Image and Vision Computing 27 (2009) 313–318

Contents lists available at ScienceDirect

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

Review article

Aircraft recognition in infrared image using wavelet moment invariants

Zhang Feng [a,*], Liu Shang-qian [a], Wang Da-bao [a], Guan Wei [b]

[a] School of Technical Physics, Xidian University, Taibai Road 2#, Xi'an 710071, China
[b] ShuangLiu Airport Swatmb CNS Division, Chengdu 610202, China

ARTICLE INFO

Article history:
Received 2 July 2007
Received in revised form 29 May 2008
Accepted 18 August 2008

Keywords:
Wavelet moment invariants
Image recognition
Moment invariants
Rotation invariant

ABSTRACT

Automatic Target Recognition (ATR) of infrared object has been taking a great interest to the researchers in recent years. ATR requires invariance of high cognition accuracy in translation, scaling and orientation, but classification of two-dimensional (2D) shapes despite of their position, size and orientation in infrared image remains a difficult problem. In this paper, a feature extraction method is proposed using Wavelet Moment Invariants (WMI). The very similar objects can be classified correctly by virtue of the wavelet moment with its multi-resolution properties. Compared with some other geometry moments, the classification rate and the recognition efficiency are improved with wavelet moments. As different wavelet basis will have different impacts to wavelet moment, it affects the efficiency of classification. Some important properties such as orthonomality, supported length and vanishing moments which affect the performance of wavelet moment are discussed in this paper. Through experimental analysis, a conclusion is obtained that symmetry, compactly supported wavelet has more high-performance, and using wavelet function with proper vanishing moments could effectively improve the efficiency of classification.

1. Introduction

ATR in infrared image has been taking a great interest to the researchers in recent years. Since there are many applications of ATR, the classification rate and the recognition efficiency become critical factors to these applications. As some aircrafts are much alike, it is very hard for the computer to classify them correctly. Moment based invariants, which provide translation, scale and rotation invariance of 2D shapes, have been widely used over recent years as the feature vectors for image analysis in many areas, as far as aircraft recognition area is concerned, it uses moment invariant features of the aircraft silhouette and silhouette border. In 1961 Hu proposed the concept of moment invariant [1], Li derived a method for constructing an arbitrary order invariant using Fourier–Mellin transform and pointed out that Hu-moment is a special example of it [2]. Teague suggested orthogonal moments be constructed by mean of using orthogonal polynomial to overcome the shortcomings of Hu-moment which includes a large number of redundant information, and Zernike-moment invariant [3] is one of the orthogonal moment invariant [4]. But such moment invariants mentioned above are calculated over the whole image space, as a result the feature vectors abstracted are global features which is not conducive to classification. In order to overcome these disadvantages, Shen and Horace put forward a proposal that using wavelet transform to construct wavelet moment so that the spatially and local features could be obtained simultaneously [5], and higher classification rate would be received in classifying similar shapes with slight difference.

Previous works have been done using moment invariant by many researchers for aircraft recognition. Dudani and Breeding utilize moment invariants to identify aircraft [6], and some researchers use curve contour for abstracting features [7,8], some research is based on shape matching of 3-D objects [9,10]. However it does not have a multi-resolution representation. We investigate the performance of wavelet moment invariants for classifying different aircrafts.

In this study, wavelet moment with its multi-resolution properties has been used for feature extraction of the aircraft in infrared image. We use radial wavelet transform to the image, and a set of discriminative wavelet moment is obtained. Also the Mallat pyramid algorithm [11] is used to accelerate the whole process.

2. Mathematical derivation of moment invariant

Let $f(x,y)$ represent the density distribution function of a 2D image in Cartesian coordinates, and the geometric moment of $(p+q)$ order of it is defined as:

$$M_{pq} = \iint f(x,y)x^p y^q \,\mathrm{d}x\mathrm{d}y \tag{1}$$

where $p, q = 0, 1, 2\ldots$ and $f(x,y)$ is the gray value of the image at x and y location, and the central moment is as follows:

* Corresponding author. Tel.: +86 1320153997.
E-mail addresses: cdmavsgprs@hotmail.com, goodfengzi@163.com (F. Zhang).

