3D Compression: from A to Zip
a first complete example

Thomas Lewiner[1]

Abstract: Images invaded most of contemporary publications and communications. This expansion has accelerated with the development of efficient schemes dedicated to image compression. Nowadays, the image creation process relies on multi-dimensional objects generated from computer aided design, physical simulations, data representation or optimisation problem solutions. This variety of sources motivates the design of compression schemes adapted to specific class of models.
The recent launch of Google Sketch'up and its 3D models warehouse has accelerated the shift from two-dimensional images to three-dimensional ones. However, these kind of systems require fast access to eventually huge models, which is possible only through the use of efficient compression schemes.
This work is part of a tutorial given at the XX^{th} Brazilian Symposium on Computer Graphics and Image Processing (Sibgrapi 2007).

Abstract: Imagens invadiram a maioria das publicações e comunicações contemporâneas. Esta expansão acelerou-se com o desenvolvimento de métodos eficientes de compressão da imagem. Hoje o processo da criação de imagens é baseado nos objetos multidimensionais gerados por CAD, simulações físicas, representação de dados ou soluções de problemas de otimização. Esta variedade das fontes motiva o desenho de esquemas de compressão adaptados a classes específicas de modelos.
O lançamento recente do Google Sketch'up com o seu armazém de modelos 3D acelerou a passagem das imagens bidimensionais às tridimensionais. Entretanto, este o tipo de sistemas requer um acesso rápido aos modelos 3D, possivelmente gigantes, que é possível somente usando de esquemas eficientes da compressão.
Esse trabalho faz parte de um tutorial ministrado no Sibgrapi 2007.

1 Introduction

Images surpassed the simple function of illustrations. In particular, artificial and digital images invaded most of published works, from commercial identification to scientific explanation, together with the specific graphics industry. Technical advances created supports, formats and transmission protocols for these images, and these contributed to this expansion. Among these, high quality formats requiring low resources appeared with the development

[1]Departamento de Matemática, PUC-Rio de Janeiro (`http://www.mat.puc-rio.br/~tomlew/`)
The author would like to thank the CNPq for financial support through projects MCT/CNPQ 02/2006.

of generic, and then specific, compression schemes for images. More recently drew on the sustained trend to incorporate the third dimension into images, and this motivates orienting the developments of compression towards higher dimensional images.

There exists a wide variety of images, from photographic material to drawings and artificial pictures. Similarly, higher dimensional models are produced from many sources: The graphics industry designers draw three–dimensional objects by their contouring surface, using geometric primitives. The recent developments of radiology make intense use of three–dimensional images of the human body, and extract isosurfaces to represent organs and tissues. Geographic and geologic models of terrain and underground consist in surfaces in the multi–dimensional of physical measures. Engineering usually generate finite elements solid meshes in similar multi–dimensional spaces to support physical simulations, while reverse engineering, archæological heritage preservation and commercial marketing reconstruct real objects from points.

Compression methods for three–dimensional models appeared mainly in the mid 1990's with [1] and developed quickly since then. This evolution turned out to be a technical necessity, since the size and complexity of the typical models used in practical applications increases rapidly. The most performing practical strategies for surfaces are based on the Edgebreaker of [35] and the valence coding of [26]. These are classified as connectivity–driven mesh compression, since the proximity of triangles guides the sequence of the surface vertices to be encoded. More recently, dual approaches proposed to guide the encoding of the triangle proximity by the geometry, such as done in [15].

Actually, the diversity of images requires this multiplicity of compression programs, since specific algorithms usually perform better than generic one (such as the popular Zip method), if they are well adapted. This tutorial aims at introducing the basic concepts of compression with examples on 3D models compression for an audience without prior knowledge in modeling or compression. The examples and algorithms are chosen from works published in the Sibgrapi.

2 Information Representation

We would like first to briefly introduce what we mean by compression, in particular the relation of the abstract tool of information theory [39, 33, 9], the asymptotic entropy of codes [9, 38, 10] and the practical performance of coding algorithms [38, 28, 27, 23]. General references on data compression can be found in [54].

2.1 Coding

Source and codes. Coding refers to a simple translation process that converts symbols from one set, called the *source* to another, this last one being called the set of *codes*. The conversion can then be applied in a reverse way, in order to recover the original sequence of symbols, called *message*. The purpose is to represent any message of the source into a more convenient way, typically a way adapted to a specific transmission channel. This coding can intend to reduce the size of the message [54], for example for compression applications, or on the contrary increase its redundancy to be able to detect transmission errors [8].

Enumeration. A simple example coder would rely on enumerating all the possible messages, indexing them from 1 to n during the enumeration. The coder would then simply assign one code for each message. In practise, the number of possibilities is huge and difficult to enumerate, and it is hard to recover the original message from its index without enumerating again all the possible messages. However, this can work for specific cases [44]. These enumerative coders give a reference for comparing performance of coders. However, in practical cases, we would like the coding to be more efficient for the most frequent messages, even if the performance is altered for less frequent ones. This reference will thus not be our main target.

Coder performance. Two different encodings of the same source will in general generate two coded messages of different sizes. If we intend to reduce the size of the message, we will prefer the coder that generated the smallest message. On a specific example, this can be directly measured. Moreover, for the enumerative coder, the performance is simply the logarithm of the number of elements, since a number n can be represented by $\log(n)$ digits. However, this performance is hard to measure it for all the possible messages of a given application. [39], [33] and [9] introduced a general tool to measure the asymptotic, theoretic performance of a code, called the *entropy*.

2.2 Information Theory

Entropy. The entropy is defined in general for a random message, which entails message generators as symbol sources or encoders, or in particular to a specific message (an observation) when the probabilities of its symbols are defined. If a random message $\mathbf{m}$ of the code is composed of n symbols $\mathbf{s}_1 \dots \mathbf{s}_n$, with probability $p_1 \dots p_n$ respectively, then its entropy $\mathfrak{h}(\mathbf{m})$ is defined by $\mathfrak{h}(\mathbf{m}) = \sum_i -p_i \log(p_i)$. As referred in [9], this definition fits communication systems, but it is only one possible measure that respects the following criteria:

1. $\mathfrak{h}()$ should be continuous in the p_i.
2. If all p_i are equal, $p_i = \frac{1}{n}$, then $\mathfrak{h}()$ should increase with n, since there are more possible messages.
3. If the random message $\mathbf{m}$ be broken down into two successive messages $\mathbf{m}_1$ and $\mathbf{m}_2$, then $\mathfrak{h}(\mathbf{m})$ should be the weighted sum of $\mathfrak{h}(\mathbf{m}_1)$ and $\mathfrak{h}(\mathbf{m}_2)$.

Huffman coder. [38] introduced a simple and efficient coder that writes each symbol of the source with a code of variable size. For example, consider that a digital image is represented by a sequence of colours $\mathbf{s}_{black}, \mathbf{s}_{red}, \mathbf{s}_{darkblue}, \mathbf{s}_{lightblue}, \mathbf{s}_{white}$. A simple coder will assign a symbol to each colour, and encode the image as the sequence of colours. This kind of coder will be called next an *order 0 coder*.

If the image is a photo of a seascape, as the one of Figure 1, the probability to have blue

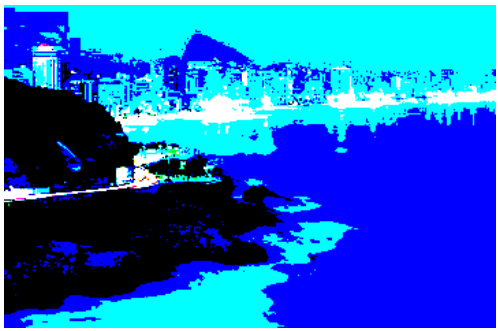


Figure 1. Huffman coding relies on the frequency of symbols, here the pixel colours.

colours in the message will be higher than for red colours. Huffman proposed a simple way to encode with less bits the more frequent colours, here blue ones, and with more bits the less frequent symbols. Consider that each of the colour probabilities is a power of 2: $p_{black} = 2^{-3}, p_{red} = 2^{-4}, p_{darkblue} = 2^{-1}, p_{lightblue} = 2^{-2}, p_{white} = 2^{-4}$.
These probabilities can be represented by a binary tree, such as each symbol of probability 2^{-b} is a leaf of depth b in the binary tree. Then each symbol is encoded by the left (0) and right (1) choices to get from the root of the tree to that symbol. The decoding is then performed by following the left and right codes until reaching a leaf, and the symbol of that leaf is a new element of the decoded message. In that context, the probability of each left and right operation is $\frac{1}{2}$, which maximises the entropy ($\mathfrak{h}(\mathbf{m}) = 1$), i.e., the theoretical performance.

Entropy coder. The original Huffman code also worked out for general probabilities, but without maximising the entropy. It uses a greedy algorithm to choose how to round off the probabilities towards powers of 2 [38]. However, Shannon proved that it is asymptotically possible to find a coder of maximum entropy [9], and that no other coder can asymptotically work better in general. This is the main theoretical justification for the definition of $\mathfrak{h}()$. [38] introduced a simpler proof of that theorem, by grouping sequence of symbols until their probability become small enough to be well approximated by a power of 2.

2.3 Levels of Information

In practise, although the entropy of a given coder can be computed, the theoretical entropy of a source is very hard to seize. The symbols of the source are generally not independent, since they represent global information. In the case of dependent symbols, the entropy would be better computed through the Kolmogorov complexity [31]. For example, by increasing the contrast of an image, as human we believe that we loose some of its details, but from the information theory point of view, we added a (mostly) random value to the colours, therefore increasing the information of the image.

An explanation for that phenomenon is that the representation of an image as a sequence of colours is not significant to us. This sequence could be shuffled in a deterministic way, it would not change the coding, but we would not recognise anymore the information of the image. In order to design and evaluate an efficient coding system, we need to represent the **exact** amount of information that is needed for our application, through an **independent** set of codes. If we achieve such a coding, then its entropy can be maximised through a universal coder, such as the Huffman coder or the *arithmetic coder* [28, 23].

