Quelques problèmes posés par la gestion des primitives géométriques en vision par ordinateur

Xavier PENNEC   Nicholas AYACHE

Projet EPIDAURE
INRIA, B.P. 93, 2004 route des Lucioles
06902 SOPHIA ANTIPOLIS Cedex, FRANCE
Email: Xavier.Pennec@sophia.inria.fr

1 Introduction

De nombreux algorithmes en vision par ordinateur, et en particulier en reconnaissance des formes, utilisent des primitives géométriques simples comme les points. D'un autre côté, les modèles géométriques du monde réel amènent souvent à considérer des primitives plus complexes, par exemple des droites [Gri92], des plans [Fau93], des points ou des plans orientés, des repères [PA94, PT96], etc. L'utilisation de telles primitives plus sélectives est par ailleurs justifiée par la réduction de complexité qu'elle apporte pour la reconnaissance (par exemple de $O(n^4)$ à $O(n^2)$ dans [PA94]), et le gain de précision au niveau du recalage.

La gestion de ces primitives soulève cependant un certain nombre de problèmes qui peuvent conduire à des paradoxes. La première partie sera consacrée aux probabilités géométriques, et le paradoxe de Bertrand nous montrera que si l'on n'y prend garde, on peut obtenir des probabilités arbitraires en fonction des méthodes de calcul utilisées. Le second point abordé dans cet article concerne la moyenne de primitives géométriques, ou plus généralement la notion d'espérance mathématique. Toutes les statistiques que l'on peut faire sur nos primitives dépendent de cette définition (ou de celle d'une valeur centrale équivalente), mais un exemple sur des droites en 2D montrera qu'un simple changement de repère fausse cette estimation !

S'il existe parfois des solutions ad-hoc pour chacun de ces problèmes, celles-ci ne sont en général valables que pour certains types de primitives dans certaines conditions. L'objectif de cet article n'est pas de présenter une théorie complète, mais d'exhiber les problèmes et d'esquisser des solutions, basées sur la notion d'invariance de l'opérateur considéré sous l'action d'un groupe de transformation. Nous renvoyons le lecteur au rapport de recherche associé [PA96] pour les détails et les compléments.

2 Probabilités géométriques classiques

L'objet des probabilités géométriques est de déterminer la probabilité d'observer un événement quand un certain nombre d'éléments géométriques sont répartis au hasard. Le problème qui se pose ici est la signification de “répartis au hasard” : le paradoxe de Bertrand est un calcul simple de probabilité où trois définitions du hasard donnent trois résultats différents. Plus précisément, on peut ramener ce problème à la notion de distribution (ou mesure) uniforme sur un ensemble, et une hypothèse raisonnable est de supposer l'invariance de cette distribution par l'action d'un groupe de transformation. Des résultats plus récents en théorie des groupes de Lie permettent de savoir si cette mesure invariante existe et de la calculer.

2.1 Le Paradoxe de Bertrand

Le problème soulevé par J. Bertrand en 1907 est de calculer la probabilité qu'une "corde aléatoire" d'un cercle ait une longueur supérieure au côté d'un triangle équilatéral inscrit dans ce cercle. En supposant un cercle de rayon 1, le côté de ce triangle est $\sqrt{3}$. Trois méthodes sont envisagées pour résoudre ce problème, illustrées figure 1.

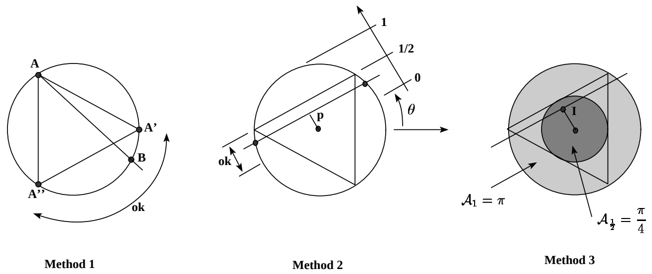

FIG. 1 – *Trois méthodes pour calculer la probabilité qu'une corde "aléatoire" ait une longueur supérieure au côté d'un triangle équilatéral inscrit. De gauche à droite, une probabilité de $\frac{1}{3}$, $\frac{1}{2}$ et $\frac{1}{4}$.*

Méthode 1: Une corde intersecte par définition le cercle en deux points que l'on peut supposer indépendants et distribués uniformément. En supposant que l'un des points soit A sur la figure 1, le second point doit se trouver entre A' et A'' pour que la longueur de la corde soit supérieure au côté du triangle. Comme ceci représente un tiers de la circonférence du cercle, la probabilité cherchée est $\frac{1}{3}$.

