Hierarchical Representation of Surfaces Using 3D Wireframes

I. Kompatsiaris **D. Tzovaras** **M. G. Strintzis**

Information Processing Laboratory
Electrical and Computer Engineering Department
Aristotle University of Thessaloniki
Thessaloniki, Greece
Email: srtintzi@eng.auth.gr

ABSTRACT

In this paper a novel procedure for the representation of 3D surfaces using 3D hierarchical adaptive wireframes is presented. The 3D surfaces are generated by dense depth maps. The procedure is based on pyramidal analysis using the Quincunx Sampling Minimum Variance Interpolation (QMVINT) filters. The use of this method minimizes the variance (which is a measure of the entropy) of the interpolation error and therefore results to optimal compression of the wireframe information transmitted. At the same time, it produces a hierarchy of meshes based on quincunx sampling where coarse meshes are as similar to their finer versions as possible. Depending on its interpolation error and the available bitrate, each filtered sample is candidate for becoming a node of the wireframe. The result is a progressive sequence of wireframes consisting of more triangles wherever large variations in depth exist and fewer in uniform regions. Experimental results demonstrate the usage and performance of the algorithm.

Keywords: visualisation of surfaces, quincunx sampling, hierarchical representation

1 INTRODUCTION

Polygonal surface approximations are an essential preprocessing step in model based video coding [1, 3] and scientific visualization [2]. In both areas, it is often required to adapt the number of triangles representing the object to the needs of specific applications.

Several adaptive triangle reduction techniques have been presented in the literature. Most attempt to establish mathematical criteria defining the importance of each particular mesh vertex, so as to remove it if unimportant and locally retriangulate the mesh. In [4], prominent feature points are first selected (e.g., as defined by edges of the object), based on which a 2-D mesh structure is constructed (e.g., by constrained Delaunay triangulation). In [5] new points are added to an initially uniform wireframe wherever large reconstruction errors of the real surface occur. Also, large triangles created by non-rigid motion of the nodes of the wireframe, are divided into smaller ones. In [6] the wavelet transform (WT) is used as an overall mathematical framework controlling the data approximation. In [7] *progressive meshes* are introduced, and complete correspondence between vertices in different levels of the hierarchy is established, something that cannot be easily achieved for triangles/faces at different levels.

In this paper a novel procedure for hierarchical representation of 3D surfaces using 3D adaptive triangular wireframes is presented. The procedure is based on pyramidal analysis using the Quincunx Sampling Minimum Variance Interpolation (QMVINT) filters. These filters are applied to the depth image of the surface and each sample used by the filters in order to predict a pixel depth value, along with the pixel whose depth was predicted, are candidates for becoming triangularly connected nodes of the wireframe. Instead of transmitting the actual depth values of the wireframe nodes, the uncorrelated prediction errors of the QMVINT pyramid are transmitted, resulting in compression of the wireframe information. The use of these methods minimizes the variance (which is a measure of the entropy) of the error transmitted and therefore results to optimal compression of the wireframe information transmitted. At the same time, it produces a hierarchy based on quincunx sampling where coarse meshes are as similar to their finer versions as possible. The triangulation is performed adaptively, placing more triangles in highly detailed areas where the larger prediction errors occur. The triangulation algorithm is integrated with a bit allocation procedure based on the variances of the prediction errors at each hierarchy level. This procedure extends from the coarser to the finer level, until the desired detail of the wireframe is reached. Furthermore, precise correspondence between triangles at each level is achieved, resulting to a fully hierarchical representation of the wireframe. The number of nodes and triangles of the resulting wireframe is not controlled through a predefined threshold, as is often the case, but rather, it is integrated with the available bitrate, either for storage or for transmission purposes, through the bit allocation procedure.

The paper is organised as follows. The mathematical framework of quincunx sampling and the QMVINT filters are introduced in Section 2. The bit allocation procedure is described in Section 3. The prioritized transmission according to the prediction error and the entropy estimation are presented in Section 4. The adaptive triangulation procedure performed at the receiver side along with the multiresolution representation of surfaces are described in Section 5. Experimental results are given in Section 6.

