

Description Invariante des Formes Picturales par des Fonctions d'Autocomparaison*

Helmut Glünder

La partie principale du traité est composée de trois chapitres. Après l'introduction vient une présentation étendue du concept nouveau et systématique de la description invariante des formes picturales. Ceci repose sur l'analyse des formes par des fonctions dites d'autocomparaison, qui forment une généralisation de l'analyse d'autocorrélation, déjà bien connue dans le domaine de la théorie du signal. Les propriétés des caractéristiques, qui sont invariantes par toutes les transformations par similitude (générale), et certaines méthodes de calculs sont présentées et discutées par la suite. Il est également démontré, que la force de description de ces caractéristiques dépasse de beaucoup celle des caractéristiques qui résultent du seul procédé comparable concernant la propriété d'invariance. Dans le chapitre suivant, il est montré que les réseaux neuronaux conviennent parfaitement pour le calcul parallèle des caractéristiques proposées. L'analyse invariante des formes peut être faite dans le même réseau que celui qui permet l'extraction indépendante des formes des paramètres de mouvement, par un nouveau procédé d'intercorrélation. Ce principe d'analyse dualiste représente une condition essentielle à la faisabilité d'une autostructuration du circuit neuronal requis, avec son schéma d'interconnexions à haut degré de spécificité. Des systèmes de l'analyse du mouvement avec les propriétés exigées ont été découvertes, il y a peu, grâce à des expériences sur des singes et des chats.

Les données et les algorithmes se trouvent dans la deuxième partie. Grâce à un processeur opto-digital, spécialement développé pour cette expérience, les caractéristiques proposées ont été évaluées à partir de 30 signes graphiques. Deux exemples simples permettent d'expliquer les algorithmes pour le traitement en série de ces caractéristiques. Une annexe termine le traité. On y trouve quelques thèmes choisis de la géométrie analytique et une analyse, rapportée à la théorie des systèmes, du zoom-corrélateur optique.

Introduction

La signification des caractéristiques invariantes pour la description des images est expliquée dans le cadre de l'analyse et de la reconnaissance des formes. En particulier, il est montré que lorsque le nombre de classes est grand et que la variance géométrique des formes, concernant l'extension des variances, n'est pas limitée – ce qui est typique dans le cas de la reconnaissance biologique des

* Résumé d'une thèse doctorale acceptée par la faculté d'Electrotechnique et des Sciences de l'Information à la Technische Universität München (version originale en langue allemande)
Traduction par S. Desmoulin

formes – la classification basée sur les caractéristiques invariantes est la meilleure comparativement aux trois autres procédés de base, ceci en regard du rapport performance/dépense. Toutefois, il faut supposer que la force descriptive des caractéristiques est dans tous les cas suffisante pour la classification. En principe, l'optimum de la force descriptive pour des caractéristiques invariantes est donnée par les invariants des transformations géométriques considérées. Par conséquent, il est proposé pour la première fois d'utiliser directement les invariants des formes en tant que caractéristiques et de les calculer conformément à la définition. De là, on soumet la forme en question à des transformations géométriques et on la compare à la forme non transformée. On détermine alors les invariants de la forme, ou plus précisément la mesure de l'invariance par des transformations. Dans le cadre du traité, des considérations sont faites sur les caractéristiques, qui sont invariantes par toutes les transformations par similitude (translation: ξ, η ; similitude: s ; rotation: ψ ; réflexion: σ). On ne peut trouver dans la littérature qu'un seul type de caractéristiques invariantes par des transformations par similitude, qui peuvent être extraites sans aucune normalisation géométrique. Il apparaît alors comme un cas particulier des caractéristiques ici présentées.

Théorie

Seules les formes qui ont une quantité considérable de liens internes permettent l'extraction des propriétés invariantes qui sont utiles. La fonction d'autocomparaison (AVF) d'ordre m représente une approche universelle de l'analyse des liens internes des formes, qui ne représentent, de préférence, que des contours ou des squelettes d'une image. L'AVF est une généralisation de la fonction d'autocorrélation classique (AKF) concernant les transformations géométriques appliquées, le genre d'opération mathématique définie entre les fonctions picturales, et le nombre (l'ordre) des versions transformées des fonctions comparées. Par définition, l'AVF d'ordre 2 est appropriée à l'extraction des invariants. Toutes les considérations suivantes sont restreintes à l'AVF d'ordre 2, qui dépend des paramètres de transformation par similitude: $c_{\ddot{a}}(\xi, \eta, s, \psi, \sigma)$. A cause des points fixes des transformations, cette fonction est en général non invariante. C'est pourquoi trois caractéristiques appelées primaires sont tirées de l'AVF $c_{\ddot{a}}$, qui sont invariantes par toutes les transformations par similitude: les fonctions $C_s(s)$, $C_\psi(\psi)$, et $C_\sigma(\sigma)$, qui sont formées à partir des coefficients de comparaison bien choisis de l'AVF $c_{\ddot{a}}$. C'est à dire, à partir des meilleures valeurs concernant la position des points fixes. Les fonctions caractéristiques impliquées,

aussi bien que celles qui en découlent, sont alors discutées en considérant particulièrement l'évaluation de la position relative des points fixes optimaux. Plusieurs alternatives de la solution proposée, concernant le problème de la variance introduit par les points fixes, sont discutées et comparées. Il apparaît que l'approche choisie est la seule qui permette des mesures de diverses symétries, de congruences et de similitudes des formes.