3 Meshes and Geometry

Geometrical objects are usually represented through meshes. Especially for surfaces in the space, triangulations had the advantage for rendering of representing with a single element (a triangle) many pixels on screen, which reduced the number of elements to store. Although the increasing size of usual meshes reduced this advantage, graphic hardware and algorithms are optimised for these representations and meshes are still predominant over point sets models. Moreover, several parts of the alternative to meshes require local mesh generation, which becomes very costly in higher dimensions. Finally, meshes describe in a unique and explicit manner the support of the geometrical object, either by piecewise interpolation or by local parameterisation such as splines or NURBS.

To a real object correspond several meshes. These meshes represent the same geometry and topology, and thus differ by their *connectivity*. The way these objects are discretised usually depends on the application, varying from visualisation to animation and finite element methods. These variations make it harder to define the geometric quality of a mesh independently of the application, even with a common definition for the connectivity. Further references on the following definitions can be found in [2, 42, 45].

3.1 Simplicial Complexes

There are various kind of meshes used in Computer Graphics, Scientific Visualisation, Geometric Modelling and Geometry Processing. However, the graphic hardware is optimised for processing triangles, line segments and points, which are all special cases of *simplices*.

We will therefore focus mainly on meshes made of simplices, called *simplicial complex*, and one of its extensions to meshes made of convex elements, which we will refer as *polytopes*. This notion can be further extended to *cell complexes* [45], but these are only used for high–level modelling and we will not use them in this tutorial.

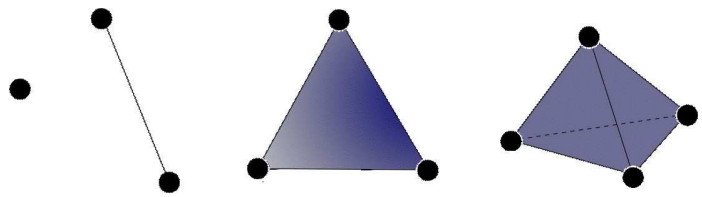


Figure 2. Simplices from dimension 0 to 3.

Simplex. A simplex is an n–dimensional analogue of a triangle. More precisely, a simplex σ^n of dimension n, or n–simplex for short, is the open convex hull of $n+1$ points $\{v_0, \ldots, v_n\}$ in general position in some Euclidean space $\mathbb{R}^d$ of dimension n or higher, i.e., such that no m–plane contains more than $(m+1)$ points. The closed simplex $\bar{\sigma^n}$ is the closed convex hull of $\{v_0, \ldots, v_n\}$. The points v_i are called the *vertices* of σ^n. For example, a 0–simplex is a point, a 1–simplex is a line segment, a 2–simplex is a *triangle*, a 3–simplex is a *tetrahedron*, and a 4–simplex is a *pentachoron*, as shown on Figure 2.

Incidence. The open convex hull of any $m < n$ vertices of σ^n is also a simplex τ^m, called an m–*face* of σ^n. We will say that σ^n is *incident* to τ^m, and denote $\sigma^n \succ \tau^m$. The 0—faces are called the *vertices*, and the 1–faces are called the *edges*. The *frontier* of a simplex σ, denoted by $\partial\sigma$, is the collection of all of its faces.

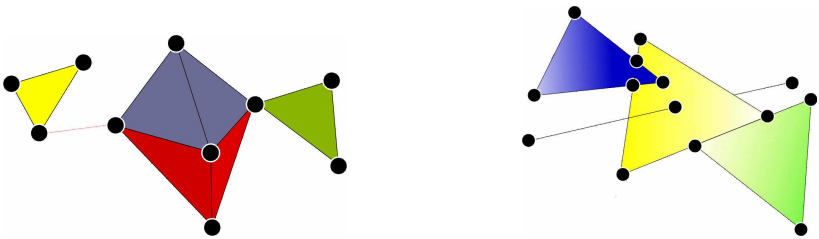


Figure 3. Simplicial complex (left) and a set of simplices not being a complex (right).

Complex. A *simplicial complex* K of $\mathbb{R}^d$ is a coherent collection of simplices of $\mathbb{R}^d$, where coherent means that K contains all the faces of each simplex ($\forall\sigma \in K, \partial\sigma \subset K$), and

contains also the geometrical intersection of the closure of any two simplices ($\forall\,(\sigma_1,\sigma_2) \in K^2, \bar{\sigma_1} \cap \bar{\sigma_2} \subset K$), as illustrated on Figure 3. Two simplices incident to a common simplex are said to be *adjacent*. The *geometry* of a complex usually refers to the coordinates of its vertices, while its *connectivity* refers to the incidence of higher–dimensional simplices on these vertices.

3.2 Pure Simplicial Complexes

Dimension. The *dimension* n of a simplicial complex K is the maximal dimension of its simplices, and we will say that K is an n–complex. A *maximal face* of a simplicial complex of dimension n is an n–simplex of K.

Euler–Poincaré characteristic. Denoting $\#_m\,(K)$ the number of m–simplices in K, the *Euler–Poincaré characteristic* $\chi\,(K^n)$ of an n–complex K^n is a topological invariant [45] defined by $\chi\,(K^n) = \sum_{m\in\mathbb{N}} (-1)^m \#_m\,(K^n)$.

Pure complexes. Roughly speaking, a complex is pure if all the visible simplices have the same dimension. More precisely, a simplicial complex K^n of dimension n is *pure* when each p–simplex of K, $p < n$, is face of another simplex of K.

Boundary. The *boundary* ∂K of a pure simplicial complex K^n is the closure of the set, eventually empty, of its $(n-1)$–simplices that are face of only one n–simplex: $\partial K^n = \left\{\sigma^{n-1} : \#\,\mathrm{lk}\left(\sigma^{n-1}\right) = 1\right\}$. The simplices of the boundary of K and their faces are called *boundary* simplices, and the other simplices are called *interior* simplices.

3.3 Simplicial Manifolds

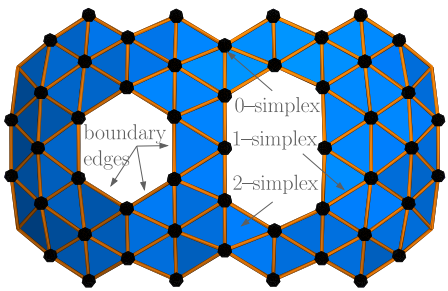


Figure 4. A surface with two bounding curves

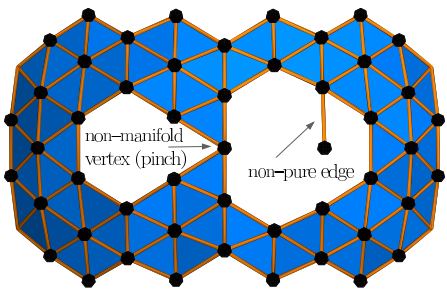


Figure 5. A non–pure 2–complex with a non–manifold vertex.

Manifolds. A simplicial n–manifold $\mathcal{M}^n$ is a pure simplicial complex of dimension n where the adjacent simplices of each interior vertex is homeomorphic to an open n–ball $\mathbb{B}^n$ and the adjacent simplices of each bounding vertex is homeomorphic to the intersection of $\mathbb{B}^n$ with an closed half–space. This implies that each $(n-1)$–simplex of $\mathcal{M}$ is the face of either one or two simplices. In particular, the boundary of an n–manifold is a $(n-1)$–manifold with an empty boundary.

Orientability. An orientation on a simplex is an ordering $(v_0, \ldots, v_n)$ on its vertices. Two orientations are equivalent if they differ by an even permutation. There are therefore two opposite orientations on a simplex. A simplicial manifold $\mathcal{M}^n$ is *orientable* when it is possible to choose a coherent orientation on all its simplices. More precisely, if $\sigma^{n-1} = (v_1, \ldots, v_n)$ is an oriented interior $(n-1)$–simplex of $\mathcal{M}^n$, face of $\rho = \sigma^{n-1} \star v$ and $\rho' = \sigma^{n-1} \star v'$, then the orientation of ρ and ρ' is coherent the orientation of ρ is equivalent to $(v, v_1, \ldots, v_n)$ and the orientation of ρ' is opposed to $(v', v_1, \ldots, v_n)$. This orientation thus defines the notion of *next* and *previous* vertex inside a triangle of a simplex.

Surfaces. For example, a 2–manifold is a surface, i.e. a simplicial complex made of only vertices, edges and triangles where each edge is in the frontier of either one or two triangles and where the boundary does not pinch. For example, Figure 4 shows an example of 2–manifold and Figure 5 illustrates a 2–complex that is neither pure nor a manifold. The topology of surfaces can be easily defined from its orientability and its Euler–Poincaré characteristic, using the Surface classification theorem [3]: Any oriented connected surface $\mathcal{S}$ is homeomorphic to either the sphere $\mathbb{S}^2$ ($\mathfrak{g}(\mathcal{S}) = 0$) or a connected sum of $\mathfrak{g}(\mathcal{S}) > 0$ tori, in both cases with some finite number $\mathfrak{b}(\mathcal{S})$ of open disks removed. The number $\mathfrak{g}(\mathcal{S})$ is called the *genus* of $\mathcal{S}$, and $\mathfrak{b}(\mathcal{S})$ its *number of boundaries*. The Euler–Poincaré characteristic $\chi(\mathcal{S})$ of $\mathcal{S}$ is equal to $\chi(\mathcal{S}) = \#_2(\mathcal{S}) - \#_1(\mathcal{S}) + \#_0(\mathcal{S}) = 2 - 2 \cdot \mathfrak{g}(\mathcal{S}) - \mathfrak{b}(\mathcal{S})$.

Dual. The dual of an n–manifold $\mathcal{M}^d$ is the manifold polytope obtained by reversing the incidence relations of its cells, i.e. creating a vertex for each n–cell of $\mathcal{M}^d$, and an m–cell for each $(n-m)$–cell of $\mathcal{M}^d$, spanning the vertices created for each n–cell of its adjacent simplices in $\mathcal{M}^d$.

4 Connectivity–Driven Compression

Since there is still no strong relation between the geometry and the connectivity of these meshes for the usual objects considered by graphics applications, dedicated compression schemes consider either that the common information can be deduced from either the connectivity or the geometry. The first option assumes that the star of a simplex has a simple geometry, which can be well approximated by simple methods such as linear interpolation. Then, the geometry can be efficiently encoded by a connectivity traversal of the mesh, leading to *connectivity–driven* compression schemes. The second option predicts the connectivity

from the geometry, and will be referred as *geometry–driven* compression schemes. In that case, the connectivity is usually better compressed, but it needs efficient geometry coding.

