

## 基于 EMD 细化四元数谱的纹理分割

崔峰, 沈滨, 彭思龙

(中国科学院自动化研究所, 北京 100080)

(fengcui@mail.ia.ac.cn)

**摘要:**提出一种基于 EMD 细化四元数谱的无监督纹理分割方法。创新之处在于:首先采用二维可分离 EMD 分解提高各图像成分的四元数 Fourier 频域特征的分离度;再使用四元数 Gabor 滤波器组进行解析特征提取;最后利用新的模糊 C 均值聚类评价函数实现无监督的纹理分割。实验结果证明了该方法的有效性。

**关键词:**纹理分割;四元数;EMD 分解;Gabor 滤波器;时频分析

**中图分类号:**TP391.41 **文献标识码:**A

## Texture segmentation using EMD refined quaternionic spectrum

CUI Feng, SHEN Bin, PENG Si-long

(Institute of Automation, Chinese Academy of Sciences, Beijing 100080, China)

**Abstract:** This paper presented an unsupervised texture segmentation algorithm utilizing EMD (Empirical Mode Decomposition) refined quaternionic spectrum. The innovations include: first, 2-D separable EMD algorithm was utilized to enhance the separability of features in frequency-domain; second, a quaternionic Gabor filter bank was designed to extract analytic features; at last, a novel validity function of FCM (Fuzzy C-Means Clustering) was introduced to realize unsupervised texture segmentation. Experiment results indicate that this method explores a reasonable way to use EMD and quaternionic spectrum analysis in image processing.

**Key words:** texture segmentation; quaternion; EMD (Empirical Mode Decomposition); Gabor filter; time-frequency analysis

### 0 引言

纹理分割<sup>[1]</sup>是机器视觉和数字图像处理领域的重要课题,其目标是判断相邻纹理区域的异同和不同纹理区域的边界。纹理分割方法按纹理建模的不同可以分为四类:基于统计的方法、基于几何结构的方法、基于模型的方法和基于信号处理的方法。心理学研究证实人脑对图像的感知具有空频局部性的特点,大部分纹理在频域上具有局部窄带特性,所以基于信号处理的空频分析方法是有效的纹理处理方法<sup>[2-3]</sup>。经验模式分解 (Empirical Mode Decomposition, EMD)<sup>[4]</sup>是一种新颖的时频分析工具。EMD 分解完全由数据驱动,可以自适应地通过筛过程 (sifting) 将信号分解为局部窄带的各个内蕴模式函数 (Intrinsic Mode Function, IMF)。EMD 分解对于非线性和非平稳信号的处理有独特的优势,已经在一维信号处理中得到了广泛的应用。虽然纹理信号符合非线性和非平稳信号的特点,但是现有的基于径向基函数插值的二维 EMD 分解方法<sup>[5]</sup>由于计算量大等原因未能用于解决其应用问题。

超复合数解析信号是由 Thomas Bulow 等人提出的高维解析信号定义<sup>[6,7]</sup>,四元数解析信号即为二维的超复合数解析信号。四元数解析信号克服了以往各种二维解析信号定义的不足,为图像处理的空频分析提供了新工具。Thomas Bulow 定义了四元数解析信号的瞬时相位与幅值,其中瞬时相位的  $\psi$  组份值与原二维实信号的局部结构相关,可以作为信号新的解析特征。虽然他们设计了四元数 Gabor 滤波器以提取信号的解析特征,但由于实际信号的四元数 Fourier 频谱分离度不高等原因尚未体现出显著效果。

本文提出一种应用纹理图像的四元数解析信号特征进行纹理分割的方法。该算法针对四元数 Fourier 频谱的特点,首先将纹理

图像进行二维可分离的 EMD 分解,将第一层 IMF 作为分析信号;然后设计了四元数 Gabor 滤波器组,提取信号对应于各个滤波器的瞬时幅值响应和四元数瞬时相位响应  $\psi$  组份值作为信号每个像素的解析特征;最后在新的聚类评价函数的基础上对信号的解析特征进行无监督的模糊 C 均值聚类 (Fuzzy C-Means Clustering, FCM) 得到分割结果。实验数据验证了本文方法的有效性。

### 1 EMD 分解与二维可分离 EMD 分解

#### 1.1 EMD 分解简介

EMD 将信号分解为多个局部窄带的 IMF 和残差趋势项的和,具体算法见文献[4]:

$$f(x) = \sum_{i=1}^n \text{imf}_i(x) + r_n(x) \quad (1)$$

其中  $f(x)$  代表信号,  $n$  代表分解层数,  $\text{imf}_i(x)$  表示第  $i$  层 IMF,  $r_n(x)$  表示残差项。IMF 之间是近似正交的,代表了数据局部的各层频率成分。IMF 主要有两个特征:1) 在整个数据集中,各个 IMF 的极值点数目等于过零点的数目,或者最多相差一个;2) IMF 的各点在由局部极大值点定义的上包络与局部极小值点定义的下包络上的对应值的和接近于零。

