

# 视觉计算中概率方法\*

陈熙霖 高文

(哈尔滨工业大学 计算机系, 150006)

## 摘 要

本文分析了近些年来计算机视觉发展较慢的原因, 指出忽略图像中不确定性干扰是视觉处理效果不好的重要原因. 分析了当前世界上在视觉中引入的随机方法, 将其归纳为理论、模型、估计准则与方法和应用几个方面, 并就存在的问题进行了讨论.

关键词 计算机视觉, 概率, 随机过程.

## 1 引 言

在人工智能方面, 计算机视觉曾经作为一个发展较快的领域而受到瞩目. 但近几年来发展速度却慢了下来. 形成这种局面的原因何在? 出路又如何?

Pavlidis<sup>[1]</sup>认为: 就目前而言, 计算机视觉作为一门科学, 不象其它领域那样有明确的而且容易达到的子目标, 而只是含糊的目标. 并且我们试图解决的问题模型是建立在若干年前提出的过于简单和不现实的基础之上的. 这个领域进展很慢, 不仅因为问题很难, 而且因为我们(至少在公开场合)拒绝承认很难.

Jain 和 Binford 等<sup>[2]</sup>则认为: 在计算机视觉系统中使用了大量的不适当的知识; 在空间和时间处理上缺乏远见以及缺少实验是视觉系统的三个严重障碍.

除了上述原因之外, 还在于我们把图像结构看得过于简单. 由于人的视觉效应, 人往往对干扰不敏感, 而机器不同. 在处理视觉问题时, 我们常忽略这些不确定的干扰, 从而使算法陷入困境. 这就是为什么很多算法在精心布置的实验室中效果很好, 而对于自然景物则难以处理的主要原因. 忽略这些不确定干扰可以使模型得到简化, 但同时也使处理不尽人意. 所以, 有必要引入不确定求解方法, 引入随机模型.

此外, 视觉问题是投影成象的逆问题, 由于三维向二维投影时往往失去了不少有用的信息, 因此逆问题的求解是病态的, 而且多是欠定的. 对于欠定问题的求解, 人们多借助于增加物理假设的方法来作为约束条件. 同时不断探索最优解或次最优解的解法. 正规化方法和概率方法就是其中主要的两种. Poggio 等人对正规化方法作了较充分的研究<sup>[3-4]</sup>, 而对于概率方法的研究还不充分, 正是由于这些原因, 近些年来不少人开始研究随机模型在计算机视觉中的应用.

本文总结了近几年该方法研究的情况, 将其归结为理论、模型、估计准则与方法和应用等几个方面, 并就存在的问题进行了讨论.

\* 本课题得到 863 高技术计划 (合同号 863-306-06-02-2) 和清华大学智能技术与系统国家开放实验室的资助, 本文于 1992 年 10 月 24 日收到

## 2 理论基础

在随机方法中, 其理论基础主要是概率论, 随机理论和熵理论.

作为一门较古老而成熟的学科, 概率论在现代技术中扮演着一个非常重要的角色, 是研究不确定性的重要工具. Bayes 公式是将先验分布与后验观测值联系的桥梁.

随机过程与概率论不同, 前者着重于研究的是随着某个参数  $t$  变化的随机函数的分布. 随机过程, 特别是 Markov 过程, 为建立一般的图像模型提供了十分有力的工具.

用  $\{S, E\}$  表示任意的图,  $X = \{X_s, s \in S\}$  表示  $S$  的数值属性集,  $\Omega$  表示所有可能的分布, 记为

$$\Omega = \{\omega = (x_{s_1}, \dots, x_{s_N}) | x_{s_i} \in L, 1 \leq i \leq N\}$$

同时记  $\{X_{s_1} = x_{s_1}, \dots, X_{s_N} = x_{s_N}\}$  为  $\{X = \omega\}$ .

在图像中,  $S$  表示像素点集,  $E$  表示各像素点之间的邻接关系,  $X_s \in L$ ,  $L$  为光强值集. 例如, 对于二值图像,  $L = \{0, 1\}$ ,  $S$  中所有像素点构成的分布的一个实现就是  $\omega$ . 在图 1 中假定  $\star$  表示 0,  $\blackstar$  表示 1, 则  $\omega = \{0, 1, 1, 0, 0, 1, 0, 1, 0\}$ .