0262-8856/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.imavis.2008.08.007

$$\mu_{pq} = \iint f(x - x_0, y - y_0)(x - x_0)^p (y - y_0)^q \, \mathrm{d}x\mathrm{d}y \tag{2}$$

where $x_0 = M_{10}/M_{00}$ and $y_0 = M_{01}/M_{00}$ are the gravity center of the image. The complex moment is given by:

$$F_{pq} = \iint (x + iy)^p (x - iy)^q f(x, y) \, \mathrm{d}x\mathrm{d}y \tag{3}$$

Let $x = r\cos\theta$, $y = r\sin\theta$ and substitute them into (3), we get

$$\begin{aligned} F_{pq} &= \iint (x + iy)^p (x - iy)^q f(x, y) \, \mathrm{d}x\mathrm{d}y \\ &\iint r^{p+q} \, \mathrm{e}^{\mathrm{i}(p-q)\theta} f(r, \theta) r \, \mathrm{d}r\mathrm{d}\theta \\ &\iint (r\cos\theta + tr\sin\theta)^p (r\cos\theta - tr\sin\theta)^q f(r, \theta) r \, \mathrm{d}r\mathrm{d}\theta \end{aligned} \tag{4}$$

Let $p + q = m$, $q - p = n$, (4) can be written as:

$$F_{mn} = \iint r^m \, \mathrm{e}^{-\mathrm{i}n\theta} f(r, \theta) r \, \mathrm{d}r\mathrm{d}\theta$$

Of cause it can be rewritten as:

$$F_{pq} = \iint r^p \, \mathrm{e}^{-\mathrm{i}q\theta} f(r, \theta) r \, \mathrm{d}r\mathrm{d}\theta$$

By substituting r^p with $g_p(r)$ into above equation, we obtain the general expression of moments:

$$F_{pq} = \iint g_p(r) \, \mathrm{e}^{-\mathrm{i}q\theta} f(r, \theta) r \, \mathrm{d}r\mathrm{d}\theta \tag{5}$$

where $g_p(r)$ represents the radial part and $\mathrm{e}^{-iq\theta}$ represents the angular part of the kernel function. Hu-moment, Li-moment and Zernike-moment can be achieved by using different $g_p(r)$ and $\mathrm{e}^{-iq\theta}$.

In order to normalize the image, we move the origin to the mass center through $f_T(x,y) = f(x - x_0,\ y - y_0)$, therefore $f_T(x,y)$ have the quality of translation invariant. Define scaling factor

$$\alpha = \sqrt{\frac{M_{00}}{\text{area}}}$$

where area is a constant which equals to the expected image size. So we get the image which has the quality of translation and scaling invariant by $f_N(x,y) = f((x - x_0)/\alpha,\ (y - y_0)/\alpha)$. After that, we transform $f_N(x,y)$ from Cartesian coordinate to polar coordinates, expressed as $f_N(r,\theta)$.

Suppose $f'_N(r, \theta)$ represent $f_N(r,\theta)$ rotated by ∂ degree and is defined as:

$$f'_N(r, \theta) = f_N(r, \theta + \partial)$$

By substituting above equation into (5), we have

$$\begin{aligned} F'_{pq} &= \iint f'_N(r, \theta) g(r) \, \mathrm{e}^{-\mathrm{i}q\theta} r \, \mathrm{d}r\mathrm{d}\theta = \iint f_N(r, \theta + \partial) g(r) \, \mathrm{e}^{-\mathrm{i}q\theta} r \, \mathrm{d}r\mathrm{d}\theta \\ &= \mathrm{e}^{\mathrm{i}q\theta} \iint f_N(r, \theta_t) g_p(r) \mathrm{e}^{-\mathrm{i}q\theta_t} r \, \mathrm{d}r\mathrm{d}\theta_t = \mathrm{e}^{\mathrm{i}q\theta} F_{pq} \end{aligned}$$

Eliminating $\mathrm{e}^{iq\partial}$ using some mathematical operation, moments of translation, scaling and orientation invariant could be obtained. For example, it is obviously that $\|F_{pq}\|$ is an invariant, the combined moments are also rotation invariant such as $F_{p_1q}F^*_{p_2q}$ where $F^*_{p_2q}$ is the conjugate of F_{p_2q} [12].

