

Séparation Aveugle de Sources appliquée à la compression du maillage 3D

M. CURILA^{1,2}, D. NUZILLARD¹, S. CURILA^{1,2}

¹Laboratoire d'Automatique et de Micro-électronique, Université de Reims Champagne-Ardenne, Moulin de la Housse, BP 1039, F-51687 Reims Cedex 2, Tel : 03 26 91 84 24, Fax : 03 26 91 31 06

²University of Oradea, 3700 Oradea, ROMANIA, str. Armatei Romane 5

Mircea.Curila@univ-reims.fr, Danielle.Nuzillard@univ-reims.fr, Sorin.Curila@univ-reims.fr

Résumé – Les méthodes récentes de compression des maillages 3D peuvent être divisées en trois : simplification polyhédrale, compression des positions 3D et des attributs et codage de la connectivité. Les travaux présentés dans ce papier concernent le codage des positions 3D (géométrie du maillage 3D), leur objectif est de contribuer aux performances de la compression en réduisant la corrélation spatiale par une méthode de Séparation Aveugle de Sources (SAS) introduite avant le codage entropique. La méthode *Eigenvalues Decomposition (EVD)* a été adaptée pour travailler en sous-blocs et ainsi réduire la taille de la matrice de transformation pour fournir une compression efficace.

Abstract – Recent methods in 3D mesh compression may be divided into three classes: polyhedral simplification, compression of 3D positions and attributes and coding of the connectivity. This work concerns the coding of 3D positions (geometry of 3D mesh), their objective is to contribute to performances of the compression while reducing the spatial interrelationship by a method of Blind Source Separation (BSS) introduced before entropic coding. The *Eigenvalues Decomposition method (EVD)* has been adapted to work with under-blocks, while reducing the size of the matrix of transformation it provides an effective compression.

1. Introduction

La norme VRML97 sert à décrire le format de fichier le plus véhiculé par Internet pour représenter les données 3D. La représentation minimale d'un maillage 3D (sans attribut - texture, couleur, vecteur normal), est décrite par un nœud *IndexedFaceSet* avec deux champs *coord* et *coordIndex*. Le premier champ décrit la géométrie (la position 3D de chaque sommet). La compression du maillage 3D permet de réduire le stockage et les temps de transmission. Des chercheurs ont essayé de trouver des algorithmes différents pour compresser et manipuler les données 3D. Les principaux travaux ainsi que les résultats proches de notre travail sont mentionnés ci-après : J. Rossignak [14], G. Taubin [14, 15], M. Deering [8], F. Bossen [9], H. Hoppe[10], F. Lazarus, M. Chow[2], C. Gostman [16], F. Preteux [7], C. J. Kuo [11], A. Guezic [9], etc.

La SAS est un outil adapté aux situations dans lesquelles le modèle des sources est inconnu ou presque ou difficilement exploitable à cause de sa complexité. Une ou plusieurs sources physiques émettent des signaux (S), après leur propagation, ceux-ci sont reçus sur un réseau de capteurs. A partir des signaux reçus (X), la SAS recherche les signaux émis par chacune des sources [1, 3].

Le modèle le plus simple de SAS suppose que les N signaux observés $X=(X_1, X_2, \dots, X_N)^T$ sont des mélanges linéaires de M signaux source inobservables $S=(S_1, S_2, \dots, S_M)^T$ où $M < N$, et A est la matrice de mélange.

$$X_i = \sum_j \frac{a_{ij}}{\alpha_j} \alpha_j S_j \quad (1)$$

L'identification de cette matrice n'est que partielle car il est

toujours possible d'échanger un facteur scalaire fixe α_j ou de permuter une colonne de A et une ligne du vecteur des sources S .

Par ailleurs, les techniques actuelles de compression des données 3D utilisent une procédure prédictive pour réduire la corrélation spatiale. Le principe du codage prédictif couramment utilisé est évoqué en section 2. En section 4, il est montré comment la géométrie du maillage 3D peut s'exprimer par une combinaison linéaire des composantes du vecteur de la géométrie décorrélée rendant possible l'application des méthodes de SAS pour la compression. Ainsi la méthode de SAS *Eigenvalues Decomposition (EVD)* [3], adaptée pour travailler en sous-blocs, est proposée pour réduire la corrélation spatiale des données 3D. Les résultats obtenus par cette méthode sont fournis en section 5.