In this section, we will focus on the connectivity part of the compression. These connectivity–driven methods improved so much in the last decade that the compression ratio for usual surface connectivity turns around 2/3 bits per vertex. We will give a general framework for handling the critical elements of the connectivity: the topological singularities We will then focus on the Edgebreaker (Figure 6 and Figure 7) scheme, and introduce two new improvements: the handling of boundary, as a consequence of this framework for singularities, and a small improvement of the decompression algorithm. We will conclude this section with compression ratios of the Edgebreaker on usual models, and we will detail the specificities of connectivity–driven compression scheme.

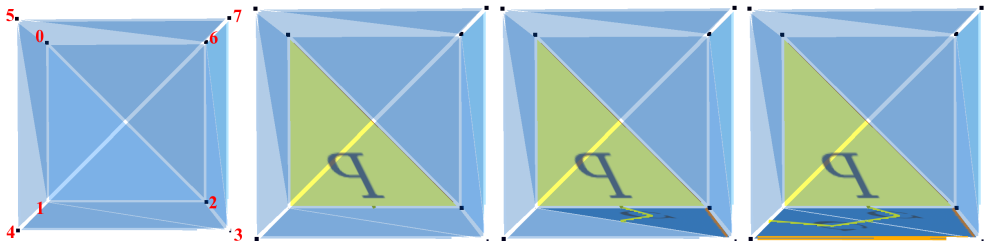
4.1 Principles

Connectivity–driven compression schemes rely on a traversal of the mesh in order to visit each vertex, and to identify it on further visits. This way, the geometry of the vertex needs to be transmitted only once, and the traversal encodes the connectivity of the mesh. This general framework suits particularly well for manifold. Most of the existing compression techniques are dedicated to surfaces, and we will focus on these algorithms. Further extensions to non–manifold cases are described in [34], while simple extensions of the most common schemes exist for solid models in [22, 29].

Connectivity–driven compression begun with cache problems in graphic cards: the rough way of transmitting triangle meshes from the main memory to the graphic card is (still) to send the three vertices of the triangle, represented by their three floating–point coordinates. Each triangle is then encoded with 96 bits! [1] proposed to represent these triangle meshes by generalised strips in order to share one or two vertices with the last triangle transmitted, reducing by at least a half the memory required previously. This mechanism uses also a small prediction scheme to optimise caching.

Then, these strips were generalised by a topological surgery approach in [5, 4]. These works introduced the most general framework for connectivity–driven compression, and has been efficiently derived into the Edgebreaker [35], and with a more flexible way into the valence coding of [26, 14]. The Edgebreaker has been extended to handle larger categories of surfaces in [46, 19], while valence coding has been tuned using the geometry in [14], discrete geometry [50]. In addition, the generated traversal of valence coding can be cleaned using [44].

With these improvements, the connectivity of usual models can be compressed with less than 3 bits per vertex. Geometry became the most expensive part, which can be reduced using prediction [26, 24] and high–quality quantisation [25, 49]. However, we will not focus here on the compression of the geometry.

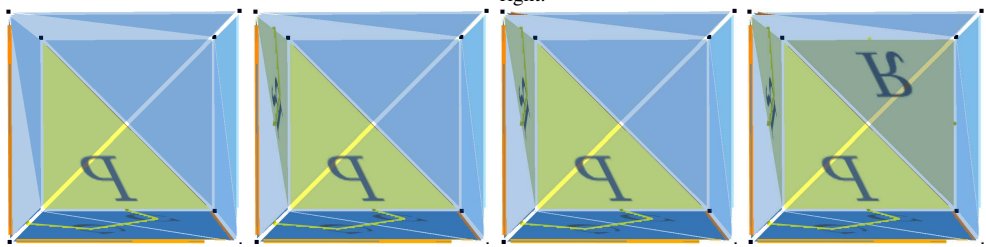


(a) Vertex labels used in the next sequence.

(b) First triangle not encoded: **P**, vertices 0, 1, 2 are marked. It will be the root of the dual tree. The traversal starts from edge 12.

(c) Since vertex 3 is unmarked, 132 is created and 3 is marked: **C**. This extends the dual tree and the primal remainder. The traversal continues on the right.

(d) Similarly, since vertex 4 is unmarked, 143 is created and 4 is marked: **C**.

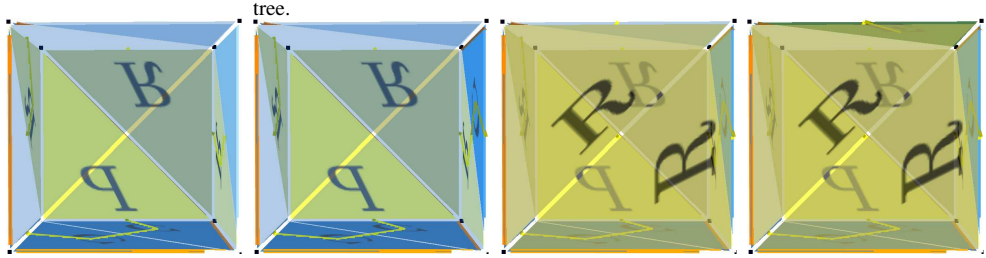


(e) Again, vertex 5 marked: **C**

(f) Since vertex 0 is marked and the right triangle is marked already, 105 is attached and the traversal continues on the left: **R**. This extends only the dual tree.

(g) **C** again: vertex 6 is marked.

(h) Again, vertex 2 is already marked and the right triangle also: **R**.



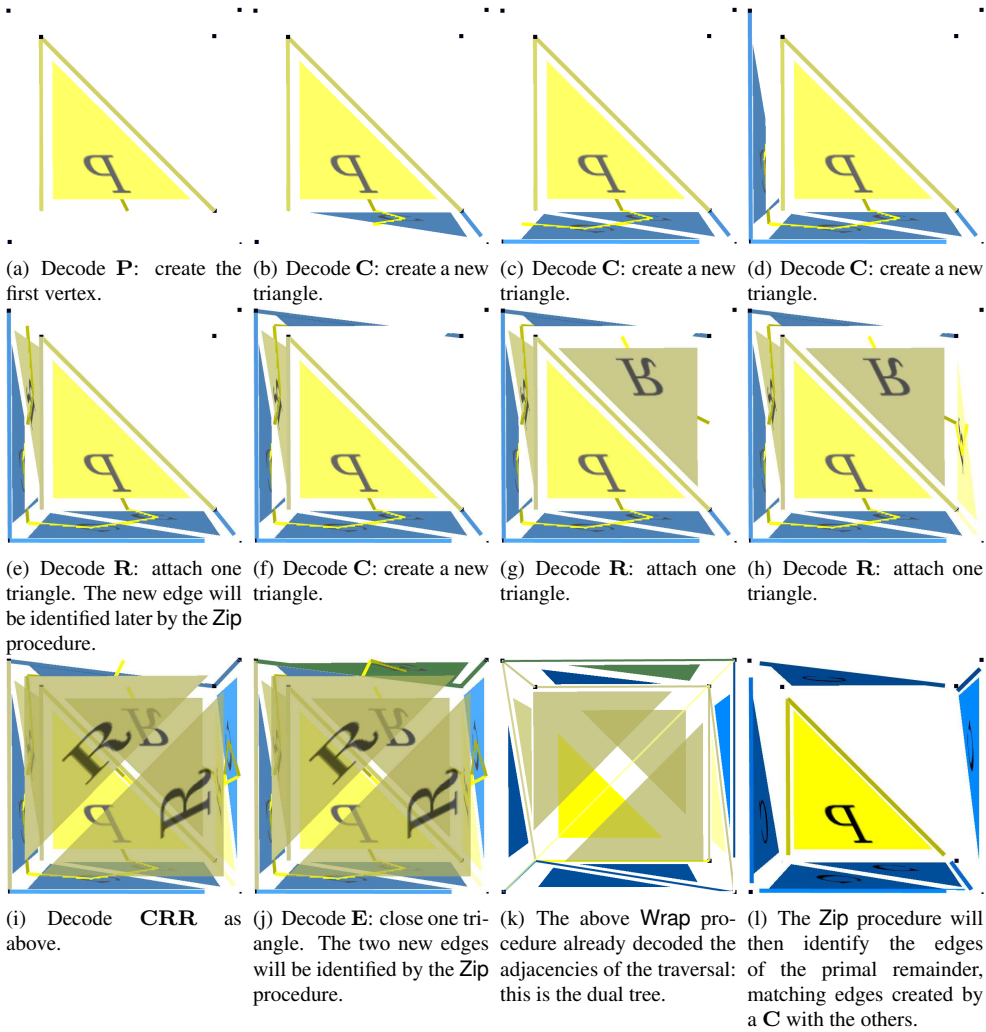
(i) Again: **R**.

(j) **C** again: vertex 7 is marked.

(k) Again, vertices 4 and then 5 are already marked, with their right triangles also: **RR**.

(l) Since vertex 6 is marked, and both the right and left triangles are marked, attach 567: **E**. This extends the dual tree only.

Figure 6. Edgebreaker compression of a triangulated cube.



(a) Decode **P**: create the first vertex.

(b) Decode **C**: create a new triangle.

(c) Decode **C**: create a new triangle.

(d) Decode **C**: create a new triangle.

(e) Decode **R**: attach one triangle. The new edge will be identified later by the Zip procedure.

(f) Decode **C**: create a new triangle.

(g) Decode **R**: attach one triangle.

(h) Decode **R**: attach one triangle.

(i) Decode **CRR** as above.

(j) Decode **E**: close one triangle. The two new edges will be identified by the Zip procedure.

(k) The above Wrap procedure already decoded the adjacencies of the traversal: this is the dual tree.

(l) The Zip procedure will then identify the edges of the primal remainder, matching edges created by a **C** with the others.

Figure 7. Wrap&Zip decompression of a triangulated cube.

4.2 Primal or dual remainders

Primal and dual graphs. The main advance of topological surgery [5] was to substitute mesh connectivity compression by graph encoding. A graph can be considered as a simplicial complex of dimension 1. Therefore, the 1–skeleton $K_{(1)}$ of any simplicial complex is a graph, called the *primal graph* of the manifold. Moreover for manifolds, the 1–skeleton of this dual manifold is also a graph, called the *dual graph* of the manifold. For example, Figure 8 represents the primal and the dual graph of a triangulated sphere.

