UDC 528.715:004.932(045)

Maryna Mukhina

CORRELATION EXTREMAL PRINCIPLE OF OBJECT DETECTION ON GROUND SURFACE

National Aviation University
Kosmonavta Komarova avenue 1, 03680, Kyiv, Ukraine
E-mail: m_mukhina@inbox.ru

Abstract. *Features of correlation extremal principle for detection of ground objects are considered for visual surveillance. The method of phase correlation is researched and variants of its optimization are proposed.*

Keywords: correlation extremal principle, phase correlation, logarithmic polar transform

1. Introduction

Correlation extremal principle is based on the comparison of current signal (image) with template images and on further determination of extremes of their correlation function. Maximum of correlation function for definite template will characterize the maximal conformity of two images. In case of presented template images of target object it is possible to provide its detection with visual surveillance of ground surface. Visual information is assumed to be so called optical geophysical field (Baklitskii 2009).

The optical field contains large amount of information including both the valid signals like images of ground objects, elements of relief, texture features, etc., and noises. The latter may include glares, weather phenomena like fog, rain, imperfection of vision sensors and others. The quality of detection significantly depends on the method of correlation analysis used to process images. To estimate the quality let's use such quantity characteristics as probability of object skipping and probability of fault detection.

The proposed solution includes the following stages:

– Pre-processing of image;
– Determination of the criterion type;
– Search of extremes of criterion function.

2. Problem statement

Continuous 2D image can be represented in discrete form by its sampling by lattice function $S(x, y, t)$ on discretization interval $\Delta x, \Delta y$ and quantization by level of image intensity. Time will be presented for video sequence where each frame is selected for definite interval of time Δt.

$$S(x, y, t) = \sum_{k=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \delta\left(x - i\Delta x, y - j\Delta y, t - k\Delta t\right),$$

where k is the number of frames in video sequence, i, j are sampling indexes.

But in practice the discrete images are represented not in the terms of delta-functions δ, but as matrices $F(i, j)$ of dimension $n \times m$. Besides such representation the 2D image can be written in vector form (Pratt 1982). Let's introduce vector $\mathbf{V}_r$ of dimension $m \times 1$ and matrix $\mathbf{H}_r$ of dimension $nm \times n$:

$$\mathbf{V}_r = \begin{array}{c} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \end{array} \begin{array}{c} 1 \\ \vdots \\ r-1 \\ r \\ r+1 \\ \vdots \\ m \end{array}, \qquad \mathbf{N}_r = \begin{array}{cc} & n \\ n & \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \end{array} \begin{array}{c} 1 \\ \vdots \\ r-1 \\ r \\ r+1 \\ \vdots \\ m \end{array}.$$

Matrix $\mathbf{F}$ will be obtained in the vector form with the help of ordering operation:

$$\mathbf{f} = \sum_{r=1}^{m} \mathbf{N}_r \mathbf{F} \mathbf{V}_r .$$

Vector $\mathbf{V}_r$ separates the r^{th} column of matrix $\mathbf{F}$, and matrix $\mathbf{N}_r$ puts this column at the place reserved for r^{th} section of vector $\mathbf{f}$. Briefly, the vector $\mathbf{f}$ contains all elements of matrix $\mathbf{F}$, sequentially read by columns. The inverse operation to convert vector $\mathbf{f}$ into matrix $\mathbf{F}$ is the following:

$$\mathbf{F} = \sum_{r=1}^{m} \mathbf{N}_r^T \mathbf{f} \mathbf{V}_r^T .$$

Let's assume that the template images will have the dimensions $n_{tem} \times m_{tem}$, and dimension of current image is $n_c \times m_c$, taking into account that $n_{tem}, m_{tem} << n_c, m_c$.

Pre-processing of current image usually include the minimization of the noise level and decreasing of the resolution of current image without loss in informativity.

Statistical properties of discrete images can be represented by their mathematical expectation, variance, correlation and covariance. The mathematical expectation of image is determined as following:

$$E\{\mathbf{F}\} = \left[E\{F(i,j)\}\right],$$

where $E\{F(i,j)\}$ is the mathematical expectation of element.