Méthode 2: Une corde est caractérisée par sa distance p au centre du cercle (entre 0 et 1) et son orientation θ (entre 0 et 2π) par rapport à une droite fixe. Si l'on trace un triangle équilatéral dont un côté est parallèle à la corde, on voit que la distance p doit être inférieure à un demi pour que la corde ait une longueur suffisante. En supposant une orientation et une distance à l'origine uniforme, on trouve une probabilité normalisée de $\frac{1}{2}$.

Méthode 3: Une corde est définie de manière unique par le pied I de la perpendiculaire joignant le centre du cercle. Ce point doit être dans un disque de rayon $\frac{1}{2}$ pour assurer une longueur suffisante à la corde. En supposant une distribution uniforme de I dans l'intérieur du cercle original, la probabilité normalisée est de $\frac{1}{4}$.

Les trois solutions présentées sont correctes, mais ne se réfèrent pas à la même notion d'uniformité dans la façon de choisir une corde. En utilisant la représentation (p, θ) décrite dans la seconde méthode, on peut calculer ([KM63]) que les mesures de probabilité associées sont respectivement :

$$d\sigma_1 = \frac{dp.d\theta}{2\pi\sqrt{1-p^2}} \qquad d\sigma_2 = \frac{dp.d\theta}{2\pi} \qquad d\sigma_3 = p.\frac{dp.d\theta}{\pi}$$

Seule la seconde est invariante par l'action des rotations, translations et réflections.

2.2 Mesure invariante sur une variété

La solution du "paradoxe" réside donc dans la définition de la notion d'uniformité à partir d'une contrainte d'invariance : soit x_0 un point fixe et x une variable de $\mathbb{R}^n$. L'équation : $d(x + x_0) = dx$ traduit essentiellement l'invariance par translation de la mesure de Lebesgue. De même, on ne peut comparer des objets géométriques que par l'intermédiaire d'un groupe de transformation (figure 2).

Notons $\mathcal{M}$ (pour manifold) la variété des objets considérés [1], $f \star x$ l'action d'une transformation f du groupe $\mathcal{G}$ sur un point x de la variété, et $d\mathcal{M}(x)$ une mesure sur la variété. On définit la mesure uniforme comme celle qui reste invariante sous l'action de n'importe quelle transformation du groupe :

1. On rappelle qu'une variété (différentielle) est un ensemble localement difféomorphe (*i.e.* en bijection différentiable) à un espace vectoriel. Ceci nous permet de dériver et d'agir *localement* comme dans un espace vectoriel.

on impose que $d\mathcal{M}(f \star x) = d\mathcal{M}(x)$ en tout point x. On peut rechercher directement cette mesure pour une représentation de la variété (voir par exemple [KM63]), mais un formalisme général est développé dans [San76], dont nous résumons ici les résultats de manière très simplifiée.

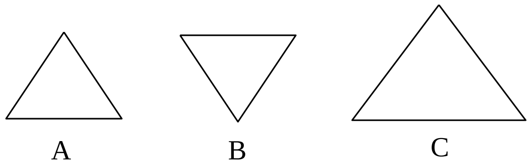

FIG. 2 – *Comparaison d'objets géométriques: on peut dire les trois triangles sont similaires (relativement au groupe des similitudes), ou bien que* **A** *et* **B** *seulement sont congruents (pour les transformations rigides), ou encore qu'ils sont tous différents (en translation).*

La variété est dite *homogène* ou *transitive* si l'on peut atteindre n'importe quel autre point de la variété à partir de l'origine grâce à une transformation du groupe, i.e. si $\mathcal{G}\star\mathcal{O} = \{g\star\mathcal{O} \ / \ g \in \mathcal{G}\} = \mathcal{M}$. Cela veut dire que les primitives considérées n'ont pas d'invariant (sous l'action de ce groupe). En fait, on suppose que l'on peut séparer les primitives en une partie invariante (que l'on ne considère pas ici) et une partie homogène.