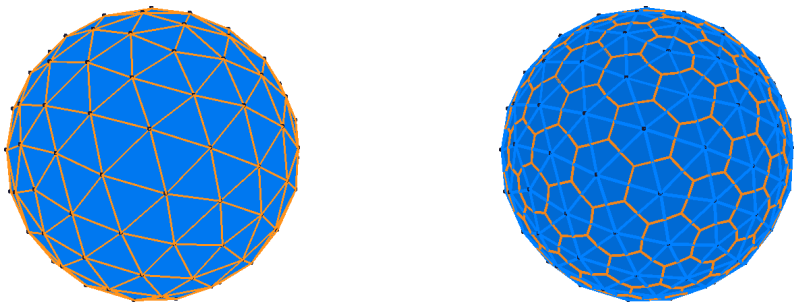


Figure 8. (left): the primal graph and (right): the dual graph of a triangulated sphere.

Tree encoding. For simplicial surfaces, the dual graph has a very nice property: each node of the graph has three incident links. Encoding the connectivity thus resumes to encoding this dual graph. In order to encode the geometry, this graph must be encoded by traversal, i.e. a spanning forest. Since each connected component can be encoded separately, we will consider only connected orientable surfaces, and the spanning forest is, in that case, a tree. This tree can be encoded from its root by enumerating for each node how many sons he has. This is the principle of both the valence coding and the Edgebreaker algorithms. The first one encodes the mesh by enumerating the valence of each node of a spanning tree in the primal graph, while the second one encodes a little more than the valence of each node of a spanning tree in the dual graph. In this last case, the valence is either 1, 2 or 3 since the nodes of the dual graph have a constant valence, which simplifies the coding.

Remainders. For clarity of the presentation, we will focus on spanning tree of the dual graph and the primal remainder, which is the focus of the Edgebreaker. What follows can be read identically by considering spanning tree of the primal graph and the dual remainder, which is the point of view of the valence coding. Consider a surface $\mathcal{S}$, with a spanning tree $\mathcal{S}^{21}$ of its dual graph. Observe that the links of $\mathcal{S}^{21}$ correspond to edges of $\mathcal{S}$. Then, consider the primal graph $\mathcal{S}^1$ (1–skeleton) of $\mathcal{S}$. Its links also correspond to edges of $\mathcal{S}$. The graph $\mathcal{S}^{01}$ having the same nodes as $\mathcal{S}^1$ and the links of $\mathcal{S}^1$ not represented in the dual spanning

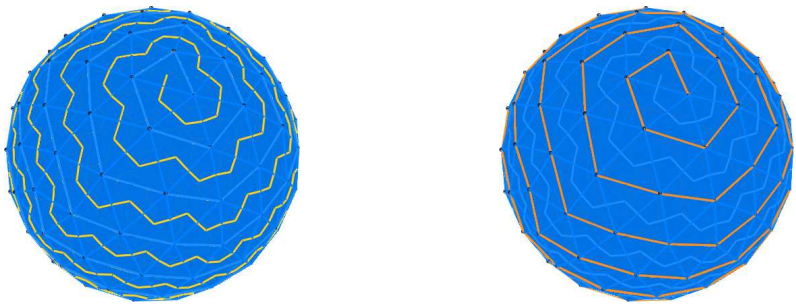


Figure 9. (left): a dual spanning tree $\mathcal{S}^{21}$ extracted from the dual graph. (right): the primal remainder $\mathcal{S}^{01}$ of $\mathcal{S}^{21}$, which is a subgraph of the primal graph.

tree $\mathcal{S}^{21}$ is called the *primal remainder* of $\mathcal{S}^{21}$. This remainder is what is left to encode after the traversal of the dual mesh, i.e. $\mathcal{S}^{21}$, has been encoded. For example, the Edgebreaker encodes this primal remainder by specific symbols for the valence 1 and 2 of the dual tree. Moreover, this primal remainder contains all the vertices of the mesh, and will therefore be used to drive the encoding of the geometry.

4.3 Topological Singularities

Topology of the remainders. If the remainder is a tree, then it can be easily encoded. The original Edgebreaker works directly in that case. However, this is not always the case, and the topology of the primal remainder actually characterises the topology of the (orientable) surface. For the dual remainder used by the valence coding, there is a detail to assert when the surface has a non–empty boundary. This process relies on a very simple calculus of the Euler characteristic of the remainder. According to Section 3.2, the Euler characteristic of a surface is given by $\chi(\mathcal{S}) = \#_2 - \#_1 + \#_0$, and according to the surface classification theorem introduces in Section 3.3, $\chi(\mathcal{S}) = 2 - 2 \cdot \mathfrak{g}(\mathcal{S}) - \mathfrak{b}(\mathcal{S})$. Since $\mathcal{S}^{21}$ is a tree with exactly one node for each of the $\#_2$ faces, it has $\#_2 - 1$ links. Therefore, the Euler characteristic of the primal remainder $\mathcal{S}^{01}$ is $\chi(\mathcal{S}^{01}) = \chi(\mathcal{S}) - \chi(\mathcal{S}^{21}) = 1 - 2 \cdot \mathfrak{g}(\mathcal{S}) - \mathfrak{b}(\mathcal{S})$. We get the same result for the case of a dual remainder.

Remainder of topological spheres. If the surface $\mathcal{S}$ is a topological sphere, then $\mathfrak{g}(\mathcal{S}) = \mathfrak{b}(\mathcal{S}) = 0$, and the remainders have Euler characteristic 1. From the Jordan curve theorem [3], the remainders are connected, since they cannot be disconnected by the corresponding spanning tree, which has no closed curve. Then, the remainder is a connected graph with Euler characteristic 1: it is a tree. This primal remainder will be easy to encode, relating topological simplicity to easy compression with connectivity–driven schemes.

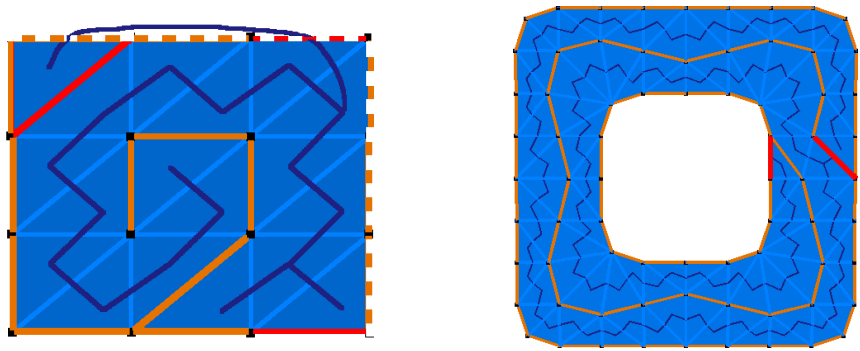


Figure 10. (left): a primal remainder on a torus (genus 1): the topmost and bottommost horizontal edges are identified, and so do the leftmost and rightmost ones. (right) a primal remainder on an annulus (two bounding curves).

Morse edges. For a generic remainder, its Euler characteristic is $1 - 2 \cdot \mathfrak{g}(\mathcal{S}) - \mathfrak{b}(\mathcal{S})$. In the case of a dual spanning tree, the primal remainder is always connected. However, for primal spanning trees on surfaces with a non–empty boundary, the dual remainder can be disconnected. This can be avoided if the primal spanning tree contains all the bounding edges of the surface, except one per boundary components to keep it as a tree. With this restriction, the remainder is a connected graph with exactly $2 \cdot \mathfrak{g}(\mathcal{S}) + \mathfrak{b}(\mathcal{S})$ independent cycles, where a cycle is a sequence of distinct adjacent links whose last one is adjacent to the first one, and where independent means that removing one link of a cycle does not break any other. For each cycle, one edge that would break it will be called a *Morse edge*, since it induces a change in the topology of the surface, Any connectivity–driven compression scheme designed for topological spheres can be extended to any orientable surface by encoding separately these Morse edges. For example, in the case of a sphere, the primal remainder is a tree, as shown on Figure 9. For a mesh with genus one or with two boundary curves, the primal remainder is a graph with two cycles, as shown on Figure 10.

5 The Edgebreaker example

The Edgebreaker scheme has been enhanced and adapted from Topological Surgery [5] to yield an efficient but initially restricted algorithm [35], which encodes the connectivity of any simplicial surface homeomorphic to a sphere with a guaranteed worst case code of 1.83 bits per triangle [21]. The Wrap&Zip algorithm introduced in [37] enhanced the original Edgebreaker decompression worst–case complexity from $O(n^2)$ to $O(n)$, where n is the number of triangles of the mesh. It decompresses the mesh in two passes, a direct and a

recursive one. It is possible to decompress it in only one pass using the Spirale Reversi algorithm of [30], but it requires to read the encoded backwards, which is not appropriate for the Huffman encoding of [21] or the arithmetic encoding. But the true value of Edgebreaker lies in the efficiency and in the simplicity of its implementations [36], which is very concise. This simple algorithm has been extended to deal with non–simplicial surfaces [51] and the compression of simplicial surfaces with handles has been enhanced in [47] using *handle data*. Because of its simplicity, Edgebreaker is viewed as the emerging standard for 3D compression [54] and may provide an alternative for the current MPEG–4 standard, which is based on the Topological Surgery approach [5].

In this section, we will enhance the Edgebreaker compression for surfaces with a non–empty boundary. [21] encoded these surfaces by closing each bounding curve with a dummy vertex. This is a very simple but expensive solution: first, it requires encoding each bounding edge with a useless triangle; second, it requires extra code to localise the dummy vertex; and third, it gives bad geometrical predictors on the boundary. The original solution of [35] however encodes bounding curves a special symbol containing their length, which solves the first item but does not describe explicitly the topology of the surface, and gave bad prediction on the boundary. As we introduced in [19], we use directly the handle data to encode the boundaries, which solves the above mentioned problems and enhances the compression ratio. We will also introduce a small acceleration to the Wrap&Zip procedure in order to avoid the recursion, accelerate the decompression and reduce the memory use.