#### 1.2 二维可分离 EMD 分解

我们将一维 EMD 分解的方法引申到二维上形成二维可分离的 EMD 分解。该方法利用简单快速的二维可分离的筛过程在水平和竖直两个正交方向上提取图像的二维包络,进而得到了二维信号的多层 IMF。记图像信号为  $f(x_1, x_2)$ ,  $x_1 = 1, \dots, X_1$ ,  $x_2 = 1, \dots, X_2$ , 则本文的二维可分离 EMD 分解算法简述如下:

1) 初始化:  $r_0(x_1, x_2) = f(x_1, x_2)$ ,  $j = 1$ 。

2) 筛选抽取第  $j$  个 IMF:

a) 初始化:  $h_0(x_1, x_2) = r_{j-1}(x_1, x_2)$ ,  $i = 1$ 。

b) 提取  $h_{i-1}(x_1, x_2)$  的上、下包络和包络均值:

① 分别提取  $h_{i-1}(x_1, x_2)$  每行的局部极大值序列和局部极小值序列, 利用三次样条插值算法对  $h_{i-1}(x_1, x_2)$  每行的局部极大值序列插值, 将插值结果排成上包络中间值矩阵  $h_{mid\_upper}(x_1, x_2)$ , 同理利用  $h_{i-1}(x_1, x_2)$  每行的局部极小值序列通过插值得到下包络中间值矩阵  $h_{mid\_lower}(x_1, x_2)$ ;

② 求取上、下包络中间值的平均:

$$m_{mid}(x_1, x_2) = (h_{mid\_upper}(x_1, x_2) + h_{mid\_lower}(x_1, x_2))/2;$$

③ 分别提取  $m_{mid}(x_1, x_2)$  每列的局部极大值序列和局部极小值序列, 利用三次样条插值算法对  $m_{mid}(x_1, x_2)$  每列的局部极大值序列插值, 将插值结果排成上包络矩阵  $h_{upper}(x_1, x_2)$ , 同理利用  $m_{i-1}(x_1, x_2)$  每列的局部极小值序列通过插值得到信号的下包络矩阵  $h_{lower}(x_1, x_2)$ ;

④ 求取上、下包络的平均:  $m_{i-1}(x_1, x_2) = (h_{upper}(x_1, x_2) + h_{lower}(x_1, x_2))/2$ 。

c)  $h_i(x_1, x_2) = h_{i-1}(x_1, x_2) - m_{i-1}(x_1, x_2)$ 。

d) 计算终止条件  $SD$ : 由于二维信号的长度一般小于一维信号, 在处理二维 EMD 分解时必须考虑减少边界效应的影响。取  $m_{i-1}(x_1, x_2)$  的中间部分  $(\frac{X_1}{2} \times \frac{X_2}{2})$ , 设该部分所含元素的绝对值的

最大值为  $Max_{mid}$ , 同理求得  $f(x_1, x_2)$  的中间部分  $(\frac{X_1}{2} \times \frac{X_2}{2})$  所含元素的绝对值的最大值  $Max_{uma}$ , 计算终止条件:

$$SD = \frac{Max_{mid}}{Max_{uma}} \quad (2)$$

e) 若  $SD$  小于给定的门限值  $\varepsilon$  (例如 0.03), 则有  $imf_j(x_1, x_2) = h_{i-1}(x_1, x_2)$ 。否则令  $i = i + 1$ , 转到 b) 步。

3)  $r_j(x_1, x_2) = r_{j-1}(x_1, x_2) - imf_j(x_1, x_2)$ 。

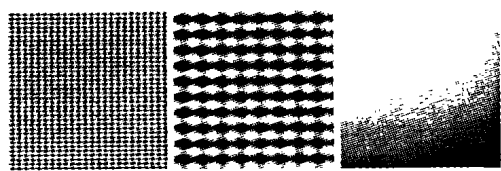
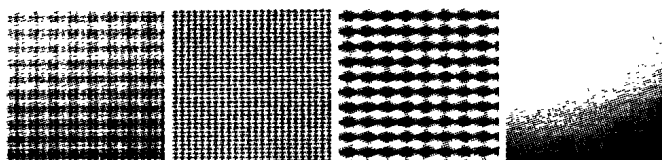
4) 若  $r_j(x_1, x_2)$  的中间部分仍有不少于两个的极值点或者分解所得的 IMF 数目未达到要求, 则令  $j = j + 1$ , 转到第 2) 步。

5) 最后得到二维信号的 EMD 分解表达:

$$f(x_1, x_2) = \sum_{j=1}^n imf_j(x_1, x_2) + r_n(x_1, x_2) \quad (3)$$

我们将该算法应用于图 1(a0) 由不同频率的余弦组份构成的合成纹理图像(200×200):

$$f(x_1, x_2) = f_{high}(x_1, x_2) + f_{middle}(x_1, x_2) + f_{low}(x_1, x_2) \quad (4)$$

