Invariant Description of Pictorial Patterns by Auto-Comparison Functions*

Helmut Glünder

The main part of this work consists of three chapters. Following the introduction an extensive exposition is given of a new and systematic concept for the invariant description of pictorial patterns. The approach is based on the analysis of patterns by so-called auto-comparison functions that are a generalization of the auto-correlation analysis, well-known in signal theory. Features are presented that are invariant under all similarity transformations. They are discussed in depth, together with methods for their extraction. It is demonstrated that the descriptive power of these features surpasses that of the only competing kind of features, comparable with respect to their invariance properties. In the following chapter it is shown that neural networks are well suited for the parallel computation of the proposed features. Invariant pattern analysis can be performed within the same network that serves the pattern-independent extraction of motion parameters based on an also new cross-correlation method. This dualistic principle of analysis is an essential condition for the feasibility of self-organization of the required neural circuitry with its highly specific interconnection scheme. Systems for the analysis of motion with the claimed properties have recently been identified in monkey and cat.
Data and algorithms are contained in a second part. The proposed features were computed from 30 characters, using a specially designed opto-digital processor. Algorithms for the serial computation of the features are presented and illustrated by two basic examples. An appendix concludes the work. It deals with several selected topics from the field of analytic geometry and gives a systems-theoretical analysis of the optical zoom-correlator.

Introduction

The relevance of invariant features for the description of pictorial patterns in the field of pattern analysis and pattern recognition is discussed. Emphasized is that in case of great numbers of classes and unrestricted geometrical variances (concerning their extent) – i.e. under conditions typical for biological pattern recognition – classification based on invariant features is best with respect to the performance/cost-ratio, when compared to the other three basic methods of invariant pattern recognition, provided the features' descriptive power is sufficient for classification. In principle, the optimum descriptive power of invariant features

* Summary of a doctoral thesis accepted by the faculty of Electrical Engineering and Information Sciences at the Technische Universität München (original version in German)

is given by the invariants of the considered transformations. Therefore, it is proposed (for the first time) to use as features the invariants of patterns and to compute them according to their definition. For this purpose, a pattern under investigation is geometrically transformed and compared with the untransformed pattern, thereby revealing its invariants or, more precisely, the extent of its transformation invariance. Within the scope of this work, features are investigated that are invariant under all similarity transformations (similitudes) (rigid translation: ξ, η; expansion: s; rotation: ψ; reflection: σ. From literature one other kind of similarity-invariant feature is known that can be extracted without geometric normalization. It turns out to be a special case of the presented feaure family.

Theory

Patterns must have sufficient inner bindings in order to permit the extraction of useful invariant properties. The auto-comparison function (AVF) of order m represents a universal approach to the analysis of inner bindings of preferably line-like patterns. The AVF is a generalization of the classic auto-correlation function (AKF) with respect to the applied geometrical transformations as well as the combining operations and the number (order) of pictorial functions. Per definition, the second order AVF is appropriate for the extraction of invariants. All further investigations are restricted to the second order AVF that depends on the parameters of the similarity transformations: $c_{ä}(\xi, \eta, s, \psi, \sigma)$. Because of the fixed points of the transformations it is generally not invariant. Therefore, three so-called primary features are derived from the AVF $c_{ä}$ that are invariant under the similitudes: the functions $C_s(s)$, $C_\psi(\psi)$, and $C_\sigma(\sigma)$ consisting of comparison coefficients of the AVF $c_{ä}$ that are optimum with respect to the position of the fixed points. The properties of these feature functions are presented. Related and derived feature functions are discussed, and emphasis is put on the evaluation of the relative position of the optimum fixed points. Different solutions of the variance problem, introduced by the fixed points, are investigated and compared. It appears, however, that the chosen approach is the only one that leads to measures of pattern symmetries, congruences, and similarities.