2 OPTIMISATION OF 2D HIERARCHICAL INTERPOLATION

Hierarchical interpolation based on quincunx sampling has been widely use in progressive image transmission terminating with a lossless version of the image. This class contains the recently proposed "Reduced Laplacian Pyramid" [9] as well as the 2D version of the "Hierarchical Interpolation" (HINT) method. Based on these schemes, at the decoder/receiver side the *synthesis* process takes place as shown in Fig. 1. Referring to the synthesis process in Fig. 1 it is assumed that the black vertices arrive. This is a low-resolution version of the original image produced by the corresponding level at the encoder side. On applying the quincunx oversampling procedure, indicated by $\uparrow \mathbf{M}$, the grey vertices are created. The value of these grey vertices is esti-

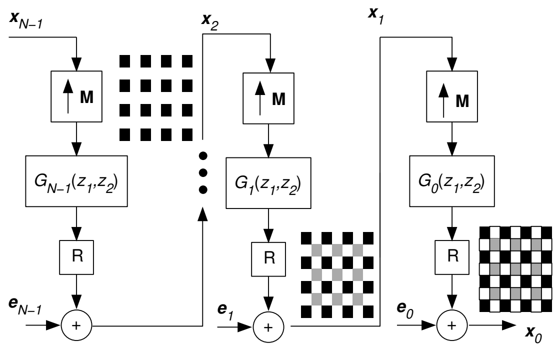

Figure 1: Reduced pyramid with quincunx sampling matrix $\mathbf{M}$ and filters $G_i(z_z, z_2)$. Black vertices are input to level 1, grey vertices are inserted as a result of the quincunx upsampling procedure and their values are estimated from filters $G_i(z_z, z_2)$. The same procedure is applied at the following level, where the white vertices are introduced. The corresponding triangulation is shown in Fig. 2.

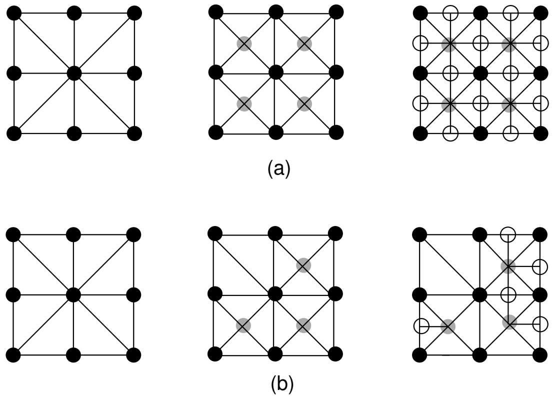

Figure 2: The corresponding triangulation of the interpolation procedure shown in Fig. 1. (a) Every vertex is triangulated, when error information for every vertex is transmitted. (b) Only some vertices are triangulated when only vertices with higher errors have been transmitted until the available for that level bitrate has been exhausted.

mated by interpolation of the previously transmitted black vertices. The interpolation procedure is performed by the linear filter with transfer function $G_1(z_z, z_2)$ as shown in the figure. The estimated value is then rounded, an operation indicated by R in the figure. The interpolation errors have been estimated at the encoder/transmitter side, as preciously described. Those transmitted errors are added to the estimated values and the real depth value for each vertex is obtained. This procedure is continued at the next level, where the white vertices are introduced by the upsampling. The associated proposed triangulation procedure, that takes place in the receiver, can be seen in Fig. 2. In Fig. 2(a) full triangulation at each level of the algorithm is shown, meaning that the error of each vertex has been transmitted. In Fig. 2(b) only some of the vertices are triangulated, meaning that only vertices with higher errors have been transmitted until the available for that level bitrate was exhausted.

Since quincunx sampling is to be used as the basis for the triangulation procedure, the optimum form of hierarchical interpolation based on quincunx sampling is used in this paper. This optimum form, presented originally in [11] guarantees the smoothness of the depth information as well as the optimization of the prediction of depth information for coding purposes, by minimizing the variance of the error in each scale of the hierarchical representation.