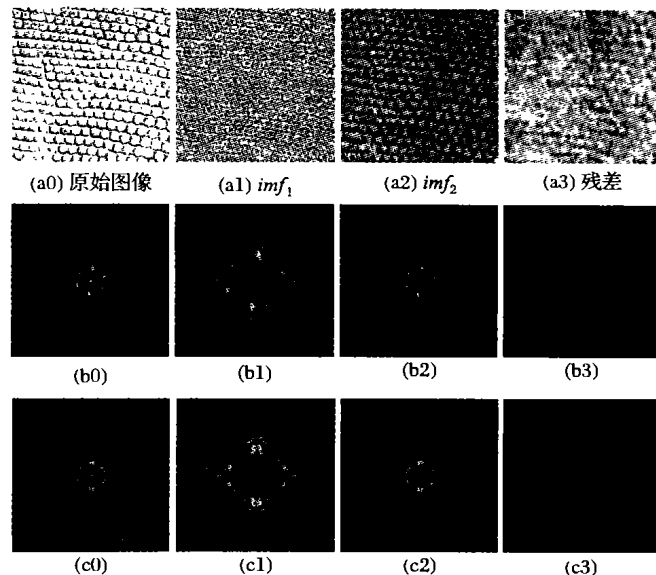


(a0~a3: 合成纹理与不同频率组份; b1~b3: EMD 分解结果)

图 1 合成纹理与 EMD 分解结果

如图 1 所示, 原信号的三层频率成分被提取出来, 说明了本文二维可分离 EMD 分解的有效性。对于纹理图像而言, 一般分解出 2~3 个 IMF 就可以提供足够的特征信息。图 2 是对 Brodatz 标准

纹理库<sup>[8]</sup>中的 D03 纹理子图(200×200) EMD 分解两层的结果及对应的 Fourier 频谱图与四元数 Fourier 频谱图, 纹理图像在处理前已先减去灰度均值以去除直流分量。可以看到,  $imf_1$  蕴含了纹理的局部高频频率特征和主要结构信息,  $imf_2$  蕴含了纹理的局部中频频率特征和较粗尺度结构信息, 残差项蕴含了纹理的局部低频信息和整体趋势。我们将看到 EMD 分解对于信号局部频带信息的自适应划分对于后续的纹理分割提供了便利。



(a0~a3: 原始图像和 EMD 分解结果; b0~b3: a0~a3 对应的 Fourier 频谱图; c0~c3: a0~a3 对应的四元数 Fourier 频谱图)

图 2 D03 的 EMD 分解结果及 Fourier 频谱图与四元数 Fourier 频谱图

## 2 四元数谱信号表达与四元数 Gabor 滤波器

### 2.1 四元数代数简介

我们称数集  $H$  为四元数集, 若其满足:

$$H = \{a + ib + jc + kd \mid a, b, c, d \in \mathbb{R}\} \quad (5)$$

其中  $i, j, k$  为虚数单位。四元数三个虚数单位的乘法运算法则定义如下:

$$\begin{aligned} i^2 = j^2 = k^2 &= -1, \quad ij = -ji = k, \\ jk &= -kj = i, \quad ki = -ik = j \end{aligned} \quad (6)$$

可见四元数的乘法不符合交换律。四元数  $q = a + ib + jc + kd$  的共轭  $\bar{q}$ 、模  $|q|$  和倒数  $q^{-1}$  的定义分别为:

$$\begin{aligned} \bar{q} &= a - ib - jc - kd \\ |q| &= \sqrt{q\bar{q}} = \sqrt{a^2 + b^2 + c^2 + d^2} \\ q^{-1} &= \bar{q}/|q|^2 \end{aligned} \quad (7)$$

四元数域  $H$  中有三个卷乘(involution)算子:

$$\alpha_i(q) = -iqi, \quad \alpha_j(q) = -jqj, \quad \alpha_k(q) = -kqk \quad (8)$$

由这些卷乘算子可以定义四元数域的 Hermit 对称。一个定义在  $\mathbb{R}^2 \rightarrow H$  的函数  $f(x_1, x_2)$  是四元数 Hermit 对称的, 如果  $f(x_1, x_2)$  满足:

$$\begin{aligned} f(x_1, -x_2) &= \alpha_i(f(x_1, x_2)) \\ f(-x_1, x_2) &= \alpha_j(f(x_1, x_2)) \\ f(-x_1, -x_2) &= \alpha_k(f(x_1, x_2)) \end{aligned} \quad (9)$$

Thomas Bülow 等人根据三维空间转动与单位四元数的对应关系定义了新的四元数表达形式<sup>[6-7]</sup>:

$$q = |q| e^{i\phi} e^{j\psi} e^{k\theta}, \quad (\phi, \psi, \theta) \in [-\pi, \pi] \times [-\pi/2, \pi/2] \times [-\pi/4, \pi/4] \quad (10)$$

