

DESCRIPTION OF PLANAR PATTERNS BY INVARIANT FEATURES - AN ATTEMPT TOWARDS THE EXPLANATION OF VISUAL PATTERN RECOGNITION

H. GLÜNDER, T. KRAMER

Lehrstuhl für Nachrichtentechnik, Technische Universität München, Arcisstr. 21, D-8000 München 2, FRG

ABSTRACT

A pattern description is proposed which is based on geometric relations between pattern elements. Well-known preprocessing techniques are used for a decomposition of the pattern into line elements of different orientations and widths. Suitably chosen relational parameters between these pattern-excerpts lead to a pattern representation from which invariant features are extracted. The way they are evaluated agrees with processes in neural circuits of the visual system. Their nature corresponds to that determined and postulated by perceptual psychologists of the Gestalt school. They indicate global shape properties such as symmetries, congruences, similarities, etc. In the presented study emphasis is put on the extraction of relational parameters and invariant features.

INTRODUCTION

In previous articles we showed that global comparisons between a planar pattern and geometrically transformed versions of it, which can be mathematically formulated by so called generalized auto comparison operations, result in highly invariant features^{1,2}. In other words, we proposed a concept considering as features answers to the following question: "Which transformations convert parts of a pattern into others"? This transformational method, which describes patterns by their invariant elements together with the corresponding transformations, results in pattern characterizations that agree well with those proposed by Gestalt psychologists³. A parallel computing structure, however, explicitly performing the required geometric transformations is not quite plausible within neural circuits. Therefore, following a suggestion of Palmer^{4,5} we no longer carried out the transformations explicitly but evaluated geometric relations between elements of a pattern. To each of these relations, or *generalized dipole moments*, defined by a pair of elements we can assign a (implicit) transformation⁶.

Obviously all global comparisons between a pattern and all of its explicitly transformed and deformed versions that can be calculated within a given pattern format can also be achieved by comparisons between appropriately chosen pixels of the original pattern, i.e., without actually performing geometric transformations. The interpretation, however, of the resulting pattern representation especially for complicated deformations is difficult and the question may arise whether the great variety of possible transformations must always be considered. Depending on the desired accuracy of a pattern description and the claims made for its invariance simplifications and data reduction can be applied. In practice, the group of affine transformations is a good approximation for the analysis of parallel projections of three-dimensional rigid objects⁷. For the treatment of rigid and planar patterns in planes parallel to the sensor array (retina) it is even sufficient to consider the subgroup of similarity transformations. On the other hand, comparisons between pattern elements (primitives) may replace those between pixels and thus a pattern must be decomposed.

This paper's goal is not to present examples of features gained from complex input patterns (this would require an enormous computation capacity on a serially working v. Neumann-type machine) but to explain useful processing steps for the extraction of "relational features" via the dipole moment approach and to exemplify its equivalence with the recently introduced explicit transformational approach. The quite moderate computer simulations at the end of this article serve for this purpose. They can, however, by no means express the potential processing power of a suitable (neural) parallel processor.

PATTERN DECOMPOSITION

There is no great choice of suitable pattern segmentations if we consider the following claims and boundary conditions:

- (i) The desired features shall be invariant at least under the similarity transformations of a pattern, i.e., rigid translations, rotations, changes in scale, and reflections (sense).
- (ii) The primitives should support the computation of the desired implicit transformations.
- (iii) Their variety (shape) should be as small as possible to minimize the computational effort.
- (iv) The elements must be simple in order to facilitate their extraction and characterization.
- (v) For the purpose of concept formation about visual pattern recognition the decomposition process must be in accordance with fundamental neurobiological findings.

The relevance of point (i) for the segmentation process consists in a claim for the extraction mechanisms. They must guarantee an *invariant evaluation* of the primitives, i.e., the whole pattern must be *processed* no matter how it was presented to the input plane. Since invariant features are computed from these primitives they themselves need not be invariant.