5.1 CLERS encoding

Gate based compression. Edgebreaker encodes the connectivity of the mesh by producing the stream of symbols taken from the set $\mathbf{C}$,$\mathbf{L}$,$\mathbf{E}$,$\mathbf{R}$,$\mathbf{S}$, called the *clers stream*. It traverses spirally the dual graph of a surface in order to generate a spanning tree. At each step, a decision is made to move from one triangle t to an adjacent triangle t' through an edge e' called the *gate*. The vertex v of t not contained in the previous gate e is called the *apex* of the gate. This decision depends on the previously visited triangles, which are marked together with their incident vertices.

Right–first traversal. The spiral traversal means that the next triangle is chosen to be the one on the right if not marked, where the right triangle means that the link of the new gate e' contains the vertex next to the apex v of the previous gate e (see Section 3.3 for the definition of next). This gives a direct construction of the dual spanning tree and an order on it.

CLERS codes. The traversal is then encoded by the valences (1, 2 or 3) of the nodes of the dual spanning tree $\mathcal{S}^{21}$, and for the valence 2 case, by the current position ($\varnothing$, left or right) of the primal remainder $\mathcal{S}^{01}$ with respect to the new triangle. The corresponding symbols are stated on Table 1. The valence of the nodes of $\mathcal{S}^{21}$ can be easily detected during the traversal, using the rules of Table 1 [35].

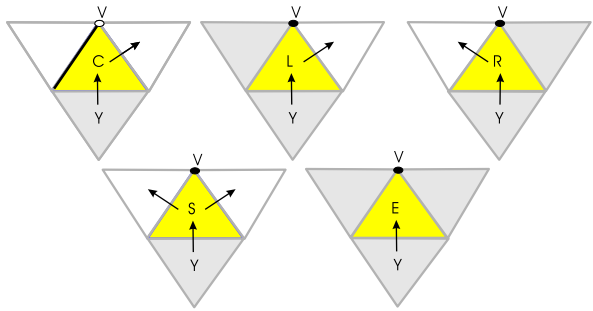
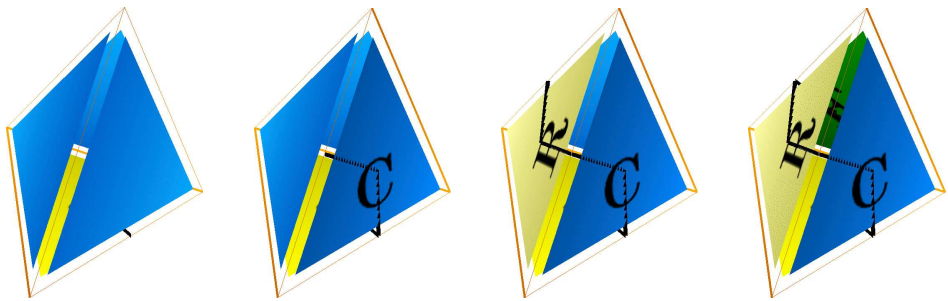


Figure 11. The Edgebreaker encoding. A **C** corresponds to a vertex **C**reation. With the outward orientation, an **L** means that the **L**eft triangle has been visited, whereas an **R** means that the **R**ight triangle has been visited. **S** stands for **S**plit, and **E** for **E**nd.

	operation	$\mathcal{S}^{21}$ val.	$\mathcal{S}^{21}$ pos.	apex	left tri.	right tri.
C	make Δ with 2 $/$ and 1 $\cdot$	2	$\varnothing$	unmarked	unmarked	unmarked
R	complete Δ with 1 right $/$	2	left	marked	unmarked	marked
L	complete Δ with 1 left $/$	right	marked	marked	unmarked	
E	complete Δ	1		marked	marked	marked
S	make Δ with 2 $/$	3		marked	unmarked	unmarked

Table 1. The CLERS codes.

Original compression. We will now describe directly the above formal presentation of the Edgebreaker. The algorithm starts by encoding the geometry of a first triangle, that will be the root of $\mathcal{S}^{21}$. In the text, we will call it a **P** triangle. The traversal begins right after with



(a) **P** (b) **C** (c) **R** (d) **E**

Figure 12. Coding of a tetrahedron: **PCRE**.

the rules of Table 1: if the apex is not marked, a **C** is encoded with the geometry of the apex, and the traversal continues on the right triangle. Otherwise, if the left triangle is marked, an **R** symbol is encoded and the traversal continues on the right triangle. Similarly, if the right triangle is marked, an **L** symbol is encoded and the traversal continues on the left triangle. If none of the triangles are marked (but the apex is), an **S** symbol is encoded. The traversal splits since the spanning tree has a branching here. The first traversed branch begins with the right triangle, and continues on the left one when the first branch ends. Finally, if both adjacent triangles are marked, the branch ends with an **E** symbol. This branching mechanism can be simply implemented with an **S** stack that stores the left triangle of each **S** triangle.

5.2 Fast decompression

Wrap&Zip decompression. The original Wrap&Zip procedure of [37] decodes the clers stream in two passes. The Wrap simply decodes the dual spanning tree, with the geometry of each vertex at each **C** symbol. It decodes the **S**/**E** branchings and positions correctly the adjacent triangles using the branching order and the distinction between the **C** or **L** symbols and the **R** symbols. Then, the Zip procedure completes this spanning tree to obtain the dual graph. If the surface has the topology of a sphere, then there is enough information to recover the entire dual graph, as we will see next. The procedure is very similar to the enumeration of [43]: it looks for the star of each vertex v, and if its star is not closed, and if the two bounding edges of its star are associated to a **C** on one side, and on another symbol on the other side, then these two edges are identified. A recursive implementation of this procedure is necessary to achieve a linear complexity, using the fact that the closure of a star usually allows closing adjacent stars, except when reaching an **L** or **E** symbol.

Fast Zip. Actually, the Zip procedure is a recursive traversal of the dual spanning tree, and it closes the stars from the leaves to the root. Actually, since the algorithm just built the spanning tree, there is no need to traverse it all to find the leaves. It is sufficient to use a **C** stack during the Wrap that stores each **C** triangle. Popping the **C** stack reads it in the reverse way, and the algorithm closes one star at each **C** symbol, and three for each **P** symbol, instead of trying all triangles. This spares half of the tests. Moreover, stars can be closed at some **R** and **E** symbols during the Wrap. This can be used to keep the size of the **C** stack small, and allows a better usage of the multiway geometry prediction of [24].

5.3 Topology encoding

Handle $\mathbf{S}^h$ symbols. As we said earlier, if the surface $\mathcal{S}$ has genus $\mathfrak{g}(\mathcal{S}) > 0$, the primal remainder $\mathcal{S}^{01}$ is not a tree anymore, as illustrated on Figure 13. For a surface with an empty boundary, $\mathcal{S}^{01}$ has $2 \cdot \mathfrak{g}(\mathcal{S})$ cycles. These cycles can be simply detected during the traversal

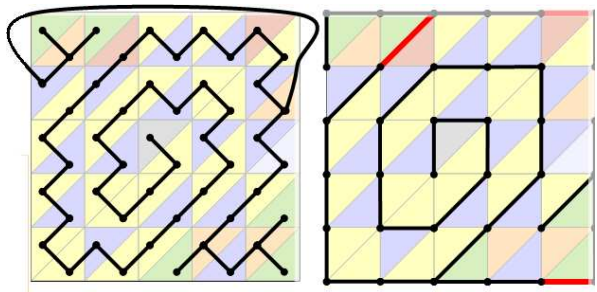
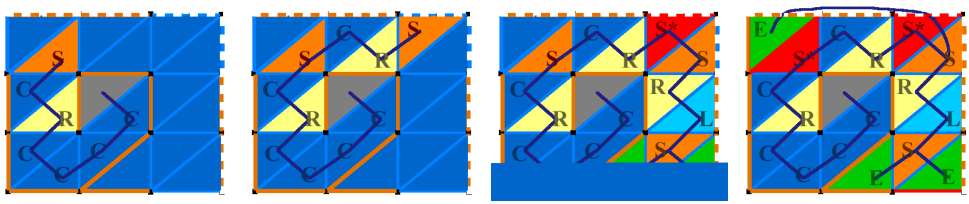


Figure 13. Dual tree generated by the Edgebrealer traversal and the primal remainder, with the two Morse edges in red.

and efficiently encoded using [47], while preserving the original Edgebreaker compression scheme. These cycles correspond to a branching, and thus to an **S** symbol. However, the two branchings induced by each genus of the surface loops back, and the left edge of the **S** triangle is visited before its right branch ends. During the execution, this is easily detected when popping the **S** stack containing the triangles left to **S** symbols: if the top of the **S** stack is not marked, the algorithm continues as normally. If the left triangle was marked, the **S** symbol actually corresponds to a handle, and will be marked as a *handle* $\mathbf{S}^h$ symbol. This symbol is encoded as a normal **S**, and special information identifying this $\mathbf{S}^h$ symbol is encoded in the *handle data*. In order to decompress handles directly, the position of the left triangle in the clers stream can be encoded, for example by the number of **S** symbol that preceded the $\mathbf{S}^h$ symbol and by the number of **R**, **L** and **E** symbols that preceded the left triangle, since *handle* $\mathbf{S}^h$ triangles are obviously closed by only these kind of triangles. These numbers can be encoded by differences to spare even more space.



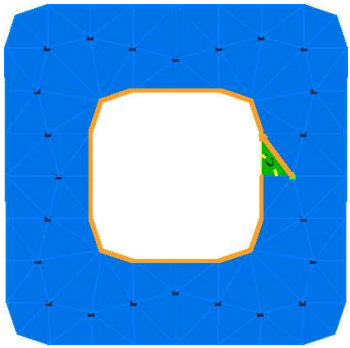
(a) Reaching first **S** triangle (b) Reaching second **S** triangle (c) The lower–right **E** triangle closes the handle. (d) The upper–left **E** triangle closes the handle.

Figure 14. Coding of a torus: the creation of two *handle* **S** triangles: the first and the second **S** symbols.

