THE CALCULATION OF THE IMPULSE RESPONSE IN THE BINAURAL TECHNIQUE

Lamberto Tronchin*,**, Angelo Farina**, Michele Pontillo*, Valerio Tarabusi*

* DIENCA - CIARM , University of Bologna, Viale Risorgimento 2 - 40136 Bologna Italy
** Ind. Eng. Dept., University of Parma, Viale delle Scienze - 43100 Parma, Italy

E-MAIL: mailto: tronchin@ciarm.ing.unibo.it

Abstract

The importance of the head of a listener to influence the sound that reaches his ears proved a not negligible component among those intervening to modify it during its way. In particular the absorption and the reflection of the skin, the diffraction of the nose and, above all, the “coding” inserted by the auricle, are able to equalise (both in frequency and in phase) the original signal, creating the presupposition for a better interpretation by the brain. The traditional recording techniques, achieved using conventional microphones, are not able to treat with respect those effects, missing so precious information at the origination. Vice versa binaural techniques (but also other, as the interesting Ambisonics), achieved by special dummy-heads, are able to preserve them, allowing their fruition during the reproduction, through the use of headphones; the advantages in terms of sound tridimensionality are considerable. In this paper the binaural representation of calculated impulse responses is mathematically analysed. By defining "azimuth" and "elevation" angles, the reciprocal orientation between sound source and receiver has been mathematically obtained and tested.

INTRODUCTION

At the basis of binaural techniques is the “Transfer Function” (TFs) concept, that is a powerful mathematical (but also data processing) instrument able to characterise the acoustical response of a specific object at a definite solicitation. To have its TFs, measured in different conditions, and to implement them in prediction software means can calculate the effects of its presence without to place it in the condition where it should be during the real event. Many measures about Head-Related Transfer Function (HRTF) was done at the M.I.T. [1].

To choose the appropriate HRTFs fundamental is the calculation of the Elevation and Azimuth angles that the sound ray coming from the source forms with the listener head; by them is possible to choose the HRTFs to use in the filtering of the signals arrived to his ears, including so the alterations caused by his own head.

GEOMETRICAL ANALYSIS

To be able to describe how to execute the calculation of the angles is necessary to decide which reference frames to consider. There are two reference frames: one absolute (**XYZ**), in which is described all the geometry of the room, and one relative (**X'Y'Z'**), integral with the receiver.

As regards the first the source is positioned in $\mathbf{O}_{source} \equiv (\mathbf{X_{source}}, \mathbf{Y_{source}}, \mathbf{Z_{source}})$, whereas the receiver is positioned in $\mathbf{O}_{rec} \equiv (\mathbf{X_{rec}}, \mathbf{Y_{rec}}, \mathbf{Z_{rec}})$ and it is directed to a target point $\mathbf{P}_{targ} \equiv (\mathbf{X_t}, \mathbf{Y_t}, \mathbf{Z_t})$.

The reference frame integral with the receiver has point of origin in $\mathbf{O}_{rec}$, ideally in the barycentre of his head. The three orthogonal axes are so defined: axis **X'** passing through $\mathbf{O}_{rec}$ and leading to the nose vertex (on this axis there is the target point $\mathbf{P}_{targ}$); axis **Y'** passing through $\mathbf{O}_{rec}$ and leading to the left ear (this axis is always parallel with the absolute plane XY, because the two ears are always at the same height) and axis **Z'** passing through $\mathbf{O}_{rec}$ and leading to the top of the head. The provenience point of the sound ray has coordinates $\mathbf{P}_{prov} \equiv (\mathbf{X_{prov}}, \mathbf{Y_{prov}}, \mathbf{Z_{prov}})$.

The two angle to calculate are so defined:

- **Elevation** angle φ (fig. 4) is the angle formed between the **X'Y'** plane and the **u** sound ray vector; it is the complementary angle of that is formed between the **Z'** axis and the **u** vector; it is considered positive if oriented clockwise respect the **Y'** axis; it varies between [-90°,90°];
- **Azimuth** angle θ is the angle formed between the axis **X'** and the projection of the vector **u** on the **X'Y'** plane; it is considered positive if oriented anticlockwise respect the **Z'** axis; it varies in the interval [0°,360°).