Dans le cas d'une variété homogène, on appelle groupe d'isotropie le sous-groupe $\mathcal{H}$ des transformations de $\mathcal{G}$ qui laisse invariant un point $\mathcal{O}$ de la variété que l'on choisit comme origine : $\mathcal{H} = \{h \in \mathcal{G} \ / \ h \star \mathcal{O} = \mathcal{O}\}$, et coset $\hat{x}$ d'un point $x \in \mathcal{M}$ l'ensemble des transformations qui amène l'origine sur x : $\hat{x} = \{g \in \mathcal{G} \ / \ g \star \mathcal{O} = x\}$. On utilisera la notation f_x pour indiquer un représentant (quelconque) de $\hat{x}$.Afin de pouvoir dériver et intégrer sans problème, on suppose que l'on utilise une représentation minimale de la variété dont le domaine de définition couvre presque toute la variété. Avec ces notations, on montre qu'il existe sur $\mathcal{M}$ une mesure invariante par le groupe $\mathcal{G}$ si et seulement si :

$$\forall h \in \mathcal{H}, \qquad \left| \left.\frac{\partial(h \star e)}{\partial e}\right|_{e=\mathcal{O}} \right| = 1 \tag{1}$$

Toute transformation qui laisse l'origine invariante doit donc également conserver l'élément de volume infinitésimal associé. Si cette condition est verifiée, la mesure invariante est unique à un scalaire près :

$$d\mathcal{M}(x) = \frac{dx}{|J(f_x)|} \qquad \text{avec} \qquad J(f_x) = \left.\frac{\partial(f_x \star e)}{\partial e}\right|_{e=\mathcal{O}} \tag{2}$$

2.3 Mesure invariante sur les repères

Les primitives que nous utilisons comme exemple d'application dans cet article sont des *repères*, i.e. des points munis d'un trièdre orthonormé direct. Notons qu'un repère peut être identifié à la transformation rigide du repère canonique au repère considéré.

Nous représentons une rotation d'angle θ autour de l'axe unitaire n par le *vecteur rotation* $r = \theta.n$ avec comme domaine la boule ouverte $\mathcal{B}^3(0, \pi)$. Cette carte représente de manière unique toutes les rotations sauf les réflections (rotations d'angle π), mais on peut l'étendre à toute la variété en identifiant les deux vecteurs rotation antipodaux $r = \pi.n$ et $r' = -\pi.n$ sur le bord du domaine.

La représentation choisie pour les repères est donc : $f = (r, x)$ où r est un vecteur rotation représentant le trièdre et x la position du point. Il est naturel de prendre l'origine à l'identité : $\mathcal{O} = (\ I_d, 0)$, et le groupe d'isotropie se réduit à l'identité. On obtient alors la mesure invariante :

$$d\mathcal{M}(f) = \frac{4\sin^2(\theta/2)}{\theta^2}\, dr\, dx$$

2.4 Sélectivité des primitives

Dans un algorithme de matching de type "plus proche voisin itéré", souvent utilisé comme étape de vérification, on considère généralement que les primitives que l'on mesure sont bruitées et on autorise en conséquence l'appariement de x dans la première image avec y dans la seconde (avec la transformation f) si les deux primitives se correspondent à une certaine erreur près. La probabilité *a priori* de faux positif est définie comme la probabilité d'accepter cet appariement si x est en fait un outlier (auquel on attribue une densité uniforme sur la première image), et elle se calcule comme le volume de la zone d'erreur normalisé par le volume de l'image. Ces volumes sont bien sûr calculés avec la mesure uniforme.

A titre d'exemple, on considère dans [PA94] que deux repères sont appariés si la distance entre leurs points est inférieure à d_0 et si l'angle de la rotation d'ajustement entre leurs trièdres est inférieur à θ_0. Si l'on suppose une image cubique de côté l (seule la position d'un repère aléatoire uniforme est contrainte, l'orientation restant libre), on obtient une probabilité de faux positif de :

$$P(f \star x \leftrightarrow y) = \left(\frac{\theta_0 - \sin\theta_0}{\pi}\right)\left[\frac{4}{3}\left(\frac{d_0}{l}\right)^3\right]$$

Nous avons isolé dans le premier terme la probabilité de faux positif due au trièdre seul, ce qui reflète le gain de sélectivité obtenu en utilisant le repère au lieu du point seul : même pour une borne très lâche de $\theta_0 = \pi/2 = 90\,\mathrm{deg}$, plus de 80% des faux appariements sont rejetés. Pour une borne (usuelle) de $\theta_0 = \pi/10 \simeq 20\,\mathrm{deg}$, la probabilité de faux positif due au trièdre chute à 0.0016 : il faudrait pouvoir diviser l'incertitude sur la position par 10 pour obtenir une telle sélectivé avec les points seuls.