其中指数函数的 Euler 方程成立,  $|q|$  为四元数  $q$  的模,  $(\phi, \psi, \theta)$  为  $q$  的相位矢量。对于下面将要介绍的四元数解析

信号来说,  $|q|$  即为四元数解析信号的瞬时幅值,  $(\phi, \psi, \theta)$  为瞬时相位。

四元数  $q$  的相位  $\psi$  组份值的算法如下, 其他两个相位组份值的求法详见文献[6~7]:

1) 单位化四元数  $q$ :  $\bar{q} = q/|q| = a + ib + jc + kd$ ;

2) 求取四元数  $q$  的相位  $\psi$  组份值:

$$\psi = -\arcsin(2(bc - ad))/2 \quad (11)$$

## 2.2 四元数解析信号

二维实函数  $f(x_1, x_2): \mathbf{R}^2 \rightarrow \mathbf{R}$  的四元数 Fourier 变换 (QFT):  $F^q(u_1, u_2): \mathbf{R}^2 \rightarrow \mathbf{H}$  和四元数 Fourier 反变换 (IQFT) 分别定义为<sup>[6~7]</sup>:

$$F^q(u_1, u_2) = \int_{\mathbf{R}} \int_{\mathbf{R}} e^{-i2\pi u_1 x_1} f(x_1, x_2) e^{-i2\pi u_2 x_2} dx_1 dx_2 \quad (12)$$

$$f(x_1, x_2) = \int_{\mathbf{R}} \int_{\mathbf{R}} e^{-i2\pi u_1 x_1} F^q(u_1, u_2) e^{-i2\pi u_2 x_2} du_1 du_2 \quad (13)$$

其中  $i, j$  不可交换, 指数函数的 Euler 方程成立, 虚单位  $k$  在乘积中产生,  $u_1, u_2$  为四元数 Fourier 变换频域的频率单位。QFT 和 IQFT 均可通过分步的 Fourier 变换与反 Fourier 变换计算。

Thomas Bülöw 等人证明了实函数  $f(x_1, x_2)$  的四元数 Fourier 变换  $F^q(u_1, u_2)$  满足四元数 Hermit 对称的定义, 因此  $f(x_1, x_2)$  的所有 QFT 信息可以由单象限的 QFT 信息得出。由此得出  $f(x_1, x_2)$  的四元数解析信号  $f_{\text{an}}^q(x_1, x_2): \mathbf{R}^2 \rightarrow \mathbf{H}$  的定义。 $f_{\text{an}}^q(x_1, x_2)$  可以通过其四元数 Fourier 变换  $F_{\text{an}}^q(u_1, u_2)$  的 IQFT 求出:

$$F_{\text{an}}^q(u_1, u_2) = F^q(x_1, x_2) (1 + \text{sign}(u_1))(1 + \text{sign}(u_2)) \quad (14)$$

$$\text{其中 } \text{sign}(u) = \begin{cases} 1, & \text{if } u > 0 \\ 0, & \text{if } u = 0 \\ -1, & \text{if } u < 0 \end{cases}, F_{\text{an}}^q(u_1, u_2): \mathbf{R}^2 \rightarrow \mathbf{H},$$

$F_{\text{an}}^q(x_1, x_2)$  为  $F_{\text{an}}^q(u_1, u_2)$  的 QFT。

四元数解析信号较以往的二维解析信号的不同定义更加符合一维解析信号的定义在二维上的延伸<sup>[6~7]</sup>, 以下是其突出的优点: 1) 四元数解析信号的四元数 Fourier 变换频谱仅在一个象限内, 而且可以由四元数解析信号得到原信号的所有信息; 2) 四元数解析信号的实部就是原二维实函数。

对于二维实信号的四元数解析信号, 其瞬时相位的  $\psi$  组份值与原信号的局部二维结构相关, 可作为表征信号局部结构的新解析特征<sup>[6~7]</sup>。

## 2.3 四元数 Gabor 滤波器

针对纹理图像的特点, 结合 Gabor 滤波器良好的空频分辨率和选择特性, Thomas Bülöw 提出了四元数域的 Gabor 滤波器  $h^q(x_1, x_2)$ <sup>[6~7]</sup>:

$$h^q(x_1, x_2) = g(x'_1, x'_2) \exp(i2\pi U_1 x_1) \exp(i2\pi U_2 x_2) \quad (15)$$

$$(x'_1, x'_2) = x_1 \cos \alpha + x_2 \sin \alpha, -x_1 \sin \alpha + x_2 \cos \alpha \quad (16)$$

$$g(x_1, x_2) = \frac{\varepsilon}{2\pi\sigma^2} \exp\left(-\frac{2x_1^2 + (\varepsilon x_2)^2}{2\sigma^2}\right) \quad (17)$$