Primitives that allow a high flexibility for the evaluation of features and that meet the above conditions are straight line elements of different orientations and widths. Since they represent merely the *data base* for the proposed feature generation we will not discuss methods for their extraction in detail. A simple and well-known method was used for the demonstrations in this article: A pattern's two-dimensional correlation with laterally bandpass-filtered straight lines or bars of suitable length. In visual sciences such functions are known as orientation selective receptive fields. A set of such correlation kernels (processing channels) contains them in different orientations and sizes. According to the above stated claims we need a set containing functions of all possible orientations and, within limits, widths. In practical systems, due to their limited overall bandwidth in conjunction with suitably applied maximum selection mechanisms, a moderate number of correlation kernels will suffice. Since we confine the present study to shape analysis we do not treat the problems arising in connection with the preservation of intensity information within such a preprocessor. A more elaborate preprocessing that leads

to a biologically more plausible data base, and represents a further development of the approaches of Marr and Hildreth⁸ as well as Watt and Morgan⁹, is introduced in a forthcoming paper⁶.

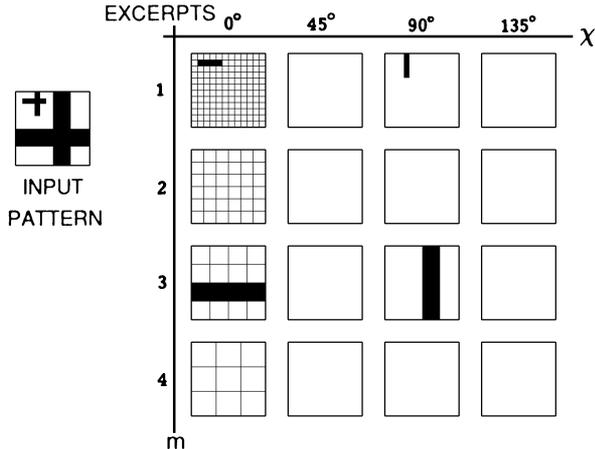


Fig. 1 Pattern excerpts of different orientations and sizes represent the data base from which the relational pattern representation is constructed

For the explanation of the processing principle we start from a data base which is structured like that sketched in Fig. 1 and from which a relational pattern representation can be constructed. We assume an unequivocal assignment of line elements to channels of different widths and equal orientations but not to channels of different orientation. The excerpts' sampling grids are chosen according to the size of the applied correlation kernels (sampling theorem).

RELATIONAL PATTERN REPRESENTATION

We introduce a pattern representation that gives a shift invariant description of a pattern. It consists of geometric relations between all *n active pixels* (indicating line elements) of the data base.

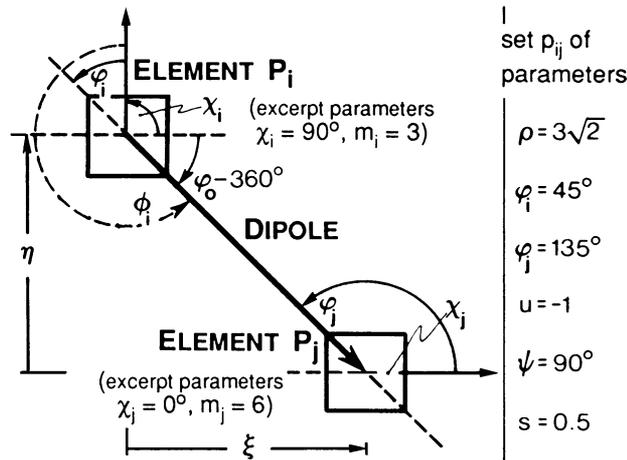


Fig. 2 Definition of dipole parameters characterizing pairs of elements in the data base

According to Fig. 2 every pair of pixels (P_i, P_j) is characterized by the following five dipole parameters:

- (i) the dipole length ρ (in multiple units of the sampling grid P_i is taken from),

$$\rho = \sqrt{\xi^2 + \eta^2}; \quad \text{with } \xi = x_j - x_i \text{ and } \eta = y_j - y_i$$

- (ii) the angle relations ϕ_i and ϕ_j between the absolute direction ϕ_0 of the dipole and the excerpt directions χ_i and χ_j ,

$$\phi_i = \begin{cases} \text{mod}(\phi_i, 180^\circ) & \text{with } \phi_i = \phi_0 - \chi_i & \text{for } \rho \neq 0 \\ \text{not defined (X)} & & \text{for } \rho = 0 \end{cases}$$

$$\phi_j = \begin{cases} \text{mod}(\phi_j, 180^\circ) & \text{with } \phi_j = \phi_0 - \chi_j & \text{for } \rho \neq 0 \\ \text{mod}(\phi_j, 180^\circ) & \text{with } \phi_j = \chi_i - \chi_j & \text{for } \rho = 0 \end{cases}$$

with $\phi_0 = \arctan(\eta/\xi)$

- (iii) the characteristic number u ,

$$u = \begin{cases} +1 & \text{for } 0 \leq \phi_i < 180^\circ \\ X & \text{if } \phi_i \text{ is not defined } (\rho = 0) \\ -1 & \text{for } 180^\circ \leq \phi_i < 360^\circ \end{cases}$$

(The characteristic number, though useful for digital computations, is not obligatory and can be avoided by slightly altered definitions of the angles ϕ_i and ϕ_j .)

