



HAL
open science

Classification de signaux multidimensionnels utilisant la distribution de Wishart : Application à la reconnaissance de mouvements

Sullivan Hidot, Christophe Saint-Jean, Jean-Yves Lafaye

► **To cite this version:**

Sullivan Hidot, Christophe Saint-Jean, Jean-Yves Lafaye. Classification de signaux multidimensionnels utilisant la distribution de Wishart : Application à la reconnaissance de mouvements. colloque GRETSI, Sep 2007, Troyes, France. p. 1321-1324. hal-00180108

HAL Id: hal-00180108

<https://hal.science/hal-00180108>

Submitted on 11 Jun 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Classification de signaux multidimensionnels utilisant la distribution de Wishart : Application à la reconnaissance de mouvements

Sullivan HIDOT¹, Christophe SAINT-JEAN², Jean-Yves LAFAYE¹

¹Laboratoire Informatique, Image et Interaction (L3i)
Avenue M. Crépeau, Université de La Rochelle, France

²Laboratoire de Mathématiques et Applications (LMA)
Avenue M. Crépeau, Université de La Rochelle, France

shidot@univ-lr.fr, csaintje@univ-lr.fr, jylafaye@univ-lr.fr

Résumé – Cet article propose une nouvelle application de la distribution de Wishart à la reconnaissance et la classification supervisée ou non de mouvements. La signature est la matrice de covariance entre les coordonnées spatiales de capteurs placés sur les principales articulations d’un exécutant. Dans les deux types de classification, on attribue à chaque classe une densité de Wishart de paramètres préalablement estimés dépendant d’une fonction de vraisemblance à maximiser. L’avantage de ces deux approches est d’une part de fournir de bons taux de reconnaissance (classification supervisée) et d’autre part d’extraire des sous-mouvements présents dans chaque classe de mouvements-type *a priori* (classification non supervisée). Nous mettons en valeur la robustesse de la méthode en étudiant la classification pour différents niveaux de lissage des variables-capteurs. Nous présentons les résultats obtenus sur un corpus de 219 mouvements typiques de la danse classique.

Abstract – This paper proposes a novel use of the Wishart distribution within the scope of supervised or unsupervised movement recognition and clustering. As a movement signature, we retain the covariance matrix between spatial coordinates of sensors, conveniently fastened on the performer joints. In both cases, each movement type is supposed to be associated with a special Wishart distribution, whose parameters are liable to be estimated with maximum likelihood technique. The assets of these two approaches are on the one hand (supervised), to provide satisfying rates of recognition, and on the other hand (unsupervised) to point out a finer taxonomy of movement. We assert the robustness of the method when clustering the data with several smoothing levels. We eventually present with the results obtained on a set of 219 classical dancing movements.

1 Introduction

La distribution de Wishart est une version multidimensionnelle de la loi du χ^2 qui caractérise les matrices de covariance estimées de lois gaussiennes multivariées. Dans la littérature, on trouve de nombreux travaux théoriques sur la distribution de Wishart dont l’inverse est extensivement utilisé comme modèle *a priori* pour les méthodes bayésiennes [8]. Cependant, pour autant que nous sachions, les applications à la reconnaissance des formes sont relativement récentes et peu nombreuses [9].

Dans cet article nous proposons une application de la loi de Wishart à la classification supervisée ou non de données spatio-temporelles, réalisations de vecteurs aléatoires, pour lesquels on adopte la décomposition commune en une tendance et une partie résiduelle : $X_t = \varphi(t) + \epsilon_t$. Dans les deux types de classification, la signature retenue est la matrice de covariance du signal analysé. Nous étudions une approche bayésienne en complément d’une méthode géométrique fondée sur l’inertie, que nous avons précédemment exposée et appliquée dans [3, 4]. Nous détaillons ici une approche stochastique, en analysant le vecteur $\epsilon_t = X_t - \varphi(t)$ et nous intéressons à des algorithmes de type moyenne mobile en explorant plusieurs niveaux de lissage. Le terme ϵ_t est supposé gaussien centré, ce qui

induit une distribution de Wishart pour la matrice de covariance empirique.

