

概率图模型及其图像与视频应用研究

王相海^{1),2)} 陈明莹¹⁾ 方玲玲¹⁾

¹⁾ (辽宁师范大学计算机与信息技术学院, 大连 116029) ²⁾ (南京大学计算机软件新技术国家重点实验室, 南京 210093)

摘要 概率图模型及其应用作为一个具有挑战性的研究领域目前已成为一个新的研究热点。概率图模型为解决智能信息领域的不确定性问题提供了重要途径。尽管目前概率图模型还处于不断发展之中,但近年来基于概率图模型的图像和视频智能信息处理的应用研究受到人们的关注,出现了许多有效的算法,这些算法为解决一些传统的图像和视频智能信息处理问题提供了新的途径。本文首先对概率图模型的3种重要表现形式、特性和主要技术进行了分析和讨论,在此基础上,以概率图模型在图像和视频中的应用为线索,对目前基于概率图模型的图像和视频智能信息处理的主要技术进行了概述和比较研究;最后对概率图模型所存在的一些问题及进一步的发展进行了展望。

关键词 概率图模型 贝叶斯网络 马尔可夫网络 隐马尔可夫网络 图像 视频

中图法分类号: TP391 **文献标识码:** A **文章编号:** 1006-8961(2009)09-1712-09

Probabilistic Graphical Model and Its Application in Image and Video Information Processing

WANG Xiang-hai^{1),2)}, CHEN Ming-ying¹⁾, FANG Ling-ling¹⁾

¹⁾ (College of Computer and Information Technology, Liaoning Normal University, Dalian 116029)

²⁾ (State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093)

Abstract Probabilistic graphical model (PGM) and its application have become a new hot-spot as a challenging research. It provides an important means for resolving the uncertainty of intelligent information field. Although PGM is still in its development, the application of intelligent information processing of images and videos based on PGM has emerged and many effective algorithms has emerged which provide a new strategy to solve for the traditional problems. Based on the three important expressions, the characteristics and the main technology of PGM, we analyze and introduce the primary technology of intelligent information processing of images and videos based on PGM. Finally, we discuss the disadvantages and development trend of PGM.

Keywords probabilistic graphical model (PGM), bayesian network, Markov network, hidden Markov network, image, video

1 引言

概率图模型 (probabilistic graphical model, PGM) 作为一类用图形模式表达基于概率相关关系的模型

总称,是基于概率论中贝叶斯规则而建立起来的一种图形结构,也是计算机人工智能领域最流行的一种图模型结构^[1-2],该模型最早由 Whittaker 基于统计学的一些理论提出的^[3],模型能够很好地根据上下文的关系进行预测,取得了很好的预测效果。这

基金项目:辽宁省自然基金项目(20072156);辽宁省教育厅科学技术研究项目(20060486);辽宁“百千万人才工程”培养经费和南京邮电学院图像处理与图像通信江苏省重点实验室开放基金(ZK207008)

收稿日期:2009-01-01;**改回日期:**2009-04-27

第一作者简介:王相海(1965~),男,教授,博士生导师,中国计算机学会高级会员。2001年于南京大学计算机科学与技术博士后流动站出站。主要研究方向为计算机图形学及多媒体信息处理。E-mail:xhwang@lnnu.edu.cn

种结合了概率论和图论的概率图模型为解决人工智能领域的不确定性问题提供了重要途径。它能够将实际问题的时间和空间等信息有机地结合到一起,依赖大量的变量独立关系,构建一种基于联合概率分布的结构模型,从而对信息数据进行有效地表示。这种紧凑而简单的信息表达方式可以节省大量的存储空间,简化知识获取和领域建模过程,降低了计算复杂度^[4]。概率图模型自提出以来,在人工智能、数据挖掘、生物信息学、语言处理文本分类和控制理论等领域得到了广泛的应用,取得了很好的应用效果。