- (iv) the angle difference ψ ,

$$\psi = \chi_i - \chi_j$$

- (v) the scale factor s by which the excerpt resolutions (m_i, m_j) are related,

$$s = m_i/m_j.$$

We arrange sets of parameters p_{ij} in an $n \times n$ -array: Pixels stemming from the same excerpt of the data base are arranged in groups $E_{k,l}$, thereby defining submatrices $M_{k,l}$. In this array each row describes the whole pattern seen from a certain pixel P_i . Figure 3 shows this representation for the pattern of Fig. 1, however, only for excerpts with $m=1$ and therefore $s=1$. (The designation of pixels by letters is arbitrary.)

	E_1	a	b	c	d	E_2	e	f	g	h						
$E_1(x=0^\circ)$	0	X	1	0°	2	0°	3	0°	2.2	27°	2	0°	2.2	153°	2.8	135°
a	X	0°	+ 0°	+ 0°	+ 0°	+ 0°	+ 117°	+ 90°	- 63°	- 45°						
b	1	0°	0	X	1	0°	2	0°	1.4	45°	1	0°	1.4	135°	2.2	117°
c	-	0°	X	0°	+ 0°	+ 0°	+ 0°	+ 135°	+ 90°	- 45°	- 27°					
d	2	0°	1	0°	0	X	1	0°	1	90°	0	X	1	90°	2	90°
e	-	0°	-	0°	X	0°	+ 0°	+ 0°	+ 0°	X	90°	-	0°	-	0°	-
f	3	0°	2	0°	1	0°	0	X	1.4	135°	1	0°	1.4	45°	2.2	63°
g	-	0°	-	0°	-	0°	X	0°	+ 45°	- 90°	- 135°	-	0°	-	153°	-
$E_2(x=90^\circ)$	2.2	117°	1.4	135°	1	0°	1.4	45°	0	X	1	0°	2	0°	3	0°
e	+ 27°	+ 45°	- 90°	- 135°	X	0°	-	0°	-	0°	-	0°	-	0°	-	0°
f	2	90°	1	90°	0	X	1	90°	1	0°	0	X	1	0°	2	0°
g	+ 0°	+ 0°	X	90°	-	0°	+ 0°	+ 0°	X	0°	-	0°	-	0°	-	0°
h	2.2	63°	1.4	45°	1	0°	1.4	135°	2	0°	1	0°	0	X	1	0°
a	+ 153°	+ 135°	+ 90°	- 45°	+ 0°	+ 0°	+ 0°	X	0°	-	0°	-	0°	-	0°	-
b	2.8	45°	2.2	27°	2	0°	2.2	153°	3	0°	2	0°	1	0°	0	X
c	+ 135°	+ 117°	+ 90°	- 63°	+ 0°	+ 0°	+ 0°	+ 0°	X	0°	-	0°	-	0°	-	0°

Fig. 3 Relational representation of a pattern as array of sets of dipole moments

Since the code is based on spatial relations and the construction of the array *must obey a once defined rule* the representation is shift invariant, though it is by no means rotation or scale invariant. For rotations and scalings the

submatrices are permuted. The extraction of features, invariant under these transformations, is explained in the next section.

INVARIANT FEATURES

A relational analysis of the pattern representation leads to invariant geometric features. It is based on the search for coincidences of dipole parameters within the array. We start with the extraction of simple features that show some invariances due to their numerical character, such as the number of line joints, intersections and corners^{10,11}. They result from elementary spatial or semantic coincidences and their quality appears to be similar to that of Julesz' "textons", i.e., to conspicuous pattern properties in preattentive vision¹².

