pp. 129-134 in Neural Networks for Computing, American Institute of Physics Publication, Conference Proceedings 151 (New York,1986)

Stochastic Spin Models for Pattern Recognition

R. Divko and K. Schulten
Physik-Department
Technische Universität München, D-8046 Garching

Abstract

We exploit for the recognition of patterns the properties of physical spin systems to assume long range order and, thereby, to establish a global interpretation of patterns. For this purposes we choose spins which can take a discrete set of values to code for local features of the patterns to be processed (feature spins). The energy of the system entails a field contribution and interactions between the feature spins. The field incorporates the information on the input pattern. The spin-spin interaction represents 'a priori' knowledge on relationships between features, e.g. continuity properties. The energy function is chosen such that the ground state of the feature spin system corresponds to the best global interpretation of the pattern. The ground state is reached in the course of local stochastic dynamics, this process beeing simulated by the method of Monte Carlo annealing[1]. Our study is related to work presented in Refs. 2, 3.

Introduction

Spin systems are characterized by a set of values for the spin variable $S_{i,j}$, a lattice on which they are defined, and by an interaction energy. In the two-dimensional Ising model the spins take the values ± 1 and the interaction energy is defined by the Hamiltonian

$$E = -J \sum_{<(i,j),(k,l)>} S_{i,j}\, S_{k,l} - \sum_{(i,j)} H_{i,j}\, S_{i,j} \qquad (1)$$

where the brackets indicate summation over next neighbors. In the ferromagnetic case ($J > 0$) the first term, the exchange interaction, gives a negative contribution if neighboring spins point into the same direction. This term creates a tendency for an alignment of all the spins. The second term describes the interaction of the spins with a local magnetic field $H_{i,j}$ tending to align the spins locally with the field. The regularizing effect of the exchange interaction will be utilized in the following to solve pattern recognition tasks under the constraint that pattern features are expected to vary continuously.

Feature Spins

For the purpose of picture processing the spins are chosen to code local features of a pattern. Examples for attributes coded by such feature spins are intensities, disparities between corresponding points in a stereogram or edges of different directions. Several different types of feature spins, interacting with each other and with external fields, may be needed to solve a specific pattern recognition problem.

At finite temperatures the feature spin system shows fluctuations like its physical counterpart. Certain values of a feature spin at a certain lattice point are more probable than others. One may consider the value of a feature spin as the hypothesis that the picture has a certain local attribute at this point.

At high temperatures all hypotheses are equally probable. After carefully cooling down to low temperatures (simulated annealing[1]) the fluctuations eventually disappear. At zero temperature the feature spins take definite values indicating the final global hypothesis about a pattern.

The final hypothesis, i.e. the ground state, achieved by the system after cooling down to low temperatures depends on the interaction among the feature spins as well as on the interaction with the external field. The interaction among the feature spins contains an 'a priori' global knowledge on relationships to be expected to hold between the features of a pattern. Correct interpretations of patterns must meet certain constraints, e.g. the constraints of continuity, which have to be realised by the feature spin configurations in the final hypothesis. Such configurations can be achieved by a properly choosen interaction between the feature spins. For example a Potts model type interaction between intensity spins yields a smooth change of brightness. The external field serves to communicate the pattern to be processed to the system of feature spins. Examples for pattern attributes coded by the external fields are local brightness or edges of various directions.

Stereo Vision

Whereas a certain degree of depth vision can be obtained from perspective distortion or from hidden parts of a scene, full stereo vision is a result of binocular perception. The projections of an object in both eyes differ slightly from each other. This difference (disparity) allows the reconstruction of the three-dimensional information. Figure 1 shows an image pair of dot patterns appearing completely random when viewed monocularly. But when viewed one through each eye the two pictures fuse showing a three-dimensional structure (square hovering over the ground). The absence of higher level structures in the patterns shows that disparity alone can be used to obtain three-dimensional information from a stereogram.

Fig. 1. Julesz pattern[5] with 50% black dots. This random-dot stereogram of 50×50 pixels is generated by copying the rigth image from the left one, shifting a square-shaped region of 30×30 pixels sligthly to the left and filling the gap caused by the shift with a new random pattern.

To obtain stereo information the disparity of corresponding points in the two retina projections of an image must be determined. The problem is to assign correspondences between points of the two pictures. This is a difficult task because of the socalled 'false target problem'[4] occuring in its extreme in Julesz patterns[5]. Every black pixel could correspond to every other black pixel. To restrict all possible combinations of points from both pictures to physically plausible correspondences the following matching conditions must hold:

- Compatibility: Black dots can only match black dots and vice versa.
- Continuity: The physical feature disparity varies smoothly almost everywhere over the image.
- Uniqueness: Except in rare cases each point from one image can match only one point from the other image.