随着数字成像技术、设备及网络和多媒体技术的迅速发展,在诸多应用领域都产生了大量的图像及视频数据信息,对这些数据信息的智能处理成为多媒体信息处理的重要研究领域。近年来,人们将概率图模型应用于图像及视频的智能信息处理中,对图像及视频深层内联和结构进行了深入地研究,提出了许多有效的算法,为解决一些传统的图像和视频智能信息处理问题提供了新的途径。本文首先分析了概率图对应的几个主要模型的原理及其特性,并对其主要技术进行了介绍;然后分别以概率图模型在图像和视频的智能应用领域为线索对其主要技术进行了概述和比较研究;最后对概率图模型所存在的一些问题及进一步的发展进行了展望。

2 概率图模型分析

目前应用在图像和视频智能信息处理领域的基本的概率图模型主要包括贝叶斯网络、马尔可夫网络和隐马尔可夫网络。在实际应用中,有时将这3种模型混合使用,形成更精确的关系网,使其更贴近对实际问题的理想描述,从而形成混合模型。常见的有模型嵌模型的2维隐马尔可夫模型(EHMM),对应时间和空间关系的马尔可夫链和马尔可夫网,贝叶斯网和隐马尔可夫网的结合等,近年来,还产生了一种结合概率图模型与逻辑理论的新的模型,马尔可夫逻辑网(MLN)^[5-6]。下面分别对3种基本的概率图模型进行介绍。

2.1 贝叶斯网

贝叶斯网是概率论与图论的有机结合,是对概率关系有向图解的一种描述,适用于处理不确定性问题。该模型经历了路径分析、因果模式、影响图等发展过程,最后经 Pearl 建立起贝叶斯基础理论体

系^[7]。贝叶斯网是表示基于概率因果关系的一种结构,其表现形式如下:首先用结点来表示起因的假设和结果数据,接下来用有向弧来表示它们之间的联系,然后用数字编码的形式来描述变量之间互相影响的程度。主要由有向无环图 G 和条件概率表 θ 两部分构成,其中, θ 是对联合概率分布的定量表示,即在已知 π_i ($\pi_i \in \{X_1, \dots, X_{i-1}\}$) 的情况下, X_i 发生的条件概率,可用条件独立关系和条件概率来表示^[8](见式(1))。在模型建立之后,还要对其结构和参数进行学习,结构学习包括基于约束和基于打分的两类方法;参数学习算法则可分为基于最大似然的 MLE 方法和基于最大后验概率的贝叶斯方法。

$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P(X_i | \pi_i) \quad (1)$$

贝叶斯网可用于有条件地依赖多种控制因素的决策当中,比如说语言理解、医疗诊断、语音识别、模式识别、数据挖掘、数据压缩、机器人行为规划等方面。贝叶斯网模型之所以能够被广泛应用并得以不断发展,主要原因在于^[9]:(1)该模型具有坚实的理论基础。贝叶斯网络是基于概率统计的理论体系和方法论,推理过程严密且推断算法灵活;(2)具有强大的知识表达和推理能力。对于不确定性问题,贝叶斯网络结合网络结构和概率参数进行有效地表达和分析,可以实现原因与结果的相互推理。(3)具有灵活的学习机制。该模型可以模拟人类的学习方式和认知过程来灵活地对结构和参数进行相应地修正与更新。(4)具有开放的决策体系。贝叶斯网络分别将经验知识和样本信息,静态和动态,主观和客观相结合,从而得到了良好反应动态变化的后验概率,并使静态概率广泛适用到动态过程中,反应了数据内在的联系与本质。

目前,贝叶斯网模型的理论研究主要体现在以下几个方面^[10]:(1)研究表达能力强且带有自动性的概率模型,即如何通过学习自动确定和优化网络的拓扑结构;(2)研究参数先验概率分布的说明;贝叶斯网络推理需要提前假设先验概率,这些先验概率可以由大量历史的样本数据统计分析得到,也可由领域专家长期的知识或经验总结主观给出,或者根据具体情况事先假设给定;(3)高效的概率推理近似算法的研究,它也是贝叶斯网的一个关键组成部分。