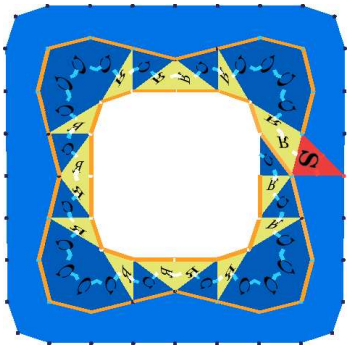
Example. To illustrate the algorithm, consider the triangulated torus of Figure 14, where the edges on the opposite sides of the rectangle are identified. This simplicial complex can be embedded in $\mathbb{R}^3$. The Edgebreaker compression algorithm encodes the connectivity of the mesh though the following clers stream: **CCCCRCSCRSSRLSEEE**, completed with the following handle data: $0\text{—}4^-$,$0\text{—}3^+$. There are four triangles labelled with an **S** symbol. The left triangles of the two last ones are visited when popping the **S** stack. On the contrary, the two first ones are visited before the being popped out of the **S** stack. These two triangles are detected as handle $\mathbf{S}^h$ symbols. This is encoded in the handle data as follows: the first handle $\mathbf{S}^h$ symbol is also the first **S** symbol, and the first number encoded is therefore 0. There are four possible matches (**R**, **L** and **E** symbols) for its left triangle before the good one, which is encoded by the 4. Since it is an **E** triangle, it can be glued on both sides, and the left side is indicated by the $\epsilon = ^-$. The encoding is done the same way for the second handle $\mathbf{S}^h$ symbol.

First bounding curve. This scheme can be extended to boundary compression, since they correspond to the same Handle symbols. Using the handle data to encode boundaries is then more coherent, gives a direct reading of the surface topology through this handle data even before decoding the mesh, and allows a specific prediction scheme for boundaries. Consider first a connected surface $\mathcal{S}$ with one bounding curve. Suppose that we close it by adding a face incident to each bounding edge of $\mathcal{S}$, called the *infinite face*. The resulted surface $\mathcal{S}^+$ has no boundary, and can almost be encoded by the previous algorithm. However, the infinite face is not a triangle. In the same way that the **P** triangles are not explicitly encoded, we will not encode this infinite face, and start the compression directly one of its adjacent triangle. As in the original Edgebreaker algorithm, we encode and mark first all its vertices, e.g., all the vertices belonging to the boundary of $\mathcal{S}$. Then, for the first boundary, we only need to know if the surface component has a boundary or not.

Boundary $\mathbf{S}^b$ symbols. Now, consider a connected surface has more than one bounding curve. Then, we distinguish arbitrarily one of them as the first boundary and the encoding uses the technique of the last paragraph. During the traversal, we label each triangle touching a new bounding curve as a boundary $\mathbf{S}^b$ triangle. As for handles, we encode it as a normal **S** symbol in the clers handle, and specify that it is a boundary $\mathbf{S}^b$ symbol in the handle data. To distinguish with handle $\mathbf{S}^h$ symbols, their first number is negative. Also, due to the orientation of the bounding curve, the left triangle is always glued on its left side, and we do not need to specify the last $\epsilon = ^+$ or $\epsilon = ^-$, and we can avoid counting the **L** symbols to localise it. From the Euler characteristic, we know that there is exactly one boundary $\mathbf{S}^b$ symbol per bounding curve. On Figure 15, the only *handle* **S** triangle is the first triangle with a vertex on the internal boundary that we encounter during the traversal. As said before, there are $2 \cdot \mathfrak{g}(\mathcal{S}) + \mathfrak{b}(\mathcal{S}) - 1$ such *handle* **S** triangles for each surface component with genus $\mathfrak{g}(\mathcal{S})$ and $\mathfrak{b}(\mathcal{S})$ bounding curves.



(a) The first triangle is chosen adjacent to a boundary. The vertices of the central infinite face are encoded.



(b) An unmarked boundary is reached: the corresponding **S** triangle is a *boundary* **S** triangle.

Figure 15. Coding of an annulus: initialisation and creation of *boundary* **S** triangles.

Multiple components. The compression processes successively each surface component. When the component has no boundary, the compression encodes explicitly the vertices of the first triangle (uncoded **P** symbol). Otherwise, it encodes the vertices of the first bounding curve. In practise, we only need to transmit the number of components with boundary of $\mathcal{S}$. Then we transmit first all the components with a non–empty boundary, and the other ones.

5.4 Compression algorithms

The compression scheme then decomposes in handling the multiple components and their first boundaries (Algorithm 2: compress), compress each component by the dual spanning tree traversal (Algorithm 5: traverse). The handles are tested along the traversal with Algorithm 1: check handle. The whole process is linear and performed in one pass only.

Algorithm 1 check handle(t): check if triangle t is left to a $\mathbf{S}^h$ triangle

1: **if not** is boundary(t.right) **and** t.right.mark $\notin \{true, false\}$ **then** *// handle $\mathbf{S}^h$ triangle to the right*
2: write$\left(handle, t.\text{right.mark} - \#^{+}_{\mathbf{RE}}\right)$ *// write the handle data*
3: **if not** is boundary(t.left) **and** t.left.mark $\notin \{true, false\}$ **then** *// handle $\mathbf{S}^h$ triangle to the left*
4: write$\left(handle, t.\text{left.mark} - \#^{-}_{\mathbf{LE}}\right)$ *// write the handle data*

Algorithm 2 compress($\mathcal{S}$): compress separately each component of $\mathcal{S}$

```
1:  𝔟* ← 0                                   // counts number of components with boundary
2:  for all vertices v ∈ S do                // reset marks
3:      v.mark ← is boundary (v)             // mark boundary vertices
4:  for all triangles t ∈ S do               // compress components with boundary first
5:      if not t.mark and is boundary(t) then   // not boundary or already encoded
6:          write boundary (t)               // encode boundary
7:          traverse(t)                      // component compression
8:          𝔟* ← 𝔟* + 1                      // one more component with boundary
9:  for all triangles t ∈ S do               // compress the other components
10:     if not t.mark then                   // not already encoded
11:         t.mark ← true                    // mark P triangle
12:         for all vertices v ∈ ∂t do        // encode the 3 vertices of the P triangle
13:             write vertex (v)             // encode the geometry of v
14:             v.mark ← true                // mark the vertex
15:         traverse(t.right)                // component compression
16: write(handle, 𝔟*)                        // write the number of components with boundary
```

Algorithm 3 decompress(streams): decompress separately each component

```
1:  repeat
2:      s − t^ϵ ← read(handle)               // read handle data
3:      if s > 0 then                        // handle S^h symbol
4:          glue(s, t, ϵ)                    // glue the handle on side ϵ before the decompression
5:      else                                 // boundary S^b symbol
6:          glue(−s, t, ⁻)                   // close the bounding curve before the decompression
7:  until end of file(handle)                // passed the last couple of data
8:  𝔟* ← s                                   // last handle data counts number of components with boundary
9:  stack Cstack ← ∅                         // stack of the C and boundary S^b triangles
10: wrap(𝔟*, Cstack)                         // wrap using the clers stream
11: fast zip(Cstack)                         // closes the stars of the primal remainder
12: read geometry(Cstack)                    // reads the geometry of the surface
```

Algorithm 4 fast zip($Cstack$): decompress one primal remainder

```
1:  stack Cstack' ← ∅                        // reverse copy of the C stack for the geometry
2:  while Cstack ≠ ∅ do                      // traverse the stack
3:      t ← Cstack.pop ()                    // pop the next element of the C stack
4:      Cstack'.push (t)                     // copy the C stack
5:      if t ⩾ 0 then                        // not a boundary triangle
6:          close star(t)                    // close the star of the next vertex
7:      Cstack ← Cstack'                     // returns the copy of the C stack
```

Algorithm 5 traverse(t): encode one component starting from triangle t

```
1: stack Sstack ← ∅                                   // stack of the triangles left to S symbols
2: repeat
3:    t.mark ← true                                   // mark current triangle
4:    v ← t.apex                                      // orient the triangle from its apex
5:    if v.mark = false then                          // C triangle
6:       write vertex (v)                             // encode the geometry of v
7:       v.mark ← true                                // mark the vertex
8:       write symbol (C)                             // encode the clers code: C
9:       t ← t.right                                  // spiral traversal to the right
10:   else if is boundary(t.right) or t.right.mark then   // right triangle visited
11:      if is boundary(t.left) or t.left.mark then   // E triangle
12:         write symbol (E)                          // encode the clers code: E
13:         check handle (t)                          // check if it is the left triangle of a S^h triangle
14:         repeat
15:            if Sstack = ∅ then                     // end of compression
16:               return                              // exit the external repeat loop
17:            t ← Sstack.pop                         // pop the S stack
18:         until not t.mark                          // skip left of a handle S^h triangle
19:      else                                         // R triangle
20:         write symbol (R)                          // encode the clers code: R
21:         check handle (t)                          // check if it is the left triangle of a S^h triangle
22:         t ← t.left                                // break in spiral traversal: to the left
23:   else if is boundary(t.left) or t.left.mark then   // L triangle
24:      write symbol (L)                             // encode the clers code: L
25:      check handle (t)                             // check if it is the left triangle of a S^h triangle
26:      t ← t.right                                  // spiral traversal to the right
27:   else                                            // S triangle
28:      write symbol (S)                             // encode the clers code: S
29:      if is boundary(v) then                       // boundary S^b triangle
30:         write boundary (t)                        // encode boundary
31:         t.mark ← −#S                              // mark for the handle data
32:      else                                         // normal S or handle S^h triangle
33:         t.mark ← #S                               // mark for the handle data
34:      Sstack.push (t.left)                         // push the left triangle on the S stack
35:      t ← t.right                                  // spiral traversal to the right
36: until true                                        // infinite loop
```