The classic AKF may either be computed by multiplying two identical, but translated (explicitly transformed) images, or by collecting the products of the values at all pixel pairs (dipole-moments) of the image (implicit approach). The same holds for the AVF $c_{ä}$. Besides the explicit method, two versions of implicit evaluation (generalized dipole-moments) are presented in detail. Each is formulated for the common spatially discrete representation of patterns but also for a representation of the patterns' oriented line-elements. For all four versions parallel-processing networks are presented that perform the feature extraction. Finally, costs and advantages of either serial or parallel implementation are quantified.

Neurobiology

For the following reasons, visual pattern analysis based on the proposed invariant features is quite likely: firstly, the common (natural) situations of extreme pattern variances, and secondly, the categories of human form perception (symmetries, congruences, similarities, etc.) that are known from Gestalt-psychology. These demands have not been met to such an extent by an other known method, yet. Basic information-theoretical considerations and the signal processing mechanisms of neurons lead to the following insights:

- Nonlinearities that are necessary for any information reduction preferably make sense as multilinear forms.
- Multilinear operations on signal values appear neurally feasible by nonlinear interactions (e.g. multiplications) of neighbouring dendritic synapses whose results (products) are summed near the cell soma.
- For cost reasons, the bilinear form is suspected dominant in large neural systems.

Commonly, the classic polynom form is proposed instead. It results from the (linear) summation of signals that is followed by a nonlinear operation (e.g. threshold). Bilinear and polynom form are compared and exemplified by use of a fundamental image processing task: contour extraction. The widely accepted bandlimited Laplacian operator, followed by a threshold operation (polynom form), is thereby opposed with a new and at least as powerful approach of the bilinear kind.

The sums created during the (spatially) discrete evaluation of single AVF values (order 2) are bilinear forms. Their neural extraction from pictorial patterns requires very many and highly specific interconnections between the pattern representation and the neurons that compute the comparison coefficients. Thus, this modelling of neural signal processing is in vain, without any biologically plausible hints to the formation of the networks. It is shown that systems for pattern-independent extraction of motion parameters – also proposed here for the first time – are structurally identical to networks that compute the desired similarity-

AVF (except reflections). The pairwise nonlinear interaction in case of motion analysis is performed by so-called (simplified) Hassenstein/Reichardt-detectors (HRDs). These elementary motion detector-units introduce a delay of ΔT for one of their two input signals before nonlinearly combining them. However, for visual fixation times $t > \Delta T$ of, HRDs lose their dynamic properties. Hence, networks constructed from such detectors can serve motion analysis and the proposed kind of invariant form analysis.

The problem of system structuring, seemingly not solvable from the point of view of form analysis, appears solved when seen from the analysis of motion: a system for motion analysis is quite likely structured by ego-motions of a creature. The pattern-independent functioning of the resulting system is a direct consequence of the crucial rôle common properties of real world motion/movement-stimuli play for self-organization processes: only great numbers of identical stimuli lead to the consolidation of neural interconnections. The proposed model structure is one of a few examples for neural processing at the signal level for which plausible conditions for self-organization can be given. The presentation and investigation of mechanisms and processes of motion-induced self-structuring is beyond the scope of this work.

Neural systems for such analyses of motion in depth and fronto-parallel rotations recently have neurophysiologically been identified in monkey and cat. The expected properties of receptive fields (RFs) of model neurons in such networks well agree with the measured. Furthermore, expected properties of RFs for static stimulations ($t > \Delta T$) are presented.

Data Sections

Extensive simulations and model experiments play a major rôle in this chapter. (About 2000 cross-correlation functions of 30 patterns were computed, each of a space-bandwidth product (OBP) of around 4000.) They serve the illustration of the properties of the primary features. The specially developed optic zoom-correlator for their evaluation is optimized with respect to the OBP/cost-ratio. Geometric-optic as well as systems-theoretical analyses of this parallel analog computer are presented.

Algorithms for the serial computation of the feature by generalized dipole-moments – with and without orientation parameters – are introduced and exemplified by the analysis of two patterns.