The following spin model is designed to find the correspondences between the pixels of both pictures of a random-dot stereogram and measure the disparities. This information will be contained in the ground state of the spin system.

The feature coded by a spin is the disparity of corresponding pixels. A disparity spin with a value $S_{i,j} \in \{0, \pm 1, \ldots, \pm N\}$ at lattice site (i,j) stands for the hypothesis of a correspondence between the pixel (i,j) in the right picture of the stereogram and the pixel $(i, j + S_{i,j})$ in the left picture shifted $S_{i,j}$ units to the rigth. Both pixels are assumed to correspond to the same original point of an object and to have the disparity $S_{i,j}$.

The Hamiltonian of the disparity spin system is split into two contributions

$$E_{total} = E_{exchange} + E_{field}. \tag{2}$$

These terms reflect the continuity and the compatibility conditions, respectively. The distance of an observer to a point on the surface of an object is a smoothly varying property. To achieve a corresponding property for the values of the disparity spin a Potts model interaction is choosen

$$E_{exchange} = -J \sum_{<(i,j),(k,l)>} F(S_{i,j}, S_{k,l}),$$
$$F(S_{i,j}, S_{k,l}) = \begin{cases} 1, & \text{if } S_{i,j} = S_{k,l} \\ q, & \text{if } S_{i,j} = S_{k,l} \pm 1; \quad q < 1. \\ 0, & \text{else} \end{cases} \tag{3}$$

If the value of a disparity spin is $S_{i,j}$ and if this hypothesis is correct, the pixel (i,j) in the right picture and the pixel $(i, j + S_{i,j})$ in the left picture have identical surroundings. If the disparity hypothesis $S_{i,j}$ is wrong the surroundings may be completely different. Therefore, the comparison of the neighborhoods of the two points assumed to correspond to eachother indicates a possible correct match.

Whereas many features can (and for real pictures must) be used for comparison, we restrict ourselves in the present application to the most simple choice and compare the pixel intensities in a square shaped region only. Comparison is established by the following energy contribution

$$E_{field} = \sum_{(i,j)} G(S_{i,j}), \qquad G(S_{i,j}) = G_0 \sum_{k=i-w}^{k=i+w} \sum_{l=j-h}^{l=j+h} | H^{left}_{k,l+S_{i,j}} - H^{right}_{k,l} |. \tag{4}$$

Here $H^{left}_{i,j}$ and $H^{right}_{i,j}$ denote the intensity of the pixel (i,j) in the left and in the right picture of the stereogram and $G_0 = [(2h+1)(2w+1)]^{-1}$ is a normalisation constant. Correct disparity spin configurations are characterized by low energy contributions.

For the input pattern shown in Figure 1 the disparity field obtained is presented in Figure 2. Starting from the temperature $T = 1.5$ the annealing process was stopped at a low temperature $T = 0.01$. For the interaction parameters in (3) we have assumed the values $J = 2$ and $q = 0$ and for the maximal disparity the value $N = 5$.

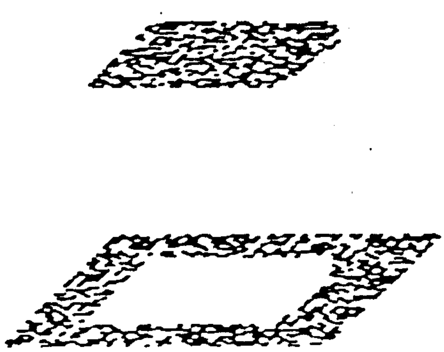


Fig. 2. Ground state of the disparity spin system corresponding to the Juless pattern in Figure 1.

A more complicated stereogram containing the three-dimensional information of an eight level pyramide is shown in Figure 3. The solutions of the disparity spin system ($N = 10$) with interaction parameters $J = 2$ and $q = 0.2$ are shown for three temperatures in Figure 4.

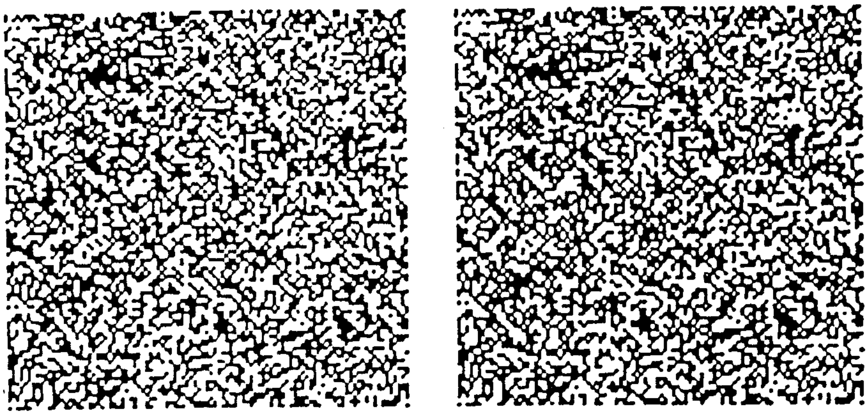


Fig. 3. Juless pattern of an eight level pyramide.

