Translation, Scale and Rotation Invariant Features Based on High-Order Autocorrelations

Shotaro AKAHO

(Accepted, 9, 27, 1993)

Local high-order autocorrelation features proposed by Otsu have been successfully applied to face recognition and many other pattern recognition problems. These features are invariant under translation, but not invariant under scale and rotation. We construct scale and rotation invariant features from the local high-order autocorrelation features.

§ 1 Introduction

In pattern recognition problem, it is important to extract invariant features from given images. The capability of statistical methods such as discriminant analysis depends on the dimensionality of inputs and the number of samples. If the dimensionality of inputs is large, a large number of samples are needed. Therefore we should construct a low dimensional feature space from a raw image that belongs to an infinite dimensional functional space. If some transformation does not change the categorical property of an image, the feature should be invariant under that transformation. The low dimensionality of features is also effective to reduce the computation time of pattern recognition tasks.

Otsu proposed a practical simplification of high-order autocorrelation features (*high-order local autocorrelation*[5)]), which is successfully applied to many kinds of pattern recognition problems such as face recognition[6), 3)]. However, since the high-order autocorrelation features are not invariant under rotation and scale, the performance of recognition decreases under those transformations. In this paper, we construct scale and rotation invariant features (approximately) from Otsu's high-order local autocorrelation features. The difficulties of constructing

KEYWORDS— Pattern recognition, Image processing, High-order autocorrelation, Invariant feature extraction

those features are mainly caused by the fact that high-order autocorrelation function is non-linear and that local autocorrelation features do not include enough information to construct invariant features.

In § 2, we describe the high-order autocorrelation features, and in § 3 we present invariant features. Rotation invariant features are given in (13). Scale invariant features are given in (15) in the case that only Otsu's local features are used. Another set of scale invariant features are given in (20) and (26). The former is in the case that extended version of local features are used, and the latter is in the case that multiresolutional local features are used. Finally, we present features that are invariant both under rotation and scale in (29), (30) and (32).

§ 2 High-order autocorrelation features

2.1 High-order autocorrelation function

Let us consider an image data $f(\boldsymbol{r}) = f(x, y)$ on two dimension Euclidean space (E^2).

The k**-th order autocorrelation** $R(B)$ of $f(\boldsymbol{r})$ is defined by

$$R(B) = R(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k) = \int_{E^2} f(\boldsymbol{r}) f(\boldsymbol{r} + \boldsymbol{b}_1) \cdots f(\boldsymbol{r} + \boldsymbol{b}_k)\, d\boldsymbol{r}, \tag{1}$$

where $B = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k)$ is a $2 \times k$ matrix.

Although $R(B)$ is clearly a translation invariant feature, it is not invariant under scale and rotation. In fact, let $f_{\mathrm{scl}}(\boldsymbol{r}) = f(\boldsymbol{r}/\lambda)$ be the image scaled by λ and let $f_{\mathrm{rot}}(\boldsymbol{r}) = f(T(\theta)\boldsymbol{r})$ be the image rotated by θ (where $T(\theta)$ is a rotation matrix of angle $-\theta$), the k-th order autocorrelation of those transformed images becomes respectively

$$R_{\mathrm{scl}}(B) = \lambda^2 R(\frac{B}{\lambda}), \tag{2}$$

$$R_{\mathrm{rot}}(B) = R(T(\theta)B). \tag{3}$$

In practical applications, images are defined on discrete lattice points (pixels). In that case, the integral $\int_{E^2} d\boldsymbol{r}$ is approximated by the sum over all pixels. Mclaughlin has proved that the second order autocorrelation function defined on pixels with compact support is a complete system except the freedom of translation[4)].

2.2 Local autocorrelation features

There are an infinite number of high-order autocorrelation features corresponding to different B's. For practical applications, we restrict the autocorrelation to be the second order and to be taken just in neighborhood of 3×3 pixels (*local autocorrelation features*[5)6)]). By this restriction, we can reduce the number of features to 25.