其中  $\varepsilon$  是二维 Gaussian 函数的高宽比,  $\sigma$  是尺度参数, 指数函数的 Euler 方程成立, 虚单位  $k$  在乘积中产生,  $\alpha$  为滤波器主轴沿  $x_1$  坐标轴方向逆时针转动的角度,  $(U_1, U_2)$  是滤波器的中心频率, 单位为  $\frac{1}{\text{图像宽度}}$ 。中心频率的方向角  $\beta = \arctan(U_2/U_1)$ 。 $h^q(x_1, x_2)$  的 QFT 形式为:

$$H^q(u_1, u_2) = \exp(-2\pi^2\sigma^2[(u_1' - U_1')^2 +$$

$$(u_2' - U_2')^2/\varepsilon^2]) \quad (18)$$

其中  $(u_1', u_2')$ ,  $(U_1', U_2')$  为  $(u_1, u_2)$ ,  $(U_1, U_2)$  与 (16) 式相似的转动形式, 转动角度也为  $\alpha$ 。可见四元数 Gabor 滤波器在四元数 Fourier 频域的形式也是由参数调整尺度、位置与形状的 Gaussian 函数。

二维实信号经四元数 Gabor 滤波器滤波的结果使得每一点成为由一个实部与三个虚部组成的四元数, 滤波结果可以看作是原信号相应的频带有限的四元数解析信号。

## 3 纹理分割算法描述

### 3.1 基本算法

四元数解析信号在纹理分割等实际应用中尚未获得满意的效果, 这是因为: 首先, 由多种纹理组成的纹理图像, 其频谱的成分往往较复杂, 无法仅通过选取几个调整过的 Gabor 滤波器得到满意的结果; 其次, 由于实信号四元数 Fourier 频谱的四元数 Hermit 对称特性, 仅在单象限频谱内设计四元数 Gabor 滤波器会遇到不同纹理的频谱信息混杂干扰问题。

由于二维 EMD 分解具有将图像分解为局部窄带信号的能力, 所以我们引入二维 EMD 分解以提高待分析信号的四元数 Fourier 频谱解析度。在实验中发现纹理图像的第一层 IMF 非常符合分割的要求, 因为第一层 IMF 保留了纹理的主要结构信息, 这些结构信息与第一层 IMF 的四元数解析信号的瞬时相位  $\psi$  组份值紧密相关。并且第一层 IMF 的频谱主要分布在中高频段, 分离度也比原图像的大, 克服了一些频谱中不同纹理的干扰因素; 此外, 我们设计了单象限的四元数 Gabor 滤波器组, 以完成对信号单象限四元数 Fourier 频谱的近似正交分解。将第一层 IMF 每个像素经各个滤波器滤波的结果所对应的瞬时幅值响应和瞬时相位响应  $\psi$  组份值排列起来作为该像素的解析特征向量。因为相邻像素属于同类纹理的概率比较大, 所以把每个像素的两个坐标值作为附加特征补充到特征向量中; 纹理类别数目用新的 FCM 聚类评价函数确定, 通过 FCM 实现纹理分割; 最后用一种简单的形态学腐蚀膨胀算法对分割结果进行后处理<sup>[9]</sup>。

### 3.2 四元数 Gabor 滤波器组

我们在四元数 Fourier 频谱的第一象限内设计近似正交的四元数 Gabor 滤波器组, 利用 2.3 节叙述的信息, 为了计算方便取  $\varepsilon = 1, \alpha = \beta$ 。当 Gabor 滤波器中心频率在  $u_1$  轴上时, Gabor 滤波器的半峰带宽定义为<sup>[2]</sup>:

$$B_r = \log_2 \left( \frac{U_1 + (2\ln 2)^{1/2} \sigma}{U_1 - (2\ln 2)^{1/2} \sigma} \right) \quad (19)$$

心理学方面的实验证明, 人类大脑视觉皮层对一定频率信号的频响带宽为一个倍频程, 即当  $U_1$  一定时, 可以取  $B_r = 1$ , 由此可以求出  $\sigma$ :

$$\sigma = \frac{U_1}{3(2\ln 2)^{1/2}} \quad (20)$$

余下的问题是如何确定在  $u_1$  轴上的滤波器中心频率  $U_1$  和滤波器沿频域原点旋转后的中心频率方向角  $\beta$  以确定其他滤波器的参数。

为了使最高频率滤波器的带宽范围覆盖在图像范围内并消除相邻滤波器的混叠影响, 我们取两个相邻滤波器的中心频率间距为一个倍频程。设图像宽度为  $N$ , 根据四元数 Fourier 频谱的特点取  $0^\circ, 45^\circ, 90^\circ$  三个中心频率方向角, 每个方向取七个尺度的滤波器。其中中心频率方向角为  $0^\circ$  的 7 个滤波器在  $u_1$  轴上的对应中心频率为:  $\{N\sqrt{2}/4, N\sqrt{2}/8, N\sqrt{2}/16, \dots, N$

$\sqrt{2}/256$  (单位为  $N^{-1}$ ), 将它们的中心频率沿频域原点旋转  $45^\circ$  和  $90^\circ$  后就可得到其他的 14 个滤波器。这 21 个滤波器就组成了四元数 Gabor 滤波器组。实验中较低频率的滤波器由于空间尺度太大对纹理信息的提取作用不明显, 多数情况下取较高频率的 12 个滤波器就可以得到满意的结果。