2.2 马尔可夫模型

马尔可夫过程是一类随机过程,其原始模型马尔可夫链是由俄国数学家马尔可夫于 1907 年提出的。该过程在已知目前状态的条件下,它未来的演变不依赖于以往的演变,即无效性,此外它还具有平稳分布性、遍历性和状态相通性等^[11]。

马尔可夫过程是研究事物的状态及其转移的理论,适合于时间序列与空间序列,此外也可以方便地求出各种状态之间相互转移的概率^[12]。目前马尔可夫性一般用分布函数来表述^[13],即设随机过程 $\{X(t), t \in T\}$ 的状态空间为 I ,如果对时间 t 的任意 n 个数值 $t_1 < t_2 < \dots < t_n (n \geq 3, t_i \in T)$,在条件 $X(t_i) = x_i (x_i \in I, i = 1, 2, \dots, n-1)$ 下, $X(t_n)$ 的条件分布函数恰等于在条件 $X(t_{n-1}) = x_{n-1}$ 下 $X(t_n)$ 的条件分布函数,即

$$P\{X(t_n) \leq x_n \mid X(t_1) = x_1, \dots, X(t_{n-1}) = x_{n-1}\} \\ = P\{X(t_n) \leq x_n \mid X(t_{n-1}) = x_{n-1}\}, x_n \in \mathbf{R} \quad (2)$$

则称过程 $\{X(t), t \in T\}$ 具有马尔可夫性或无后效性,并称此过程为马尔可夫过程。

马尔可夫模型是马尔可夫过程的模型化描述,用一系列状态的不断转移表示一个总的随机过程。在该模型中,若状态 q_i 出现的概率仅取决于前一个状态,即 $q_i = P_r(q_i \mid q_{i-1})$,则称为一阶马尔可夫过程;若取决于之前出现的状态,则可表示为

$$q_i = P_r(q_i \mid q_{i-1}, q_{i-2}, \dots, q_1) \quad (3)$$

马尔可夫模型分为马尔可夫链和马尔可夫网。时间和状态都是离散的马尔可夫过程称为马尔可夫链,简称马氏链,记为 $\{X_n = X(n), n = 0, 1, 2, \dots\}$ 上对离散状态的马氏过程相继观察的结果。此时,马尔可夫性通常用条件分布律来表示^[13],即对任意的正整数 n, r 和 $0 \leq t_1 < t_2 < \dots < t_r < m; t_i, m, n+m \in T_1$,有

$$P\{X_{m+n} = a_j \mid X_{t_1} = a_{i_1}, \dots, X_{t_r} = a_{i_r}, X_m = a_i\} \\ = P\{X_{m+n} = a_j \mid X_m = a_i\}, a_i \in I \quad (4)$$

另一类是马尔可夫网(Markov networks)又可称为马尔可夫随机域(Markov random fields, MRFs),由 Pearl 提出^[15]。一个马尔可夫网是一个无向图,同样可以用 $G(V, E)$ 来表示,其中 V, E 分别表示无向图顶点的集合和边的集合;进一步让 $\{X_v: v \in V\}$ 代表图中结点的随机变量集合。通过无向图进行推理计算时,需要借助马尔可夫网的联合概率分布,而联合概率则采用所有相容函数乘积规范化表示^[16],即

$$P(x_v) = \frac{1}{Z} \prod_{c \in C} \psi_c(x_c) \quad (5)$$

式中, Z 是规范因子, C 表示图中所有团集的集合, $\psi_c(x_c)$ 表示一个非负相容函数。

不同马尔可夫模型间的差异主要体现在状态数以及状态转移概率的差别上。在图像和视频智能信息处理应用中利用马尔可夫模型来建模通常可以根据先验知识构造一个初始的图像模型,然后通过模拟退火等算法对其进行重构,进一步将重构之后的结果与原来的图像进行比较,最后根据差异不断对模型进行修正。

