

# 彩色图象分割

郭国栋 马颂德

(中科院自动化所模式识别国家重点实验室, 北京 100080)

**摘要** 提出一种在特征空间进行非监督学习的新技术。以信息理论观点,把特征空间看成是两个不同的源所组成,即“峰(mode)”和“谷(valley)”。一个熵门限被用来自动区分特征空间中的不同单元。那些标号为“峰”的单元被连接起来形成峰的区域。提出一个修改的 Akaike 信息准则来求解相应的聚类有效化问题。当所有必需的参数都估计出来以后,将一个基于多数博弈论演化而来的标号算法用于求解分割所对应的优化问题。新方法应用到彩色图象的分割问题中,整个分割过程是自动进行的。

**关键词** 彩色图象,非监督学习,参数估计,图象分割,特征空间,熵门限,修改的 Akaike 信息准则,聚类有效化,多数博弈论,优化

## 1 引言

在图象理解问题中,特征空间分析是十分有用且非常必要的。通常,我们假设那些在图象上有明显意义的显著特征,对应于特征空间中高的密度区。而特征空间分析是一个恢复这些高密度区中心的过程,即找显著的图象特征代表<sup>[1]</sup>。

多个基于直方图的门限技术已经用来分割多维数据,如彩色图象。这类方法是通过每个特征轴进行直方图门限分割,来划分特征空间的。但此传统的门限技术主要是对灰度图象的,并不适合彩色图象分析<sup>[2]</sup>。

我们提出一种特征空间分析的新算法。该算法基于信息论,把特征空间中的样本或模式(patterns)看成是来自两个不同的信息源,即“峰(mode)”和“谷(valley)”。使用熵门限技术,我们先把所有样本点划分成峰和谷两类。然后,我们提出一个修改的 Akaike 信息准则,来决定特征空间中究竟有几个峰,即解决聚类有效化问题。假定复原出的峰(对应特征空间中的各高密度区域中心)就是特征空间中的聚类中心,我们执行一个基于博弈论的新标号算法,来进行彩色图象分割。

## 2 基于信息准则的聚类

为获得特征空间中的聚类,我们采取两步走的办法:第一步,使用熵门限把特征空间中的样本划分成“峰”和“谷”;第二步,确定正确聚类数和各聚类中心。

### 2.1 使用熵门限来检测峰

图象中象素点的数量通常非常巨大,首先需要把图象的特征空间离散化,变成一个个离散的单元(cells)。然后我们对这些离散的单元进行划分,归属于峰或谷。设特征空间的维数为 $d$ ,特征空间离散化是用一个间隔值 $g_r$ 来划分每一个特征轴 $r, r=1, \dots, d$ ,从而把连续的特征空间变成一个多维的规则晶格 $L$ 。每个格子就是一个单元,用它的中心 $s_i$ 来表示。记 $Y_i = [y_{i,1}, y_{i,2}, \dots, y_{i,d}]^T$ 是特征空间中的第 $i$ 个样本或模式,经过离散化以后,它将位于单元:

$$s_i = (\lfloor y_{i,1}/g_1 + 0.5 \rfloor, \lfloor y_{i,2}/g_2 + 0.5 \rfloor, \dots, \lfloor y_{i,d}/g_d + 0.5 \rfloor)^T$$

其中 $\lfloor x \rfloor$ 是实数 $x$ 的整数部分。对大的样本数,可使用大的来减小特征空间的单元数。

记落入单元  $s_i$  的样本数为  $n_i$ 。直觉上,如果单元  $s_i$  是在特征空间的“峰”区,则相应的  $n_i$  值将比较大;如果单元  $s_i$  是在“谷”区,则相应的  $n_i$  值将比较小。基于此来给出一个峰检测(mode detection)算法。