2. Le codage prédictif de la géométrie du maillage 3D

Le principe du codage uni-résolution de la géométrie du maillage 3D exige les étapes suivantes [11, 14]:

- les coordonnées des sommets sont uniformément quantifiées, le pas de quantification est à choisir en utilisant un algorithme de recherche itérative pour contrôler le nombre de bits par sommet ;
- un pas de décorrélation est exécuté par prédiction directement sur le modèle quantifié ;
- les résidus résultants sont codés sans perte en utilisant un algorithme d'approximations successives suivi par un codage entropique.

Soit une séquence de sommets P_i , $1 \leq i \leq N$, chacune de leurs coordonnées sur chaque direction est quantifiée par une valeur qP_i . Grâce à une règle de prédiction linéaire, il est possible d'en obtenir une valeur estimée $q\hat{P}_i$. La différence entre la valeur du sommet courant qP_i et son estimation $q\hat{P}_i$ est appelée « erreur prédictive » dP_i . Les techniques de codage prédictif consistent à coder seulement cette erreur. L'efficacité de la compression est fortement basée sur la décorrélation. Les techniques actuelles de compression utilisent la règle prédictive polygonale très efficace. Celle-ci a d'abord été proposée par C. J. Kuo [11], puis implantée en schéma hybride par F. Preteux *et al* [7]. Son principe consiste à prédire pour chaque sommet courant, k positions de ce sommet à partir de chacun de ses k polygones adjacents, dont les sommets sont les plus proches voisins. La position finale résulte de la valeur moyennée de ces prédictions.

Considérons un point P , sommet d'un polygone régulier comportant n sommets ($P_1 P_2 \dots P_n$), (avec $n > 3$), ce point P dans un plan de l'espace Euclidien 3D ne passant que par trois points non-colinéaires P_i, P_j et P_k , peut être exprimé en utilisant des coordonnées barycentriques comme il suit:

$$P = \alpha P_i + \beta P_j + \gamma P_k, \text{ with } \alpha + \beta + \gamma = 1, \quad (2)$$

La règle prédictive polygonale utilise les coefficients α , β et γ , de la prédiction précitée [11], pour prédire le sommet courant P à partir des trois sommets déjà traversés qui appartiennent au même polygone que P . A noter que les trois sommets utilisés pour la prédiction ne sont pas nécessairement consécutifs. Si pour chaque polygone adjacent à P il y a moins de trois sommets traversés, la prédiction polygonale ne peut pas être appliquée comme décrite au-dessus. Le sommet P est prédit comme la moyenne des sommets voisins traversés. Le premier sommet est traité comme le résidu d'une prédiction nulle.

3. Quantification des données et codage arithmétique

Une grille uniforme, définie par les valeurs minimales et maximales des données et le nombre de niveaux de quantification, est d'abord déterminée. Les données en virgule flottante du modèle sont uniformément quantifiées sur cette grille par l'équation suivante :

$$qP = \text{int} \left(\frac{P - IC}{\max S} (2^{b_{pv}} - 1) \right) \quad (3)$$

où P est une donnée en virgule flottante, IC est la valeur minimale des données, $\max S$ est déterminé par la différence des valeurs maximale et minimale et b_{pv} représente le nombre de niveaux de quantification.

Les concepts d'Information et d'Entropie sont très importants en compression de données. Le nombre de bits d'information pour un symbole est égal au nombre de bits du code optimum de ce symbole. L'efficacité de transmission est améliorée en utilisant des mots de code plus courts pour les symboles les plus probables et des mots de code plus longs pour les symboles les moins probables. Le principe est d'associer à chaque symbole, un nombre de bits qui va dépendre de sa probabilité d'apparition. Les symboles binaires produits par l'algorithme de quantification sont codés en utilisant un codeur arithmétique binaire comportant 113 états de Markov pour l'estimation de la probabilité [13].