L’intérêt de notre étude réside dans la discussion des hypothèses sous-jacentes et la gestion de contraintes contradictoires liées à l’estimation de la tendance $\varphi(t)$ et du résidu ϵ_t . Les méthodes géométriques sont plus naturellement dédiées à l’étude des termes tendanciels tandis que les méthodes stochastiques requièrent des hypothèses de stationnarité vérifiées seulement par les termes résiduels. L’étude pratique s’appuie sur l’analyse d’un corpus de données obtenu par capture du mouvement de danseurs du Ballet Atlantique Régine Chopinot (BARC). Le but est de développer une méthode permettant d’aider à reconnaître automatiquement le type d’un mouvement quelconque dont on ignore *a priori* la nature. L’analyse du mouvement dansé est un domaine reliant l’Art et la Science et a fait l’objet de publications récentes [1, 10].

Le plan de l’article est le suivant. Dans la section 2, nous explicitons le modèle choisi et détaillons les deux modes de classification. Nous présentons en section 3 les résultats obtenus sur un corpus de 219 mouvements par niveau de lissage puis nous concluons en section 4.

2 Modèle proposé

2.1 Suppression de la tendance

Le lissage d'une variable s'effectue par moyenne mobile où la largeur de la fenêtre est contrôlée par un paramètre θ . Ce paramètre correspond au nombre de réalisations pour lesquelles on calcule une moyenne locale. La tendance est supprimée en retranchant localement de la variable les moyennes obtenues. Par conséquent, plus θ est faible, plus la variable perd de sa structure et le résidu obtenu est erratique. Si θ est maximum, cela revient à centrer la variable puisqu'on lui retire sa moyenne globale.

2.2 Classification supervisée

D'une manière générale, on fixe C le nombre de classes utilisées pour reconnaître les mouvements. On considère la classe k constituée de N_k mouvements $(\epsilon_{ik})_{i=1}^{N_k}$ privés de leur tendance. Chacun est constitué de n_{ik} captures et identifié par un tableau de taille $n_{ik} \times p$. La matrice de covariance correspondante est donnée par :

$$V_{ik} = \frac{1}{n_{ik}} \epsilon'_{ik} \epsilon_{ik} \in \mathbb{R}^{p \times p} \quad (1)$$

On fait l'hypothèse que les opérateurs $M_{ik} = n_{ik} V_{ik}$ suivent une loi de Wishart de matrice de covariance Σ_k et de degré de liberté n_{ik} :

$$M_{ik} = n_{ik} V_{ik} = \epsilon'_{ik} \epsilon_{ik} \sim W_p(\Sigma_k, n_{ik}) \quad (2)$$

avec $\epsilon_{ik} \sim \mathcal{N}_p(0, \Sigma_k)$. On estime Σ_k en maximisant le logarithme de la vraisemblance et en admettant que les mouvements aient été exécutés de manière indépendante :

$$\log L(\Sigma_k) = \sum_{i=1}^{N_k} \log f(M_{ik}; \Sigma_k, n_{ik}) \quad (3)$$

où f est la densité de la loi de Wishart :

$$f(M; \Sigma, n) = \frac{|M|^{\frac{n-p-1}{2}} \exp\left\{-\frac{1}{2} \text{trace}(\Sigma^{-1} M)\right\}}{2^{\frac{np}{2}} \pi^{\frac{p(p-1)}{4}} |\Sigma|^{\frac{n}{2}} \prod_{j=1}^p \Gamma\left(\frac{1}{2}(n-j+1)\right)} \quad (4)$$

En utilisant les formules de dérivées matricielles [7], on obtient :

$$\frac{\partial \log L(\Sigma_k)}{\partial \Sigma_k} = \sum_{i=1}^{N_k} \left(\frac{1}{2} \Sigma_k^{-1} M_{ik} \Sigma_k^{-1} - \frac{n_{ik}}{2} \Sigma_k^{-1} \right) = 0 \quad (5)$$

En simplifiant par $\frac{1}{2} \Sigma_k^{-1}$, on aboutit à :

$$\widehat{\Sigma}_k = \frac{\sum_{i=1}^{N_k} M_{ik}}{\sum_{j=1}^{N_k} n_{jk}} = \frac{\sum_{i=1}^{N_k} (n_{ik} V_{ik})}{\sum_{j=1}^{N_k} n_{jk}} \quad (6)$$