☆ ★ ★ ☆ ☆ ★ ☆ ★ ☆

图 1

用  $N = \{N_s, s \in S\}$  定义  $S$  的领域系统, 如果对  $\forall s, r \in S$ , 满足

- 1)  $s \in N_s$ ,
- 2)  $s \in N_r$  当且仅当  $r \in N_s$ ,

$X$  定义为  $N$  上的 Markov 随机场 (MRF), 如果对  $\forall \omega \in \Omega$ , 和  $s \in S$ ,  $(x_{s_1}, \dots, x_{s_N}) \in \Omega$ , 满足

$$\begin{aligned} P(X = \omega) &> 0 \\ P(X_s = x_s | X_r = x_r, s \neq r) &= P(X_s = x_s | X_r = x_r, r \in N_s) \end{aligned} \quad (1)$$

在各种分布中, 由于 Gibbs 分布能直接和 MRF 找到对应关系, 因此得到广泛使用. 定义在  $\Omega$  上的与  $\{S, N\}$  相关的 Gibbs 分布为概率测度  $\pi$

$$\pi(\omega) = 1/Z \exp(-U(\omega)/T) \quad (2)$$

其中,  $T, Z$  为常数,  $T$  为温度.  $Z \approx \sum_{c \in C} \exp(-U(\omega)/T)$ ,  $c$  为基团.

$U(\omega) = \sum_{c \in C} V_c(\omega)$ , 称为能量函数.  $\{V_c, c \in C\}$  称为势函数.

下述定理<sup>[5]</sup> 奠定了 MRF 与 Gibbs 分布联系的纽带.

定理: 设  $N$  为一个领域系统, 则  $X$  是对应于  $N$  的 MRF 当且仅当  $\pi(\omega) = P(X = \omega)$  是对应于  $N$  的 Gibbs 分布.

MRF 与 Gibbs 分布的等价为方便地进行研究提供了一个非常便利而有效的工具. 特别是 Bayes 公式的使用, 使得很多计算化为对能量函数的处理.

熵是一个概率分布散乱程度的度量. 对于有若干个结果的离散分布  $1, 2, \dots, i, \dots$ , 相应的概率为  $P_1, P_2, \dots, P_i, \dots$ , 则该分布的熵为

$$q = - \sum_i P_i \log P_i \quad (3)$$

### 3 假设与模型

在随机处理方法中关于图像结构的模型是处理的基础,模型是在假设的基础上建立的.此外,由于投影的逆问题是病态的,为了得到最优解,必须引入约束条件,这些约束条件构成了假设的一部分.假设是求解问题的重要条件,必须满足

- 1 客观性,即符合客观世界的要求和限制.
- 2 计算的方便性,即根据假设建立的模型应易于在计算机上实现,否则没有应用的价值.

对于不同类型的问题,大致可以认为主要有三类基本假设:

- 1 关于原始图像:原始图像是 MRF,即图像在时间或空间上具有相关性.
- 2 关于成象过程:得到的图像是原始图像经过某种模糊化(镜头作用)之后,再经非线性传输,最后迭加上噪声而形成的.噪声的迭加一般是加性的或乘性的.
- 3 关于噪声:噪声一般是均值为  $\mu$ 、方差为  $\sigma$  的 Gauss 噪声.

随机方法中,由于 MRF 与 Gibbs 分布的等价性,可以方便地建立原始图象模型和噪声模型.

例如,在运动估计<sup>[15]</sup>中,假定原始运动场  $u(x,t)$  为 MRF,噪声是均值为 0 的高斯白噪声,成象过程是原始图像迭加上噪声.

### 4 估计准则与方法

估计准则,一般多采用最大后验概率(MAP),最小均方差准则(MMSE),最大熵估计准则(ME)等.

对于一幅  $M \times M$  的灰度级为  $L$  的图像,其可能的分布可达  $L^{M^2}$  种,用一般方法求解是不现实的,并且有时会陷入局部极小能量点,因此常常采用一些其它方法.这些方法主要有以下两种:

#### 1 Metropolis 算法及其变形

Metropolis<sup>[6]</sup>提出的模拟退火(Simulated Annealing)算法是模拟经过加热的物理系统,在冷却过程中粒子的迁移行为.在设定能量函数后,通过多次迭代使能量达到全局极小值.

