

Chaos dans la modélisation des tissus mous

Michel Pitermann

Laboratoire Parole et Langage
Université de Provence, 29 av. R. Schuman - 13621 Aix-en-Provence Cedex 1, France
Tél. : +33 (0)4 42 95 36 26 - Fax : +33 (0)4 42 59 50 96
Email : mpiter@lpl.univ-aix.fr

ABSTRACT

Masses connected by springs are often used in biomechanical models of soft tissue, but chaotic noise is expected with this technique. A method is presented here to analyze this type of noise. The method is illustrated by an analysis of a biomechanical model of the face but the method can be used for almost any type of nonlinear model of soft tissue. The results indicated that (i) the noise properties of the face model were similar in stationary and dynamic state ; (ii) noise amplitude ($\approx 10\%$ of synthetic movement amplitude) was independent of computing accuracy ; (iii) the face model was sensitive to initial conditions. These properties indicated chaotic noise that could not be decreased by means of improving computing accuracy. The noise was visible in the animations and was perceived as skin vibration. This undesirable property may be found in other biomechanical models of soft tissues used in speech production research (tongue, lip or face models).

1. INTRODUCTION

La modélisation des tissus mous est devenu un sujet important en animation faciale (la peau du visage [7]), en production de la parole (la langue [6]) et en perception audiovisuelle de la parole (synthèse de stimuli [1]).

Parmi les nombreuses techniques de modélisation, la *modélisation biomécanique*¹ mène aux meilleurs modèles pour la recherche en production de la parole et permet de produire les meilleures animations. Par exemple, un treillis de points massiques connectés par des ressorts non linéaires peut approcher de nombreuses propriétés des tissus mous [7, 3]. Malheureusement, cette technique entraîne des difficultés dont la possibilité de solutions d'équations différentielles peu précises, instables ou chaotiques.

Le sujet de cet article est l'analyse des solutions chaotiques. Une solution chaotique entraînera une vibration du tissu mou qui peut être visible dans les animations. Par exemple, l'amplitude de cette vibration atteint un tiers de l'amplitude du mouvement des lèvres dans une expérience de *bite-block*² pour un modèle de visage. Une méthode est proposée ici pour détecter ce type de bruit dans n'importe quelle composante d'un modèle et pour quantifier l'amplitude de ce bruit. La méthode s'appuie sur

¹Les modèles biomécaniques de visage décrivent le comportement physique de la langue, de la peau, de la mâchoire et des muscles.

²Expérience au cours de laquelle la mâchoire du sujet est maintenue immobile, par exemple en demandant au sujet de parler en mordant dans un stylo à bille.

l'analyse du bruit intrinsèque au modèle dans des analyses où le bruit est généré pour différents niveaux de précision des calculs et pour différentes conditions initiales. Plutôt que de détailler techniquement la méthode, l'analyse d'un modèle de visage est présentée en régime stationnaire puis en régime dynamique en guise d'illustration. Les résultats montrent que (1) l'amplitude de la vibration de peau était similaire pour les deux régimes ; (2) cette amplitude était indépendante de la précision des calculs ; (3) le modèle de visage était sensible aux conditions initiales. La technique exposée ici peut être utilisée pour détecter la présence de chaos dans presque n'importe quel type de modèle de tissu mou.

2. MÉTHODE

Les analyses ont été réalisées en deux étapes. Premièrement, le modèle a été étudié pour un état stationnaire, c.à.d. avec une activation musculaire constante. Deuxièmement, le modèle a été examiné pour un état dynamique, c.à.d. avec une activation musculaire variant au cours du temps. Les caractéristiques communes aux deux expériences sont décrites ici, les particularités de chacune des expériences seront détaillées dans des sections séparées.

2.1. Le modèle de visage

Le modèle de visage décrit dans [2] a été utilisé dans toutes les expériences. Comme le modèle a été détaillé dans [2], seules ses principales caractéristiques sont résumées ici.

Le modèle était composé d'un module de mâchoire, d'un module de muscle et d'un module de peau. La mâchoire était décrite par une simple charnière contrôlée cinématiquement par un angle. Les muscles étaient modélisés par une formulation standard de type Hill. La peau était composée d'un treillis de points massiques à trois couches aux propriétés mécaniques isotropiques. Les 1435 masses du treillis étaient connectées par près de 6000 ressorts amortis non linéaires.

Pour toutes les expériences, huit paires de muscles (un élément de chaque paire pour chaque profil du modèle de visage) étaient activées simultanément avec le même niveau d'énergie. Les huit paires étaient le levator labii superior, le levator anguli oris, le zygomatic major, le depressor anguli oris, le depressor labii inferior, le mentalis, l'orbicularis oris superior et l'orbicularis inferior.