马尔可夫模型刻画数据的规律及模拟结构的能力强、表述问题清晰、直观,且在保证准确分类的情况下能够提高分类的效率,减少盲目性^[17],因而在物理、化学、生物、医学、多媒体应用等诸多领域得到广泛的应用。近年来在基础模型的基础上为了适应不同领域的需求,又出现了 Ising 模型、potts 模型、Auto-logistic 模型、MLL (multi-level logistic) 模型等^[18]。

2.3 隐马尔可夫模型

隐马尔可夫模型(hidden Markov model, HMM)作为一种统计分析模型,创立于 20 世纪 70 年代,之后作为信号处理的一个重要工具被成功地应用于语音识别、行为识别、文字识别及故障诊断等领域,近年来在移动通信核心技术“多用户的检测”、生物信息科学、故障诊断等领域得以应用^[19]。

HMM 与马尔可夫模型的区别在于前者是一个双重随机过程,即具有一定状态数的隐马尔可夫链和显示随机函数集,它由两层组成:一层是隐含层,包括一定量的状态,该层不能被直接观察到;另一层是观测层(实际观测值),在该层能观测到的只是与状态对应的一些观察值(observations),观测向量是由一个具有响应概率密度分布的状态序列所产生,即它是通过某些概率密度分布表现各种状态。

一个 HMM 可记为 $\lambda = (\pi, A, B)$ ^[14],其中用以描述该模型的元素含义如下:(1)模型中 Markov 链的状态数目 N 。通常记 N 个状态为 $S = \{1, 2, \dots, N\}$,并记 t 时刻 Markov 链所处的状态为 q_t ,其中 $q_t \in S$;(2)每个状态对应的可能观察值数目 M 。通常记 M 个观察值为 $V_1 \sim V_M$,记 t 时刻观察到的观察值为 O_t ,其中 $O_t \in \{V_1, V_2, \dots, V_M\}$;(3)初始状态概率矢量 π 。 $\pi = \{\pi_1, \pi_2, \dots, \pi_N\}$,其中 $\pi_i = P(q_1 = i)$, $1 \leq i \leq N$;(4)状态转移概率矩阵 A 。 $A = (\alpha_{ij})_{N \times N}$,

其中 $\alpha_{ij} = P(q_{t+1} = j | q_t = i)$, $1 \leq i, j \leq N$; (5) 观察值概率矩阵 $\mathbf{B}$ 。 $\mathbf{B} = (b_{jk})_{N \times M}$, 其中 $b_{jk} = P(O_t = V_k / q_t = j)$, $1 \leq j \leq N, 1 \leq k \leq M$ 。

将 HMM 应用于实际问题通常要经历如下 3 个环节: (1) 估计问题, 即给出一个 HMM 模型 λ 如何才能有效地计算产生观察序列 X 的条件概率 $P(X | \lambda)$; (2) 解码问题, 即如何在给出观测序列 X 和模型 λ 下选择能最好解释观测序列的状态序列; (3) 学习问题, 即给定观测序列 X 如何调整模型 $\lambda = (\pi, A, B)$ 的参数, 使条件概率 $P(X | \lambda)$ 达到最大。这 3 个问题可分别采用前向-后向算法、Viterbi 算法、期望最大 (EM) 算法或 Baum-Welch 算法来求解^[20-21]。

HMM 具有强大的时序模式分类能力, 适合于动态过程时间序列建模, 并且从理论上讲, 这些时间序列可满足无限性。此外从统计特性上看, 比较大的时间尺度上, HMM 依旧能够反应观测特征矢量的统计特征变换, 同时不需要精确观测所有的状态就能实现模式分类^[14, 22-23]。

3 PGM 在图像信息处理方面的应用

尽管概率图模型作为一个新兴的研究领域还处于不断发展中, 但近年来人们将概率图模型应用于图像处理应用领域, 在诸如图像分割、图像检测和识别等智能信息应用中提出了一些有效的方法, 取得了很好的应用效果。下面就概率图模型在这些领域中的应用情况进行分析和讨论。