Let $\mathbf{M}$ be the sampling matrix, $M = det\mathbf{M}$, $\mathbf{p}_i$ the coset vectors of $\mathbf{M}$, and $\mathbf{q}_i$ the coset vectors of $\mathbf{M}^T$, $i = 0, \ldots, M-1$. It can be shown [11, 8] that the optimal post-filter minimising the corresponding mean-square error variance is obtained by

$$G_r(e^{j\mathbf{w}}) = \frac{\Phi_r(e^{j\mathbf{w}})}{\frac{1}{M}\sum_{i=0}^{M-1} \Phi_r\big(e^{j(\mathbf{w}+2\pi\mathbf{M}^{-T}\mathbf{q}_i)}\big)},$$

or by solving in the time domain the following equation

$$R_r[\mathbf{Mt}+\mathbf{p}_i] = \sum_{\mathbf{k}} g_r^{(i)}[\mathbf{Mk}]R_r[\mathbf{M}(\mathbf{t}-\mathbf{k})],\ \forall \mathbf{t}\ ,$$

For the design of these filters, it is necessary to specify the spectral densities describing the class of depth maps which are of interest. Two spectral density models, the "separable" and the "isotropic", appear sufficient for the accurate characterization of the great majority of depth maps encountered in practice. The optimal post-filters are determined separately for each of these models [11].

3 OPTIMUM BIT ALLOCATION

Depending on the capacity of the transmission channel, a total size of B bits/vertex is assumed to be transmitted to the receiver for the creation of the mesh. This information consists of the prediction errors along with the position information of the vertices. The available bitrate is allotted to each level depending on the error variances and the number of vertices at this level. At each level $r = 0, \dots N-1$, where N is the total number of levels, $P_i^{(r)}, i = 0, \dots, M^{(r)}$ depth values are predicted from depth values of vertices at the previous level $r-1$. Specifically, if σ^2 is the sum of the error variances of all levels $\sigma^2 = \sum_{i=0}^{r} \sigma_r^2$ then B_r is $B_r = \frac{\sigma_r^2}{\sigma^2}\ B$.

4 ERROR PRIORITIZATION AND ENTROPY ESTIMATION

At each hierarchy level the prediction errors corresponding to all predicted vertices are calculated and sorted with the vertices corresponding to higher errors being put first on the list. We shall assume that entropy coding (e.g. Huffman or arithmetic coding) is used, with an adaptive probability model. In this case, the number of bits needed for error transmission is the entropy R_e of the errors of the $L^{(r)}$ nodes transmitted. Using the quincunx sampling geometry at the receiver, there is no need to transmit the exact coordinates of the position of each transmitted vertex. The transmission of only a bit indicating whether a vertex was transmitted or not, suffices for the generation of the mesh at the receiver side. Let the additional overhead for the transmission of these bitmaps be the entropy R_p. The final cost of the transmission of $L^{(r)}$ vertices is the sum of the error entropy R_e and the position entropy R_p.

5 ADAPTIVE TRIANGULATION PROCEDURE

At the receiver side, a proper triangulation representing the 3D surface is created using the received information, the QMVINT pyramid and the quincunx sampling geometry. For each level r, $P_i^{(r)}, i = 0, \dots, L^{(r)}$ vertices are received. The result of the triangulation procedure is a valid mesh at each level r : $W^{(r)} = \{P_i^{(r)}, T_i^{(r)}\}$, where $T_i^{(r)}, i = 0, \dots, K^{(r)}$ are the triangles of the mesh. Basic characteristics of the mesh created are the hierarchical representation of the surface and the avoidance of creation of "cracks" as defined in [6]. A crack occurs whenever vertices appear, which are not properly connected in a triangular mesh with neighbouring vertices. In this case the surface breaks up, holes appear and the consistency required for normal interpolation is lost.

At each level, for each new vertex received, the following triangulation procedure is applied.