3. Translation, scaling and orientation normalization

In traditional way, the scaling factor of the present aircraft size compared with the expected size is as follows:

$$\alpha = \sqrt{\frac{M_{00}}{\text{area}}}$$

where area is a constant which represents the expected size. In order to facilitate normalizing in polar coordinates, here we normalize the image by utilizing the distance to the mass center (x_0,y_0).The scaling factor α is defined as:

$$\alpha = \frac{\max\sqrt{(x - x_0)^2 + (y - y_0)^2}}{N} \tag{6}$$

where N is the normalized radius of the image with its value we set 64. So we obtain the expected image though changing the coordinate and size of the present image as follows:

$$\begin{cases} X = (x - x_0)/\alpha \\ Y = (y - y_0)/\alpha \end{cases} \tag{7}$$

we use bilinear interpolation to complete the above process. Fig. 1 shows the original image and the normalized image.

After that, let $f(\theta,r)$ represents the corresponding form in polar coordinates of $f(x,y)$, and is defined as:

$$\begin{cases} x = r\cos(\theta) \\ y = r\sin(\theta) \end{cases} \tag{8}$$

where

$$\begin{cases} r = \sqrt{(x - x_0)^2 + (y - y_0)^2} \\ \theta = \arctan\frac{(y - y_0)}{(x - x_0)} \end{cases}$$

$$f(r\cos\theta, r\sin\theta) = f(x, y) \qquad 0 \leqslant r \leqslant R, \quad 0 \leqslant \theta \leqslant 2\pi$$

$$R = \max\left(\sqrt{(x - x_0)^2 + (y - y_0)^2}\right)$$

In our operation, we take (x_0,y_0) as the center with $(r_i = \frac{i \times R}{N} \quad i = 0, 1, \ldots, N - 1)$ as radiuses and draw concentric circles, and take horizontal direction as origination with $2\pi/M$ as step and go on angle splitting. Then discrete grids (r_i,θ_i) in polar coordinates have obtained, where $(\theta_i = \frac{j \times 2\pi}{M} \quad j = 0, 1, \ldots, M - 1)$, M and N are radial and directional discrimination indices, respectively. After that, we calculate the average gray value of the area $S_{ij} = \{(r_i,\theta_j)|r_i \leqslant r < r_{i+1},\ \theta_j \leqslant \theta < \theta_{j+1}\}$ in which the corresponding pixels in Cartesian coordinates are, and give this value to the $f(r_i,\theta_j)$.

For digital image, it is somewhat difficult to transform the image from Cartesian coordinates to polar coordinates, for the reason that the size of single pixel in both coordinates is $\Delta x\Delta y$ and $r_i\Delta r\Delta\theta$, respectively, as is show in Fig. 2. To keep the accuracy of the transformation, we have an expansion as large as 64 times of the original image, which equals that we divide each pixel into 64 sub-pixels. Every sub-pixel can be mapped to the only grid in polar coordinates. This ensures us that we can get a sufficient accuracy.

After that, the normalized image in polar coordinates has been obtained. Fig. 3 shows the normalized images. Image (a) is the image which image (c) rotates right by 90°. It is obviously that (b) and (d) do have the quality of being periodic.

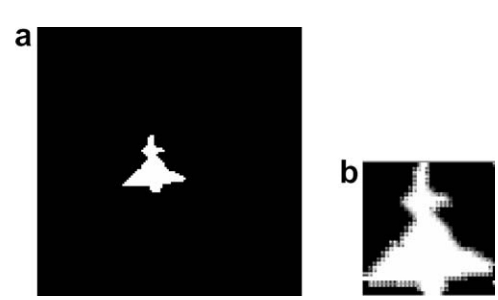

Fig. 1. Images before and after translation, scaling normalization, (a) original image, (b) normalized image.