我们用落入每个单元的样本数  $n_i$  与总的样本数  $N$  的比值来近似每个单元的概率密度函数值,记作

$$p_i = \frac{n_i}{N}, i=1, 2, \dots, S$$

其中  $S$  是晶格的单元数,  $N = \sum_{i=1}^S n_i$  是总样本数。把  $p_i$  以升序排列,我们就得到了关于  $p_i$  的直方图(histogram)。从这个直方图上,我们来求得两个不同的源的集合:一个是 1 到  $t$  的分布(即下面的  $A$ ),另一个则是  $t+1$  到  $S$  的分布(即下面的  $B$ ),  $1 \leq t \leq S$ :

$$A: \frac{p_1}{P_t}, \frac{p_2}{P_t}, \dots, \frac{p_t}{P_t} \tag{1}$$

$$B: \frac{p_{t+1}}{1-P_t}, \frac{p_{t+2}}{1-P_t}, \dots, \frac{p_s}{1-P_t} \tag{2}$$

其中:  $P_t = \sum_{i=1}^t p_i, 1-P_t = \sum_{i=t+1}^S p_i$  与两个不同的源相关的熵<sup>[3]</sup>是:

$$H_A(t) = - \sum_{i=1}^t \frac{p_i}{P_t} \ln \frac{p_i}{P_t} - \frac{1}{P_t} [ \sum_{i=1}^t p_i \ln p_i - P_t \ln P_t ] = \ln P_t + \frac{H_t}{P_t} \tag{3}$$

$$H_B(t) = - \sum_{i=t+1}^S \frac{p_i}{1-P_t} \ln \frac{p_i}{1-P_t} - \frac{1}{1-P_t} [ \sum_{i=t+1}^S p_i \ln p_i - (1-P_t) \ln(1-P_t) ] = \ln(1-P_t) + \frac{H_s - H_t}{1-P_t} \tag{4}$$

其中  $H_t = - \sum_{i=1}^t p_i \ln p_i, H_s = - \sum_{i=1}^S p_i \ln p_i$

记  $\psi(t) = H_A(t) + H_B(t)$ , 由等式(1)至(4), 我们得到:

$$\psi(t) = \ln P_t (1-P_t) + \frac{H_t}{P_t} + \frac{H_s - H_t}{1-P_t} \tag{5}$$

使得  $\psi(t)$  取极大值, 就得到特征空间中的两个源, 即“峰” $H_B(t)$  和“谷” $H_A(t)$  的最大信息, 而对应的离散值  $t$  就是区别“峰”和“谷”的最优阈值。高于此阈值的那些单元将属于“峰”, 低于此阈值的那些单元将属于“谷”。那些空的单元一律看作“谷”。

经过熵门限划分后, 所有的单元都有了相应的归属。下一步就是把那些属于“峰”的单元连接起来, 形成一个个被“谷”单元隔离开的“峰区域”。区域个数记为  $m$ 。而落入每个“峰区域”的样本数记为  $h_k, k=1, 2, \dots, m$ 。把  $h_k$  以降序排列。

## 2.2 修改的 AIC 信息准则来获取聚类

由上一步获得的“峰区域”数比真正的聚类数通常要大  $m$  的大小, 在某种程度上与离散化的间隔值大小有关。

Akaike 信息准则(AIC)符合最大熵原则, 文献[4]中用来求解聚类有效化问题:

$$AIC(K) = -2 \log \{ p_K(Y | \hat{a}^{(K)}) \} + 2K' \tag{6}$$

其中  $Y$  是样本集,  $\hat{a}^{(K)}$  是模型参数向量  $a^{(K)}$  的估计值,  $K'$  是  $K$  类数据模型的独立可调节参数的个数。AIC 选择聚类数为  $K_0$ , 如果:

$$K_0 = \arg \min_{1 \leq K \leq K_{\max}} AIC(K) \tag{7}$$

其中  $K_{\max}$  是最大可允许的聚类数。以上的 AIC 准则把所有的样本点都考虑在内, 而且需要样本分布的假设(以便把每一类样本分布的概率密度参数化)。但是, 在这里, 我们只考虑“峰区域”的样本, 而且我们只用均值向量(即中心点)来表示每一个聚类。我们的最大相似性度量是基于最小距离测量。因此, 我们提出另一个信息准则, 定义如下:

$$MAIC(k) = \sum_{i=1}^k D^2(s_i, \mu_j) + TU(k) \tag{8}$$

其中  $\mu_j, j=1, 2, \dots, k$ , 是在  $k$  个聚类中距离单元最近的一个聚类的均值向量; 而  $D^2(s_i, \mu_j)$  是一个(非空)单元  $s_i$  与均值向量  $\mu_j$  的欧氏距离平方。把位于每个特征轴  $r, r=1, 2, \dots, d$  上的所有“峰”单元的均值和方差记为  $\mu_r$  和  $\sigma_r^2$ , 则:

$$\mu_r = \frac{1}{m} \sum_{i=1}^m s_i^r, \sigma_r^2 = \frac{1}{m} \sum_{i=1}^m (s_i^r - \mu_r)^2 \tag{9}$$