4. La décorrélation spatiale de données 3D par SAS

La description de la géométrie du maillage 3D dans le fichier VRML présente une forte corrélation spatiale dans chaque direction du système de coordonnées cartésien. Afin de décorréler des données, il est envisageable de substituer la règle de prédiction linéaire par une méthode de Séparation Aveugle de Sources (SAS). Pour cela, il est nécessaire d'exprimer le vecteur à N composantes de la géométrie initiale $\mathbf{g}^T = [P_1, P_2, \dots, P_N]^T$ par un vecteur $d\mathbf{g}^T$ à M composantes décorrélées $d\mathbf{g}^T = [dP_1, dP_2, \dots, dP_M]^T$ associé à une matrice de mélange $A[N \times M]$ d'éléments $A_{i,j}$, où $M \ll N$. En utilisant la méthode prédictive, mentionnée ci-dessus il est possible de ré-exprimer le vecteur de la géométrie \mathbf{g} du maillage 3D comme il suit :

$$\begin{cases} P_1(x, y, z) = P_1 \\ P_2(x, y, z) = P_1 + dP_2 \\ P_3(x, y, z) = \alpha_1 P_1 + \alpha_2 P_2 + dP_3 \\ \dots \\ P_s(x, y, z) = \alpha \cdot P_i + \beta \cdot P_j + \gamma \cdot P_k + dP_s, \quad s \in [4 \dots N] \end{cases} \quad (4)$$

où P_i, P_j et P_k sont les trois sommets déjà traversés voisins du sommet courant et α , β et γ sont les coefficients de la prédiction mentionnée par l'équation 2. En remplaçant les sommets P_i, P_j et P_k par leur description en fonction de $P_1 = dP_1$ et les différences dP_2, dP_3, \dots , la géométrie du modèle 3D s'exprime par :

$$\begin{cases} P_1(x, y, z) = dP_1 \\ P_2(x, y, z) = A_{2,1} dP_1 + dP_2 \\ P_3(x, y, z) = A_{3,1} dP_1 + A_{3,2} dP_2 + dP_3 \\ \dots \\ P_s(x, y, z) = A_{s,1} dP_1 + A_{s,2} dP_2 + \dots + A_{s,s-1} dP_{s-1} + dP_s, \quad s \in [4 \dots N] \end{cases} \quad (5)$$

La géométrie initiale s'exprime comme une combinaison linéaire des composantes du vecteur de la géométrie décorrélée $d\mathbf{g}$. Ainsi la géométrie initiale est décrite par la matrice des observations notée \mathbf{X} où chaque colonne contient les N observations des valeurs des sommets sur une direction de l'espace cartésien (colonnes 1, 2, 3 respectivement direction x, y, z) :

$$\mathbf{X} = \begin{pmatrix} P_1(x) & P_1(y) & P_1(z) \\ P_2(x) & P_2(y) & P_2(z) \\ \dots & \dots & \dots \\ P_N(x) & P_N(y) & P_N(z) \end{pmatrix} \quad (6)$$

La matrice \mathbf{S} des sources correspond à celle de la géométrie décorrélée codée par la méthode de compression prédictive :

$$\mathbf{S} = \begin{pmatrix} s_1(1) & s_1(2) & s_1(3) \\ s_2(1) & s_2(2) & s_2(3) \\ \dots & \dots & \dots \\ s_M(1) & s_M(2) & s_M(3) \end{pmatrix} = \begin{pmatrix} dP_1(x) & dP_1(y) & dP_1(z) \\ dP_2(x) & dP_2(y) & dP_2(z) \\ \dots & \dots & \dots \\ dP_M(x) & dP_M(y) & dP_M(z) \end{pmatrix} \quad (7)$$

La matrice de mélange \mathbf{A} est de la forme :

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ A_{2,1} & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ A_{N,1} & A_{N,2} & \dots & A_{N,M-1} & 1 \end{pmatrix} \quad (8)$$

Puisque les voisins du sommet courant, qui ont été déterminés par l'arbre des triangles dans le processus de codage de la connectivité [14], sont les mêmes indifféremment des coordonnées x , y ou z considérées, les coefficients α , β et γ ne diffèrent pas dans les trois situations ($\alpha_x = \alpha_y = \alpha_z$, $\beta_x = \beta_y = \beta_z$, $\gamma_x = \gamma_y = \gamma_z$). Cette remarque explique l'unicité de la matrice \mathbf{A} sur chaque direction x , y , z ($A_x = A_y = A_z$).