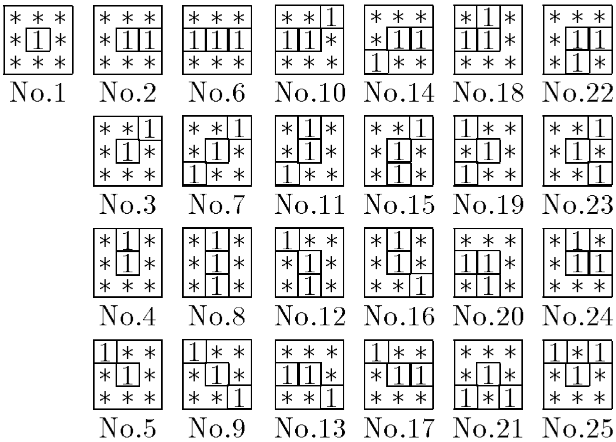

Fig. 1 Local autocorrelation mask ($*$: 'don't care')

Especially if an image is binary, we only scan the image by 25 masks shown in **Figure 1** and count the number of matching pixels in order to get the local autocorrelation features.

If an image is not binary, for degenerated masks (No. 1, 2, 3, 4 and 5), the value of the centered point is multiplied as many times as the degeneration. For example, the feature corresponding to No. 1 mask is given by

$$R(B_1) = \sum_{\boldsymbol{r}} f(\boldsymbol{r})^3, \tag{4}$$

and the feature of No. 2 is given by

$$R(B_2) = \sum_{\boldsymbol{r}} f(\boldsymbol{r})^2 f(\boldsymbol{r} + \begin{pmatrix}1\\0\end{pmatrix}).$$

In general, the local autocorrelation feature corresponding to i-th mask is given by

$$R(B_i) = \sum_{\boldsymbol{r}} f(\boldsymbol{r}) f(\boldsymbol{r} + \boldsymbol{b}_1) f(\boldsymbol{r} + \boldsymbol{b}_2), \tag{5}$$

where $B_i = (\boldsymbol{b}_1, \boldsymbol{b}_2)$. Clearly, all $R(B_i)$'s are translation invariant.

§ 3 Invariant features

3.1 Rotation invariant features

From the equation (3), we can easily find the following $Q(B)$ is rotation invariant:

$$Q(\chi_1, \ldots, \chi_k, B) = \int_{-\pi}^{\pi} R(T(\psi)B) R(T(\psi + \chi_1)B) \cdots R(T(\psi + \chi_k)B) d\psi. \tag{6}$$

This is the same form as the k-th order autocorrelation function, and let us call it the **k-th order rotation autocorrelation function**. Since the second order autocorrelation function is a complete system, the second order rotation autocorrelation function is also complete.

In order to approximate $Q(\chi_1, \ldots, \chi_k, B)$ by a finite number of local autocorrelations, let us divide the masks shown in figure 1 into subsets so that in each subset any mask can be transformed to other mask by rotation. Approximately, the mask of No.3 is considered to be the mask of No.2 rotated by $\pi/4$, and the mask of No.11 is the mask of No.12 rotated by $\pi/4$, etc. Consequently, the masks are divided into the following five subsets:

$$M_1 = \{B_1\}, \tag{7}$$
$$M_2 = \{B_2, B_3, B_4, B_5\}, \tag{8}$$
$$M_3 = \{B_6, B_7, B_8, B_9\}, \tag{9}$$
$$M_4 = \{B_{10}, B_{11}, B_{12}, B_{13}, B_{14}, B_{15}, B_{16}, B_{17}\}, \tag{10}$$
$$M_5 = \{B_{18}, B_{19}, B_{20}, B_{21}, B_{22}, B_{23}, B_{24}, B_{25}\}, \tag{11}$$

where B_i corresponds to the No. i mask. Using this division, 0-th order rotation autocorrelation $Q(B)$ can be approximated by

$$\tilde{Q}_i = \sum_{B \in M_i} R(B), \qquad i = 1, \ldots, 5, \tag{12}$$

This invariant features can be also obtained in an algebraic way[7)].