3 Espérance d'une primitive aléatoire

On modélise habituellement l'incertitude sur la mesure d'une donnée par une densité de probabilité, dont on ne conserve, pour des raisons d'efficacité calculatoire, que quelques paramètres. La méthode classique ([DW88, Aya91, ZF92]) considère ainsi la représentation des primitives géométriques aléatoires comme un *vecteur* aléatoire $\mathbf{x}$ de densité $p_\mathbf{x}$ dont on ne conservera que les moments du premier et du second ordre. On se concentre ici sur le moment du premier ordre : l'espérance, et sa mesure statistique : la moyenne. Si $\mathcal{D}$ est le domaine de définition de notre représentation, on écrira donc :

$$\bar{x} = \mathbf{E}(\mathbf{x}) = \int_\mathcal{D} y.p_\mathbf{x}(y).dy \qquad \text{et} \qquad \mathbf{E}(\{x_i\}) = \frac{1}{n}\sum x_i$$

Nous pensons que ces opérateurs sont mal définis dans le cadre des variétés. La première raison est que le résultat de l'intégrale ou de la somme n'est pas obligatoirement dans le domaine de définition : une somme de matrices de rotations n'a pas de raison d'être orthogonale dans le cas général. Le seconde raison est que cet opérateur d'espérance ne commute pas (en général) avec un changement de repère ou l'application d'une transformation : $\mathbf{E}(f \star \mathbf{x}) \neq f \star \mathbf{E}(\mathbf{x})$ (voir exemple ci-dessous). Nous sommes donc amenés à considérer la notion d'espérance de Fréchet, moins forte que l'espérance classique (ni existence ni unicité dans le cas général) mais toujours proprement définie dans un espace métrique.

3.1 Espérance d'une droite 2D

Considérons par exemple une droite vectorielle 2D orientée : elle peut être représentée par un vecteur unitaire ou plus simplement par son angle $\theta \in \mathcal{D} =]-\pi\,;\,\pi[$ par rapport à un axe donné. On définit donc une droite aléatoire par une densité $\rho(\theta)$, et sa moyenne par $\quad \bar{\theta} = E(\theta) = \int_\mathcal{D} \theta.\rho(\theta)\frac{d\theta}{2\pi}$

Notons que $d\theta$ est la mesure uniforme pour les droites soumises à l'action des rotations. L'action d'une rotation d'angle λ est simplement l'addition des angles (modulo 2π) et peut également être vue comme un changement de repère. Nous avons tracé en figure 3 une densité de moyenne $\bar{\theta}_0 = 0$. Si nous changeons de référentiel, c'est à dire appliquons une rotation d'angle λ, on peut voir que l'espérance de la droite tournée devient $\bar{\theta}(\lambda) = \sin(\lambda)$, alors que la rotation de l'espérance est $\underline{\theta}(\lambda) = \lambda$. En particulier, pour une rotation de π, on trouve que $\underline{\theta}_\pi = \pi$ et $\bar{\theta}_\pi = 0$!