Step 1 The synthesis stage of the QMVINT pyramid and the received error information are used to reconstruct the

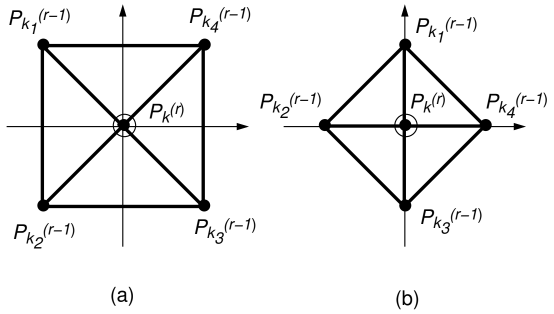

Figure 3: Depth value of $P_k^{(r)}$ at level r is predicted from $P_{k_1}^{(r-1)}$, $P_{k_2}^{(r-1)}$, $P_{k_3}^{(r-1)}$, $P_{k_4}^{(r-1)}$ and in case of triangulation, the triangles shown at the scheme are created, in (a) odd-numbered and (b) even-numbered stages of the algorithm.

initial value of the depth of the vertex,

$$I(P_k^{(r)}) = \hat{I}(P_k^{(r)}) + err(P_k^{(r)}) \ .$$

Step 2 The vertex $P_k^{(r)}$ along with the vertices used to predict the value of $P_k^{(r)}$ are added to the mesh. There are two ways in which the depth value of a vertex can be predicted in quincunx sampling and the corresponding triangulation for each is shown in Fig. 3. In both cases, the vertices added to $W^{(r)}$ are : $P_k^{(r)}$ and $P_{k_j}^{(r-1)}, j = 1, \ldots, 4$, as shown in Fig. 3. Equivalently, the triangles added to $W^{(r)}$ are the ones shown at the scheme.

Step 3 If all $P_{k_j}^{(r-1)}, j = 1, \ldots, 4$ are already part of the mesh $W^{(r-1)}$, then no cracks have been created by the triangulation.
Else, for all $P_{k_j}^{(r-1)}$ which are not part of $W^{(r-1)}$, the triangulation procedure starting from Step 1 is performed, with $r = r - 1$ and $err(P_k^{(r)}) = 0$.

Step 4 The next vertex received is then triangulated starting from Step 1.

The result of the procedure described above is a mesh with more triangles over highly detailed areas and fewer triangles over uniform regions. The problems of cracks and holes of the mesh are solved by the algorithm described above. The triangulation mechanism results in an accurate and parsimonious representation of the surface. An important advantage of the algorithm is that all triangles created at each level are of uniform shape and no triangles of irregular shapes are created.

5.1 Hierarchical Representation of Surfaces

An important advantage of this algorithm is that each parameter of the surface triangle estimated at level r (e.g. rigid motion) corresponds to the same parameter of a unique triangle at level $r-1$. This is useful in many applications such as when articulation of a mesh based on the motion of each triangle at a coarser level is estimated and must be carried over to a finer level ([3]) for use in rigid or non-rigid motion estimation. This articulation procedure is based on the homogeneity of parameters, such as rigid 3D motion, color and depth, estimated for each sub-object, which consists of a number of interconnected triangles of the 3D model.

For each triangle $T_l^{(r)}$ at level r the corresponding triangle $T_m^{(r-1)}$ at level $r-1$ is found using the hierarchy function :

$$H_{r-1}^{r}\left(T_l^{(r)}\right) = T_m^{(r-1)}$$

This is tabulated every time a new vertex $P_k^{(r)}$ is added to the mesh $W^{(r)}$ for all new triangles : $P_{k_a}^{(r-1)}, P_k^{(r)}, P_{k_b}^{(r-1)}$ where $(a,b) = \{(1,2),(2,3),(3,4),(4,1)\}$:

$$H_{r-1}^{r}\left\{P_{k_a}^{(r-1)} P_k^{(r)} P_{k_b}^{(r-1)}\right\} = \left\{P_{k_a}^{(r-1)} P_{k_b}^{(r-1)} P_{k_c}^{(r-1)}\right\} \ ,$$

where $c = 1,2,3,4$ and in any case $a \neq c$ and $b \neq c$. Triangle $\{P_{k_a}^{(r-1)} P_k^{(r)} P_{k_b}^{(r-1)}\}$

belongs to $W^{(r)}$, and $\{P_{k_a}^{(r-1)} P_{k_b}^{(r-1)} P_{k_c}^{(r-1)}\}$ belongs to $W^{(r-1)}$. Clearly, complete correspondence between meshes created at each level is established in this manner.