Les données ont d'abord été séparées avec les algorithmes : transformation de *Karhunen-Loève* (KLT), *Eigenvalues Decomposition* (EVD) [3], *Second Order Blind Identification* (SOBI) [1], qui fournissent des taux de compression proches [6]. Les vecteurs d'observations ($N \times 1$) (constitués par les 3 colonnes de la matrice \mathbf{X}) sont soumis à l'algorithme de séparation pour fournir trois vecteurs décorrelés ($M \times 1$) et une matrice de séparation \mathbf{B} unique ($M \times N$), \mathbf{B} est la pseudo-inverse de \mathbf{A} . Après la séparation, les M lignes de la matrice de séparation \mathbf{B} ont la même dimension que la géométrie du modèle original (N). Pour réduire la taille de cette matrice, la géométrie est divisée en b blocs, notée \mathbf{bg} , de dimension N/b (approchée à l'entier supérieur) :

$$\mathbf{bg}(i) = \mathbf{g}\left(\left(i-1\right)\frac{N}{b} + 1 : i\frac{N}{b}\right), i=1, \dots, b \quad (9)$$

Cette démarche a été testée avec KLT [6]. Maintenant, elle est mise en œuvre en association avec EVD. Chaque bloc fournit une matrice de covariance $\mathbf{R}_{x_i}(0)$, ces matrices sont

moyennées pour fournir une matrice de covariance globale. La diagonalisation de cette dernière fournit une matrice de blanchiment qui est ensuite appliquée sur chaque bloc de la géométrie :

$$\underline{\mathbf{bg}} [M \times K] = \mathbf{W} [M \times \frac{N}{b}] \mathbf{bg} [\frac{N}{b} \times K] \quad (10)$$

Pour chaque bloc de la géométrie blanchie une matrice de covariance décalée $\mathbf{R}_{y_i}(p)$, $p \neq 0$ est aussi déterminée, le moyennage de ces matrices fournit une matrice de covariance globale décalée. La diagonalisation de cette matrice conduit à une matrice unitaire globale dont la transposée est appliquée à chaque bloc de la géométrie blanchie pour fournir une géométrie décorrelée :

$$\mathbf{bdg}[M \times K] = \mathbf{U}[M \times M] \underline{\mathbf{bg}}[M \times K] \quad (11)$$

Ainsi, la dimension de la matrice de séparation est réduite de $M \times N$ à $M \times N/b$, et celle de la géométrie décorrelée est augmentée de M à bM . Le nombre des blocs doit être choisi de manière judicieuse pour diminuer la dimension de la matrice de transformation et augmenter faiblement celle de la géométrie décorrelée. Ce choix apporte une bonne décorrélation et donc une bonne compression avec de faibles erreurs de reconstruction. Si le nombre de blocs est trop grand l'erreur de reconstruction augmente. S'il est trop faible, le nombre M de composantes de la géométrie transformée

entraîne une augmentation de la dimension de la matrice de séparation et une plus faible compression du maillage 3D.

Parce que le nombre des composantes des vecteurs de la géométrie décorrelée est petit, celles-ci sont incluses dans l'en-tête du fichier comprimé. La matrice de mélange, de dimension importante, représente les mots-codes du fichier cité. La valeur « zéro » de chaque élément au-dessus de la diagonale principale justifie l'introduction des méthodes de SAS pour la compression des maillages 3D. La matrice de transformation \mathbf{A} est ensuite comprimée par codage entropique.