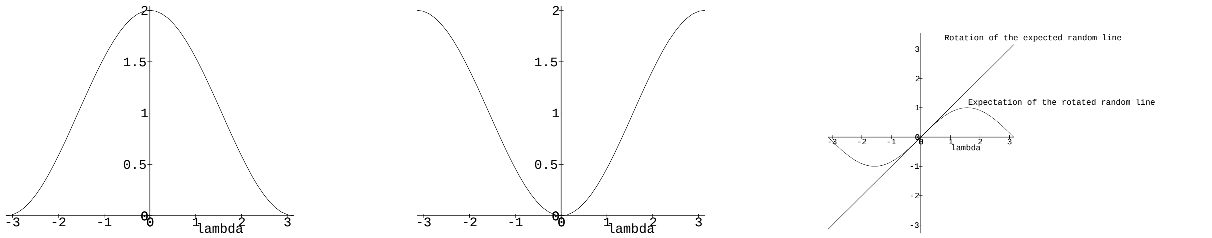

FIG. 3 – *Gauche: Densité originale $\rho_0(\theta) = 2\cos(\theta/2)^2$. L'espérance est $\bar{\theta}(0) = 0$. Milieu: densité après rotation d'un angle $\lambda = \pi$. L'espérance $\bar{\theta}(\pi)$ est encore 0, alors que la rotation de l'espérance est $\underline{\theta}_\pi = \pi$. Droite : comparaison de l'espérance de la droite tournée et de la rotation de la droite espérée.*

3.2 Espérance au sens de Fréchet

La plupart des solutions pour résoudre ce problème tournent autour de l'idée de centrer la densité ou le nuage de mesures avant d'en prendre l'espérance ou la moyenne. On parle d'ailleurs aussi de *valeur centrale* pour la moyenne ou ses équivalents. L'espérance au sens de Fréchet est un formalisme bien posé qui réalise cette idée : étant donné une primitive aléatoire $\mathbf{x}$, la mesure de centralité d'un point y de la variété est définie par la variance de $\mathbf{x}$ par rapport au point y, et on appelle valeur moyenne de $\mathbf{x}$ le (ou les) point(s) minimisant ce critère:

$$\mathbb{E}(\mathbf{x}) = \arg\min_{y \in \mathcal{M}} \left(\mathbf{E}_{\mathbf{x}}(\ \mathrm{dist}(y, \mathbf{x})^2) \right) \tag{3}$$

Notons que, dans le cadre d'une variété métrique, l'espérance d'une observable (fonction de $\mathcal{M}$ dans $\mathbb{R}$) est indépendante de la représentation choisie et bien définie par

$$E_{\mathbf{x}}(\ \mathrm{dist}(y, \mathbf{x})^2) = \int_{\mathcal{M}} \mathrm{dist}(y, z)^2 . p_{\mathbf{x}}(z) . d\mathcal{M}(z)$$

De manière similaire, on définit l'ensemble des valeurs moyennes empiriques d'un ensemble discret

$$\mathbb{E}(\{x_i\}) = \arg\min_{y \in \mathcal{M}} \left(\sum \ \mathrm{dist}(y, x_i)^2 \right) \tag{4}$$

et plus généralement encore les *valeurs centrales d'ordre* α par

$$\mathbb{E}^\alpha(\mathbf{x}) = \arg\min_{y \in \mathcal{M}} (\mathbf{E}_{\mathbf{x}}(\ \mathrm{dist}(y, \mathbf{x})^\alpha)^{1/\alpha}$$

On obtient ainsi les modes de la densité pour $\alpha = 0$, les “primitives médianes” pour $\alpha = 1$ et bien sûr les valeurs ou primitives moyennes pour $\alpha = 2$.

En dehors du cas particulier des espaces vectoriel, où cette définition de l'espérance coïncide avec la définition classique, on n'a généralement plus l'existence ni l'unicité de la valeur moyenne. Toutefois, nous avons rarement observé plusieurs valeurs moyennes en pratique, sauf dans les cas où la densité n'est plus vraiment en rapport avec une erreur de mesure (densités multimodales symétriques sur la sphère par exemple).

3.3 Distance invariante et “invariance” de l'espérance

Nous avons obtenu avec l'espérance au sens de Fréchet une définition propre des primitives moyennes, qui ne dépend pas de la représentation choisie pour la variété. Toutefois, nous ne pouvons assurer la propriété “d'invariance” de l'espérance suivante

$$\mathbb{E}(g \star \mathbf{x}) = g \star \mathbb{E}(\mathbf{x}) \qquad \text{et} \qquad \mathbb{E}(\{g \star x_i\}) = g \star \mathbb{E}(\{x_i\})$$

que si la distance définie sur la variété est invariante :

$$\forall f \in \mathcal{G}, \qquad \forall (x, y) \in \mathcal{M}^2 \qquad \mathrm{dist}(x, y) = \ \mathrm{dist}(f \star x, f \star y)$$

Nous renvoyons le lecteur à [PA96] pour la détermination d'une telle distance.