L'estimée $\widehat{\Sigma}_k$ coïncide avec la somme des éléments de la classe k , pondérée par la proportion de la taille des mouvements. Étant donné un mouvement ϵ de taille n , on cherche à calculer sa probabilité d'appartenance à la classe k , notée $P(k|\epsilon)$. D'après la règle de Bayes :

$$P(k|\epsilon) \propto P(k)P(\epsilon|k) \quad (7)$$

où $P(\epsilon|k)$ est la probabilité d'observer le mouvement ϵ dans la classe k , i.e. que le mouvement soit de covariance conforme à une densité de Wishart $W_p(\widehat{\Sigma}_k, n)$ de l'opérateur $\epsilon' \epsilon$ associé à ϵ :

$$P(\epsilon|k) = f(\epsilon' \epsilon; \widehat{\Sigma}_k, n) \quad (8)$$

La probabilité $P(k)$ est le poids de la classe k . Puisque les fréquences dans l'échantillon ne sont pas nécessairement représentatives, nous considérons l'équiprobabilité des classes plutôt que la fréquence empirique de leur effectif :

$$P(k) = \frac{1}{C}, \quad 1 \leq k \leq C \quad (9)$$

Le critère d'affectation d'une classe pour le mouvement ϵ consiste à calculer l'indice de la classe pour lequel le logarithme de la probabilité $P(k|\epsilon)$ est maximum :

$$\arg \max_{k=1, \dots, C} \log P(k|\epsilon) = \arg \max_{k=1, \dots, C} \log P(\epsilon|k) \quad (10)$$

2.3 Classification non supervisée

Dans le cas non supervisé, on détermine *a posteriori* quels éléments constituent chacune des classes. Celles-ci sont construites à l'aide de la procédure EM [2]. On choisit K un nombre de clusters C_1, \dots, C_K à estimer pour la classification et on dispose d'un nombre N d'opérateurs $(M_i)_{i=1}^N$ de mouvements sans tendance de taille $(n_i)_{i=1}^N$ où $M_i = n_i V_i$. Pour le cluster C_k , on pose $\Theta_k = (\pi_k, \Sigma_k)$, où π_k est le poids du cluster et Σ_k la matrice de covariance de la distribution de Wishart associée. On introduit dans le modèle une indicatrice d'appartenance d'un mouvement à un cluster :

$$z_{ik} = \begin{cases} 1 & \text{si } M_i \in C_k \\ 0 & \text{sinon} \end{cases} \quad (11)$$

La vraisemblance est donnée par l'expression suivante où l'on introduit de façon naturelle les termes z_{ik} :

$$L(\Theta) = \prod_{i=1}^N \prod_{k=1}^K (\pi_k P(M_i|\Theta_k))^{z_{ik}} \quad (12)$$

$$\log L(\Theta) = \sum_{i=1}^N \sum_{k=1}^K z_{ik} \log(\pi_k P(M_i|\Theta_k)) \quad (13)$$

La quantité $P(M_i|\Theta_k)$ est la probabilité d'appartenance du mouvement M_i à C_k , connaissant les paramètres Θ_k :

$$P(M_i|\Theta_k) = f(M_i; \Sigma_k, n_i) \quad (14)$$

La procédure EM consiste à maximiser l'espérance du logarithme de la vraisemblance :

$$E \log L(\Theta) = \sum_{i=1}^N \sum_{k=1}^K E(z_{ik}) \log(\pi_k P(M_i|\Theta_k)) \quad (15)$$

Le terme $E(z_{ik})$ est estimé par \widehat{z}_{ik} . À chaque étape et ce jusqu'à convergence, on ré-estime les paramètres z_{ik} (étape 'Expectation') puis π_k et Σ_k (étape 'Maximisation') de la manière suivante. En entrée :

- η : Seuil de convergence.
- $(V_i)_{i=1}^N$: Matrices de covariance des mouvements.
- $(n_i)_{i=1}^N$: Tailles des mouvements.
- K : Nombre de clusters à estimer.
- $\Theta^{(0)}$: Valeurs initiales des Θ_k .

