Dynamic Attention Map by Ising Model for Human Face Detection

Masaru Tanaka† mtanaka@etl.go.jp
Kazuhiro Hotta‡ hotta@me.ics.saitama-u.ac.jp
Takio Kurita† kurita@etl.go.jp
Taketoshi Mishima‡ mishima@me.ics.saitama-u.ac.jp

† Electrotechnical Laboratory
1-1-4, Umezono, Tsukuba
Ibaraki 305-8568, JAPAN

‡ Saitama University
255, Shimo-Okubo, Urawa
Saitama 338, JAPAN

Abstract

We present a method to narrow down the search space for scale-invariant human face detection, which uses Dynamic Attention Map implemented by Ising dynamics. Combining the proposed method and the scale-invariant face detection method which is based on both Higher-Order Local Autocorrelation (HLAC) features of Log-Polar image and Linear Discriminant Analysis for "face" and "not face" classification, it is shown that the "face" region in the image can be detected faster with some experiments.

1. Introduction

For the face recognition, there are many methods presented [1, 2, 3, 4, 5, 6].. But their performance varies depending on illumination, a size of a target, a facial expression and/or background etc. Previously, we presented the scale and rotation invariant face recognition method [7, 8] based on both Higher-Order Local Autocorrelation (HLAC) features of Log-Polar image and Linear Discriminant Analysis for "face" and "not face" classification. In our scale and rotation invariant face recognition method, the searching for the "face" region was performed randomly or sequencially on the image. Therefore its searching performance was not satisfiable. In this paper, we propose a method to narrow down the search space by dynamically using the information obtained at the previous search point through constructing the Dynamic Attention Map. The Dynamic Attention Map is the distribution of the "face" and "not face" regions updated as the searching step goes on. To implement the Dynamic Attention Map, we make use of the Ising model [9, 10], which is the simplest model of magnetization. The Ising model consists of spins interacting each other in some manner, each of which can take only one state of "up" or "down" at a time. The spins will be flipped through a certain dynamics depending on both the states of their neighbor spins and an external magnetic field. It is natural to look upon two states as "face"and "not face" in face detection. In our face detection method [7], the distance in the discriminant space between the feature vector at a search point and the mean vector of "face" class represents the likelihood of face. The shorter distance in the discriminant space, the higher likelyhood of face. Each spin represents the state of the corresponding search window, which shows whether it is in the "face" state or not. The neighboring region (spins) of the search point (the selected spin) is expected to have the almost same distance in the discriminant space as that of the search point. Then, if the measured distance is integrated into the energy function of the Ising model, then the states of the spins in the neighboring region of the search point can be estimated through the Ising dynamics.

2. Search Method for Face Detection

2.1. Ising model

The Ising model is the simplest model of magnetization. It consists of Ising spins, each of which takes one of two states "up" and "down", with some interactions among the nearest neighboring spins. Originally, The Ising model was proposed as a simplified version of Heisenberg model, which consists of two-state spins and interactions between all the spins [9, 10]. Now this Ising model is quite famous for its usefulness in the fields of physics and neural networks. The dynamics of the Ising model decreases the total energy of a system, which is the sum of the interaction energy among the nearest neighboring spins and the interaction energy with an external magnetic field.

2.2. Dynamic Attention Map

We introduce the Ising dynamics to utilize information obtained at the previous search point effectively. Ising dynamics works to minimize the energy function which includes the interactions of the neighboring spins and the interaction with an external magnetic field. If we integrate the measured distance in the discriminant space (the likelihood of "face") into the energy function as an external magnetic field, then the state of the neighboring spins to the selected spin can be estimated whether they are in the "face" states or not, through the Ising dynamics. Then this can be used to reduce the search space dynamically. In the face detection, there are two states; "face" and "not face". Here we set the "face" state to -1 (a down spin) and the "not face" state to 1 (an up spin). The direction of an external magnetic field (H), which is a measured distance in the discriminant space, is assumed to point the "not face" direction, because the regions of the "not face" in the image are usually wider than that of the "face". When a spin s_a is selected to evaluate a likelihood of "face", then the energy E_i of the neareset neiborghing spins $\{s_i | i \in nn(a)\}$, $nn(a)$ means the nearest neighboring spins to the spin s_a, is given by