Pour animer le modèle, une image était calculée tous les 1/60 s. Les équations de mouvement étaient résolues à

l'aide de l'algorithme de Runge-Kutta standard en utilisant un pas constant de 50 itérations par image [5, section 16.1].

2.2. Analyses statistiques

Les trajectoires 3D de 11 points du modèle de visage ont été suivies au cours du temps afin d'analyser statistiquement la vibration de peau (cf. figure 1 pour les positions approximatives de ces 11 points).

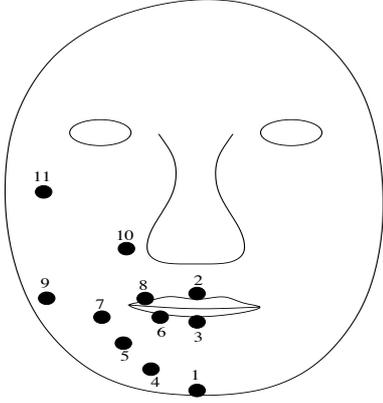


FIG. 1: Positions approximatives et numéros des points analysés (ellipses noires).

Quelques indices statistiques unidimensionnels ont été généralisés à trois dimensions afin de caractériser les trajectoires 3D des 11 points analysés. La position moyenne d'une trajectoire 3D v composée de n échantillons (x_i, y_i, z_i) était son centroïde μ_v :

$$\mu_v = \left(\frac{1}{n} \sum_{i=1}^n x_i, \frac{1}{n} \sum_{i=1}^n y_i, \frac{1}{n} \sum_{i=1}^n z_i \right) \quad (1)$$

L'écart-type σ_v d'une trajectoire 3D v était estimée par :

$$\sigma_v = \sqrt{\frac{1}{n-1} \sum_{i=1}^n |v_i|^2 - |\mu_v|^2} \quad (2)$$

où $||^2$ était l'opérateur « valeur absolue » d'un vecteur, c.à.d. la somme des carrés des coordonnées du vecteur. La différence entre deux séries chronologiques v et w était définie par :

$$\Delta_{vwi} = v_i - w_i \quad \forall i \quad (3)$$

où i représentait n'importe quelle coordonnée temporelle des séries chronologiques.

3. EXP. 1 : ÉTAT STATIONNAIRE

Le but de la première expérience était de caractériser la vibration de peau générée par le modèle lorsqu'il était dans un état stationnaire, c.à.d. pour une activité musculaire constante au cours du temps. Ensuite, l'amplitude de la vibration a été comparée à l'amplitude d'un mouvement extrême de la peau du modèle de visage.

3.1. Méthode

Les 16 muscles sélectionnés ont été simultanément activés par une fonction « marche » composée de 10 échantillons valant 0 suivis de 1400 échantillons valant 1. Ensuite, les

amplitudes du mouvement et de la vibration de peau résultants ont été mesurées et comparées. La valeur 0 correspondait à une absence d'activité musculaire, la valeur 1 correspondait au maximum de l'activité musculaire.

L'amplitude de mouvement d'un point du treillis était estimée par la distance euclidienne comprise entre la position moyenne du point avant le saut d'activation musculaire et celle estimée après le saut. Afin de laisser le modèle atteindre son équilibre après le saut d'activation musculaire, seuls les 1200 derniers échantillons de la trajectoire 3D du point ont été utilisés dans l'estimation de sa position moyenne pour l'activation musculaire maximale.

L'amplitude de vibration de la peau a été estimée de trois manières différentes. Premièrement, par l'écart-type 3D [Équation (2)] de la trajectoire de chaque point. Deuxièmement, par l'écart-type 3D de la différence Δ_{vw} [Équation (3)] séparant deux trajectoires v et w d'un point calculées successivement pour la même activation musculaire mais à partir de deux conditions initiales différentes de 0.001 mm vers la droite à la dixième image pour la seconde animation, les positions des autres points n'ayant pas été perturbées. Troisièmement, par l'écart-type 3D de la différence Δ_{vw} séparant deux trajectoires d'un point calculées successivement pour la même activation musculaire, mais avec deux niveaux de précision numérique différents dans les calculs. Les calculs ont en effet été réalisés en simple, double ou quadruple précision, ce qui correspondait aux types C++ « float », « double » et « long double », c'est-à-dire pour des représentations des nombres réels en 32, 64 ou 128 bits pour notre implantation C++ de l'époque (gcc 2.95.2, Solaris 2.7, UltraSparc II). Pour les trois méthodes, les écarts-types 3D ont été estimés à partir des 1200 derniers échantillons de la fonction « marche ».