6 EXPERIMENTAL RESULTS

The proposed hierarchical adaptive triangulation algorithm of 3D surfaces was evaluated for 3D mesh adaptive representation and compression of real surfaces. The "Venus" dense map used was created from range data obtained from a laser scan of a statue. The original depth map is shown in Fig. 4. The original surface visualised in the 3D space is shown in Fig. 5.

The algorithm was tested for the creation of adaptive meshes coded at 0.5 bits per vertex. This resulted in adaptive 3D models able to reproduce with sufficient accuracy the original surfaces, with much fewer vertices and triangles than the originals, as well as very low bitrate transmission requirements, compared with the original entropy of the depth maps. The meshes and the reconstructed surfaces for levels 0, 1 and 2, for the "Venus" sequence are in Fig. 6, 7, 8, 9, 10 and 11.

ACKNOWLEDGMENT

This work was supported by the AMEA and PANORAMA project of the Greek Secretariat of Research and Technology.

REFERENCES

[1] H. G. Musmann, M. Hotter, and J. Ostermann, "Object-oriented analysis-synthesis coding of moving images," *Signal Processing: Image Communication*, vol. 1, no. 2, pp. 117–138, Oct. 1989.

[2] G. M. Nielson, *"Modeling and Visualizing Volumetric and Surface-On-Surface Data"*, pp. 191–242, Springer, 1993.

[3] I. Kompatsiaris, D. Tzovaras, and M. G. Strintzis, "3D Model Based Segmentation of Videoconference Image Sequences," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 8, no. 5, Sept. 1998.

[4] Y. Altunbasak and A. M. Tekalp, "Content-based mesh design for 2-D object-based video coding," in *Symp. on Multimedia Comm. and Video Coding*, New York City, New York, Oct. 1995.

[5] W. C. Huang and D. B. Goldgof, "Adaptive-Size Meshes for Rigid and Non-Rigid Shape Analysis and Synthesis," *IEEE Trans. Pattern Anal. and Mach. Intell.*, vol. 15, no. 6, pp. 611–616, 1993.

[6] M. Gross, O. Staadt, and R. Gatti, "Efficient Triangular Surface Approximations Using Wavelets and Quadtree Data Structures," *IEEE Trans. on Visualisation and Computer Graphics*, vol. 2, no. 2, pp. 130–143, June 1996.

[7] Hugues Hoppe, "Progressive meshes," in *SIGGRAPH 96 Conference Proceedings*, Holly Rushmeier, Ed. ACM SIGGRAPH, August 1996, Annual Conference Series, pp. 99–108, Addison Wesley, held in New Orleans, Louisiana, 04-09 August 1996.

[8] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice-Hall, 1993.

[9] B. Aiazzi, L. Alparone, and B. Baronti, "A Reduced Laplacian Pyramid for Lossless and Progressive Image Communication," *IEEE Trans. Communications*, vol. 44, no. 1, pp. 18–23, Jan 1996.

[10] P. Roos, M. A. Viergever, M. C. .A. Dijke, and J. H. Peters, "Reversible Intraframe Image Compression of Medical Images," *IEEE Trans. Med. Imaging*, , no. 7, pp. 328–336, 1988.

[11] D. Tzovaras and M. G. Strintzis, "Optimal Reduced Pyramid Interpolation for Lossless and Progressive Image Coding," in *IEEE First Workshop on Multimedia Signal Processing*, Princeton, USA, June 1997, pp. 151–156.

[12] M. G. Strintzis, "Optimal Biorthogonal Wavelet Bases for Signal Representation," *IEEE Trans. Signal Processing*, vol. 44, no. 6, pp. 1406–1418, Jun. 1996.

Figure 4: Original dense depth map of the "Venus" data.