The points to obtain the two angles are:

- $\mathbf{P}_{prov} \equiv (\mathbf{X_{prov}}, \mathbf{Y_{prov}}, \mathbf{Z_{prov}})$ = provenience point of the ray;
- $\mathbf{O}_{source} \equiv (\mathbf{X_{source}}, \mathbf{Y_{source}}, \mathbf{Z_{source}})$ = source origin;
- $\mathbf{O}_{rec} \equiv (\mathbf{X_{rec}}, \mathbf{Y_{rec}}, \mathbf{Z_{rec}})$ = receiver origin;
- $\mathbf{P}_{targ} \equiv (\mathbf{X_t}, \mathbf{Y_t}, \mathbf{Z_t})$ = receiver target point. (obs: if the reflection number is zero $\Rightarrow \mathbf{P}_{prov} = \mathbf{O}_{source}$).

Their co-ordinates are related with the absolute reference frame; to calculate the two angles is suitable to execute a co-ordinates transformation, expressing them in the relative frame. The basis versors in the absolute frame is considered as $\{\underline{e}_1, \underline{e}_2, \underline{e}_3\}$ and those in the relative frame as $\{\underline{e}'_1, \underline{e}'_2, \underline{e}'_3\}$, orthogonal Cartesian terns with the mutual point of origin (XYZ reference frame is translated in $\mathbf{O}_{rec}$).

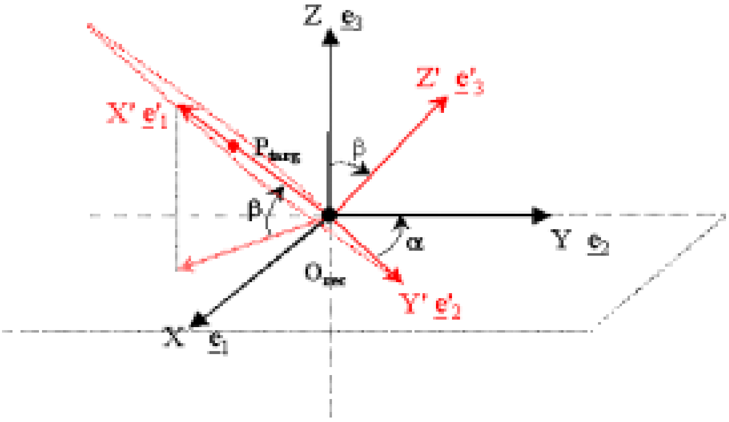


Fig. 1 - α and β angles, formed between the two terns of the reference frames

The possible angles that the two terns can form between them are two (they can be interpreted as a particular case of Eulero angles):

1. *the angle between the XY plane* (translated in $\mathbf{O}_{rec}$) and the X'Y' plane; indicate as β;
2. *the angle between the projection of the Y' axis on the XY plane* (translated in $\mathbf{O}_{rec}$) *and the Y axis* (obs.: owing to the translation the projection of Y' coincide with the same axis); it is indicate α.

Calculus of the α and β angles The two angles are obtained by the target point $P_{targ} \equiv (X_t, Y_t, Z_t)$ coordinates, expressed in the absolute reference frame.
Considering the XYZ absolute frame (translated in $\mathbf{O_{rec}}$), the expression of **β** becomes:

$$\beta = \mathrm{arctg}\left[\left(Z_t - Z_{rec}\right) \Big/ \sqrt{\left(X_t - X_{rec}\right)^2 + \left(Y_t - Y_{rec}\right)^2} \right] \tag{1}$$