Pour les deux dernières méthodes d'estimation de l'amplitude de vibration de la peau impliquant une différence Δ_{vw} entre deux trajectoires d'un même point du treillis, la variance de Δ_{vw} a été divisée par 2. Par conséquent, les écarts-types correspondant ont été divisés par $\sqrt{2}$ pour l'estimation de l'amplitude de vibration de la peau. La raison était qu'après le calcul d'une différence Δ_{vw} , il ne restait plus que du bruit provenant des deux trajectoires. Dans ce cas, la variance de la différence Δ_{vw} devait être la somme des variances du bruit de chaque trajectoire pour autant que les bruits des deux trajectoires fussent indépendants. Les résultats montrant plus loin l'origine chaotique du bruit, les bruits de deux trajectoires pouvaient être considérés comme indépendants. Par conséquent, la variance de Δ_{vw} devait être égale à deux fois la variance de v ou de w .

3.2. Résultats et discussion

La figure 2 présente les trois types d'estimation des écarts-types 3D des trajectoires pour chacun des 11 points sélectionnés du treillis après que le modèle de visage ait atteint un état stationnaire. Les écarts-types de la figure constituaient notre estimation de l'amplitude de vibration de la peau en différents endroits du visage.

Une analyse de la variance à deux facteurs (« numéro du point » et « précision numérique ») de l'écart-type pour les trois premiers cas de la légende de la figure 2 a montré

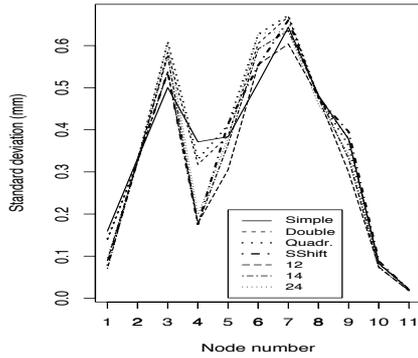


FIG. 2: Trois types d'estimation de l'amplitude de vibration de la peau pour un état stationnaire du modèle : (i) écarts-types 3D de chaque trajectoire de point pour trois niveaux de précision des calculs (trois premiers cas de la légende); (ii) différence entre les deux trajectoires d'un point calculées pour deux conditions initiales différentes (cas « SShift »); (iii) différences entre deux trajectoires d'un point calculées pour deux niveaux de précision différents (trois derniers cas de la légende où 1, 2 et 4 représentent simple, double et quadruple précision).

un effet significatif au niveau 0.05 pour « numéro du point » [$F(10, 20) = 175.33; p < 1e-9$], mais pas pour « précision numérique » [$F(2, 20) = 1.22; p = 0.32$]. Cela signifie que l'amplitude de vibration de la peau n'était pas la même en tous les endroits du visage et qu'elle ne dépendait pas de la précision des calculs. Par conséquent, l'amplitude de ce « bruit » indésirable ne pouvait pas être atténuée par une augmentation de la précision des calculs.

Les résultats précédents suggéraient la présence de chaos. Le modèle de visage devait donc être sensible aux conditions initiales. Calculer successivement la trajectoire d'un point pour la même activité musculaire mais pour deux conditions initiales différentes ou pour deux niveaux de précision numérique différents devait conduire à deux trajectoires différentes. Comme expliqué dans la section 3.1, le bruit chaotique pouvait être estimé à partir de la différence entre ces deux trajectoires. Le bon accord entre ces nouvelles estimations de l'amplitude de vibration de la peau et les premières peut être observé sur la figure 2 (les quatre derniers cas de la légende contre les trois premiers). Une analyse de la variance à deux facteurs (« numéro du point » et « méthode » d'estimation de l'amplitude de vibration de la peau avec ou sans différence) de l'estimation de l'amplitude de vibration de la peau a montré un effet significatif au niveau 0.05 pour « numéro du point » [$F(10, 55) = 468.87; p < 1e-9$], pour « méthode » [$F(1, 55) = 23.55; p < 2e-4$] et pour leur interaction [$F(10, 55) = 5.62; p < 1e-05$]. Mais comme l'indique la figure 2, les résultats produits par chaque méthode étaient si proches qu'il pouvaient être considérés comme équivalents pour toute application pratique.