L'AKF classique ne peut pas seulement être calculée en multipliant deux fonctions picturales identiques, mais translatées l'une par rapport à l'autre (explicitement transformées), mais aussi en calculant tous les produits entre les valeurs des pixels pris deux à deux (moments dipolaires) de la fonction picturale elle-même (approche implicite). Et c'est la même chose pour l'AVF $c_{\ddot{a}}$. En plus de la méthode explicite, deux versions d'un calcul implicite (moments dipolaires généralisés) sont présentées en détail. Chacune d'elle est formulée pour une représentation des formes spatialement discrètes, aussi bien que pour une représentation contenant les orientations des éléments linéaires des formes. Pour chacune des quatre versions d'évaluation des caractéristiques, des réseaux à calcul parallèle sont présentés. Les dépenses et les avantages d'une implémentation soit série, soit parallèle sont décrits en détail à la fin du chapitre.

Neurobiologie

L'analyse visuelle des formes basée sur les caractéristiques invariantes qui ont été proposées est probable pour les raisons suivantes: Premièrement à cause des situations communes (naturelles) des variances extrêmes des formes et deuxièmement à cause des catégories humaines de la perception des formes (symétries, congruences, similitudes, etc.), qui sont bien connues de la Gestalt-psychologie, c'est à dire, des demandes qui ne sont pas satisfaites dans une mesure comparable par d'autres méthodes connues. A partir des considérations de base de la théorie de l'information avec la connaissance sur les mécanismes du traitement du signal dans des neurones, on en déduit ceci:

- Les nonlinearités qui sont nécessaires à n'importe quelle réduction de l'information donnent, avant tout, un sens aux formes multilinéaires.
- Les formes multilinéaires construites à partir des valeurs du signal apparaissent comme réalisables d'une manière neuronale grâce à des interactions non linéaires (e.g. multiplicatives) entre des synapses dendritiques voisins et ensuite grâce à une sommation des résultats près du soma cellulaire.

- Pour des raisons de dépense, on peut supposer que la forme bilinéaire est l'opération non linéaire dominante dans de vastes systèmes neuronaux.

La forme polynomiale classique est communément proposée à la place. Elle résulte de la sommation (linéaire) des signaux suivie d'une nonlinéarité (i.e. un seuillage). Les formes bilinéaires et polynomiales sont comparées en utilisant un problème classique du traitement d'images: l'extraction des contours. L'opération de Laplace limitée en bande, couramment utilisée, suivie d'une opération de seuillage (forme polynomiale) est opposée à une nouvelle approche de type bilinéaire au moins aussi puissante.

Les sommes nécessaires pour l'évaluation discrète (spatiale) de chaque valeur de la AVF (d'ordre 2) sont des formes bilinéaires. Leur calcul neuronal à partir des formes picturales nécessite un grand nombre d'interconnexions extrêmement spécifiques, entre la représentation des formes et les neurones qui codent les coefficients de comparaison. Ainsi, l'approche pour la modélisation du traitement du signal neuronal ne vaut rien sans les idées sur la formation des réseaux nécessités qui sont plausibles d'un point de vue biologique. Il est montré que les systèmes pour l'extraction indépendante des formes des paramètres de mouvement – qui sont aussi proposés ici pour la première fois – sont structurellement identiques aux réseaux nécessités pour l'évaluation de l'AVF de similitude générale (exceptés les réflexions). Pourtant, dans le cas de l'analyse du mouvement, les interactions non linéaires entre les paires de pixels sont réalisées grâce aux détecteurs (simplifiés) de type Hassenstein/Reichardt (HRDs). Ces détecteurs élémentaires de mouvement introduisent un retard de ΔT pour un des deux signaux entrés avant qu'ils soient combinés non linéairement. Pour des durées de fixation visuelle $t > \Delta T$, les HRDs perdent leurs propriétés dynamiques et ainsi les réseaux construits à partir de tels détecteurs peuvent être utiles pour l'analyse de mouvement comme pour l'analyse invariante des formes proposée.

Le problème de la formation du système, apparemment insoluble du point de vue de l'analyse des formes, apparaît résoluble quand on considère l'analyse du mouvement: On peut supposé raisonnablement qu'un système pour l'analyse du mouvement est structuré à partir des mouvements propres de l'être. Le fonctionnement indépendant des formes du système résultant est une conséquence directe du rôle cruciale que jouent les propriétés communes de ses stimuli naturels dans les procédés d'autoorganisation: Seul un grand nombre de stimuli identiques permet une consolidation des interconnexions neuronales. La structure de calcul proposée est un premier exemple d'un modèle de calcul neuronal au niveau du signal pour laquelle on peut spécifier des conditions directement liées à la faisabilité de sa formation. La présentation et les investigations des mécanismes et

des procédés de l'autoorganisation induite par les mouvements ne font pas partie de ce développement.

Les systèmes neuronaux pour l'analyse des mouvements de profondeur et des rotations fronto-parallèles du type postulé ont été récemment identifiés de manière neurophysiologique chez les singes et les chats. Les propriétés des champs réceptifs (RFs) postulées pour des neurones modèles dans les réseaux s'accordent bien avec celles réellement mesurées. De plus, les attentes des propriétés des RFs pour les stimulations statiques ($t > \Delta T$) sont formulées.

Sections de Données

Les expériences de simulation occupent la plus grande partie de ce chapitre. (Environ 2000 fonctions d'intercorrélation de 30 signes graphiques ont été calculées. Chacune d'elles a environ 4000 degrés de liberté (OBP)). Ces résultats sont donc utiles pour l'illustration des propriétés des caractéristiques primaires. Le zoom-corrélateur optique, développé spécialement pour leur évaluation est optimisé suivant le critère du rapport OBP/dépense. Les analyses de ce calculateur faites à partir de l'optique géométrique et de la théorie des systèmes sont présentées.

Les algorithmes pour le calcul série des caractéristiques à partir de moments dipolaires généralisés – avec ou sans paramètres d'orientation – sont introduits et illustrés par l'analyse de deux signes graphiques.