需要注意的是, 经滤波器滤波的初始特征不能直接用于聚类分割, 因为当某一纹理的频率波动范围超过任何一个 Gabor 滤波器的频率覆盖范围时滤波的结果就不太理想。因此我们把信号经每个滤波器滤波的初始特征经对应的同形状但尺度更大的 Gaussian 滤波器平滑处理后作为用于聚类的解析特征<sup>[3]</sup>。下式为 Gaussian 平滑滤波器的形式, 每个平滑滤波器对应的  $\sigma$  由 (20) 式得出, 一般取  $\gamma = 2/3$ :

$$g(\gamma x_1, \gamma x_2) = \frac{\varepsilon}{2\pi\sigma^2} \exp\left(-\frac{\gamma x_1^2 + (\varepsilon\gamma x_2)}{2\sigma^2}\right) \quad (21)$$

每个像素对应各个滤波器的经平滑处理后的瞬时幅值响应与瞬时相位响应  $\psi$  组份值及两个空间坐标值就组成了下述聚类算法中每个像素即每个样本的特征向量。

### 3.3 纹理类别数的确定

确定纹理类别数目在实际应用中十分重要。针对聚类算法的特点已有一些聚类有效性的判别方法<sup>[10]</sup>, 但在实验中经常会引起过分类的结果。FCM 计算代价低而且可以同时得到每个样本隶属于各个类别的置信度, 所以我们结合 FCM 的性质引入新的 FCM 聚类评价函数来确定样本类别数目。

因为分类的目的是使所分的类别具有类内距离小、类间距离大的特点, 所以好的聚类评价函数不仅要考虑各个类别内部的性质, 还要顾及到各个类别之间的差别, 另外可以充分利用每个样本的类别置信度信息。借鉴 Hausdorff 距离测度的思想, 我们采用类间最小距离和类内最大距离的比值作为 FCM 的聚类评价函数值  $Valid_{fun}$ :

$$Valid_{fun} = \frac{\min_{i \neq j, i, j=1, \dots, c} \{dist(S_i, S_j)\}}{\max_{k=1, \dots, c} \{dist(S_k)\}} \quad (22)$$

其中  $c$  为聚类类别数目,  $S_i$  为各个聚类集合。由文献[10]得类间距离  $dist(S_i, S_j)$  如下:

$$dist(S_i, S_j) = 1 - \frac{\sum_{n=1}^N \min[u_{in}, u_{jn}]}{\sum_{n=1}^N \max[u_{in}, u_{jn}]} \quad (23)$$

其中  $N$  为样本总数,  $u_{in}$  表示样本  $n$  属于类别  $i$  的置信度。

一个类中所有样本到类别中心的距离之和能够代表该类的类内距离, 因此类内距离  $dist(S_k)$  可由模糊体积来表示:

$$dist(S_k) = \frac{\sum_{n=1}^N u_{kn} (x_n - v_k) (x_n - v_k)^T}{\sum_{n=1}^N u_{kn}} \quad (24)$$

其中  $x_n$  表示样本  $n$  的特征向量,  $v_k$  表示第  $k$  类的聚类中心矢量。

实验中采用不同的聚类类别数目  $c = 2, 3, \dots, c_{max}$  对特征向量进行聚类, 再由 (22) 式得出相应的评价函数值。我们把评价函数最大值所对应的聚类类别数目作为纹理图像类别数, 将其对应的聚类结果进行后处理, 取得了比较满意的效果。

## 4 试验结果分析

我们对由 Brodatz 纹理库中的纹理组成的 4 类纹理图像 ( $256 \times 256$ ) 的实验结果如图 3 所示, 误分类率为 1.31%。未

应用 EMD 分解直接使用四元数 Gabor 滤波器组对该图像分割的实验结果误分类率为 1.62% (本文采用的误分类率定义为分类错误的像素数目与像素总数之比)。图 4 是 5 类纹理图像 ( $256 \times 256$ ) 的分割结果, 误分类率为 1.50%, 而应用经典 Gabor 滤波器组方法<sup>[2]</sup>的分割结果的误分类率为 1.91%。