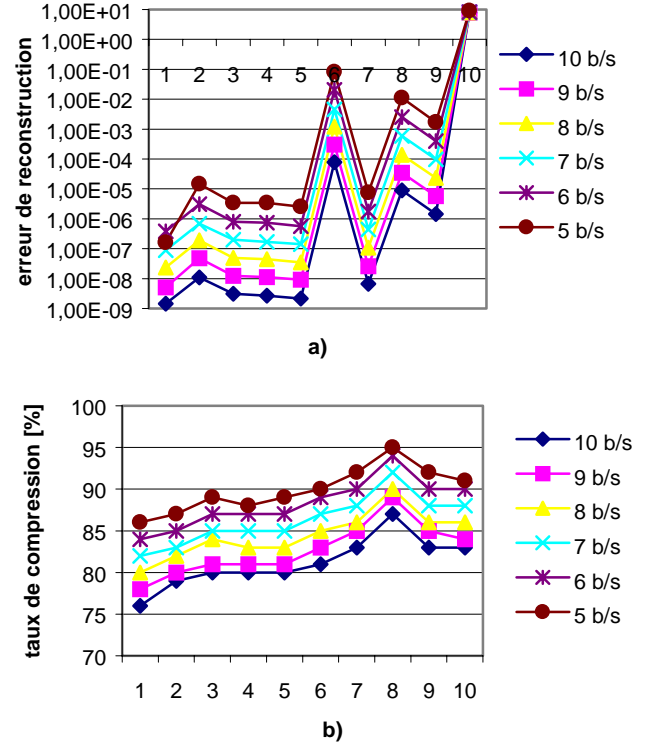


FIG. 1: a) Erreur de reconstruction et b) Taux de compression pour différents « bit rates » (1. ship2.wrl nr. sommets = 114 ; 2. romulan.wrl nr. sommets = 225 ; 3. banana.wrl nr. sommets = 271 ; 4. cat.wrl nr. sommets = 352 ; 5. satell.wrl nr. sommets = 360 ; 6. buggy.wrl nr. sommets = 485 ; 7. rose.wrl nr. sommets = 850 ; 8. yoda.wrl nr. sommets = 853 ; 9. fiero.wrl nr. sommets = 993 ; 10. chopper.wrl nr. sommets = 1072)

5. Résultats expérimentaux

Nous avons testé les deux algorithmes de décorrélation de la géométrie (le codage prédictif et le codage basé sur la décorrélation par SAS) pour plusieurs maillages 3D. La figure 1 montre l'erreur de reconstruction (équation 13) et le taux de compression (équation 12) obtenus par le deuxième algorithme pour dix maillages complexes. Le taux de compression s'écrit :

$$r = \frac{do - dc}{do} \cdot 100 [\%] \quad (12)$$

où do , dc représentent respectivement la taille du fichier original et celle du fichier comprimé. Les maillages ont été triés par ordre croissant du nombre des sommets. L'erreur de reconstruction est définie ici comme la distance entre les sommets des maillages originaux et ceux reconstruits. La

mesure de la distorsion est estimée par la valeur moyenne de ces distorsions pour tous les sommets :

$$e_r = \frac{1}{N} \sum_{i=1}^N \sqrt{(X_i - X_{ri})^2 + (Y_i - Y_{ri})^2 + (Z_i - Z_{ri})^2} \quad (13)$$

où N indique le nombre de sommets, X_i , Y_i , et Z_i sont les valeurs originales de la géométrie et X_{ri} , Y_{ri} , et Z_{ri} sont les valeurs de la géométrie reconstruite.

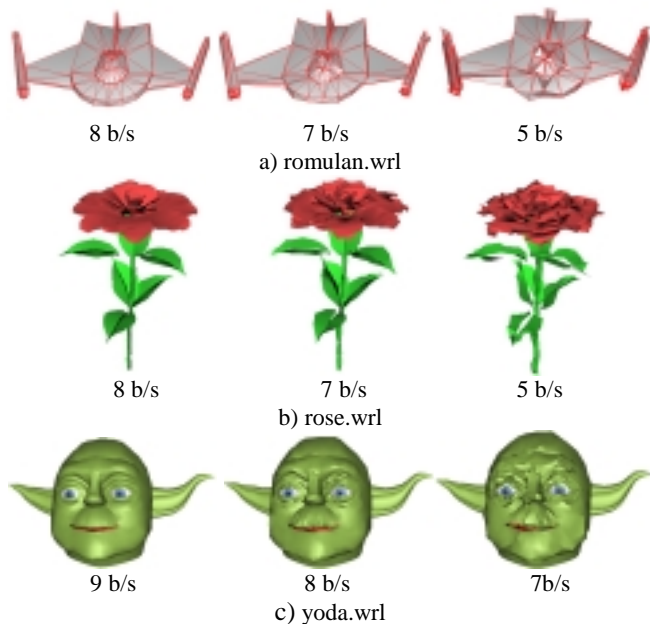


FIG. 2 : Maillages 3D reconstruits en utilisant la méthode EVD

Une analyse comparative des performances des deux méthodes, pour trois maillages représentatifs de la liste expérimentée en figure 1, est présentée dans le tableau 1. Chaque ligne du tableau, pour un maillage donné, correspond approximativement à la même erreur de reconstruction.

TAB. 1 :

a) romulan.wrl		b) rose.wrl				c) yoda.wrl					
EVD – diviz.		Met. predictiv.		EVD – diviz.		Met. predictiv.		EVD – diviz.		Met. predictiv.	
b	r	b	r	b	r	b	r	b	r	b	r
7	83	8	83	8	86	8	85	11	85	8	85
6	85	7	87	7	88	7	88	10	87	7	88
5	87	6	89	6	90	6	91	9	89	6	91
4	89	5	92	5	92	5	93	7	92	5	94

b = bits/sommet ; r = taux de compression [%].