In general, k-th order rotation autocorrelation $Q(\chi_1, \ldots, \chi_k, B)$ can be approximated by

$$\tilde{Q}_{i;I_k} = \sum_{B \in M_i} R(B) R(T(\frac{\pi}{4} i_1) B) \cdots R(T(\frac{\pi}{4} i_k) B), \tag{13}$$
$$i = 1, \ldots, 5, \qquad I_k = i_1, \ldots, i_k : \text{ integers.}$$

Remark that $T(\frac{\pi}{4} i_1)B, \ldots, T(\frac{\pi}{4} i_k)B$ are also included in M_i and $\tilde{Q}_{i;I_k}$ can be calculated by just using Otsu's local masks.

3.2 Scale invariant features using local features

In this section, we construct scale invariant features by only using the features extracted by the local autocorrelation masks.

From the equation (2), the autocorrelation of the image scaled by λ is a function of $R(B/\lambda)$, which is not included in the local masks. Let us interpolate it by linear approximation with respect to $1/\lambda$. Namely,

$$\begin{aligned} R_{\text{scl}}(B) &= \lambda^2 R(\frac{B}{\lambda}) \simeq \lambda^2 \{(1 - \frac{1}{\lambda}) R(O) + \frac{1}{\lambda} R(B)\} \\ &= \lambda(\lambda - 1) R(O) + \lambda R(B), \end{aligned} \tag{14}$$

where O denotes the zero matrix.

From this approximation, we obtain scale invariant features,

$$\tilde{U}(B_i, B_j, B_l, B_m) = \frac{R(B_i) - R(B_j)}{R(B_l) - R(B_m)}, \qquad i, j, l, m \in \{1, \ldots, 25\}. \tag{15}$$

We can also derive the same approximated invariant features from the interpolation by Shannon's sampling theorem[1)].

3.3 Scale invariant features using extended local features

In this section, we extend the local autocorrelation features, namely we consider the case that the autocorrelation features $R(B_i), R(2B_i), R(3B_i), \ldots$ $(i = 1, \ldots, 25)$ are given, where B_i corresponds to the local autocorrelation masks. This extension makes it possible to get a better scale invariant features than (15) in such sense that the set of invariant features is a complete system except the freedom of scale.

First, we shall prove that the following value is scale invariant:

$$V(\mu, B) = (R(O))^{-\frac{3}{2}} \int_0^{\mu\sqrt{R(O)}} R(tB)dt, \tag{16}$$

where μ is a real-value parameter. We can also show that $V(\mu, B)$ is a complete system except the freedom of scale.

Proof: From the equation (2), $V(\mu, B)$ for scaled image is given by

$$\begin{aligned} V_{\mathrm{scl}}(\mu, B) &= (R_{\mathrm{scl}}(O))^{-\frac{3}{2}} \int_0^{\mu\sqrt{R_{\mathrm{scl}}(O)}} R_{\mathrm{scl}}(tB)dt \\ &= (\lambda^2 R(O))^{-\frac{3}{2}} \int_0^{\mu\sqrt{\lambda^2 R(O)}} \lambda^2 R(\frac{t}{\lambda}B)dt. \end{aligned} \tag{17}$$

Substituting t by $t' = t/\lambda$, we get

$$V_{\mathrm{scl}}(\mu, B) = \lambda^{-3}(R(O))^{-\frac{3}{2}} \int_0^{\mu\sqrt{R(O)}} \lambda^3 R(t'B)dt' = V(\mu, B). \tag{18}$$

Thus $V(\mu, B)$ is scale invariant.

