The sound fields within the global region and the high priority sub-regions are all represented using spherical harmonics, then the spherical harmonic coefficients associated with the sub-regions are mapped into the global coordinate using a translation matrix. A LMS algorithm is then performed on the resulting coefficients to derive the appropriate loudspeaker driving signals. The resulting driving signals are optimized primarily for the sound field reproduction of the high priority sub-regions, while the reproduction in the global region is also taken care of, though with slightly reduced accuracy. It is shown that the proposed method yields significantly higher accuracy in the high priority zones compared to the normal LMS method, this is especially true when only a limited number of secondary loudspeakers are available.

REFERENCES

1. Gerzon M. Periphony: With-height sound reproduction. Journal of the Audio Engineering Society. 1973;21(1):2–10.

2. Ward DB, Abhayapala TD. reproduction of a plane-wave sound field using an array of loudspeakers. IEEE Trans Speech Audio Process. 2001;9(4):697–707.

3. Betlehem T, Abhayapala TD. Theory and design of sound field reproduction in reverberant rooms. J Acoust Soc Amer. 2005;117(4):2100–2111.

4. Wu Y, Abhayapala TD. Theory and design of sound field reproducion using continuous loudspeakers concept. IEEE Trans Audio, Speech, Lang Process. 2009;17(1):107–116.

5. Polleti MA. Three-dimensional surround sound systems based on spherical harmonics,. JAudio Eng Soc. 2005;53(11):1004–1025.

6. Ahrens J, Spors S. An analytical approach to sound field reproduction using circular and spherical loudspeaker distributions. Acta Acust united with Acust. 2008;94:988–999.

7. Berkhout AJ. A holographic approach to acoustic control. J Audio Eng Soc. 1988;36(12):977–995.

8. Berkhout AJ, de Vires D, Vogel P. Acoustic control by wave field synthesis. J Acoust Soc Amer. 1993;93(5):2764–2778.

9. Williams EG. Fourier Acoustics: Sound Radiation and Near field Acoustical Holography. USA: Academic; 1999.

10. Spors S, Rabenstein R, Ahrens J. The theory of wave field synthesis revisited. In: 124th Conv. Audio Eng. Soc.,Amsterdam,Netherlands; 2008. .

11. Boone MM, Verheijen NG, Tol PFV. Spatial sound field reproduction by wave field synthesis. J Audio Eng Soc. 1995;43(12):1003–1012.

12. Boone MM. Multi-actuator panels (MAPs) as loudspeaker arrays for wave field synthesis. J Audio Eng Soc. 2004;52(7-8):712–728.

13. Spors S, Ahrens J. Analysis and improvement of pre-equalization in 2.5-dimensional wave field synthesis. In: 128th Conv. Audio Eng. Soc., London, U.K.; 2010. .

14. Epain N, Friot E. Active control of sound inside a sphere via control of the acoustic pressure at the boundary surface. J Sound Vibr. 2007;299(3):587–604.

15. Naoe M, Kimura T, Katumoto M. Performance evaluation of 3D sound field reproduction system using a few loudspeakers and wave field synthesis. In: 2nd Int. Symp. Universal Commun., Osaka, Japan; 2008. .

16. Fazi FM, Nelson PA, N CJE, Seo J. Surround system based on three-dimensional sound field reconstruction. In: 125th Conv. Audio Eng. Soc., San Francisco, CA, USA; 2008. .

17. Zhang W, Abhayapala TD. Three Dimensional Sound Field Reproduction using Multiple Circular Loudspeaker Arrays: Functional Analysis Guided Approach. IEEE Trans Audio, Speech, Lang Process. 2014;22(7):1184–1194.

18. Wu YJ, Abhayapala TD. Spatial Multizone Sound Field Reproduction: Theory and Design. IEEE Trans Audio, Speech, Lang Process. 2011;19(6):1711–1715.

19. Martin PA. Multiple scattering: interaction of time harmonic waves with N obstacles. Cambridge Univ.; 2006.

20. Samarasinghe PN, Abhayapala TD, Poletti MA. 3D spatial soundfield recording over large regions. In: Acoustic Signal Enhancement; Proceedings of IWAENC; 2012. .