3.1 PGM 在图像分割中的应用

图像分割作为从图像处理到图像分析的关键步骤, 其分割结果的好坏直接影响到对图像的理解。在图像分割中, 分割函数的获取作为一个重点和难点问题至今未得到很好的解决。近年来, 人们将 PGM 应用于图像分割中, 通过概率推理分析解决势函数难以获取的问题, 取得了很好的效果, 同时还可得到目标之间的空间位置关系, 大大提高了执行效率。文献[24]提出了一种基于组合技术的近似方法来构建混合模型, 采用数据简约、竞争学习和变异期望-最大化 (EM) 算法等来估计混合模型的参数, 然后基于 SIC 准则确定图像分割区域数, 再利用贝叶斯规则把像元分类到适当的区域当中。同时采用简单随机采样 (SRS) 技术、竞争学习 (CL) 算法变异和 EM 算法获取简约子集、初始化均值矢量初值。

所提出的方法在保证分割质量的前提下, 节省了计算时间; 文献[25]探讨了一种新的图像自动分割方法。该方法首先结合高斯有限混合模型与期望对图像特征空间的数据进行聚类, 然后采用信息理论准则 (ITC) 确定要分割的图像区域, 最后通过贝叶斯概率来分割图像, 取得了很好的分割效果。该方法对一些具有相干和连续区域特征的复杂图像需要采用精心设计的特征空间以达到有效的分割效果, 对复杂区域对象分割的特征空间的构造还有待于进一步研究。

近年来, 基于小波的图像处理得到了很好的发展, 基于小波域概率图模型的图像分割技术同样受到关注, 郭松涛等人在文献[26]中提出了一种基于改进小波域隐马尔可夫树 (HMT) 模型的遥感图像分割方法, 该方法利用基于希尔伯特变换对的 2 维方向小波以及拓展 HMT 对小波域尺度间的小波系数相关性进行建模, 同时结合多背景融合技术, 取得了很好的分割结果, 并降低了算法的时间复杂度。此外, 针对医学图像分割的特点, 文献[27]引入高阶邻域空间约束信息解决了医学图像分割的精度和分割效率问题。并提出了一种基于广义模糊 Gibbs 随机场模型的医学图像多类模糊分割的理论, 同时在模型中提出了一种广义模糊边界图和一组方便的参数来适应不同性质的图像分割, 解决了传统边界泄漏问题, 为图像分割提供了一种新的思路。

3.2 PGM 在图像检测和去噪中的应用

近几年, 基于概率图模型的图像检测、识别或跟踪方法的研究已成为一个热点问题受到关注, 在指纹和人脸检测与识别等应用领域取得了非常好的应用效果, 表现出了很好的发展趋势。文献[28]根据指纹分类及指纹的方向场能有效描述不同指纹类型的纹理结构差异的特点, 利用嵌入式隐马尔可夫模型进行建模分析, 并合理地抽取指纹的类型特征构造观察向量对模型进行学习, 最后根据产生最大概率序列对应的模型进行匹配, 所提出的方法具有很强的鲁棒性。方法中由于运用了指纹分类信息, 不仅使指纹的预处理简单, 而且使提取出的指纹特征更加可靠。鉴于图像小波变换子带间所形成的小波树结构具有数据压缩的特性, 文献[29]把隐马尔可夫模型引入到低频子带系数为根的特征小波树中, 并对其进行降维和去相关等处理, 使得隐马尔可夫模型通过这些小波树所构成的观测向量进行学习, 最后根据训练后的参数进行人脸识别, 与传统