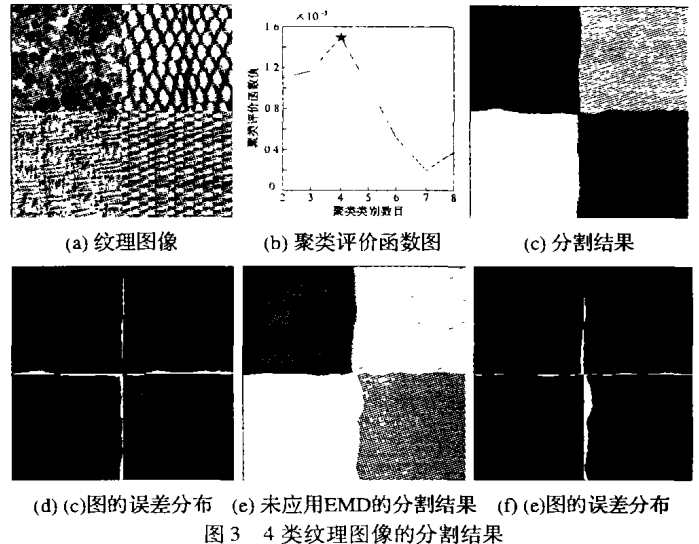


图 3 4 类纹理图像的分割结果

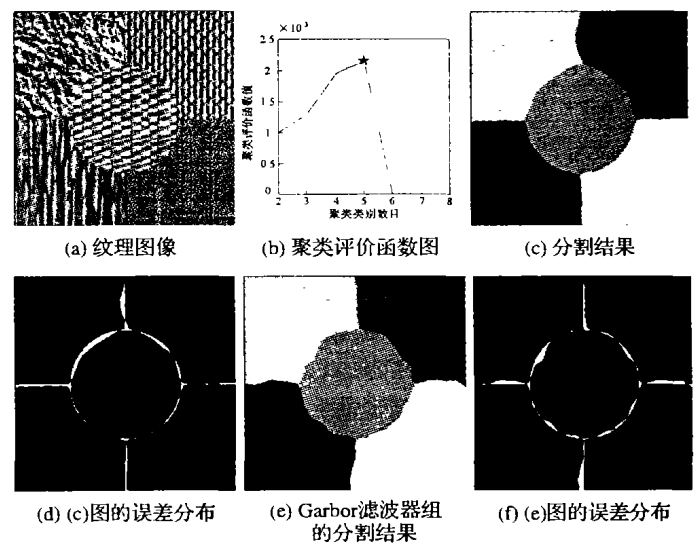


图 4 5 类纹理图像的分割结果

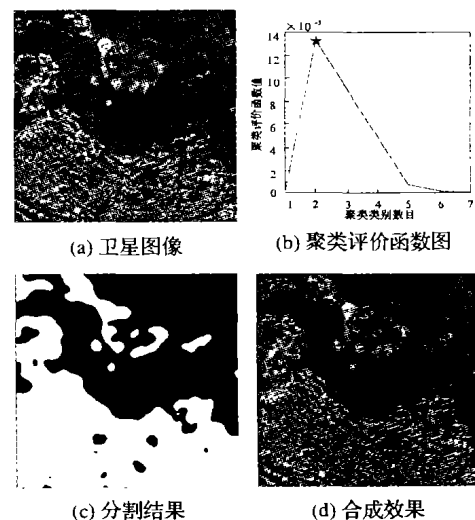


图 5 卫星图像的分割结果

值为  $\Delta x_i = i / (\frac{H}{X_{off}} + 1)$ 。因此  $x' = x_i + \Delta x_i, y' = y_i + \Delta y_i$ 。该方法仅需  $2W \times H$  次加法运算,因此可大大节省校正时间。

#### 4 实验结果及分析

选择几幅倾斜角度各异的图像作实验。图像为 A4 大小的页面,以 200dpi 的分辨率扫描输入。本文算法在 PIII 800 上实现,为了进行对比实验,我们将本文算法、Hough 算法和交叉相关算法分别应用于不同倾斜角的同一文档。本文算法倾斜角检测的平均处理时间为 2.03s, Hough 算法的评价处理时间为 21.70s, 交叉相关算法为 5.69s。普通 Hough 变换的最大绝对误差为 0.30, 交叉相关算法为 0.21, 本算法仅为 0.11。标准偏差值反映算法的精确度, 值越小表示偏差越小, 精确度越高。Hough 算法的标准偏差为 0.06701, 交叉相关算法为 0.04011, 而本文算法仅为 0.03036, 可见本算法的精确度较高。

对于 A4 大小的倾斜页面, 本算法的平均校正时间为 5.4s, 校正结果如图 1 所示。

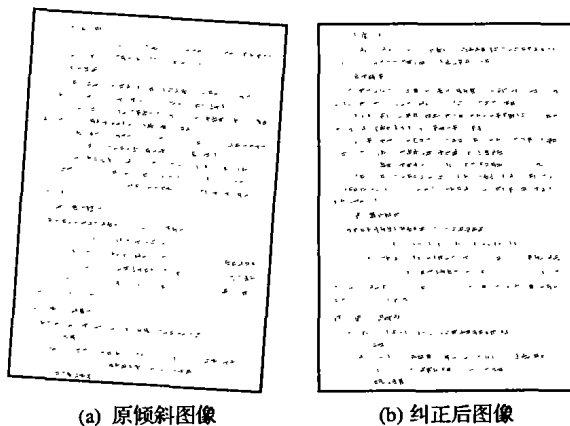


图1 倾斜图像校正示例

#### 5 结语

本文介绍了一种文档图像倾斜角检测与校正的算法。该算法首先选择图像的一个子区域, 然后提取水平边缘, 对提取出的水平边缘进行二级 Hough 变换。同时, 采用直线拟合的方法确定像素点的水平、竖直方向偏移值对倾斜页面校正。实验结果表明, 利用本文算法对图像倾斜角进行检测和校正, 速度较快, 准确度较高, 并且有很好的应用适应性。