$$E_i = -J \sum_{j \in nn(i)} s_i s_j - H(m(a) - \theta) s_i \ , \ i \in nn(a), \quad (1)$$

where J is the strength of the interaction between spins, $m(a)$ is the likelihood of "face" (the distance of the spin s_i and the mean vector of "face" class in the discriminant space), and θ is the threshold which is used to decide whether the region is the "face" or the "not face". Then the state of each spin is updated according to the probability which is proportional to

$$\exp(-\beta \Delta E_i) \ , \quad (2)$$

where ΔE_i is the energy change caused by flipping the spin s_i, that is,

$$\Delta E_i = 2J \sum_{j \in nn(i)} s_i s_j + 2H(m(a) - \theta) s_i \ , \quad (3)$$

and β is a reciprocal of the temperature. This spin flip dynamics in the Ising model creates the Dynamic Attention Map. In the following, we show the meta algorithm for the search method and call this search method Ising search.

1. Set all the spins to -1 ("face") and make the "face" list for search. The list consists of the spins in the "face" state, that is, the set of the pixel coordinates on the image and the corresponding spin state.

2. Select one spin randomly from the "face" list.

3. Extract the HLAC [7, 8] features from the region centered the selected spin through the Log-Polar transformation. Then measure the likelihood of "face" in the discriminant space.

4. Apply the Ising dynamics for suitable times.

5. Remove the spins flipped from the "face" to the "not face" from the "face" list and add the spins flipped from the "not face" to the "face" to the "face" list. Note that, if the selected spins to measure the likelihood of "face" are found to be in the "not face" state, they are removed from the "face" list and never flipped again.

6. Repeat from 2 to 5 until the region, in which the likelihood of "face" of spins is below the threshold, is found.

3. Experiments

In the following experiments, the size of an input image is 160×120 and that of a search window is 59×59. One spin is assigned at the center pixel of a search window. The other parameters of the Ising search are set to $\beta = 0.5$, $H = 0.25$. Monte Carlo steps (MCS) are performed 5 times in the neighboring 5×5 lattice centered the selected spin. For each MCS, the Ising dynamics is performed 24 times. Comparing the performance of the Ising search and the random search, we experimented for four cases: The case 1 is one person in an image with the "face" to the "not face" pixel-ratio 27/5657. The case 2 is one person with the "face" to the "not face" pixel-ratio 24/5660. The case 3 is two persons at the different depths with the "face" to the "not face" pixel-ratio 113/5571. The case 4 is one person and his hand with the "face" to the "not face" pixel-ratio 92/5592. See Figure 1. The face detection experiment was repeated 100 times for each case. Table 1 shows the means and the medians over 100 trials. Figure 2 shows how the length of the "face" list changes. The random search makes the length of the "face" list decrease gently. But the Ising search makes

	mean		median	
	Ising	random	Ising	random
case 1	125.65	258.99	124	190
case 2	139.94	206.23	133	172
case 3	44.23	53.19	37	44
case 4	41.89	66.83	29	55

Table 1. The Ising search vs. the random search

the length of the "face" list decrease steeply. Thus we can see that the Ising search make the search space narrower efficiently than the random search. Both Table 1 and Figure 2 show that the Ising search works effectively for the face detection.