At the high temperature $T = 0.8$ the fluctuations of the disparity spin values are large. This is demonstrated by a snapshot of the dynamics shown in Figure 4.a. Figure 4.b illustrates that at the intermediate temperature $T = 0.3$ the system still fluctuates; however, the disparity field already indicates the presence of different disparity planes. Figure 4.c shows that at the low temperature $T = 0.1$ the fluctuations almost disappeared and that the disparity spin system achieves the correct interpretation of the Juless pattern, an eight level pyramide.

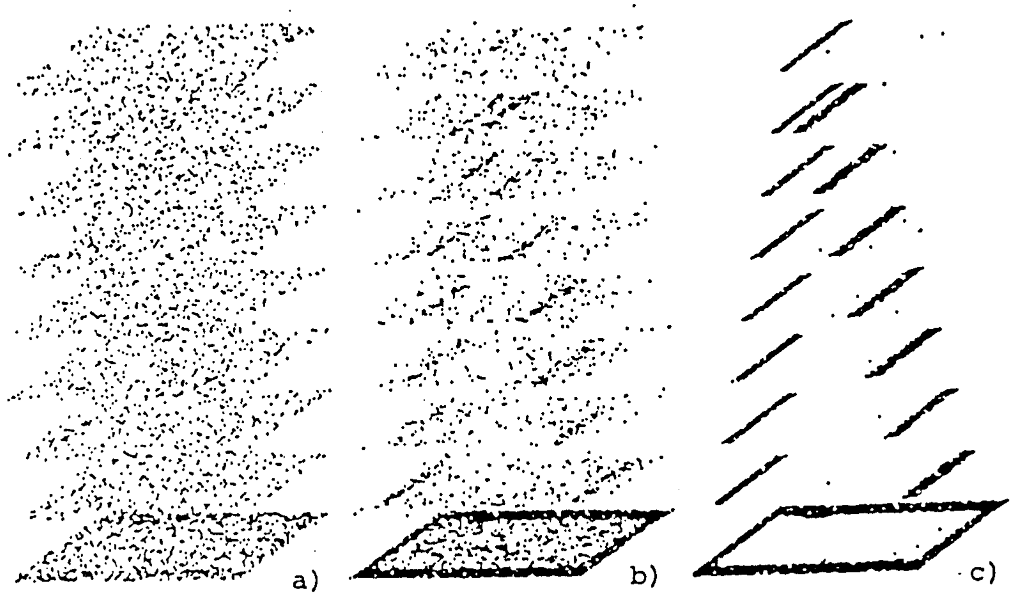


Fig. 4. Behaviour of the disparity spin system for the Juless pattern in Figure 3.

A Spin Model for Picture Restoration

Restoration of noisy pictures can be simplified if expected relations between picture attributes are known. As an example we consider a chessboard-like pattern as input for a picture restoring system. There are several 'a priori' qualities present in such pattern: the intensity in a square is constant, at a square's border are straight edges in vertical or in horisontal direction, edges are continuous. A system of feature spins instructed with this knowledge can restore noisy chessboard patterns. Such system entails three kinds of feature spins

- Intensity spins which take the values ± 1 for black and white colours, respectively.
- Horisontal edge spins which take the values $+1$ for intensity changes from white to black, the value -1 from black to white and the value 0 in the case of an absence of any edges.
- Vertical edge spins which follow corresponding rules.

The intensity spins are defined on a square lattice. The edge spins are located between neighboring intensity spins.

The Hamiltonian of the picture restoring spin system can be written

$$E_{total} = E_{i-ifield} + E_{h-hfield} + E_{v-vfield} + E_{i-i} + E_{h-h} + E_{v-v} + E_{i-h} + E_{i-v} \quad (5)$$

where the indices i, h, v refer to intensity, horisontal and vertical edge spins,respectively, with the corresponding fields $ifield$, $hfield$, $vfield$. To implement the continuity condition for the intensity spins $I_{i,j} \in \{-1, +1\}$ an Ising-like interaction is assumed. To implement the continuity property of edges an attractive interaction in the proper direction is employed for the horisontal edge spins $H_{i,j} \in \{-1, 0, +1\}$ as well as for the vertical edge spins $V_{i,j} \in \{-1, 0, +1\}$.