Numerical and line features

For the extraction of such simple features it is useful to group line elements within each excerpt in order to form straight lines of constant line width. Only submatrices M_{kl} with $k=1$ must be considered, i.e., $\psi=0^\circ$ and $s=1$. Line elements with $\phi_i=\phi_1=0^\circ$ are "in line" and can be assumed to be parts of a straight line. Whether this line is continuous or broken depends on the configuration of the parameter ρ which is also essential for the definition of end of line elements. The length of a line is given by the number of aligned elements. For $\phi_i=\phi_1=\phi$ we deal with parallel line elements. Their orthogonal distance, or that of their productions, is defined by $d = \rho \cdot \sin\phi$. Parallel line elements that are not staggered are identified by $\rho=d$ or $\phi=90^\circ$. These investigations can also be extended to line elements of different widths. Then, however, the distance measures must be corrected by the corresponding scale factors s .

Dipoles with $\rho=0$, $\phi_i=X$, $\phi_j \neq 0^\circ$ and $u=X$ indicate characteristic pattern elements. If both line elements, P_i and P_j , are end of line elements a corner is present. A line joint is detected if one of both line elements, P_i or P_j , is an end of line element, and a line intersection (crossing) if neither of them is of this type. The acute angle α of either of these configurations is given by $\alpha = \text{mod}(\phi_i, 90^\circ)$. For the detection of these characteristic pixels it is sufficient to investigate only one half of the matrix (above or below the diagonal).

Rotational symmetry

A feature is introduced that describes a pattern by its degree of self-congruence as a function of the rotation angle. Since rotation is bound to a centre (fixed point) we propose to search for the optimum centre for each angle, i.e., for the maximum degree of congruence. Because the feature shall be rotation invariant we must get rid of the *absolute* orientation angles χ . This is achieved by using the angle differences ψ for the abscissa of the desired feature function c_ψ . For the evaluation we consider groups of rows $E_k(\chi, m=\text{const})$ and dipoles with $m=m_i=m_j$ (columns) thus defining a resolution level. Then, the values of $c_\psi^m(\psi)$ are given by the maximum number of identical dipoles (parameters ρ, ϕ_i, ϕ_j, u and s) that can be found between all rows of angle groups $E_k(\chi_i)$ and all rows of angle groups $E_k(\chi_j) = E_k(\chi_i + \psi)$ within the matrix representation. (The positions of the code words in the rows are not relevant.) This method allows the determination of $c_\psi^m(\psi)$ for $0^\circ \leq \psi < 180^\circ$. To obtain the value for $\psi=180^\circ$ we compare dipoles that are identical except for the parameter u (which must have the opposite sign) in rows of groups $E_k(\chi)$ with $\chi=\chi_i=\chi_j$. In the example of Fig.3 the maximum number of code words (7) for $\psi=90^\circ$ is found, for instance, by comparing rows "d" and "e". This number can be related to the total number of pixels (8) if the relative angular self-congruence (7/8) is desired.

Speaking in terms of the explicit transformational approach² this means comparisons between the original pattern and its rotated versions that are shifted in order to reach the best match for each angle ψ . In the above mentioned example a

configuration is investigated where the pattern is rotated by 90° with respect to the original and shifted so that pixel "d" falls onto pixel "e".

Obviously the feature function c_ψ is invariant under similarity transformations of the input pattern². A consequence of its independence from a rotation centre is the loss of inner positional relations. Examples demonstrating the resulting ambiguities are given in the experiment section of this paper. This loss, however, can easily be compensated for by other features, e.g., those described in the previous paragraph.

Similarity

A pattern's self-congruence is described as a function of scale factor. Again we suffer from the shift variance of the transformation (centre of zoom) and therefore we search again for maxima. To achieve scale invariance we do not use the *absolute* size m but the scale factor s for the abscissa of the feature function c_s . Their values are given by the maximum number of identical dipoles (parameters ρ, ϕ_i, ϕ_j , and u) found between all rows of size groups $E_k(m_i)$ and all rows of size groups $E_k(m_j) = E_k(s \cdot m_i)$ within the array representation. There are no restrictions concerning the dipoles (columns) to be investigated. It is, however, convenient to perform the comparisons separately for each orientation angle χ (rows).

Again, considering the explicit transformational approach, we realize that the original pattern is compared with either reduced or enlarged versions that are shifted in order to reach the best match for each scale factor s . The feature function c_s has the same invariance properties as c_ψ .