##### 参考文献:

- [1] CIARDIELLO G, SCAFUR G, DEGRANDI M, et al. An Experimental System for Office Document Handling and Text Recognition [A]. Proceedings of Ninth International Conference on Pattern Recognition [C], 1998. 739 - 743.
- [2] BAIRD HS. The Skew angle of printed documents [A]. SPSE 40th Annual Conference and Symposium on hybrid Imaging System [C], 1987. 739 - 743.
- [3] ILLINGWORTH J, KITTLER J. The adaptive Hough Transform [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1987, 9(5): 691 - 698.
- [4] LI H, LAVIN MA, Le MASTER RJ. Fast Hough Transform [R]. Research Report RC11080 (#49754), IBM T. J. Watson Research Center, 1985.
- [5] RISSE T. Hough Transform for Line Recognition: Complexity of Evidence Accumulation and Cluster Detection [J]. Computer Vision, Graphics and Image Processing, 1989, 46(3): 327 - 345.
- [6] 潘武模, 焦阳, 王庆人. Hough 变换在中文名片图像倾斜校正中的应用 [J]. 中文信息学报, 2001, 15(3): 50 - 56.
- [7] 林纯青, 徐立亚, 戚飞虎. 汽车图像中字符目标的提取算法 [J]. 上海交通大学学报, 1998, 32(10): 1 - 3.
- [8] OKUN O, YU Y, ROBUST M. Text Detection from Binarized Document Images [A]. Proceedings of the 16th International Conference on Pattern Recognition (ICPR2002) [C], 2002. 61 - 64.

(上接第 576 页)

对 Brodatz 纹理库中的不同纹理样本进行组合实验, 都取得了良好的效果。图 5 是对卫星图像 (512 × 512) 的分割结果。由此可见, 四元数 Gabor 滤波器组对于标准纹理和自然纹理的分割都具有良好效果, 引入 EMD 分解并采用第一层 IMF 作为分析信号后, 分类准确率又得到了提高。图 3 至图 5 的各个 (b) 图中的 “\*” 号所对应的列坐标值即为最优聚类类别数目, 行坐标值为相应的聚类评价函数值。

#### 5 结语

本文提出一种基于 EMD 细化四元数谱的无监督纹理分割方法。利用二维可分离的 EMD 分解提高纹理图像信息在频域上的分离度并提取出纹理的主要结构, 使得可检测图像二维局部结构的四元数 Gabor 滤波器组的应用变得更加有效, 并由 FCM 聚类评价函数确定纹理类别数, 通过聚类得到最终分割结果。实验结果验证了本文方法的有效性, 该方法还可以引入到纹理分类检索等领域, 为 EMD 分解和四元数谱分析应用于图像处理做出了合理探索。

##### 参考文献:

- [1] TUCERYAN M, JAIN AK. Texture analysis [A]. The Handbook of Pattern Recognition and Computer Visions (2nd Edition) [C]. New Jersey: World Scientific Publishing Co, 1998. 207 - 248.
- [2] JAIN AK, FARROKHNI F. Unsupervised texture segmentation u-

sing Gabor filters [J]. Pattern Recognition, 1991, 24(12): 1167 - 1186.

- [3] BOVIK AC, MARIANNA C. Multichannel texture analysis using localized spatial filters [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990, 12(1): 55 - 73.
- [4] HUANG NE, SHEN Z, LONG SR, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis [A]. Proceeding of Royal Society [C]. London, 1998. 903 - 995.
- [5] NUNES JC, BOUAOUNE Y, DELECHELLE E, et al. Image analysis by bidimensional empirical mode decomposition [J]. Image and Vision Computing, 2003, 21(12): 1019 - 1026.
- [6] Bülow Th. Hypercomplex spectral signal representations for the processing and analysis of images [D]. Kiel, Germany: Christian Albrechts University, 1999.
- [7] Bülow Th, SOMMER G. Hypercomplex signals - a novel extension of the analytic signal to the multidimensional case [J]. IEEE Transactions on Signal Processing, 2001, 49(11): 2844 - 2852.
- [8] BRODATZ P. Textures - a photographic album for artists and designers [Z]. Dover, New York, 1966.
- [9] LUO JB, SAVAKIS AE. Self-supervised texture segmentation using complementary types of features [J]. Pattern Recognition, 2001, 34(11): 2971 - 2982.
- [10] BEZDEK JC, PAL NR. Some new indexes of cluster validity [J]. IEEE Transactions on Systems, Man and Cybernetics, 1998, 28(3): 301 - 315.