Next we show the completeness of $V(\mu, B)$ except the freedom of scale. Differentiating $V(\mu, B)$ by μ, it follows

$$\frac{\partial}{\partial\mu}V(\mu, B) = (R(O))^{-1}R(\mu\sqrt{R(O)}B). \tag{19}$$

This equation shows the function $R(B)$ can be found from $V(\mu, B)$ except the unknown freedom of $R(O)$ that is equivalent the freedom of scale. □

If the values $R(B_i), R(2B_i), \ldots$ are given, $V(\mu, B)$ can be approximated by

$$\tilde{V}(\mu, B_i) = (R(O))^{-\frac{3}{2}} \sum_{j=0}^{\left[\mu\sqrt{R(O)}\right]} R(jB_i), \qquad i = 1, \ldots, 25, \tag{20}$$

where $\left[\mu\sqrt{R(O)}\right]$ denotes the rounded integer of $\mu\sqrt{R(O)}$.

3.4 Scale invariant features using multiresolution local features

In this section, we introduce scale invariant features based on the local autocorrelation features for a set of images of different resolutions. The local autocorrelation features for multiresolution have been used for face recognition[3]. Images scaled by $1, 1/2, 1/4, \ldots, 1/2^k, \ldots$ can be constructed easily by a pyramidal structure[2]. Let $R^{(\nu)}(B_i)$ be the local autocorrelation feature of an image of scale $1/2^\nu$ corresponding to No. i mask. Especially, $R^{(0)}(B_i) = R(B_i)$.

In a similar way to the preceding section, we can show the following value is scale invariant:

$$W(\mu_1, \mu_2, B) = \int_{g_1}^{g_2} R^{(\nu)}(B) d\nu, \tag{21}$$

where

$$g_1 = \frac{1}{2}\log_2(\mu_1 R(O)), \qquad g_2 = \frac{1}{2}\log_2(\mu_2 R(O)). \tag{22}$$

Proof: From the equation (2), $W(\mu_1, \mu_2, B)$ for the image scaled by λ is given by

$$W_{\text{scl}}(\mu_1, \mu_2, B) = \int_{\frac{1}{2}\log(\mu_1 R_{\text{scl}}(O))}^{\frac{1}{2}\log(\mu_2 R_{\text{scl}}(O))} R_{\text{scl}}^{(\nu)}(B) d\nu \tag{23}$$

$$= \int_{\frac{1}{2}\log(\mu_1 \lambda^2 R(O))}^{\frac{1}{2}\log(\mu_2 \lambda^2 R(O))} \frac{\lambda^2}{2^{2\nu}} R(\frac{2^\nu}{\lambda} B) d\nu. \tag{24}$$

Letting $\xi = 2^\nu/\lambda$, it follows

$$W_{\text{scl}}(\mu_1, \mu_2, B) = \int_{\sqrt{\mu_1 R(O)}}^{\sqrt{\mu_2 R(O)}} \frac{1}{\xi^2} R(\xi B) \cdot \frac{d\xi}{\xi \log 2}, \tag{25}$$

which does not depend on λ, hence $W(\mu_1, \mu_2, B)$ is scale invariant. □

Using $R(B_i), R^{(1)}(B_i), \ldots, R^{(k)}(B_i), \ldots$, we can approximate $W(\mu_1, \mu_2, B)$ by

$$\tilde{W}(\mu_1, \mu_2, B_i) = \sum_{k=\tilde{g}_1}^{\tilde{g}_2} R^{(k)}(B_i), \qquad i = 1, \ldots, 25, \tag{26}$$

where

$$\tilde{g}_1 = \left[\frac{1}{2}\log_2(\mu_1 R(O))\right], \qquad \tilde{g}_2 = \left[\frac{1}{2}\log_2(\mu_2 R(O))\right]. \tag{27}$$

3.5 Rotation and scale invariant features

Using the fact described in the preceding sections, we can easily construct invariant features both under rotation and under scale.