E-step : Estimation des $z_{ik}^{(t)}$

$$\hat{z}_{ik}^{(t)} = \frac{\pi_k P(M_i | \hat{\Theta}_k^{(t)})}{\sum_{l=1}^K \pi_l P(M_i | \hat{\Theta}_l^{(t)})} \quad (16)$$

M-step : Estimation des $\Theta_k^{(t+1)}$

$$\hat{\pi}_k^{(t+1)} = \frac{\sum_{i=1}^N \hat{z}_{ik}^{(t)}}{N} \quad (17)$$

$$\hat{\Sigma}_k^{(t+1)} = \frac{\sum_{i=1}^N \hat{z}_{ik}^{(t)} n_i V_i}{\sum_{j=1}^N \hat{z}_{jk}^{(t)} n_j} \quad (18)$$

$t \leftarrow t + 1$
 Jusqu'à $|E \log L(\hat{\Theta}^{(t+1)}) - E \log L(\hat{\Theta}^{(t)})| < \eta$

La ré-estimation des z_{ik} donnée par (16) se simplifie après élimination des constantes communes et s'exprime de la manière suivante :

$$z_{ik} = \frac{\pi_k |\Sigma_k|^{-\frac{n_i}{2}} \exp \left\{ -\frac{n_i}{2} \text{trace}(\Sigma_k^{-1} V_i) \right\}}{\sum_{l=1}^K \pi_l |\Sigma_l|^{-\frac{n_i}{2}} \exp \left\{ -\frac{n_i}{2} \text{trace}(\Sigma_l^{-1} V_i) \right\}} \quad (19)$$

L'expression (17) est obtenue en maximisant (13) par dérivation avec la contrainte $\sum_k \pi_k = 1$. On établit la formule (18) de façon analogue à celle obtenue pour la classification supervisée.

3 Expérimentations

3.1 Données

Les mouvements ont été enregistrés au Cyberdome et exécutés par des danseurs du BARC. L'enregistrement d'un mouvement est obtenu de la manière suivante. On place 15 capteurs sur les principales articulations du danseur et l'on enregistre leurs coordonnées spatiales avec une fréquence d'acquisition de 25 captures par seconde ; soit un ensemble de $p = 15 \times 3 = 45$ variables aléatoires discrètes réelles. Nous disposons d'un corpus de $N = 219$ mouvements se répartissant dans l'un des 14 mouvements type : Marche, Enveloppé, Chute Pliée, Enroulé Tête, Enroulé Bassin, Grand Plié, Marche Glissée, Saut Attitude, Chute Repliée, Chute Allongée, Jeté, Tour, Suspension ou Saut Temps Levé. Chaque mouvement type a été réalisé à plusieurs reprises par des danseurs de différentes corpulences. Les classes ne sont pas homogènes et vont de 6 éléments pour la classe 'Saut Temps Levé' à 25 pour la classe 'Jeté' ($6 \leq N_k \leq 25$). De même, les tailles des mouvements vont de 52 à 338 captures ($52 \leq n_k \leq 338$).

3.2 Résultats expérimentaux

Pour l'ensemble des résultats présentés, nous avons considéré deux cas de figures. Le premier consiste à sélectionner un nombre de classes *a priori* égal au nombre de mouvements type ($C = 14$). Le second prend en compte les différentes variantes possibles pour certains types de mouvements. Par exemple, nous disposons de tours sautés ou non sautés, de marches orientées différemment ... La matrice de covariance entre variables-capteur n'est pas robuste sur les différentes variantes d'exécution mais également sur l'orientation et la position initiale de l'exécutant.

Nous avons noté 5 variantes possibles, qui s'additionnent aux 14 classes initiales, soit $C = 19$ classes.

Le but de la classification supervisée est de calculer le taux de reconnaissance en fonction du niveau de lissage et de déterminer dans quelle mesure se manifeste la perte d'information due à la suppression de la tendance. Le tableau 1 présente les taux de reconnaissance en pourcentage selon différentes tailles de la fenêtre de lissage. La procédure choisie est la validation croisée avec le plus proche voisin.

TAB. 1: Classification supervisée en supprimant la tendance en fonction du niveau de lissage.