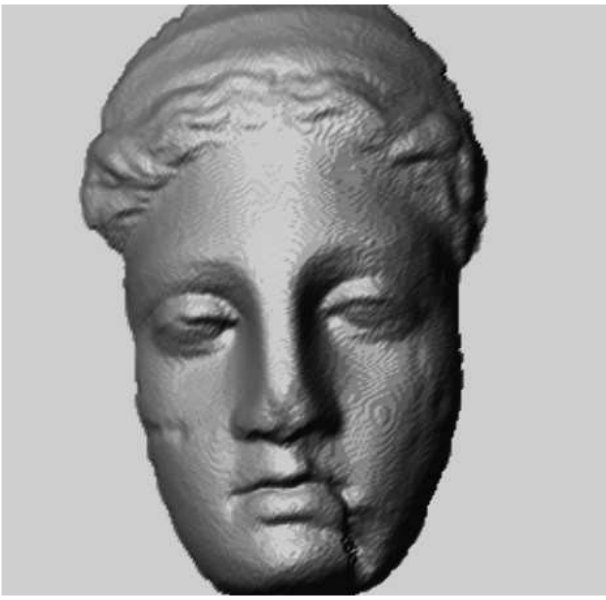

Figure 5: Original "Venus" surface.

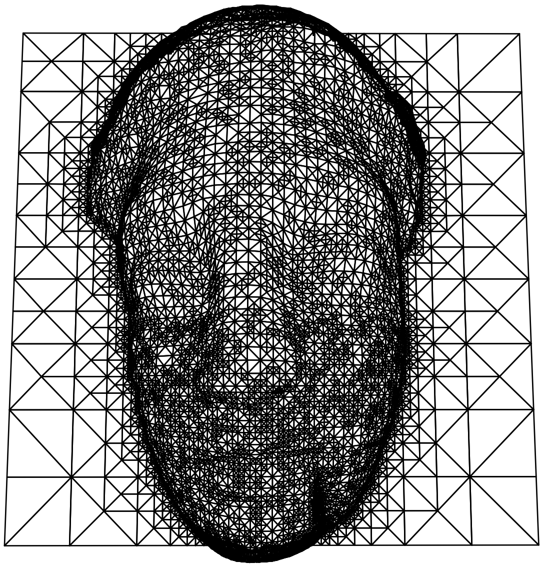

Figure 6: The resulting mesh with 11416 vertices and 15827 triangles at level 0.

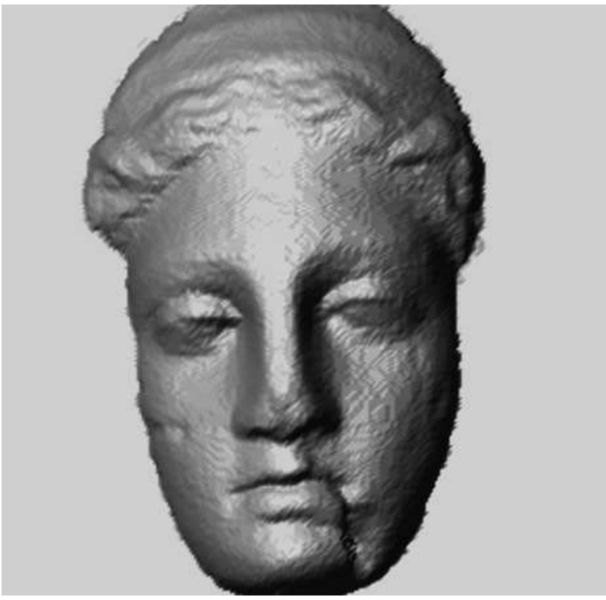

Figure 7: 3D visualisation of the surface with MSE 0.12.

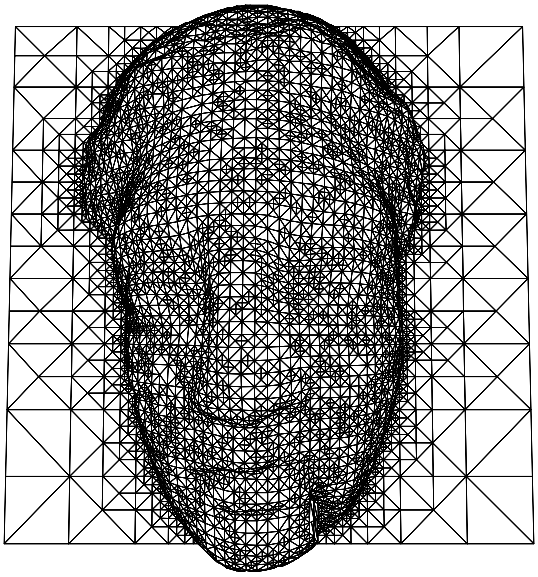

Figure 8: The resulting mesh with 7661 vertices and 11135 triangles at level 1.