Other features

A combined, two-dimensional feature function $c_2(\psi, s)$ can be calculated which contains much more information than the two one-dimensional features. The axial reflection feature results from comparison of a pattern with its mirrored version. Again one has to eliminate the fixed points (symmetry axes). This feature $c_\sigma(\sigma=\pm 1)$ is not extracted by detection of equal dipole properties but of dipoles and their mirrored counterparts, expressed by the appropriate parameters. Thinking again of similarity transformations we realize that a feature function expressing shift congruences is still missing. Such a feature c_0 is known since long: the auto correlation function. It can be gained from the array representation by counting dipoles that have equal properties with respect to the parameters ρ and ϕ_0 , and thus coincidence detections are obsolete. This feature, however, is merely shift invariant.

EXPERIMENTS

The angle feature c_ψ was digitally computed for some simple binary patterns of constant line width. The patterns were given in a 64×64 field from which four orientation excerpts ($\chi=0^\circ, 45^\circ, 90^\circ, 135^\circ$) were computed using an orientation selective correlation kernel of 5×5 pixels and optimum width. A suitable threshold operation applied to the correlation functions led to what we called the data base of the implicit transformational concept.

Figure 4 shows the orientation excerpts (upper rows) and the feature functions (dark fields) for a circle, a cross, as well as L- and T-shaped patterns (lower left corners). The latter patterns cannot be distinguished by this feature function due to the mentioned loss of positional information. These experimental results corroborate the equivalence between the new concept and the explicit transformational approach, which is mathematically evident. The corresponding results obtained from the latter method are shown as graphs in the white fields. The maxima are converted into minima due to coincidence detection by taking the absolute values of differences; furthermore a three times finer angle increment was applied.

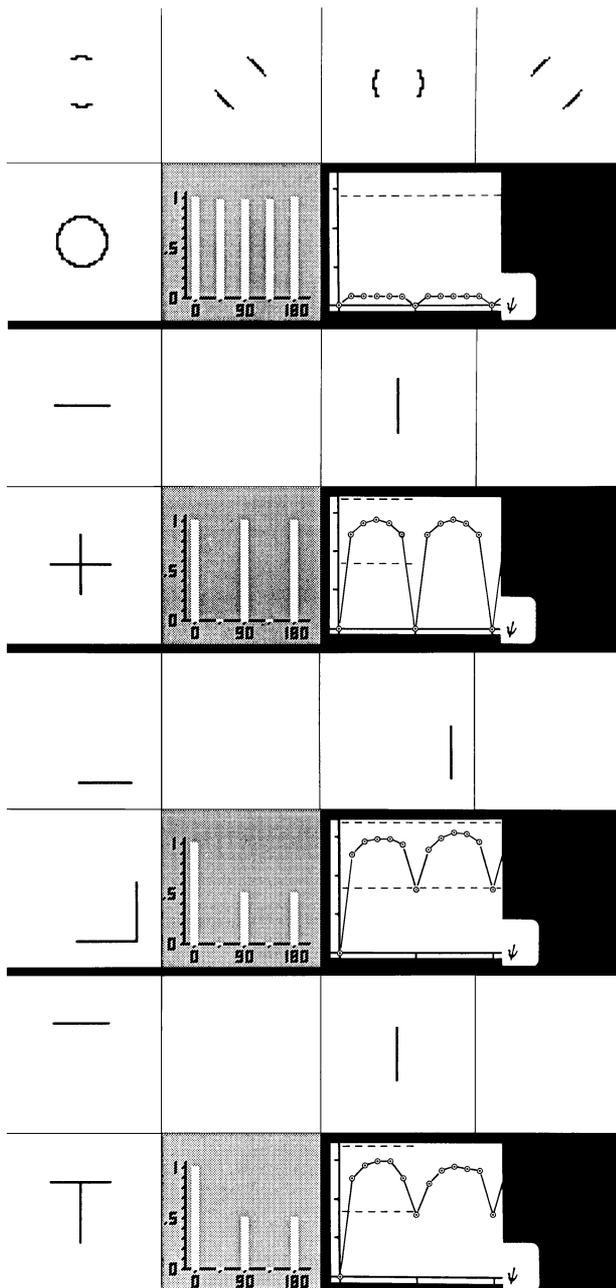


Fig.4 Four patterns (lower left corners), their orientation excerpts (upper rows) and angle feature functions obtained via implicit (dark field) and explicit transformations (white field). For the latter method a three times finer angle increment was used and the maxima were converted into minima.