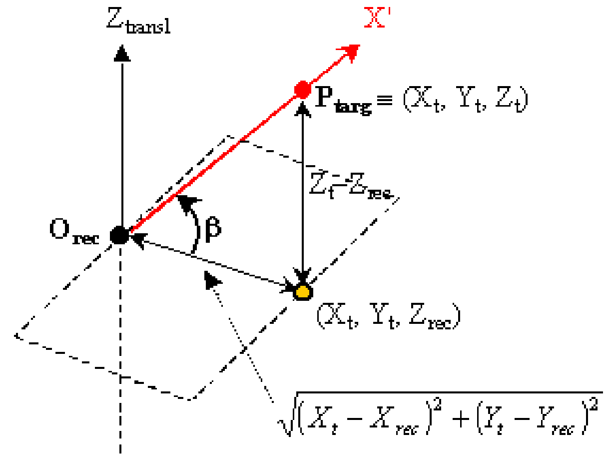
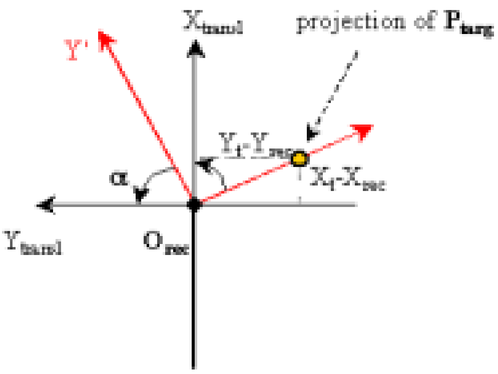


Fig. 2 – Angle β

In the same way, using the function arctg(x), it is possible to calculate the angle **α**, observing in which quadrant the projection of $\mathbf{P_{targ}}$ fall in the XY plane (translated in $\mathbf{O_{rec}}$).



⟵ quadrant where $x\text{-}X_{rec} > 0$, $y\text{-}Y_{rec} \leq 0$

⇓

$\boldsymbol{\alpha = -\,\mathrm{arctg}\,[(Y_t - Y_{rec})/(X_t - X_{rec})]}$

Fig. 3 – Angle α

likewise:

quadrant where fall the projection of $\mathbf{P_{targ}}$

- $x\text{-}X_{rec} \leq 0$, $y\text{-}Y_{rec} < 0$ ⇒ $\boldsymbol{\alpha = \mathrm{arctg}\,[(X_t - X_{rec})/(Y_t - Y_{rec})] + 90°}$;
- $x\text{-}X_{rec} < 0$, $y\text{-}Y_{rec} \geq 0$ ⇒ $\boldsymbol{\alpha = -\,\mathrm{arctg}\,[(Y_t - Y_{rec})/(X_t - X_{rec})] + 180°}$;
- $x\text{-}X_{rec} \geq 0$, $y\text{-}Y_{rec} > 0$ ⇒ $\boldsymbol{\alpha = \mathrm{arctg}\,[(X_t - X_{rec})/(Y_t - Y_{rec})] + 270°}$;

(if $x\text{-}X_{rec} = 0$, $y\text{-}Y_{rec} = 0 \Rightarrow \alpha = 0$).

After obtaining **α** and **β,** the rotation matrix **R**^{T} that describes the components of the basis {$\underline{e}_1$, $\underline{e}_2$, $\underline{e}_3$} on the basis {$\underline{e}'_1$, $\underline{e}'_2$, $\underline{e}'_3$} becomes:

$$\mathbf{R}^T = \begin{bmatrix} \cos\beta \cos\alpha & -\cos\beta\, \mathrm{sen}\alpha & \mathrm{sen}\beta \\ \mathrm{sen}\alpha & \cos\alpha & 0 \\ -\mathrm{sen}\beta \cos\alpha & \mathrm{sen}\beta\, \mathrm{sen}\alpha & \cos\beta \end{bmatrix}. \quad (2)$$

To consider the translation too, from the point of origin **O** = (0,0,0) to $\mathbf{O}_{\mathbf{rec}}$ = (X_{rec}, Y_{rec}, Z_{rec}), the lines-columns product:

$$\begin{bmatrix} x'' \\ y'' \\ z'' \end{bmatrix} = \mathbf{R}^T \begin{bmatrix} x - X_{rec} \\ y - Y_{rec} \\ z - Z_{rec} \end{bmatrix}; \quad (3)$$

gives the transfer from the (x,y,z) coordinates of a point P, expressed in the absolute reference frame, to the respective (x', y', z'), expressed in the relative reference frame.

CALCULUS OF THE ELEVATION φ AND AZIMUTH θ ANGLES

By the lines-columns product (3) is possible to have the $\mathbf{P}_{\mathbf{prov}}$ coordinates in the relative reference frame, indicated as (x'_{prov}, y'_{prov}, z'_{prov}).

The **ELEVATION φ** angle (fig. 4) is simple to obtain, because it is the complementary of the angle between the vector **u** of the sound ray and the Z' axis.