L'utilisation d'une distance invariante permet par ailleurs d'obtenir un comportement cohérent lorsqu'on cherche à déterminer une transformation à partir d'appariements de primitives, en minimisant par exemple le critère des moindres carrés : $C(f) = \sum_i \operatorname{dist}^2(f \star x_i, y_i)$. En effet, si F est la transformation minimisant ce critère, on s'attend à ce que $F^{(-1)}$ minimise le critère symétrique $C'(f) = \sum_i \operatorname{dist}^2(x_i, f \star y_i)$, et à ce que $F' = g^{(-1)} \circ F \circ g$ (changement de repère) soit la solution trouvée si les données ont subi la même transformation $g : x'_i = g \star x_i$ et $y'_i = g \star y_i$. Ces propriétés sont assurées par l'invariance de la distance.

3.4 Utilisation pratique sur les repères

En utilisant le vecteur rotation pour représenter les trièdres (et les rotations), on peut montrer que l'angle de la rotation d'ajustement entre les deux trièdres $\theta(r_1, r_2) = \|r_2^{(-1)} \circ r_1\|$ est une distance. On obtient alors une distance invariante sur les repères (et les transformations rigides) avec

$$\operatorname{dist}^2(f_1, f_2) = \lambda_\theta^2 . \|r_2^{(-1)} \circ r_1\|^2 + \lambda_x^2 . \|x_2 - x_1\|^2$$

On pondère habituellement chacun des deux termes par l'inverse de grandeurs caractéristiques : domaine de variation ou variance. Une application pratique de cette distance et de la moyenne de repères est présentée dans [PA96] avec la détermination statistique de la stabilité des points extrémaux extraits sur 24 images IRM 3D de la tête d'un même patient.

4 Conclusion

Nous avons soulevé dans cet article quelques problèmes posés par la généralisation au cas de primitives géométriques quelconques des algorithmes utilisées classiquement sur les points, et nous avons montré que la définition rigoureuse des opérateurs de base en suivant les règles suivantes offrait une solution cohérente : invariance par rapport à la représentation choisie pour la variété, et "invariance" (ou commutativité) des opérateurs par rapport à l'action d'un groupe de transformation.

Nous travaillons actuellement sur l'extension de ce formalisme pour les moments du second ordre (matrices de covariance) et certaines techniques statistiques très usitées en vision : distance de Mahalanobis, test de confiance, filtre de Kalman, etc. Les travaux futurs pourraient s'orienter sur les primitives ne possédant pas de distance invariante et des groupes de transformations plus complexes.

Références

[Aya91] N. Ayache. *Artificial Vision for Mobile robots - Stereo-vision and Multisensor Perception.* MIT-Press, 1991.

[DW88] H.F. Durrant-Whyte. *Integration, Coordination and Control of Multi-Sensor Robot Systems.* Kluwer Academic Publishers, 1988.

[Fau93] O. Faugeras. *3-dimensionnal Computer Vision: a geometric viewpoint.* MIT press, 1993.

[Gri92] W.E.L. Grimson. *Object Recognition by Computer - The role of Geometric Constraints.* MIT Press, 1992.

[KM63] M.G. Kendall and P.A.P. Moran. *Geometrical probability.* Number 10 in Griffin's statistical monographs and courses. Charlkes Griffin & Co. Ltd., 1963.

[PA94] X. Pennec and N. Ayache. An $\mathcal{O}(n^2)$ algorithm for 3D substructure matching of proteins. In A. Califano, I. Rigoutsos, and H.J. Wolson, editors, *Shape and Pattern Matching in Computational Biology.* Plenum Publishing, 1994.

[PA96] X. Pennec and N. Ayache. Randomness and geometric features in computer vision. Research Report 2820, INRIA, March 1996.

[PT96] X. Pennec and J.P. Thirion. A framework for uncertainty and validation of 3D registration methods based on points and frames. *Int. Journal of Computer Vision*, to appear, 1996.

[San76] L.A. Santalo. *Integral Geometry and Geometric Probability*, volume 1 of *Encyclopedia of Mathematics and its Applications.* Addison-Wesley Publishing Company, 1976.

[ZF92] Z. Zhang and O. Faugeras. *3D Dynamic Scene Analysis: a stereo based approach*, volume 27 of *Springer series in information science*, chapter 2: Uncertainty Manipulation and Parameter Estimation, pages 9–27. Springer Verlag, 1992.