θ	1	2	3	5	10	<i>max</i>
$C = 14$	77.17	84.93	87.21	93.60	94.06	97.71
$C = 19$	76.26	80.82	85.39	90.41	89.95	92.69

Le pourcentage de reconnaissance augmente globalement en fonction de la taille de la fenêtre. Les taux vont de 76.26% ($C = 19$, 167 mouvements bien classés) jusqu'à 97.71% ($C = 14$, 214 mouvements bien classés). On constate la robustesse de la méthode, car les résultats sont satisfaisants dans des conditions s'éloignant des hypothèses du modèle théorique pour les fortes valeurs du paramètre de lissage θ (stationnarité et modèle gaussien). Pour $\theta = 1$, les variables résiduelles obtenues correspondent à la plus grande perte d'information propre au mouvement. À l'inverse, les variables centrées ($\theta = \text{max}$) fournissent toute l'information apportée par le mouvement, jouant ainsi un rôle d'autant plus discriminatoire pour les classes. Lorsque l'on retire la tendance, on accède à d'autres critères caractéristiques de typologies moins évidentes que celles liées au côté spatial. Cela peut être lié soit à des aspects proches du bruit comme les erreurs de mesure ou l'imprécision des capteurs, soit à la morphologie du danseur, mise en valeur une fois que l'on a supprimé l'information propre à l'exécution. Globalement, les pourcentages sont plus faibles pour $C = 19$. Ce phénomène s'explique par un inconvénient de la validation croisée, qui agit sur des classes dont certaines ont été subdivisées en deux sous-classes. Le fait de retirer un élément à une classe d'effectif faible risque de modifier fortement la nature de la classe en question. Cependant, l'approche validation croisée a l'avantage de donner une estimation asymptotiquement non biaisée du taux d'erreur [6].

Pour la classification non supervisée, nous n'estimons pas de taux de reconnaissance puisque les clusters n'ont pas d'étiquetage *a priori*. Nous proposons de comparer les classements obtenus de la manière suivante. Supposons P et Q deux graphes de partition, *i.e.* $P(i, j) = 1$ si les mouvements i et j sont dans la même classe, et 0 sinon (idem pour Q). On adopte l'indice de cohérence suivant :

$$A(P, Q) = \frac{1}{N^2} \sum_{i, j=1}^N (P(i, j)Q(i, j) + (1 - P(i, j))(1 - Q(i, j))) \quad (20)$$

La quantité $D(P, Q) = 1 - A(P, Q) \in [0, 1]$ correspond à une différence symétrique et s'interprète comme une dis-

tance entre P et Q . Plus D est proche de 0, plus les partitions sont homogènes. Ainsi, pour chaque estimation, nous calculons la distance entre la partition obtenue et la partition de référence des mouvements connue *a priori*. Notons que la comparaison des valeurs pour $K = 14$ ou 19 n'est pas strictement rigoureuse car la valeur de l'indice dépend du type des partitions et en particulier du nombre de classes. Une approche rigoureuse supposerait un indice plus élaboré comme par exemple la vraisemblance de la valeur de la différence symétrique sous une hypothèse d'indépendance des partitions comparées.

Nous fixons le seuil de convergence η égal à 10^{-5} . Pour la première itération, nous choisissons aléatoirement les probabilités d'appartenance $z_{ik}^{(0)}$ afin d'estimer $\Theta_k^{(1)}$ via l'étape de maximisation. Afin de rendre la procédure moins sensible aux minima locaux, nous rajoutons une étape intermédiaire (entre la ré-estimation et la maximisation) qui consiste à faire une classification par tirage aléatoire des z_{ik} selon une loi multinomiale. Pour éviter des problèmes numériques, nous supprimons les classes vides apparaissant à chaque nouvelle itération. Le tableau 2 contient les moyennes et écarts-type des distances entre partitions après 50 itérations pour chacun des cas, en fonction des valeurs de θ , de C et de K .

TAB. 2: Classification non supervisée : Moyennes et écarts-type des différences symétriques en fonction de K et de C (50 itérations).