Algorithm 6 wrap($\mathfrak{b}^*, Cstack$): decompress the dual trees

```
1:  #2 ← 0                                                  // initialisation
2:  stack Sstack ← ∅                                        // stack of the triangles left to S symbols
3:  repeat                                                  // components loop
4:    if b* > 0 then                                        // component with boundary
5:      b* ← b* − 1; t ← ∅                                  // first boundary
6:      Cstack.push (−#2)                                   // push the boundary triangle for the geometry
7:    else                                                  // component with an empty boundary
8:      t ← #2                                              // M0 triangle
9:      Cstack.push (t)                                     // push the first triangle for the zip
10:     for all vertices v ∈ ∂t do                          // decode the 3 vertices of the M0 triangle
11:       read vertex (v)                                   // decode the geometry of v
12:     t ← t.right                                         // spiral traversal to the right
13:     #2 ← #2 + 1                                         // initialisation
14:   repeat                                                // decompress one component
15:     glue(t, #2)                                         // glue the next triangle eventually to the boundary
16:     s ← read symbol (clers)                             // reads the next symbol
17:     if s = C then                                       // C triangle
18:       Cstack.push (#2)                                  // push the C triangle for the zip
19:       t ← t.right                                       // orient the new triangle to the right
20:     else if s = R then                                  // R triangle
21:       t ← t.left                                        // orient the new triangle to the left
22:       tryclose star (t.apex)                            // eventually zip the right edge
23:     else if s = L then                                  // L triangle
24:       t ← t.right                                       // orient the new triangle to the right
25:     else if s = S then                                  // S triangle
26:       if not t.right.mark then                          // not a handle or boundary S symbol
27:         Sstack.push (#2.left)                           // push the S triangle for the next E
28:       else if is boundary(#2) then                      // boundary triangle
29:         Cstack.push (−#2)                               // push the boundary triangle for the geometry
30:       t ← t.right                                       // orient the new triangle to the right
31:     else if s = E then                                  // E triangle
32:       tryclose star (t.apex)                            // eventually zip the right and left edges
33:       if Sstack = ∅ then                                // end of the component
34:         break                                           // exits the component loop
35:       t ← Sstack.pop ()                                 // pop the next element of the S stack
36:     #2 ← #2 + 1                                         // next triangle
37:   until true                                            // infinite loop
38: until end of file(clers)                                // end of the clers stream
```

Model	$\|\#_0\|$	$\|\#_2\|$	[37, 21]	[35]	[19]	[35]/[19]	[37, 21]/[19]
sphere	1 848	926	3.39	3.39	3.45	0.98	0.98
violin	1 508	1 498	3.16	2.21	2.25	0.98	1.41
pig	3 560	1 843	3.26	3.24	3.13	1.03	1.04
rose	3 576	2 346	3.37	2.95	2.64	1.12	1.28
cathedral	1 434	2 868	2.25	1.00	0.19	5.27	11.86
blech	7 938	4 100	3.25	3.18	2.40	1.33	1.35
mask	8 288	4 291	3.19	3.12	1.93	1.62	1.65
skull	22 104	10 952	3.51	3.51	3.30	1.06	1.06
bunny	29 783	15 000	3.36	3.34	1.27	2.62	2.64
terrain	32 768	16 641	3.03	3.00	0.40	7.43	7.51
david	47 753	24 085	3.45	3.85	3.07	1.25	1.12
gargoyle	59 940	30 059	3.28	3.27	2.11	1.55	1.55

Table 2. Comparative results on different models. The sizes are expressed in bit per vertex.

5.5 Decompression algorithms

The decompression is performed in 3 passes, controlled by Algorithm 3: decompress. The first pass decodes the dual spanning tree (Algorithm 6: wrap), which is further zipped using the backward sequence of $\mathbf{C}$ symbols (Algorithm 4: fast zip). The compression described here encodes boundary curves, which improves prediction for the interior. However this means that the size of the boundary is not known to the decoder at the first pass, and the geometry must be decoded in a posterior step. This pass could be done at the wrap stage if we encode the boundaries when they are closed, or if we encode the geometry of the bounding curves in a separate stream.

6 Performances

We presented in this section the fundamental concepts of connectivity–driven compression. In particular, we focused on an extension of the Edgebreaker algorithm, which handles manifold surfaces of arbitrary topology. The complexity of the compression and the decompression are both linear in execution time and memory footprint, independently of the maximal number of the active elements during the execution. However, the decompression still requires two passes, which makes it harder to stream.

There are various ways of representing a geometrical object, even for simplicial surfaces. For specific type of meshes, some algorithms show better performances than other ones. This distinction is one of the main shifts from the MPEG compression [17] to the MPEG–4 one [41], which for example encodes differently human faces than landscapes. Although it is difficult to distinguish with precision classes of meshes and to predict exactly

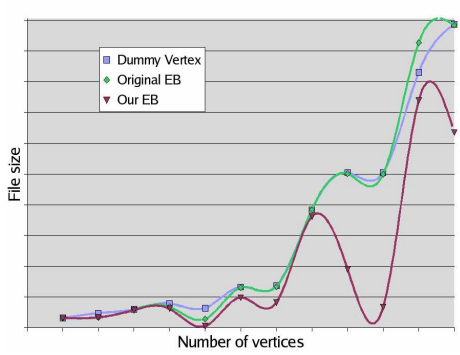
the behaviour of compression algorithms on these, we will try to get an intuition of which characteristics of a mesh are well suited for connectivity–driven compression schemes, and in particular for the Edgebreaker.

6.1 Compression Rates

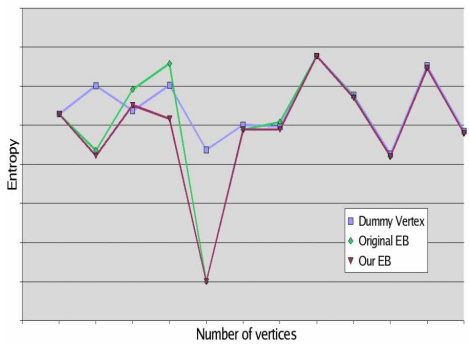
Experimental results for the Edgebreaker are recorded on Table 2 and Figure 16. We compared with the original Edgebreaker implementation with the Huffman encoding of [21] and the border handling of [37], and the encoding of [19] with the simple arithmetic coder of [53]. However, the entropy of [19] is always better than the other implementations of Edgebreaker, as shown on Figure 16(b). A compression ratio of a few bits per vertex, or even less, is a general order for efficient connectivity–driven compression schemes.

6.2 Good and bad cases

Topology–dependent applications. For the extended Edgebreaker of [19], the separate handle data informs directly the application of the topology of the mesh. Many simple parameterisations, texturing or remeshing applications work only for closed surfaces without handle. The handle data can be used to call a preprocessing step for simplifying the topology before using these kind algorithms. For the Edgebreaker, this handle data is not an overhead, since encoding the handle and boundary S symbols as a true/false code on the clers string is in the best case logarithmic, which is equivalent to the handle data.



(a) Size of the compressed file vs complexity of the model.



(b) Entropy vs complexity of the model.

Figure 16. Comparison of the final size and entropy: for the range encoder, those parameters depends more on the regularity than on the size of the model.

Regular connectivity. The valence coding of [26, 50] encodes particularly well meshes where the vertices have a uniform valence. This can be obtained by subdivision [52, 6] or remeshing [12, 13]. Remeshing can be done also to improve the Edgebreaker compression using the Swingwrapper of [40]. Without these regularisations, valence coding based algorithms have better performance when the connectivity is locally regular, whereas the Edgebreaker performs better on irregular meshes or meshes with a global regularity, such as those obtained by subdivision algorithms or with some self–similar connectivity. Meshes with a very irregular connectivity would be better encoded by enumeration methods of [43, 44].

Regular geometry. The geometry of the mesh is not directly considered in connectivity–driven compression, and therefore geometry–based compression will outperform these methods for the connectivity compression of meshes with a regular geometry. However, the geometry can be used to predict the connectivity, which works specifically when the geometry is regular. This has been done for the valence coding in [14, 11] and in [32] for the Edgebreaker.

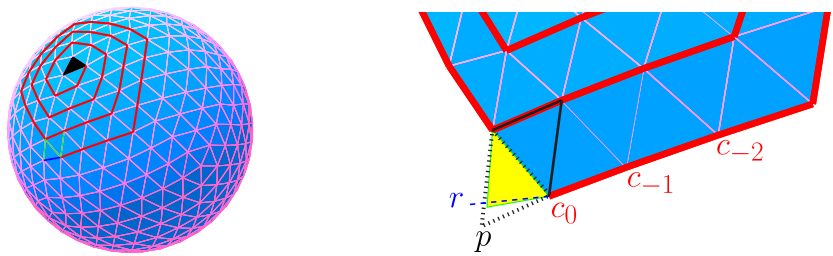


Figure 17. The Edgebreaker cuts the compressed surface along a curve in the space. An extrapolation of this curve is used to enhance the parallelogram predictor. The predictor uses the parallelogram predictor to guess the distance from the last vertex of the curve, and rotates this estimation according to the approximating curve.

Geometry prediction. Geometry prediction uses already decoded vertices to estimate the next vertex to be decoded, asserting that the geometry is locally regular. For connectivity–driven schemes are usually based on the parallelogram predictor of [26]. It can be enhanced by using more than one parallelogram to estimate the new position, as described in [24]. This is particularly well adapted to the valence coding since the traversal can be adapted to the prediction. For the Edgebreaker, the parallelogram can be distorted to adapt to local mean curvature of the surface, as in [11], or to torsion and curvature of the primal remainder, as described in [18] and on Figure 17.

Low resource applications. The Edgebreaker uses a deterministic traversal, independent of geometry considerations. Although this is less flexible for geometry prediction enhancements, it gives a very simple algorithm. Moreover, compared to the valence coding schemes that needs to maintain sorted active boundaries along compression and decompression, the Edgebreaker just needs a stack of past **S** symbols. The Edgebreaker thus requires much less memory for the execution, and spares a constant sort, which can become expensive. More generally, connectivity–driven compression schemes are easy to implement and quick to execute.

7 Next steps

The diversity of images requires a multiplicity of compression programs, since specific algorithms usually perform better than generic one (such as the popular Zip method), if they are well adapted. In particular, the simple example of Edgebreaker can be extended in many ways, to address specific issues of particular applications. Even with the few notions of this tutorial, it is feasible to improve the state-of-the art in 3D compression. In particular, the following directions may be promising:

Non–simplicial meshes. Connectivity–driven compression schemes are easier on simplicial meshes, since the dual graph has a constant valence. Most of the mesh compression algorithms for surfaces can be interpreted as a simplicial encoding preceded by a triangulation of each face. This triangulation is done in a canonical way from the traversal, and the decoder just need to know the degree of the triangulated faces. For example, the valence coding can be extended by encoding simultaneously the vertex valences and the face degrees, as in [14], and the Edgebreaker codes can be combined in a predictable way using the codes of [51].