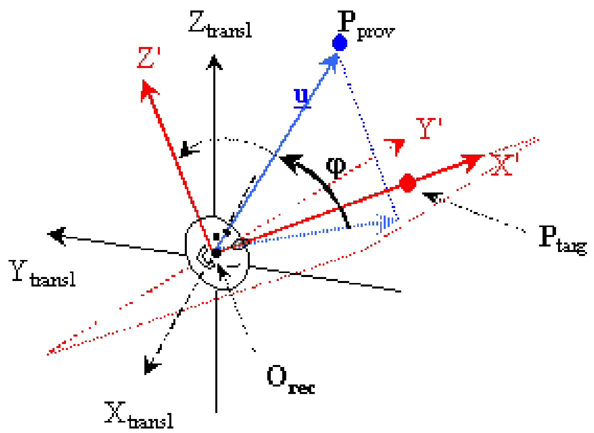


Fig. 4 – Elevation angle φ

Using the $\mathbf{P}_{\mathbf{prov}}$ coordinates in the relative frame, the **φ** angle is the complementary of the angle formed between the vector (x'_{prov}, y'_{prov}, z'_{prov}) and the Z' axis versor (0,0,1):

$$\varphi = 90^\circ - \arccos\left[z'_{prov} \Big/ \sqrt{\left(x'_{prov}\right)^2 + \left(y'_{prov}\right)^2 + \left(z'_{prov}\right)^2} \right] \quad (4)$$

The projection of $\mathbf{P_{prov}}$ on the plane X'Y' (that is Z' = 0) is given putting $z'_{prov} = 0$ (x'_{prov} and y'_{prov} remain the same). The **AZIMUTH θ** angle (fig. 5) is therefore the angle between the vector (1,0,0), versor of the X' axis in the relative reference frame, and the vector (x'_{prov} , y'_{prov}, 0), positive oriented anticlockwise respect the **Z'** axis.

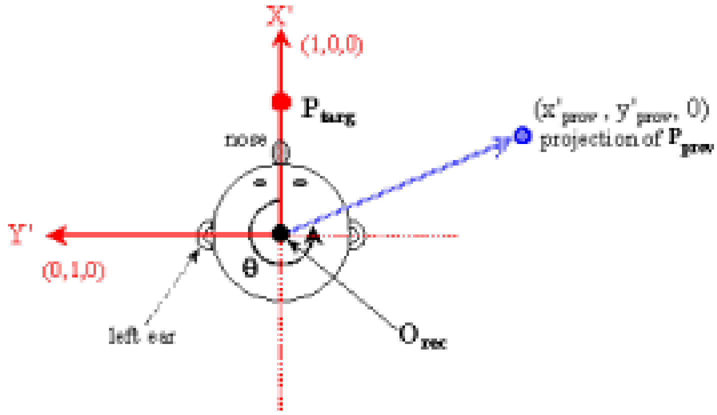


Fig. 5 – Azimuth angle θ

Since the angle can vary from 0° to 360° and the arccos(x) function give only the smaller angle between the two vectors, that is:

$$\theta^* = \arccos\left[x'_{prov} \Big/ \sqrt{(x'_{prov})^2 + (y'_{prov})^2} \right] \quad , \qquad (5)$$

The angle θ* found by the (5) is the right one and not the complementary, provided its direction, given by the product $(1,0,0)\wedge(x'_{prov}\ ,\ y'_{prov},\ 0)$, is concordant or discordant with the Z' axis versor; that means to analyse the sign of the determinant of the matrix:

$$M = \begin{bmatrix} 1 & 0 \\ x'_{prov} & y'_{prov} \end{bmatrix}$$

– if $\mathbf{y'_{prov} > 0}$ $\Rightarrow$ $\boldsymbol{\theta = \theta^*}$**;**
– if $\mathbf{y'_{prov} < 0}$ $\Rightarrow$ $\boldsymbol{\theta = 360° - \theta^*}$**;**
– if $\mathbf{y'_{prov} = 0}$ $\Rightarrow$ $\begin{cases} \text{if } \mathbf{x'_{prov} \geq 0} \Rightarrow \boldsymbol{\theta = 0°}; \\ \text{if } \mathbf{x'_{prov} < 0} \Rightarrow \boldsymbol{\theta = 180°}. \end{cases}$

THE CHOICE AND THE INTERPOLAZION OF THE HRTFs

After having computed the Elevation **φ** and Azimuth **θ** angles with which the sound ray arrives at the listener ears, the appropriate HRTFs to interpolate to use in the signal filtering must be chosen. After the observation of a large number of measurements made at M.I.T.[1] with a KEMAR dummy-head, a triangular interpolation have been preferred, because the number of measurements about the Azimuth angles is not constant, as described in table 1.