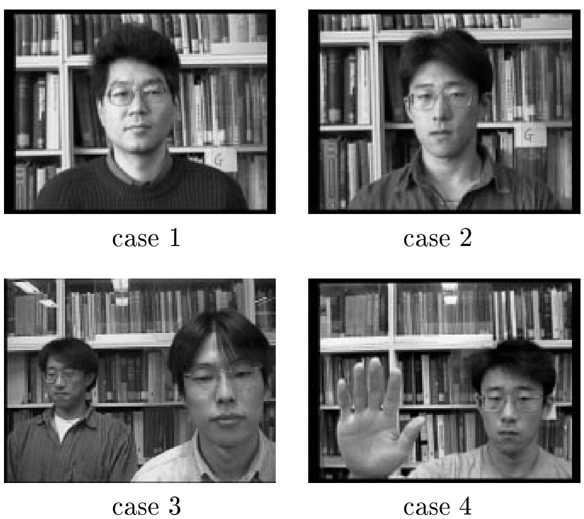

Figure 1. The images for the face detection experiments.

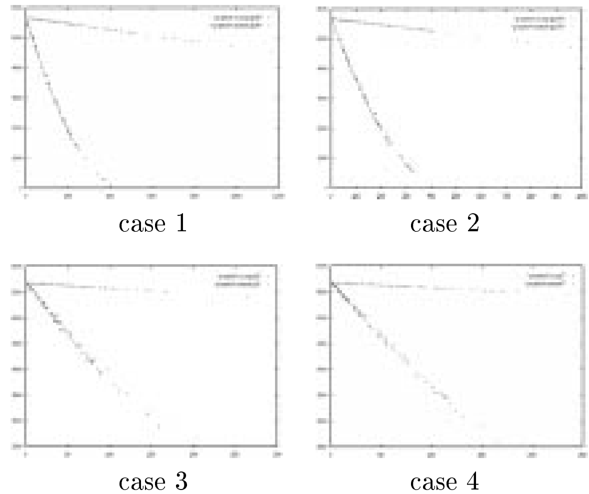

Figure 2. The change of the length of the "face" list

References

[1] R.Chellappa, C.L.Wilson, and S.Sirohey, "Human and machine recognition of faces: a survey", Proc. of IEEE, Vol.83, No.5, pp.705-740, 1995.

[2] M.Kirby and L.Sirovich, "Application of the Karhunen-Lov̀e procedure for the characterization of human faces", Trans. on Pattern Analysis and Machine Intelligence, Vol.12, No.1, pp.103-108, 1990.

[3] M.A.Turk and A.P.Pentland, "Face recognition using eigenfaces", Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, pp.586-5 91, 1991.

[4] T.Kurita, N.Otsu, and T.Sato, "A face recognition method using higher order local autocorrelation and multivariate analysis", Proc. 11th IAPR International Conf. on Pattern Recognition, pp.213-216, 1992.

[5] K.Sung and T.Poggio, "Example-based learning for view-based human face detection", A.I. Memo 1521,CBCL Paper 112, MIT, December 1994.

[6] Henry A. Rowley, Shumeet Baluja, and Takeo Kanade, "Human face detection in visual scenes", CMU-CS-95-158R, CMU, November 1995.

[7] K.Hotta, T.Kurita, T.Mishima, "Scale Invarianct Face Detection Method using Higher-Order Local Autocorrelation features extracted from Log-Polar Image", Proc. of the third IEEE International Conference on Automatic Face and Gesture Recognition, pp.70-75, Nara,1998.

[8] T.Kurita, K.Hotta, T.Mishima, "Scale and Rotation Invariant Recognition Method using Higher-Order Local Autocorrelation features of Log-Polar Image", Proc. of third Asian Conference on Computer Vision, Vol.II, pp.89-96, Hong Kong,1998.

[9] Mezard M., Parisi G., "Spin Glass Theory and Beyond", Singapore. World Scientific, 1987.

[10] Harvery Gould and Jan Tobochnik, "An Introduction to Computer Simulation Methods - APPLICATIONS TO PHYSICAL SYSTEMS 2nd EDITION", Addison -Wesley, 1996.