La vibration de peau était intrinsèque au modèle de visage et était indépendante du niveau de précision des calculs. Son amplitude a alors été comparée à celle d'un mouvement de peau extrême. La figure 3 montre le déplacement des 11 points lorsque les 16 muscles sélectionnés ont été excités par la fonction « marche ». Une comparaison entre l'échelle verticale de cette figure et celle de la figure 2 montre que le rapport entre les amplitudes d'un mouve-

ment de peau et de la vibration intrinsèque de la peau valait approximativement 10. Par conséquent, même si elle n'était pas négligeable, la vibration de peau ne dominait pas les mouvements de visage.

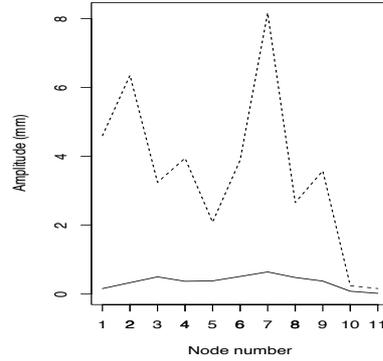


FIG. 3: Amplitude d'un mouvement de peau (trait pointillé) et de vibration de peau (trait plein) en différents endroits du visage en fonction du numéro du point (cf. figure 1 pour la position de chaque point sur le visage). Le trait plein correspond au trait plein de la figure 2.

4. EXP. 2 : ÉTAT DYNAMIQUE

Après avoir étudié le modèle en état stationnaire, l'analyse a été répliquée pour un état dynamique (activité musculaire variant au cours du temps) car la parole est intrinsèquement dynamique. De nouvelles propriétés émergeront-elles de la dynamique ?

4.1. Méthode

Les seize mêmes muscles que dans l'expérience précédente ont été activés simultanément par une onde triangulaire composée de la séquence (0, 1/6, 2/6, 3/6, 4/6, 5/6, 6/6, 5/6, 4/6, 3/6, 2/6, 1/6) répétées 100 fois afin de créer une série chronologique de 1200 échantillons composée de la répétition de 100 triangles identiques.

L'amplitude du « bruit » de la peau ne pouvait plus être estimée par l'écart-type 3D [Equation (2)] d'une trajectoire d'un point du treillis car le modèle n'était plus dans un état stationnaire. En effet, l'écart-type de la trajectoire d'un point contiendrait alors une composante liée au « bruit » de peau ainsi qu'une composante liée aux mouvements de visage. Par conséquent, seule les estimations s'appuyant sur la différence Δ_{vw} [Equation (3)] entre deux trajectoires d'un même point étaient utilisables.

4.2. Résultats et discussion

La figure 4 présente les estimations du bruit de la peau en fonction du numéro des points étudiés lorsque le modèle de visage était excité par une activité musculaire en forme d'onde triangulaire.

Les quatre derniers cas de la figure 4 montrent qu'utiliser des conditions initiales différentes ou utiliser des niveaux de précision numérique différents dans les calculs pour estimer l'amplitude du bruit de la peau menaient à des résultats similaires. Une analyse de la variance à deux facteurs (« numéro du point » et « méthode » utilisée pour estimer l'amplitude du bruit) de l'estimation de l'amplitude du bruit a montré un effet significatif pour « numéro du

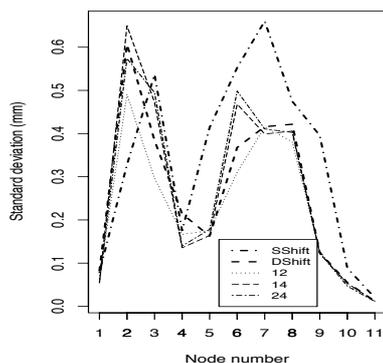


FIG. 4: Estimations de l’amplitude du bruit de la peau lorsque le modèle était excité par une activité musculaire en forme d’onde triangulaire en utilisant deux conditions initiales différentes (cas « DShift ») ou deux niveaux de précision numérique différents (cas « 12 », « 14 » et « 24 »), ou lorsque le modèle était en état stationnaire (cas « SShift », données provenant de l’expérience 1, figure 2). Cf. légende de la figure 2 pour les détails.

point » [$F(10, 30) = 81.67; p < 1e - 9$], mais pas pour « méthode » [$F(3, 30) = 2.27; p = 0.10$] au niveau 0.05. Ainsi, aucune dépendance de l’amplitude du bruit n’a pu être mise en évidence par rapport à la précision numérique des calculs ou au type d’estimateur utilisé.