其中  $s_i^r$  是单元  $s_i$  的第  $r$  个特征。记  $\sigma_{\max}^2 = \max_r \sigma_r^2$ , 则取  $T = 2\sigma_{\max}^2$ ; 而  $U(k)$  是  $k$  个聚类的独立参数的个数(即均值向量个数)。

对  $k=1, 2, \dots, m$ , 计算  $MAIC(k)$ , 并根据下式确定最佳的聚类数  $c$ :

$$c = \arg \min_{1 \leq k \leq m} MAIC(k) \tag{10}$$

一旦  $c$  确定, 那么对应于  $c$  个聚类的均值向量就作为我们所估计的参数向量。

## 3 基于多数博弈论的图象标号

在特征空间中确定了所有参数, 再返回图象域来对每个象素进行标号。

我们提出一种基于多数博弈论(majority game)(见文献[5]中的第 10 章)的松弛算法。多数游戏是

一种简单的合作的游戏(cooperative game)。目的是成功地形成一些决策性的联盟(decisive coalition)。由选手们组成的一个非空子集称为一个联盟。在游戏中,每个选手都想与它的邻域形成一个胜利联盟(winning coalition)。

基于重要的先验知识,即距离很近的象素趋向于具有相同的标号(label),我们的决策性联盟这样形成:一个选手将采取它所在的联盟(由该选手和它的最近邻域形成)中占多数的那种策略,作为它在下一次游戏中的策略。这个游戏并行进行,迭代直到收敛。所有选手的初始策略则是由每个象素与所有聚类中心的最小距离来决定。

### 4 彩色图象分割

我们的聚类算法(估计参数)和多数游戏策略(求解标号)应用到彩色图象分割中。彩色图象通常是用 RGB 彩色空间来存储和显示的。但 RGB 空间不是均匀彩色空间(uniform color space,或 UCS)。为了保证特征空间分析的各向同性(isotropy),需要把 RGB 空间转化成 UCS<sup>[6]</sup>。我们选择  $L^*u^*v$  空间来作为我们的 UCS,它是通过对 RGB 值作非线性变

换来得到。白光标准  $D_{65}$  作参考照明。

图 1(a)是真实的彩色图象“woman”,大小为  $116 \times 261$ 。运行我们的算法,找到 5 个聚类。图 1(b)是分割结果。为清楚起见,我们把区域边界显示在图 1(c)中。图 2(a)是真实的彩色图象“house”,大小为  $255 \times 192$ 。运行我们的算法,找到 17 个聚类。图 2(b)是分割结果。区域边界显示在图 2(c)中。