由于 SA 算法的时间复杂性较高,因此不少人在以下几方面作了改进:

- 1) 在原 SA 算法中搜索的控制过程是融于最优解更新过程的,可以分开进行;
- 2) 改进 SA 中的抽样过程;
- 3) 改进参数的设定;
- 4) 研究并行 SA 算法(如 Cauchy 机).

理论上已经证明<sup>[6]</sup>,只要初始温度足够高,并且降温速度满足

$$T(k) > C / \log(1+k) \quad (4)$$

其中,  $C > 0$ , 为常数,  $k$  为迭代次数.则算法必收敛.

在实现上除了一般方法外,还可用 Boltzmann 机实现.

#### 2 利用变分法

用变分法往往先把问题化为求

$$E(u) = E_d(u, d) + \lambda E_p(u) \quad (5)$$

极小值的问题. 其中,  $d$  是观测值,  $u$  是估计值,  $E_d$  是估计值与观测值间的吻合程度度量,  $E_p$  是某种约束稳定函数,  $\lambda$  是约束强度.

例如在内插中, 可以取

$$E_d(u, d) = 1/2 \sum_i c_i (u_i - d_i)^2$$

$$E_p(u) = 1/2 \int \int (u_x^2 + u_y^2) dx dy$$

## 5 应用

随机方法在视觉计算中目前主要用于以下几方面:

1 稀疏值的内插. 内插可以看作是一个独立于其它处理过程的后处理. 特别是近年来激光测距仪的广泛使用, 人们对处理深度信息的兴趣日渐增长, 都对这方面的发展起了积极的作用. Marroquin<sup>[7]</sup> 采用 MRF, Fathima<sup>[8]</sup> 采用最大熵方法来处理内插. 并和其它低层处理连接起来取得了良好的效果.

2 图像恢复是低层处理中长期以来引人注目的方面. 由于摄象系统、传输过程以及其它原因, 我们所得到的图像往往是不清晰的, 而一般线性恢复又不够理想. 这样探索非线性恢复的方法成为必然, 随机方法就是其中的一种. 通过研究图像的一般退化过程, 建立的一般退化模型<sup>[9]</sup> 为:

$$g = T[\varphi(f)] \odot N \quad (6)$$

其中,  $f$  为原图像,  $g$  为得到的退化图像,  $\varphi$  为模糊化过程,  $T$  为传输函数,  $\odot$  表示某种迭加运算, 一般是加性或乘性的,  $N$  为噪声.

从目前工作看, 这方面取得了一定的进展: Geman 等<sup>[9]</sup> 采用边缘和区域模型进行有噪声的二值图像恢复. Simehony 等<sup>[10]</sup> 采用 MAP 估计和松弛算法对混有乘性噪声的灰度图像的恢复进行了研究. Johnson 等<sup>[11]</sup> 讨论了图像恢复的边界模型以及参数的选取. Geiger 等<sup>[12]</sup> 则考虑了恢复的并行性算法. Amir Waks 等<sup>[13]</sup> 对未知参数下的图像恢复进行了研究, 并取得了较好的效果.

3 立体视觉是低层视觉处理的一个较成熟而经典的方法. 最初的体视算法是 Marr 等<sup>[14]</sup> 用立体点对进行的, 后来又有不少人以零交、边缘、片段等作基元进行体视研究. 但这些算法在处理自然景物时效果往往不佳, 因此近年来有些人开始用随机方法进行研究. 由于体视对应可以看作一个能量极小化问题, 因此广泛采用了松弛技术、层次技术等, 并取得了良好的效果.

4 运动分析. 运动视觉是近年来计算机视觉研究的一个热点. 由于其在军事上的特殊意义以及较好的形式化, 不少人都投向这方面的研究. 和其它低层视觉同样的原因, 运动分析在自然景物下同样也存在困难. 故而有人转向去探索不确定情况下的解. J. Konrad, 等<sup>[15]</sup> 利用随机松弛方法处理运动图像, 并引入多格 (Multigrid) 分层算法, 加快了运动场估计的处理速度.