Elevation **φ** (°)	Number of Measurements	Azimuth Increment (°)
± 90	1	***
± 80	12	30
± 70	24	15
± 60	36	10
± 50	45	8
± 40	56	6.43
± 30	60	6
± 20	72	5
± 10	72	5
0	72	5

Tab. 1 - Number of measurements and azimuth increment at each elevation (at M.I.T.).

Supposing to fix on a spherical surface (with unit radius) the points sampled, performing a triangulation among them the sound ray will fall in a specific triangle (fig. 6).

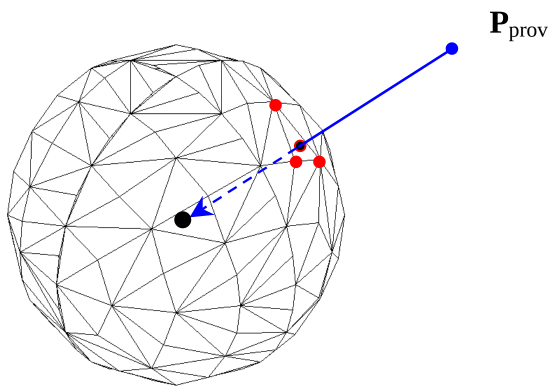
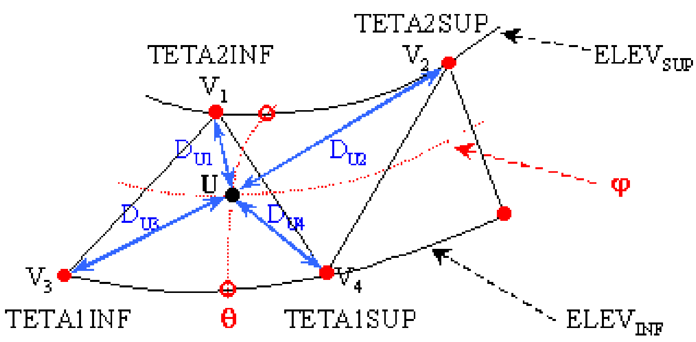


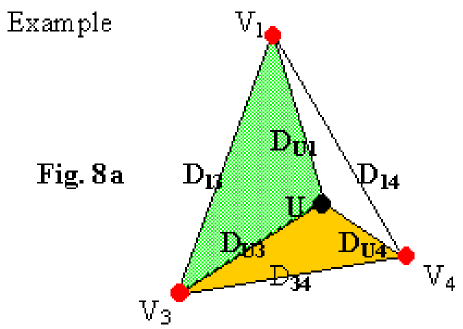
Fig 6 - Schematisation of a triangulation operated on a spherical surface.

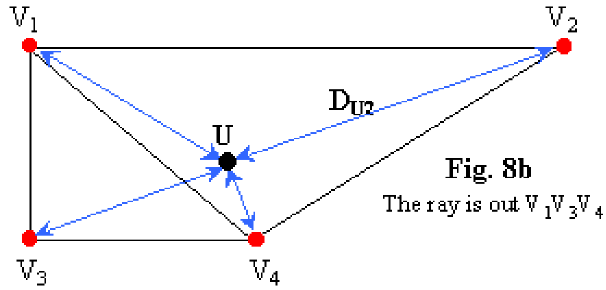
Consider the lower and the higher then **φ** elevation angles at which was sampled the measurements; after having selected the two corresponding circumferences and, likewise, the lower and the higher then **θ** elevation angles, four couples of angles (φ_i, θ_i) have therefore been selected, placed in the vertices V_1,V_2,V_3,V_4 of two adjacent triangles. In one of these falls the sound ray (fig. 7).