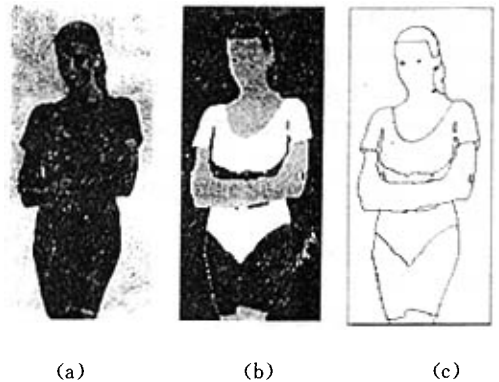


图 1 分割实验:(b)分割结果;(c)区域边界

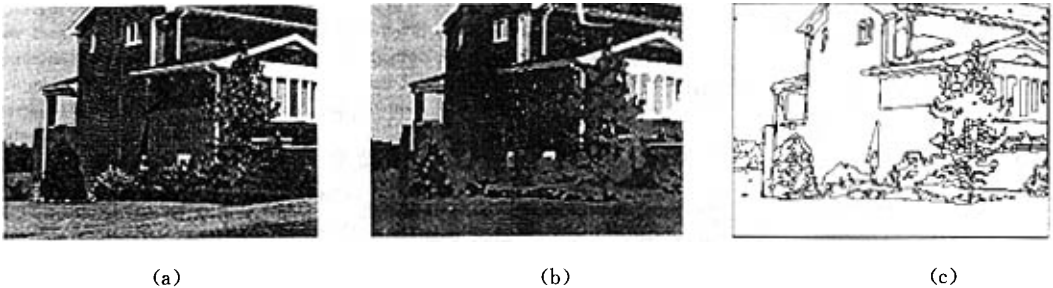


图 2 分割实验

### 5 结论

我们提出了一种新的数据驱动的彩色图象分割技术。由新的特征空间分析方法和标号策略两部分组成。进一步的研究是分析特征空间的离散化对聚类过程的影响。我们的聚类方法还可应用到更高维的特征空间分析中,如纹理图象分析。

### 参考文献

- 1 Comaniciu D, Meer P. Robust Analysis of Feature Spaces; Color Image Segmentation, CVPR'97, 750~755.
- 2 Uchiyama T, Arbib M A. Color Image Segmentation Using Competitive Learning. *IEEE Trans. Pattern Anal. Mach. Intel. PAMI-16*, 1994, 1197~1206.
- 3 Kapur J N, Sahoo P K, Wong A K C. A New Method for Gray-Level Picture Thresholding Using the Entropy of the Histogram, *CVGIP*, 1985,29:273~285.
- 4 Zhang J, Modestino J W. A Model-Fitting Approach to Cluster

Validation with Application to Stochastic Model-Based Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intel. PAMI-12*, 1990, 1009~1017.

- 5 Von Neumann, Morgenstern J. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton. 1944.
- 6 Wyszecski G, Stiles W S. *Color Science: Concepts and Methods, Quantitative Data and Formulae*. Second Ed. New York; Wiley, 1982.
- 7 Yu S, Berthod M. A game strategy approach for image labeling. *Computer Vision and Image Understanding*, 1995, 6(1): 32~37.
- 8 Touzani A, Postaire J G. Mode Detection by Relaxation, *IEEE Trans. Pattern Anal. Mach. Intel. PAMI-10*, 1988, 970~978.



**郭国栋** 1968年生,1991年毕业于清华大学,现为中科院自动化所模式识别国家重点实验室博士生,研究方向为模式识别、图象处理和计算机视觉。



**马颂德** 1946年生,IEEE高级会员,1968年毕业于清华大学,1986年获法国国家博士学位,现为中科院自动化所所长,模式识别国家重点实验室主任,博士生导师,长期从事计算机视觉、模式识别和图象处理等方面的研究工作。

## Color Image Segmentation

Guo Guodong, Ma Songde

(National Lab of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100080)

**Abstract** A novel technique for unsupervised learning in feature space is presented. The feature space is considered as composed of two distinct sources, "mode" and "valley", in the point of view of information theory. An entropy-based threshold is taken to distinguish the discrete cells in the feature space. The cells labeled as "mode" are then chained to form mode areas. Thereafter a modified Akaike's information criterion is proposed to solve the cluster validity problem. After all the parameters are estimated, a labeling algorithm is developed based on the majority game theory. The method is applied to color image segmentation. The segmentation process is completely autonomous.

**Keywords** Color image, Unsupervised learning, Parameter estimation, Image segmentation, Feature Space, Entropy threshold, Modified Akaike information criterion, Cluster validation, Majority game theory, Optimization

## 《计算机工程与应用》

● **简介**:本刊由华北计算技术研究所主办,创刊于1964年,属中国计算机学会会刊、中国电子学会一级会刊、计算机工程与应用学会学报、计算机中文核心期刊、中国科学论文统计用刊。是一本集学术与应用为一体的有较高权威性的计算机科技刊物。

● **主要栏目**:博士论坛、数据库、网络与通信、研究探讨、开发设计、工程应用等。

● **读者范围**:科研院所;大专院校;计算机工程项目规划、设计、实施人员;计算机应用人员。

● **订阅办法**:本刊1999年为月刊,大16开,每期正文128页,定价12.5元,全年150元,邮发代号为82—605,每月9日出版,全国各地邮局均可订阅。

联系地址:北京市619信箱26分箱

邮编:100083

电话:(010)62323668