θ	$C = K = 14$	$C = 14, K = 19$	$C = K = 19$
max	0.101 ± 0.014	0.087 ± 0.013	0.079 ± 0.013
10	0.093 ± 0.010	0.085 ± 0.008	0.075 ± 0.007
5	0.097 ± 0.010	0.088 ± 0.009	0.081 ± 0.009
3	0.103 ± 0.012	0.094 ± 0.010	0.088 ± 0.009
2	0.105 ± 0.013	0.098 ± 0.008	0.092 ± 0.009
1	0.110 ± 0.011	0.101 ± 0.009	0.094 ± 0.009

Les moyennes diminuent globalement suivant que le nombre de classes ou de clusters augmente. Dans les 3 cas considérés, on constate un minimum local pour $\theta = 10$, les moyennes et écarts-type étant plus élevés pour $\theta = max$. En effet, les covariances analysées avec la tendance tiennent compte des périodes d'immobilité du danseur, ce qui n'est pas le cas si l'on supprime la tendance [5]. Il y a donc des confusions possibles pour deux mouvements différents ayant des périodes statiques. Pour $\theta = 10$, il s'agit d'un bon compromis où l'immobilité est suffisamment négligée tout en conservant la structure du mouvement.

4 Conclusion

La distribution de Wishart est typiquement utilisée pour caractériser les matrices de covariance d'échantillons d'un vecteur aléatoire gaussien. Nous en avons proposé une nouvelle application, la reconnaissance et la classification de mouvements. Un mouvement est caractérisé par sa matrice de covariance inter-capteurs, information de type multi-dimensionnelle. La classification supervisée, four-

nissant de bons taux de reconnaissance même en dehors du cadre classique, permet de comparer les résultats en fonction du niveau de lissage choisi pour supprimer la tendance des variables-capteurs. Pour la classification non supervisée, nous avons appliqué la procédure EM et avons analysé les résultats en calculant la différence symétrique entre partitions. La classification non supervisée extrait des sous-mouvements, issus des éventuelles variantes possibles pour un même mouvement type.

Nous avons choisi arbitrairement le nombre de clusters sans nous intéresser à des critères permettant de fournir une valeur pertinente pour l'estimation. Des approches pour sélectionner le modèle existant, telles que l'adaptation du critère d'information bayésienne (BIC) ou d'Akaike (AIC), mais feraient à eux seuls l'objet d'un article à part entière.

References

- [1] F. Chenevière, S. Boukir and B. Vachon. *A HMM-based dance gesture recognition*. International Workshop on Systems, Signals and Image Processing (IWSSIP), pages 322-326, Manchester, UK, 2004.
- [2] A.P. Dempster, N.M. Laird and D.B. Rubin. *Maximum Likelihood from Incomplete Data via the EM algorithm*. Journal of the Royal Statistical Society B, 39(1): 1-38, 1977.
- [3] S. Hidot, J.-Y. Lafaye et C. Saint-Jean. *Analyse factorielle d'opérateurs pour l'étude du mouvement: Application à la danse*. Atelier "Fouille de données", Journées francophones Extraction et Gestion des Connaissances (EGC), pages 21-30, Lille, 2006.
- [4] S. Hidot, J.-Y. Lafaye and C. Saint-Jean. *Discriminant factor analysis for movement recognition : application to dance*. To appear in International Journal of Machine Graphics and Vision, Vol. 15, Warszawa, Poland, 2007.
- [5] S. Hidot, J.-Y. Lafaye et C. Saint-Jean. *Propriétés et interprétation de la covariance relationnelle en ACP*. Revue Traitement du Signal, 24(1), 2007.
- [6] A.K. Jain, R.P.W. Duin and J. Mao. *Statistical pattern recognition : A review*. IEEE Trans. Pattern Analysis and Machine Intelligence, 22(1) :4-37, 2000.
- [7] K.B. Petersen and M.S. Pedersen. *The matrix cookbook*. Version 16 février 2006, <http://www.est.uc3m.es/daejin/files/MatrixCookBook.pdf>.
- [8] D.B. Rowe. *Multivariate Bayesian statistics: models for source separation and signal unmixing*. Chapman and Hall/CRC, 2003.
- [9] H. Skriver, A.A. Nielsen and K. Conradsen. *Evaluation of the Wishart test statistics for polarimetric SAR data*. International Geoscience and Remote Sensing Symposium (IGARSS), Vol. 2: 699-701, 2003.
- [10] F. Thullier and H. Moufti. *Multi-joint coordination in ballet dancers*. Neuroscience Letters, 369(1): 80-84, 2004.