DISCUSSION

A parallel working processor designed for the extraction of relational features mainly consist of coincidence detectors and quite a lot of wiring between them⁶. Both is in good agreement with the anatomy of neural structures and neural processing principles. A great advantage over the explicit transformational approach is the data reducing preprocessing which, in general, agrees with neurophysiological and psychophysical findings. Contrary to the majority of the present theories about the

purpose of the orientation and size selectivity of the visual system we neither follow the Fourier analysis, nor the Perceptron approach¹³.

The features c_w , c_s , and c_o express global pattern properties corresponding to perceptual categories found and used in Gestalt psychology³ and are well suited for pattern description, association and classification purposes¹³. Local features such as endpoints, corners, line joints and intersections are important as well, though only their numbers, or relative positions are highly invariant features. An indispensable property of systems producing invariant features is also met: The preservation of variance information. It is needed to restrict the invariance, for instance in contextual situations². This information can be taken from the data base (full variance) or from the relational representation (all except shift variance). The latter nicely indicates changes in size and angular presentation of a pattern by defined permutations of groups E in the array.

This sketch of implicit transformational processing in which we put emphasis on basic computational aspects can only give a slight impression of the potential processing power. Surely, the method needs extensions and refinements but in contrast to other biologically relevant concepts, however, a *mathematical* formulation of the processing principle could be presented.

ACKNOWLEDGEMENT

We thank Prof. H. Marko and H. Platzer for their interest and support.

REFERENCES

- [1] Glünder H., Gerhard A., Platzer H., and Hofer-Alfeis J., "A geometrical - transformation - invariant pattern recognition concept incorporating elementary properties of neural circuits," in: *Proc. of the 7th Int. Conf. on Pattern Recognition* (edited by Wein M.), pp.1376-1379. Silver Spring/MD: IEEE Comp. Soc. Press, 1984.
- [2] Glünder H., "Relational processing: A concept for the explanation of invariant visual pattern recognition," in: *Processing structures for perception and action* (edited by Marko H., Hauske G., Struppler A.), in the press. Weinheim: Verlag Chemie, 1986.
- [3] Helson H., "The fundamental propositions of Gestalt psychology," *Psych. Rev.* **40**, pp.13-32, 1933.
- [4] Palmer S.E., "Symmetry, transformation, and the structure of perceptual systems," in: *Organization and representation in perception* (edited by Beck J.), pp.95-144. Hillsdale/NJ: Erlbaum, 1982.
- [5] Palmer S.E., "The psychology of perceptual organization: a transformational approach," in: *Human and machine vision* (edited by Beck J., Hope B., Rosenfeld A.), pp.269-339. New York: Academic Press, 1983.
- [6] Glünder H., "Neural computation of inner geometric pattern relations," *Biol. Cybern.* in the press, 1986.
- [7] Ullman S., *The interpretation of visual motion*. Cambridge/MA: MIT Press, 1979.
- [8] Marr D. and Hildreth E., "Theory of edge detection," *Proc. R. Soc. Lond.* **B207**, pp.187-217, 1980.
- [9] Watt R.J. and Morgan M.J., "A theory of the primitive spatial code in human vision," *Vision Res.* **25**, pp.1661-1674, 1985.
- [10] Marko H. and Giebel H., "Recognition of handwritten characters with a system of homogeneous layers," *Nachrichtentechnische Zeitschrift* **23**, pp.455-459, 1970.
- [11] Marko H., "A biological approach to pattern recognition," *IEEE Trans. SMC-4*, pp.34-39, 1974.
- [12] Julesz B. and Bergen J.R., "Textons, the fundamental elements in preattentive vision and perception of textures," *Bell Syst. Techn. J.* **62**, pp.1619-1645, 1984.
- [13] Glünder H., "On functional concepts for the explanation of visual pattern recognition," *Human Neurobiol.* **5**, pp.37-47, 1986.



afcet

EIGHTH
INTERNATIONAL CONFERENCE
ON

PATTERN RECOGNITION

PARIS, FRANCE
OCTOBER 27-31, 1986

PROCEEDINGS

Sponsor : IAPR, International Association for Pattern Recognition
Organizer : AFCET, Association Française pour la Cybernetique Economique
et Technique

Volume 2

IEEE Computer Society Order Number 742
Library of Congress Number 86-81419
IEEE Catalog Number 86CH2342-4
ISBN 0-8186-0742-4

IEEE
COMPUTER
SOCIETY
PRESS 