If image matrix is converted to the vector, than

$$E\{\mathbf{f}\} = \sum_{r}^{m} \mathbf{N}_r E\{\mathbf{F}\} \mathbf{V}_r .$$

Correlation between two elements of image with coordinates (i_1, j_1) and (i_2, j_2) is determined as follows:

$$R(i_1, j_1; i_2, j_2) = E\{F(i_1, j_1) F(i_2, j_2)\} . \qquad (1)$$

Covariance of two image elements can be found as

$$K(i_1, j_1; i_2, j_2) = \\ = E\left\{ \begin{matrix} \left[F(i_1, j_1) - E\{F(i_1, j_1)\}\right] \times \\ \times \left[F(i_2, j_2) - E\{F(i_2, j_2)\}\right] \end{matrix} \right\}.$$

Variance of image element is the following:

$$\sigma^2(i,j) = K(i,j;i,j).$$

If the image matrix is converted into vector **f**, then correlation matrix of the vector can be expressed in terms of correlations of elements in matrix **F**:

$$\mathbf{R_f} = E\{\mathbf{f}\mathbf{f}^T\} = E\left\{\left[\sum_{s=1}^{m} \mathbf{N}_s \mathbf{F} \mathbf{V}_s\right]\left[\sum_{r=1}^{m} \mathbf{V}_r^T \mathbf{F}^T \mathbf{N}_r^T\right]\right\}$$

or

$$\mathbf{R_f} = \sum_{s=1}^{m}\sum_{r=1}^{m} \mathbf{N}_s E\{\mathbf{F}\mathbf{V}_s \mathbf{V}_r^T \mathbf{F}^T\} \mathbf{N}_r .$$

Expression $E\{\mathbf{F}\mathbf{V}_s \mathbf{V}_r^T \mathbf{F}^T\} = \mathbf{R}_{s,r}$ is the correlation matrix of s^{th} and r^{th} columns of matrix **F** and has the dimension of $n \times n$. And finally $\mathbf{R_f}$ can be represented in a form of block matrix:

$$\mathbf{R_f} = \begin{bmatrix} \mathbf{R}_{1,1} & \mathbf{R}_{1,2} & \cdots & \mathbf{R}_{1,m} \\ \mathbf{R}_{2,1} & \mathbf{R}_{2,2} & \cdots & \mathbf{R}_{2,m} \\ \vdots & \vdots & & \vdots \\ \mathbf{R}_{m,1} & \mathbf{R}_{m,2} & \cdots & \mathbf{R}_{m,m} \end{bmatrix}. \qquad (2)$$

Problem to find extremes of correlation function represented either as (1) or as (2) is quite trivial task (Alpatov et al. 2008).

3. Related works

Detailed analysis of existent correlation extremal methods is given in (Baklitskii et al. 1986). For practical application, especially in real-time mode, methods with small computational costs will be considered.

Correlation algorithm in frequency domain uses Fast Fourier Transformation (FFT) to calculate the spectral correlation functions. Its insignificant disadvantage is unreasonable high computational time for small images. Pair correlation algorithm is based on the calculation of number of elements pair which have the same intensity.

The algorithm is simple but its realization is possible only for signals with normal distribution. Phase correlation algorithm uses Inverse Fourier Transform (IFT) of phase component of power cross spectrum. It is insensitive to narrow-band noise, but in case of high frequency noise the quality degrades rapidly. The hierarchical algorithm with informative feature extraction includes two levels.

At the first the current image is compared with the template by the most informative features.

At the second level the full template image is used but the comparison is done only in the points with highest correlation selected at previous level. The main difficulty of realization is the proper extraction of informative feature to have the representative sampling. Also it requires large amount of memory to be used.

Algorithm of amplitude rankings is based on hierarchical algorithm with rankings of algorithms for small images. It provides the decreasing of calculation time only for large images.

Algorithm of linear features overlapping is based on the extraction of structural features of image and their correlation with template image but at the assumption of low level of noise. Invariant moment algorithm uses the integral transform of images. It is the most useful for scaled and rotated images.

4. Phase correlation algorithm for the task of object detection

1. Let's have two images $F_1(i, j)$ and $F_2(i, j)$. To compensate Gibbs phenomenon the images are multiplied by Kaiser function (Oppenheim et al. 1999):

$$w_N = \begin{cases} \dfrac{I_0\left(\pi\alpha\sqrt{1-\left(\dfrac{2N}{M}-1\right)^2}\right)}{I_0(\pi\alpha)}, 0 \le N \le M \\ 0, \qquad \text{otherwise.} \end{cases}$$

where I_0 is the zero[th] order modified Bessel function of the first kind;

α is an arbitrary real number that determines the shape of the window. In the frequency domain, it determines the trade-off between main-lobe width and side lobe level, which is a central decision in window design.

M is an integer, and the length of the sequence is $S = M + 1$.

When S is an odd number, the peak value of the window is $w_{M/2} = 1$. And when S is even, the peak values are $w_{S/2-1} = w_{S/2} < 1$.