$$E_{i-i} = -J_i \sum_{<(i,j),(k,l)>} I_{i,j}\, I_{k,l} \quad (6)$$

$$E_{h-h} = -J_h \sum_{(i,j)} \delta(H_{i,j+1}, H_{i,j}); \qquad E_{v-v} = -J_v \sum_{(i,j)} \delta(V_{i+1,j}, V_{i,j}) \quad (7)$$

To obtain compatibility between the hypotheses of edge spins and intensity spins an interaction energy favoring consistent configurations is added

$$E_{i-h} = -J_{i-h} \sum_{(i,j)} T(I_{i,j}, I_{i+1,j}, H_{i,j}), \qquad E_{i-v} = -J_{i-v} \sum_{(i,j)} T(I_{i,j}, I_{i,j+1}, V_{i,j}). \quad (8)$$

Here T contains the compatibility condition listed in Table 1.

i	j	k	T
1	1	0	1
-1	-1	0	1
1	-1	-1	1
-1	1	1	1
1	1	1	0
-1	-1	1	0
1	1	-1	0
-1	-1	-1	0
1	-1	0	0
-1	1	0	0
1	-1	1	0
-1	1	1	0
1	-1	-1	0
-1	1	-1	0

Table 1: Compatibility condition table $T(i, j, k)$ where i, j denote pairs of intensity spins and k denotes the edge spin inbetween.

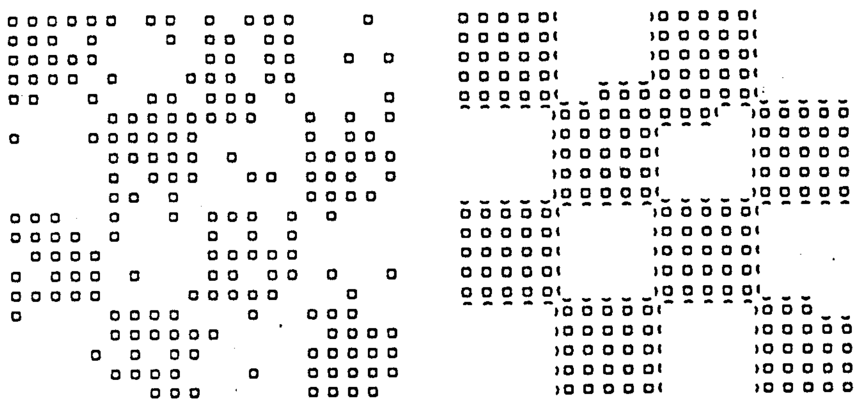
The pattern to be processed is coded as a field $F_{i,j} \in \{+1, -1\}$ corresponding to black and white pixels at position (i, j). The interaction between intensity spins and the field is chooosen like in the Ising model.

$$E_{i-ifield} = -J_{ifield} \sum_{(i,j)} I_{i,j}\, F_{i,j} \quad (9)$$

The field for edge spins codes intensity changes. The interaction between edge spins and the corresponding fields is

$$
\begin{aligned}
E_{h-hfield} &= -J_{hfield} \sum_{(i,j)} \delta(H_{i,j}, (F_{i+1,j} - F_{i,j})/2). \\
E_{v-vfield} &= -J_{vfield} \sum_{(i,j)} \delta(V_{i,j}, (F_{i,j+1} - F_{i,j})/2).
\end{aligned}
\tag{10}
$$

As input for the picture restoration spin system we choose a distorted chessboard pattern, measuring 20×20 pixels. 20 percent of the pixels were randomly reversed from black to white and vice versa. The resulting pattern is presented in Figure 5.a. The aim of the restoration process is to find the chessboard closest to this picture. The interaction paparameters assumed for the restoration are $J_i = 1$, $J_h = J_v = 4.5$, $J_{i-h} = J_{i-v} = 4.5$, $J_{ifield} = 1$, $J_{hfield} = J_{vfield} = 3$. The temperature was lowered in 12 steps from an initial value of $T = 8$ to the final value of $T = 0.05$. Figure 5.b shows the result: the chessboard pattern has been restored to a very large degree.



a) b)

Fig. 5. Input pattern and restored chessboard pattern for the feature spin system described by Eqs. (5)-(10).

References

1. J. Kirkpatrick, C.D. Gelatt, M.P. Vecchi, Science, 220, 671 (1983).
2. S. Geman, D. Geman, IEEE Tran. Pattern Anal. Machine Intell., vol. PAMI-6, 721, (1984).
3. P. Kienker, T.J. Sejnowski, G.E. Hinton, L.E. Schuhmacher, Separating Figure from Ground with a Parallel Network, preprint (1986).
4. D. Marr, Vision (Freeman, 1982).
5. B. Julesz, Foundations of Cyclopean Perception (Univ. Chicago Press, 1971).