## 6 问题与展望

随机方法用于视觉处理是近几年才发展起来的, 虽然在处理效果上较好, 但也存在一些问题, 如一般需反复迭代, 迭代次数达十几次至几百次, 个别情况下甚至达上千次, 计算的复杂度

高。此外，目前这些算法还处于初期，不论在假设的提出、模型的建立，还是估计方法上，都还有待发展与完善。

今后在随机视觉方法方面需进一步的工作包括：建立一个统一的随机视觉方法框架；将随机方法用于其它一些视觉处理过程，如 Shape From X 等；寻求新的估计；将关于景物的知识融于处理过程之中，使处理的复杂度下降；研究开发并行随机处理算法等等。

## 参 考 文 献

- [ 1 ] Pavlidis T. Why Progress in Machine Vision is So Slow. *Pattern Recognition Letter*, 1992, 13(4): 221—225
- [ 2 ] Jain R C, Binford T O. Ignorance, Myopia, and Naivete in Computer Vision System. *CVGIP: Image Understanding*, 1991, 53(1): 112—117
- [ 3 ] Poggio T, Torre V. Ill-Posed Problem and Regularization Analysis in Early Vision. *Image Understanding Workshop*. 1984: 257—263
- [ 4 ] Poggio T, Torre V, Koch C. Computational Vision and Regularization Theory. *Nature*, 1985, 317: 314—310
- [ 5 ] Beseg J. Spatial Interaction and the Statistical Analysis of Lattice System. *Journal of the Royal Statist Society, Series B*, 1974, 36: 192—326
- [ 6 ] Kirkpatrick S. Optimization by Simulated Annealing: Quantitative Studies. *Journal of Statistical Physics*, 1984, 34: 975—986
- [ 7 ] Marroquin J L. Surface Reconstruction Preserving Discontinuities. A I. Memo. 792, AI Lab, MIT, 1984
- [ 8 ] Fathima S T, Yegnanarayana B. A Maximum Entropy Approach to Interpolation. *Sigal Processing*, 1990, 21(1): 17—24
- [ 9 ] Geman S, Geman D. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Trans on PAMI*, 1984, PAMI—6(6): 721—741
- [ 10 ] Simchony T, Chellappa R, Lichtenstein Ze'ev. Relaxation Algorithm For MAP Estimation of Gray-level Images with Multiplication Noise. *IEEE Trans on Information Theory*, 1990, IT—36(3): 608—613
- [ 11 ] Johnson V E, Wong Wing H, Hu Xiaoping, Chen Chintu. Image Restoration Using Gibbs Priors: Boundary Modeling, Treatment of Blurring, and Selection of Hyperparameter. *IEEE Trans on PAMI*, 1991, PAMI—13(5): 413—425
- [ 12 ] Geiger D, Girosi F. Parallel and Deterministic Algorithms From MRF's: Surface Reconstruction. *IEEE Trans on PAMI*, 1991, PAMI—13(5): 401—412
- [ 13 ] Waks A, Tretiak O J, Gregoriou G K. Restoration of Noisy Regions Modeled by Noncausal Markov Random Fields of Unknown Parameters. *10th International Conference on Pattern Recognition*, 1990
- [ 14 ] Marr D, Poggio T. Cooperative Computation of Stereo Disparity. *Science*, 1976, 194: 283—287
- [ 15 ] Konrad J, Dubois E. Multigrid Bayesian Estimation of Image Motion Fields Using Stochastic Relaxation. *International Conference on Computer Vision*, 1988

## PROBABILITY APPROACH IN COMPUTER VISION

Chen Xilin      Gao Wen

(Dept. of Computer Science & Engineering, HIT 150006)

### ABSTRACT

This paper analyses the reason why computer vision progress so slow and points out that the uncertainty noise in image is the important reason of inefficient processing results. The stochastic approaches, used in computer vision, are discussed, and also are classified into several aspects: theory, model, estimation critical and method, and application. At last, the existing problems are argued.

**Key Words** Computer Vision, Probability, Stochastic Process.