2. New images are noted as $F_1^w(i, j)$ and $F_2^w(i, j)$. Let's find the FFT from functions $F_1^w(i, j)$ and $F_2^w(i, j)$:

$$L_1(u,v) = \sum_{(i,j)} F_1(i, j)\exp\left\{-\frac{2\pi}{n}\bar{j}\left(ui+vj\right)\right\},$$

$$L_2(u,v) = \sum_{(i,j)} F_2(i, j)\exp\left\{-\frac{2\pi}{n}\bar{j}\left(ui+vj\right)\right\}.$$

where u, v are spatial frequencies of image.

3. The next step is to find the logarithms of amplitude spectrums. Their notation will be

$$B_1(u,v) = \ln\left|L_1(u,v)\right|,$$

$$B_2(u,v) = \ln\left|L_2(u,v)\right|.$$

The transformation to polar logarithmic coordinates is used for its rotation invariant and scale invariant properties. The scaling and rotation in Cartesian domain corresponds to pure translation in polar logarithmic domain.

The polar coordinates (ρ, θ) correspond to radial distance from the center and angle from the center respectively. Taking logarithm of radial distance ρ, the polar logarithmic coordinates are obtained.

Such transformation is a conformal mapping from the points on the Cartesian plane (x, y) to points in the polar logarithmic plane $(\log(\rho), \theta)$. Considering a polar coordinate system, where ρ is the radial distance from the center of the image (x_c, y_c) and θ denotes the angle. Any point (x, y) can be represented in polar coordinates and is given by

$$(\rho, \theta) = \sqrt{\left(x-x_c\right)^2-\left(y-y_c\right)^2}, \tan^{-1}\frac{\left(y-y_c\right)}{\left(x-x_c\right)}.$$

If the image is scaled by factor α, then the coordinates (x, y) in Cartesian domain will become α_x, α_y. Introduction of logarithms will simplify the procedure, the coordinates in logarithmic domain will be determined as

$$\begin{aligned} &\left(\log\left(\alpha_x\right), \log\left(\alpha_y\right)\right) = \\ &= \left(\left(\log\alpha + \log x\right), \left(\log\alpha + \log y\right)\right). \end{aligned} \tag{3}$$

The reflection of scaling as translation in logarithmic polar domain is clear from (3). The effects of distortions are expressed by logarithmic polar image translation on ρ axis and θ axis, respectively in the logarithmic polar coordinates. However, when the original image is translated by (Δx, Δy), the corresponding coordinates are represented by

$$\begin{aligned} &\rho' = \log\sqrt{\left(e^\rho\cos\theta - \Delta x\right)^2 + \left(e^\rho\sin\theta - \Delta y\right)^2}, \\ &\theta' = \tan^{-1}\frac{e^\rho\sin\theta - \Delta y}{e^\rho\cos\theta - \Delta x}. \end{aligned} \tag{4}$$

The slight translation produces a modification of the logarithmic polar image. Therefore, such image is not suitable for extracting translation parameters of images. That's why Fourier transform was first applied to images and only then the logarithmic polar transformation is used for cross power density to recover scale and rotation by using phase correlation in logarithmic polar space.

Functions (4) are obtained using the bipolar interpolation of image intensity.

4. Calculation of discrete Fourier transform from functions $B_1\left(\rho_n, \theta_m\right), B_2\left(\rho_n, \theta_m\right)$

5. Calculation of power spectral density of functions $B_1(\rho_n, \theta_m)$ and $B_2(\rho_n, \theta_m)$ is done as follows:

$$R_B(u,v) = B_1(u,v)B_2*(u,v),$$

where sign "*" means complex conjugate.

6. To calculate the cross-correlation function it is necessary to use the IFT:

$$p_B(B_1(\rho_n,\theta_m),B_2(\rho_n,\theta_m)) = \Phi^{-1}\{R_B(\rho_n,\theta_m)\}.$$

7. Then it is necessary to find the maximum of cross-correlation function

$$\hat{k} = \arg\max_k \{\rho_B(0,\theta_m)\}.$$

Here $\hat{k}$ is number of sample where there is a maximum of cross-correlation function.

The estimation of angle is done by formula

$$\hat{\varphi} = \hat{k}\Delta_\theta.$$

8. Then the initial image is repeatedly rotated in angle $\hat{\varphi}$. The resulting image has the form

$$F_1^{rot}(i,j) = F_1(i\cos\hat{\varphi} + j\sin\hat{\varphi}, -i\sin\hat{\varphi} + j\cos\hat{\varphi}).$$

9. Then process of finding the correlation function starting from step 1 is repeated for the whole range of angle φ.

10. Search of maximum for the function $p(i,j)$ in the domain of possible values of shift.

The algorithm has been implemented using MATLAB software.

5. Conclusions

Therefore, a rotation angle can be derived by phase correlation based shift estimation in polar coordinates. Similarly, a change in scale can be determined based on a phase shift in the frequency domain presented in logarithmic coordinate units.