Fig. 7 – Selection of the vertices



Consider the distances among the point **U,** projection of the intersection between the ray and the sphere on the obtained figure, and the four vertices: D_{U1}, D_{U2}, D_{U3}, $D_{U4.}$ Consider the triangle having the vertices corresponding the three lowest distances. The ray could fall in it, but could happen as in fig 8b.





Suppose that the ray falls into the triangle (fig. 8a). The three HRTFs (in the figure related at the vertices V_1,V_3,V_4) are regarded as the "extremes" of the interpolation; the weight $\mathbf{P_i}$ (i = 1,3,4) relative each one is obtained calculating the opposite triangle area (in fig 8a the area V_1V_3U is referred to the vertex 4) and dividing it for the total initial triangle area; so we have $\mathbf{P_1+P_3+P_4} = 1$.

When **U** is out the chosen triangle (fig. 8b) it results $\mathbf{P_1+P_3+P_4} > 1$. The adjoining triangle have to be selected. It has the vertices so defined: Reconsider the vertex with the longest distance and indicate it with k; put j = 5–k; remove the vertex with j index. The three remaining vertices forms the new triangle.

THE WEIGHTS CALCULATION

Consider $\mathbf{V_1} = \mathbf{V_1(\varphi_1,\theta_1)}$, $\mathbf{V_2} = \mathbf{V_2(\varphi_2,\theta_2)}$, $\mathbf{V_3} = \mathbf{V_3(\varphi_3,\theta_3)}$ the vertices of the chosen triangle (where falls **U**) and calculate their Cartesian coordinates:

$$\left.\begin{array}{l} X = \cos(\varphi)*\cos(\theta) \\ Y = \cos(\varphi)*\text{sen}(\theta) \\ Z = \text{sen}(\varphi) \end{array}\right\}\text{vertex } \mathbf{U}\,; \qquad \left.\begin{array}{l} X_i = \cos(\varphi_i)*\cos(\theta_i) \\ Y_i = \cos(\varphi_i)*\text{sen}(\theta_i) \\ Z_i = \text{sen}(\varphi_i) \end{array}\right\}\text{vertex } V_i \quad (i=1..3).$$

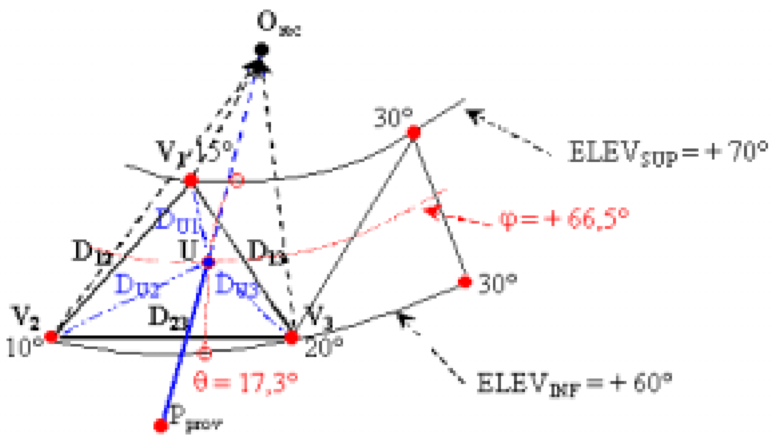


Fig. 9 – Example: $\boldsymbol{\varphi}$ = 66,5°, $\boldsymbol{\theta}$ = 17,3°; $\mathbf{V_1}$=$\mathbf{V_1}$(70°,15°); $\mathbf{V_2}$=$\mathbf{V_2}$(60°,10°); $\mathbf{V_3}$=$\mathbf{V_3}$(60°,20°).