The k-th order rotation autocorrelation function of the image scaled by λ is given by

$$Q_{\mathrm{scl}}(\chi_1, \ldots, \chi_k, B) = \lambda^{2(k+1)} Q(\chi_1, \ldots, \chi_k, \frac{B}{\lambda}). \tag{28}$$

Thus 0-th order rotation autocorrelation $Q(B)$ is transformed in the same manner as $R(B)$, and we only replace $R(B)$ by $Q(B)$ in (15), (20) and (26) in order to make them invariant under rotation. In general, making a slight modification, we obtain rotation and scale invariant features $\tilde{X}$, $\tilde{Y}$ and $\tilde{Z}$ below.

From (15),

$$\tilde{X}_{i,j,l,m;I_k} = \frac{\tilde{Q}_{i;I_k} - \tilde{Q}_{j;I_k}}{\tilde{Q}_{l;I_k} - \tilde{Q}_{m;I_k}}, \qquad i, j, l, m \in \{1, \ldots, 5\}. \tag{29}$$

From (20),

$$\tilde{Y}_{i;I_k}(\mu) = (R(O))^{-\frac{2k+3}{2}} \sum_{j=0}^{\left[\mu\sqrt{R(O)}\right]} \tilde{Q}_{i;I_k;j}, \qquad i = 1, \ldots, 5, \tag{30}$$

where

$$\tilde{Q}_{i;I_k;j} = \sum_{B \in M_i} R(jB) R(jT(\frac{\pi}{4} i_1)B) \cdots R(jT(\frac{\pi}{4} i_k)B). \tag{31}$$

From (26),

$$\tilde{Z}_{i;I_k}(\mu_1, \mu_2) = \sum_{n=\tilde{g}_1}^{\tilde{g}_2} \tilde{Q}^{(n)}_{i;I_k}, \qquad i = 1, \ldots, 5, \tag{32}$$

where $\tilde{Q}^{(n)}_{i;I_k}$ is the value $\tilde{Q}_{i;I_k}$ of the image scaled $1/2^n$.

§ 4 Conclusion

We have constructed rotation and scale invariant features from high-order local autocorrelation features. In practice, there is a trade-off between the acuracy and the computability. If we can get an infinite number of $R(B)$'s, we can construct a set of invariant features that is acurate and complete. However, it enlarges computation time and the dimensionality of feature space. In order to reduce the dimensionality of feature space furthermore, we should construct invariant features under other transformations or select effective features adaptablly for each problems. They remain as future works.

Acknowledgments

The author would like to thank M. Suwa, Director of the Information Science Division, for affording an opportunity of this study. He is also deeply indebted to N. Otsu, Director of the Machine Understanding Division, and all members of the Mathematical Informatics Section for their helpful discussions.

References

1) S. Akaho, T. Sato, I. Sekita, S. Umeyama, and N. Otsu: Position, scale and rotation invariant featues based on high-order autocorrelations. In *1991 Spring National Convention Record, IEICE*, D–509, Tokushima, 1991. (In Japanese).

2) D.H. Ballard and C.M. Brown: *Computer Vision*. Prentice-Hall, 1982.

3) T. Kurita, N. Otsu, and T. Sato: A face recognition method using higher order local autocorrelation and multivariate analysis. In *Proc. of 11th International Conf. on Pattern Recognition*, volume II, pp. 213–216, The Hague, 1992.

4) J.A. Mclaughlin and J. Raviv: N-th order autocorrelations in pattern recognition. *Information and Control*, Vol. 12, pp. 121–142, 1968.

5) N. Otsu: Mathematical studies on feature extraction in pattern recognition. Researches of the ETL 818, 1981. (In Japanese).

6) N. Otsu and T. Kurita: A new scheme for practical flexible and intelligent vision systems. In *Proc. IAPR Workshop on Computer Vision*, Tokyo, 1988.

7) S. Umeyama and T. Kurita: Rotation invariant features based on higher-order autocorrelations. In *45th National Convention Record, IPSJ*, 2–323, Tokushima, 1992. (In Japanese).