It should be pointed out that the proposed robust 2D fitting phase correlation technique has some limitation in the estimation of large shift, rotation and scale change duo to the fringe filtering operation used in the technique.

Generally speaking, the robust phase correlation technique is able to achieve sub-pixel frame registration with rotation angle less 10 degree.

References

Alpatov, B.A.; Babaian, P.V.; Balashov, O.E.; Stepashkin, A.I. 2008. *Methods of automatic detection and guidance of objects. Image processing and control*. Moscow, Radiotekhnika. 176 p. (in Russian).

[*Методы* автоматического обнаружения и сопровождения объектов. Обработка изображений и управление / Б.А. Алпатов, П.В. Бабаян, О.Е. Балашов, А.И. Степашкин. – Москва: Радиотехника, 2008. – 176 с.]

Baklitskii, V.K.; Bochkarev, A.M.; Musiakov, M.P. 1986. *Methods of signal filtering in correlation extremal navigation systems*. Edited by V.K. Baklitskii. Moscow, Radio i sviaz. 216 p. (in Russian).

[*Баклицкий В.К.* Методы фильтрации сигналов в корреляционно-экстремальных системах навигации / В.К. Баклицкий, А.М. Бочкарев, М.П. Мусьяков; под редакцией В.К. Баклицкого. – Москва: Радио и связь, 1986. – 216 с.]

Baklitskii, V.K. 2009. *Correlation extremal methods of navigation and guidance*. Tver, TO “Knizhnyi klub”. 360 p. (in Russian).

[*Баклицкий В.К.* Корреляционно-экстремальные методы навигации и наведения / В.К. Баклицкий. – Тверь: ТО Книжный клуб, 2009. – 360 с.]

Oppenheim, A.V.; Schafer, R.W.; Buck J.R. 1999. *Discrete-time signal processing.* Upper Saddle River, N.J. Prentice Hall.

Pratt, W. 1982. *Digital image processing*. Moscow, Mir. 283 p.] (in Russian).

[*Прэтт У.* Цифровая обработка изображений / У. Прэтт. – Москва: Мир, 1982. – 283 с.]

Received 5 December 2012.

М.П. Мухіна. Кореляційно-екстремальний принцип виявлення об’єктів на земній поверхні

Національний авіаційний університет, проспект Космонавта Комарова, 1, Київ, Україна, 03680
E-mail: m_mukhina@inbox.ru

Розглянуто особливості виявлення об’єктів шляхом обробки зображень земної поверхні. Показано, що подання в матричній формі дозволяє використати методи спектрального аналізу двовимірних сигналів, зокрема використання швидкого Фур’є перетворення. Проаналізовано наявні методи кореляційно-екстремальної обробки зображень. Зазначено, що саме спектральний метод, зокрема метод фазової кореляції, задовольняє вимоги, що пред’являються до навігаційних задач (працює в режимі реального часу, стійких до шумів та ін.). Запропоновано використання фазової кореляції з додатковим логарифмічно-полярним перетворенням, що усуває чутливість методу до зсувів або спотворень зображення. Визначено, що точність методу виявлення об’єкта зберігається навіть при значних поворотах зображень.

Ключові слова: кореляційно-екстремальний принцип, логарифмічно-полярне перетворення, фазова кореляція.

М.П. Мухина. Корреляционно-экстремальный принцип обнаружения объектов на земной поверхности
Национальный авиационный университет, проспект Космонавта Комарова, 1, Киев, Украина, 03680
E-mail: m_mukhina@inbox.ru

Рассмотрены особенности выявления объектов путем обработки изображений земной поверхности. Показано, что представление изображений в матричной форме позволяет использовать методы спектрального анализа двухмерных сигналов, в частности использование быстрого Фурье преобразования. Проанализированы существующие методы корреляционно-экстремальной обработки изображений. Отмечено, что именно спектральный метод, в частности метод фазовой корреляции, удовлетворяет требованиям, которые предъявляются к навигационным задачам (работает в режиме реального времени, устойчив к шумам и т.д.). Предложено использование фазовой корреляции с дополнительным логарифмически-полярным преобразованием, которое устраняет чувствительность метода к сдвигам или искажениям изображения. Установлено, что точность метода сохраняется даже при значительных поворотах изображения.

Ключевые слова: корреляционно-экстремальный принцип, логарифмически-полярное преобразование, фазовая корреляция.

Mukhina Maryna (1979). Candidate of Engineering. Associate Professor.
Aviation Computer-Integrated Complexes Department, National Aviation University, Kyiv, Ukraine.
Education: National Aviation University, Kyiv, Ukraine (2002).
Research area: image processing, navigation and control.
Publications: 22.
E-mail: m_mukhina@inbox.ru