Calculate the distances among vertices $\mathbf{V_1,V_2,V_3}$ and those among vertex **U** e $\mathbf{V_1,V_2,V_3}$:

$$\left\{\begin{array}{l} \mathbf{D_{12}} = [(X_1-X_2)^2+(Y_1-Y_2)^2+(Z_1-Z_2)^2]^{1/2} \\ \mathbf{D_{13}} = [(X_1-X_3)^2+(Y_1-Y_3)^2+(Z_1-Z_3)^2]^{1/2} \\ \mathbf{D_{23}} = [(X_2-X_3)^2+(Y_2-Y_3)^2+(Z_2-Z_3)^2]^{1/2} \end{array}\right. \qquad \left\{\begin{array}{l} \mathbf{D_{U1}} = [(X-X_1)^2+(Y-Y_1)^2+(Z-Z_1)^2]^{1/2} \\ \mathbf{D_{U2}} = [(X-X_2)^2+(Y-Y_2)^2+(Z-Z_2)^2]^{1/2} \\ \mathbf{D_{U3}} = [(X-X_3)^2+(Y-Y_3)^2+(Z-Z_3)^2]^{1/2} \end{array}\right.$$

To determine the triangle areas the Erone' formula can be utilised:

$$Area = \sqrt{S*(S-a)*(S-b)*(S-c)} \tag{6}$$

with a,b,c length of the sides and S the half-perimeter.

The half-perimeter S_{max} of the bigger triangle and those S_i of the opposite vertex triangles are:

$S_{max} = (\mathbf{D_{12}} + \mathbf{D_{13}} + \mathbf{D_{23}})/2$; $S_1 = (\mathbf{D_{U2}} + \mathbf{D_{U3}} + \mathbf{D_{23}})/2$;
$S_2 = (\mathbf{D_{U1}} + \mathbf{D_{U3}} + \mathbf{D_{13}})/2$; $S_3 = (\mathbf{D_{U1}} + \mathbf{D_{U2}} + \mathbf{D_{12}})/2$;

their areas are:

- $\mathbf{A_{TOT}} = [S_{max}*(S_{max} - \mathbf{D_{12}})*(S_{max} - \mathbf{D_{13}})*(S_{max} - \mathbf{D_{23}})]^{1/2}$;
- $\mathbf{A_1} = [S_1*(S_1 - \mathbf{D_{U2}})*(S_1 - \mathbf{D_{U3}})*(S_1 - \mathbf{D_{23}})]^{1/2}$;
- $\mathbf{A_2} = [S_2*(S_2 - \mathbf{D_{U1}})*(S_2 - \mathbf{D_{U3}})*(S_2 - \mathbf{D_{13}})]^{1/2}$;
- $\mathbf{A_3} = [S_3*(S_3 - \mathbf{D_{U1}})*(S_3 - \mathbf{D_{U2}})*(S_3 - \mathbf{D_{12}})]^{1/2}$.

The weights are so given by the expressions: $\mathbf{P_1 = A_1/A_{TOT}}$; $\mathbf{P_2 = A_2/A_{TOT}}$; $\mathbf{P_3 = A_3/A_{TOT}}$.
The final HRTF to use in the filtering (one of two, the other has complementary Azimuth angle) is:

$$HRTF(\varphi,\theta) = P_1 * HRTF(\varphi_1,\theta_1) + P_2 * HRTF(\varphi_2,\theta_2) + P_3 * HRTF(\varphi_3,\theta_3) \tag{7}$$

meaning that symbolic script an interpolation in the time domain, witch separately mediates the modulus and the phase of the complex spectrum.

CONCLUSIONS

The methods to preserve the sound “tridimensionality” of an audio event become always more important. Besides in the recording and the reproduction, they are quite important in data processing software also. Furthermore, these methods allows to observe (and often to listen to) the results obtainable in the real event without the need to carry out it.

Among the various techniques one of the more effective is the “binaural”: it uses a special dummy-head for the sound recording and a headphone for the reproduction. The formulas set described in this article form a mathematical model utilisable in prediction software. By the calculation of the Elevation and Azimuth angles between the sound ray and the relative reference frame the more appropriate HRTFs for signal filtering can be chosen, including the effects generated by the head.

The triangular interpolation chosen is a very good approximation of that (exact) “curvilinear triangles” formed on the spherical surface. This methods allows much better results than considering only the orthogonal projections on the plane containing the three extreme HRTFs. Because of the unit radius of the sphere the differences are almost null, with the advantage to have calculation times considerably reduced.

The use of a model as that here described allow to come close the sound truth with prerogatives not consented to conventional methods.

REFERENCES

[1] Bill Gardner and Keith Martin: *HRTF Measurements of a KEMAR Dummy-Head Microphone*, MIT Media Lab., May, 1994.