Une comparaison entre la courbe en traits gras « SShift » et les quatre autres courbes de la figure 4 indique que l’amplitude du bruit était du même ordre de grandeur lorsque le modèle était en état stationnaire ou dynamique. Une analyse de la variance à deux facteurs (« numéro du point » et « état du modèle ») de l’estimation de l’amplitude du bruit de la peau a montré un effet significatif pour « numéro du point » [$F(10, 33) = 87.90; p < 1e - 9$], pour « état du modèle » [$F(1, 33) = 30.88; p < 1e - 5$] et pour l’interaction entre les deux facteurs [$F(10, 33) = 8.81; p < 1e - 6$] au niveau 0.05. Néanmoins, les similarités entre les amplitudes du bruit de la peau estimées lorsque le modèle était en état stationnaire ou dynamique suggère que des niveaux de bruit analogues devraient être observés dans la plupart des situations de parole.

5. CONCLUSION

Les résultats montrent que le modèle de visage génère une vibration de peau, même pour une activité musculaire exempte de bruit. Ce bruit provient de la discrétisation de la peau en un treillis de points massiques reliés par des ressorts non linéaires. L’amplitude du bruit de la peau vaut approximativement 10 % de l’amplitude d’un mouvement de peau provenant d’une contraction musculaire maximale. Ce bruit est indépendant de la précision numérique utilisée dans les calculs, il ne peut donc pas être atténué par une augmentation de la précision numérique. De plus, le modèle de visage est sensible aux conditions initiales, et les résultats sont similaires en régime stationnaire ou dynamique. Ces résultats pourraient être généralisables à d’autres modèles de type « matelas-ressorts » utilisés en production de la parole.

Puisque ce bruit peut ne pas être négligeable par rapport à certains mouvements de visage, il pourrait avoir un impact sur des études utilisant des modèles biomécaniques

des tissus mous. En perception audiovisuelle, des modèles de visage sont utilisés pour synthétiser des stimuli audiovisuels [1]. Les résultats présentés ici suggèrent d’étudier certaines questions telles : le bruit de peau peut-il modifier une catégorisation phonétique ? Le bruit de peau peut-il changer le poids des indices visuels utilisés en intégration de l’information audiovisuelle ? En contrôle-moteur [3], un bruit chaotique est attendu tant pour les calculs directs d’un modèle que pour son inversion [4]. Il est important d’estimer le bruit chaotique intrinsèque au modèle utilisé afin d’éviter de l’attribuer erronément à d’autres causes tels une mauvaise stratégie de contrôle ou un défaut d’une méthode d’inversion. La méthode présentée ici est un moyen simple à mettre en œuvre et économique pour atteindre cet objectif.

REMERCIEMENTS

Cette recherche a été réalisée au Department of Psychology, Queen’s University, Kingston, Ontario, Canada. Elle a été financée par une bourse NIH Grant No. DC-00594 du National Institute of Deafness and other Communication Disorders and NSERC. Doug J. K. Mewhort m’a donné accès à son équipement informatique financé par une bourse d’équipement NSERC et par une bourse Academic Equipment Grant de Sun Microsystems of Canada. ATR Human Information Processing Research Laboratories, Kyoto, Japan, m’a gracieusement offert ses données morphologiques utilisées dans le modèle de visage.

RÉFÉRENCES

- [1] A. Adjoudani, T. Guiard-Marigny, Bertrand Le Goff, L. Reveret, and Christian Benoît. A multimedia platform for audio-visual speech processing. In *Eurospeech’97 Proceedings*, volume 3, pages 1671–1674, Rhodes, Greece, 1997. European Speech Communication Association.
- [2] Jorge C. Lucero and Kevin G. Munhall. A model of facial biomechanics for speech production. *The Journal of the Acoustical Society of America*, 106(5) :2834–2842, 1999.
- [3] Yohan Payan and Pascal Perrier. Synthesis of V–V sequences with a 2D biomechanical tongue model controlled by the Equilibrium Point Hypothesis. *Speech Communication*, 22 :185–205, 1997.
- [4] Michel Pitermann and Kevin G. Munhall. An inverse dynamics approach to face animation. *The Journal of the Acoustical Society of America*, 110(3) :1570–1580, 2001.
- [5] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C*. Cambridge University Press, second edition, 1992.
- [6] Maureen Stone, Moise H. Goldstein Jr, and Yongquing Zhang. Principal component analysis of cross sections of tongue shapes in vowel production. *Speech Communication*, 22 :173–184, 1997.
- [7] Demetri Terzopoulos and Keith Waters. Physically-based facial modeling, analysis, and animation. *The Journal of Visualization and Computer Animation*, 1(2) :73–80, 1990.