Le taux de compression est comparable à celui proposé par C. J. Kuo [11], les différences sont de 2 à 3 %.

La qualité visuelle du maillage 3D reconstruit peut être estimée grâce à la figure 2. Pour chaque ligne, la première image correspond à une bonne reconstruction, le nombre de bits par sommet est optimum et l'erreur de reconstruction est faible (figure 1). Pour la troisième image, la dégradation est due à un trop petit nombre de bits par sommet.

6. Conclusion

En conclusion, les méthodes de SAS ont été appliquées en tant que pré-traitement pour décorrélérer la géométrie d'un maillage 3D candidate à la compression à l'aide d'un codeur entropique. Les résultats obtenus sont équivalents à ceux découlant d'une méthode utilisant la prédiction linéaire offrant ainsi une voie alternative de traitement. Les performances de ce dernier pourraient être accrues en

optimisant le découpage de la géométrie du maillage. Par extension, une voie de recherche novatrice consiste à associer les techniques de SAS à la multi-résolution et au codage entropique pour améliorer les taux de compression à qualité égale.

Références

- [1] A. Belouchrani, K. Abdel-Meraim, J. F. Cardoso, E. Moulines, *A blind source separation technique using second-order statistics*, IEEE Trans. Signal Processing, 45, pp. 434-444, 1997.
- [2] M. Chow, *Optimized Geometry Compression for Real-time Rendering*, In Proceedings of IEEE Visualization '97, Phoenix AZ, pp. 347-354, 1997.
- [3] A. Cichocki, S. Amari, *Adaptive Blind Signal and Image Processing. Learning Algorithms and Applications*, John Wiley and Sons, Ltd Baffins Lane, Chichester West Sussex, PO19 1UD, 2002.
- [4] S. Costa and S. Fiori, *Image Compression Using Principal Component Neural Networks*, Image and Vision Computing Journal, 19(9-10), pp. 649-668, Aug. 2001
- [6] M. Curila, S. Curila, D. Nuzillard, *Decorrelation techniques for geometry coding of 3D mesh*, Fourth International Symposium on Independent Component Analysis and Blind Signal Separation, pp. 885-890, Nara, Japan, April 1-4, 2003.
- [7] S. Curila, M. Curila, T. Zaharia, G. Mozelle, F. Preteux, *A new prediction scheme for geometry coding of 3D meshes within the MPEG-4 framework*, Nonlinear Image Processing X, Proceedings of SPIE'99, San Jose, California, January, pp. 240-250, 1999.
- [8] M. Deering, *Geometric Compression*, Computer Graphics (SIGGRAPH '95 Proceedings), pp. 13-20, 1995.
- [9] A. Guezic, F. Bossen, G. Taubin, C. Silva, *Efficient Compression of Non-Manifold Polygonal Meshes*, IEEE Visualization'99, 5(4), pp. 349-359, 1999.
- [10] H. Hoppe, *Progressive Meshes*, Computer Graphics (Proc. SIGGRAPH), pp. 99-108, August 1996.
- [11] J. Li, C. J. Kuo, *Embedded Coding of Mesh Geometry*, Research Report ISO/IEC JTC1/SC29/WG11, MPEG98/M3325, Tokyo, Japan, March, 1998.
- [12] J. Li, C. J. Kuo, *Progressive Coding of 3D Graphic Models*, Proceedings of IEEE, 86, pp. 1052-1063, June 1998.
- [13] A. Moffat, R. Neal, I. Witten, *Arithmetic Coding Revisited*, In IEEE Data Compression Conference, Snowbird, 16(3), pp. 256-294, Utah, 1995.
- [14] G. Taubin, W. P. Horn, F. Lazarus, and J. Rossignac, *Geometric Coding and VRML*, Proceedings of the IEEE, 86(6), pp. 1228-1243, June 1998.
- [15] G. Taubin, *3D Geometry Compression in MPEG-4*, speaker, in Course on 3D Geometry Compression, Siggraph'2000, New Orleans, Louisiana, July 2000.
- [16] C. Touma, C. Gotsman, *Triangle mesh compression*, In Graphics Interface Conference Proceedings, Vancouver, pp. 26-34, June, 1998.