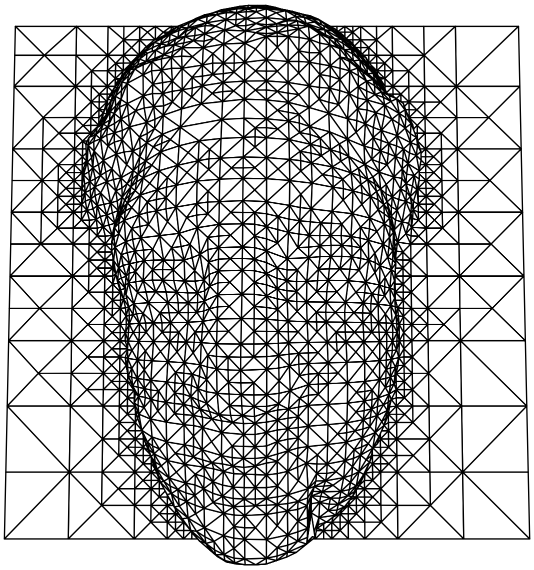

Figure 10: The resulting mesh with 2569 vertices and 4006 triangles at level 2.

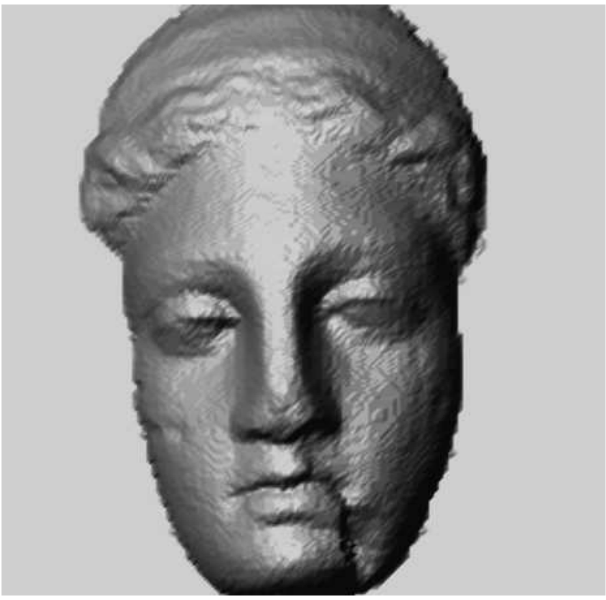

Figure 9: 3D visualisation of the surface with MSE 0.53.

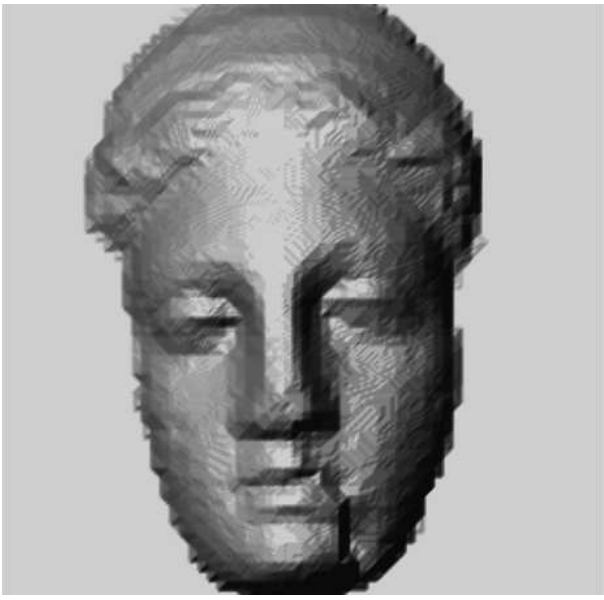

Figure 11: 3D visualisation of the surface with MSE 1.30.