Non–manifold meshes. Extending these methods to non–manifold meshes directly is a hard task. The usual method consists in cutting the non–manifold surface into manifold pieces, using the techniques of [34], encoding the manifold parts as separate components, and then encoding the cut operations that were performed. The encoding of cut operations can be done directly as in the handle data, or more carefully by propagating the curves formed by the non–manifold edges.

Higher dimensions. For solid meshes, the Edgebreaker compression has been directly extended to tetrahedral meshes in [48, 22], and the valence coding has been extended in [29]. The principles are the same, but the encoding needs some extra information to complete the intermediate dimension between the spanning tree and the remainders. This extra information has necessarily some expensive parts to encode, similar to the handle **S** symbols that are necessary to glue distant parts of the traversal. Minimising this extra information is an NP–hard

problem, as proved in [20]. For higher dimensions, the combinatory of mesh connectivity makes it difficult to find a concise set of symbols for coding, or a good statistical model for them as was done for surfaces in [21]. However, for high codimensions, the connectivity remains simple while the geometry can be efficiently predicted. Seen from the other side, this means that for low codimension, geometry–based coding can be very efficient, which is where isosurface compression outperforms any connectivity–based compression.

Robustness. The Edgebreaker is robust in the sense that it handles general manifold surfaces. However, it is not particularly robust with a noisy transmission, where the clers codes can be altered. In that case, the grammar inherent to these codes can be used to detect transmission errors, but not directly to correct them.

Deformable meshes. For animation purposes, the Edgebreaker can be used directly to compute the deformed mesh when its connectivity is constant, and using for example [25] to interpolate the geometry. Local changes in the connectivity can be further encoded using the explicit identification of vertices and triangles provided by the Edgebreaker, similarly to the description of [16].

A References

[1] M. F. Deering. Geometry compression. In *Siggraph*, pages 13–20. ACM, 1995.

[2] J. R. Munkres. *Elements of algebraic topology*. Addison-Wesley, Menlo Park, 1984.

[3] M. A. Armstrong. *Basic topology*. McGraw–Hill, London, 1979.

[4] G. Taubin, W. P. Horn, F. Lazarus, and J. Rossignac. Geometry coding and VRML. *Proceedings of the IEEE*, 86(6):1228–1243, 1998.

[5] G. Taubin and J. Rossignac. Geometric compression through topological surgery. *Transactions on Graphics*, 17(2):84–115, 1998.

[6] L. Velho and D. Zorin. 4–8 subdivision. *Computer Aided Geometric Design*, 18(5):397–427, 2001. Special issue on Subdivision Techniques.

[7] J. W. Alexander. The combinatorial theory of complexes. *Annals of Mathematics*, 31:219–320, 1930.

[8] R. W. Hamming. Error-detecting and error-correcting codes. *Bell System Technical Journal*, 29(2):147–160, 1950.

[9] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 1948.

[10] W. T. Tutte. *Graph theory as I have known it*. Oxford University Press, New York, 1998.

[11] H. Lee, P. Alliez, and M. Desbrun. Angle-analyzer: A triangle-quad mesh codec. In *Eurographics*, volume 21. Blackwell, 2002.

[12] P. Alliez, D. Cohen–Steiner, O. Devillers, B. Levy, and M. Desbrun. Anisotropic polygonal remeshing. In *Siggraph*. ACM, 2003.

[13] P. Alliez, É. Colin de Verdière, O. Devillers, and M. Isenburg. Isotropic surface remeshing. In *Shape Modeling International*. IEEE, 2003.

[14] P. Alliez and M. Desbrun. Valence–driven connectivity encoding of 3D meshes. In *Eurographics*, pages 480–489. Blackwell, 2001.

[15] P.-M. Gandoin and O. Devillers. Progressive lossless compression of arbitrary simplicial complexes. In *Siggraph*, volume 21, pages 372–379. ACM, 2002. Siggraph.

[16] A. W. Vieira, T. Lewiner, L. Velho, H. Lopes, and G. Tavares. Stellar mesh simplification using probabilistic optimization. *Computer Graphics Forum*, 23(4):825–838, 2004.

[17] D. le Gall. MPEG: a video compression standard for multimedia applications. *Communications of the ACM*, 34(4):46–58, 1991.

[18] T. Lewiner, J. Gomes Jr., H. Lopes, and M. Craizer. Curvature and torsion estimators based on parametric curve fitting. *Computers & Graphics*, 2005.

[19] T. Lewiner, H. Lopes, J. Rossignac, and A. W. Vieira. Efficient Edgebreaker for surfaces of arbitrary topology. In *Sibgrapi*, pages 218–225, Curitiba, Oct. 2004. IEEE.

[20] T. Lewiner, H. Lopes, and G. Tavares. Applications of Forman's discrete Morse theory to topology visualization and mesh compression. *Transactions on Visualization and Computer Graphics*, 10(5):499–508, 2004.

[21] D. King and J. Rossignac. Guaranteed 3.67v bit encoding of planar triangle graphs. In *Canadian Conference on Computational Geometry*, pages 146–149, 1999.

[22] A. Szymczak and J. Rossignac. Grow & Fold: compressing the connectivity of tetrahedral meshes. *Computer Aided Design*, 32(8/9):527–538, 2000.

[23] A. Moffat, R. Neal, and I. H. Witten. Arithmetic coding revisited. In *Data Compression*, pages 202–211, 1995.

[24] D. Cohen–Or, R. Cohen, and T. Ironi. Multi–way geometry encoding. Technical report, Tel Aviv University, 2001.

[25] O. Sorkine, D. Cohen–Or, and S. Toledo. High-pass quantization for mesh encoding. In *Symposium on Geometry Processing*, pages 42–51. ACM/Eurographics, 2003.

[26] C. Touma and C. Gotsman. Triangle mesh compression. In *Graphics Interface*, pages 26–34, 1998.

[27] A. Lempel and J. Ziv. A universal algorithm for sequential data compression. *Transactions on Information Theory*, 23(3):337–343, 1977.

[28] J. Rissanen. Generalized Kraft inequality and arithmetic coding. *IBM Journal of Research and Development*, 20:198–203, 1976.

[29] M. Isenburg and P. Alliez. Compressing hexahedral volume meshes. In *Pacific Graphics*, pages 284–293. IEEE, 2002.

[30] M. Isenburg and J. Snoeyink. Spirale reversi: reverse decoding of the Edgebreaker encoding. In *Canadian Conference on Computational Geometry*, pages 247–256, 2000.

[31] M. Li and P. M. B. Vitanyi. *An introduction to Kolmogorov complexity and its applications*. Springer, 1997.

[32] V. Coors and J. Rossignac. Delphi: geometry-based connectivity prediction in triangle mesh compression. *The Visual Computer*, 20(8–9):507–520, 2004.

[33] R. V. L. Hartley. Transmission of information. *Bell System Technical Journal*, 7:535, 1928.

[34] A. Guéziec, G. Taubin, F. Lazarus, and W. P. Horn. Converting sets of polygons to manifold surfaces by cutting and stitching. In D. Ebert, H. Hagen, and H. Rushmeier, editors, *Visualization*. IEEE, 1998.

[35] J. Rossignac. Edgebreaker: connectivity compression for triangle meshes. *Transactions on Visualization and Computer Graphics*, 5(1):47–61, 1999.

[36] J. Rossignac, A. Safonova, and A. Szymczak. 3D compression made simple: Edgebreaker on a corner–table. In *Solid Modeling International*, pages 278–283. IEEE, 2001.

[37] J. Rossignac and A. Szymczak. Wrap&zip decompression of the connectivity of triangle meshes compressed with edgebreaker. *Computational Geometry*, 14(1-3):119–135, 1999.

[38] D. A. Huffman. A method for the construction of minimum redundancy codes. In *I.R.E*, pages 1098–1102, 1952.

[39] H. Nyquist. Certain topics in telegraph transmission theory. *Transactions of the American Institute of Electrical Engineers*, 47:617–644, 1928.

[40] M. Attene, B. Falcidieno, M. Spagnuolo, and J. Rossignac. SwingWrapper: retiling triangle meshes for better Edgebreaker compression. *Transactions on Graphics*, 22(4):982–996, 2003.

[41] F. Pereira and T. Ebrahimi. *The MPEG–4 Book*. Prentice Hall, Upper Saddle River, 2002.

[42] J.-D. Boissonnat and M. Yvinec. *Algorithmic geometry*. Cambridge University Press, 1998.

[43] D. Poulalhon and G. Schaeffer. Optimal coding and sampling of triangulations. In *ICALP*, pages 1080–1094, 2003.

[44] L. Castelli Aleardi and O. Devillers. Canonical triangulation of a graph, with a coding application. INRIA, 2004.

[45] A. Hatcher. *Algebraic topology*. Cambridge University Press, 2002.

[46] H. Lopes, J. Rossignac, A. Safonova, A. Szymczak, and G. Tavares. Edgebreaker: a simple compression for surfaces with handles. In C. Hoffman and W. Bronsvort, editors, *Solid Modeling and Applications*, pages 289–296, Saarbrücken, 2002. ACM.

[47] H. Lopes, S. Pesco, G. Tavares, M. G. M. Maia, and Á. Xavier. Handlebody representation for surfaces and its applications to terrain modeling. In *Shape Modeling International*, volume 9. IEEE, 2003.

[48] S. Gumhold, S. Guthe, and W. Straßer. Tetrahedral mesh compression with the Cut–Border machine. In *Visualization*, pages 51–58. IEEE, 1999.

[49] N. A. Gumerov, R. Duraiswami, and E. A. Boroviko. Data structures, optimal choice of parameters, and complexity results for generalized multilevel fast multipole methods in d dimensions. Technical report, University of Maryland, 2003.

[50] F. Kälberer, K. Polthier, U. Reitebuch, and M. Wardetzky. Freelence — coding with free valences. In *Eurographics*, volume 24, pages 469–478. Blackwell, 2005.

[51] B. Kronrod and C. Gotsman. Efficient coding of nontriangular mesh connectivity. *Graphical Models*, 63:263–275, 2001.

[52] C. T. Loop. Smooth subdivision surfaces based on triangles. Master's thesis, 1987.

[53] G. Martin. Range encoding: an algorithm for removing redundancy from a digitised message. In *Video & Data Recoding*, 1979.

[54] D. Salomon. *Data compression: the complete reference*. Springer, Berlin, 2000.