的人脸识别算法相比,所提出的基于隐马尔可夫模型的算法识别效率高,对人脸的变化具有更强的鲁棒性。

为了更有效地利用时间和空间的特征,文献[30]提出了一种基于时、空域的概率图模型来跟踪拥有多表情的人脸特征。在时间域上,考虑到使用多个跟踪器描述人脸特征的空间约束不够准确,这样每个特征都各自形成一条马氏链,且对应每个特征都用几个相互独立的粒子滤波来跟踪;在空间域上,使用马尔可夫网描述人脸各个特征相互影响和作用,并使用贝叶斯推理-信任度传播算法计算人脸表情变化时的轮廓位置。该方法即使在帧间运动较大时,所跟踪的人脸多特征效果也很鲁棒。具体模型参见图 1^[30],其中沿着时间轴的方向,眉、眼、鼻和嘴 4 个特征各自形成一条马尔可夫链用来表示其在时间上的变化;而在空间中的每一时刻里,用这 4 个特征的相互联系所形成的马尔可夫网络来表示人脸模型中人脸特征的相互关系。此外,与以往采用一个高斯概率密度函数处理观察符号概率不同,文献[31]在利用连续隐马尔可夫模型对人脸进行识别时,通过 2 个高斯概率密度函数处理观察符号概率,并且采用具有稳定性、转置不变性的奇异值来分解、抽取人脸特征矢量,并通过用同一个人的多幅人脸图像来进行模型训练,也在一定程度上减少了计算量,提高了识别率。

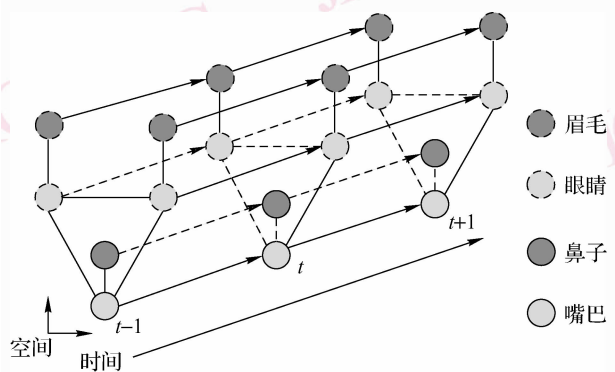


图 1 人脸多特征跟踪的时空图模型
Fig. 1 Spatio-temporal graphical model of tracking multiple face features

除了上述 PGM 在图像检测方面的应用外,近年来,也出现了一些基于 PGM 的图像去噪算法,文献[32]针对超声检测图像中的噪声进行处理,在降噪过程中结合马尔可夫随机场、Gibbs 分布函数以及

模拟退火等算法,并进行了优化和改进,提高了运算效率。文献[33]提出了一种基于多小波隐马尔可夫的图像去噪算法,该算法将多小波表达和通用隐马尔可夫(uHMT)模型相结合,在 uHMT 模型中利用 9 个参数来表述真实图像的 HMT 模型,在保证降噪效果的同时,减小了计算复杂度。

4 PGM 在视频信息处理方面的应用

与图像智能信息处理一样,近年来基于 PGM 的视频智能信息处理受到了很大的关注,在视频分割、视频检测和跟踪、视频挖掘等方面提出了很多有效的算法,为解决传统的视频处理问题提供了新的途径,同时也为进一步拓宽视频信息处理的应用领域奠定了基础。

4.1 PGM 在视频分割中的应用

在复杂的视频场景中进行背景的动态建模,准确地反映出背景的变化是许多视频智能信息处理的基础,同时也是视频处理的一个难点问题。近年来,基于 PGM 的视频分割方法的提出为该问题的进一步解决提供了重要的思路。文献[4]提出了一种基于混合概率模型的视频分割算法,该算法通过统计时间和空间相关信息对变化的背景建模,同时利用颜色同现性和变化同现性以及混合模型中的信念传播算法来解决模型的学习和估计问题,具体模型参见图 2 和图 3^[4],其中图 2 给出了视频中图像序列的 3 维网格图表示,图 3 则给出了描述同一位置像素点随时间变化的 HMM 示意图,其中在 t 时刻下,结点的观察值由其所有隐含状态(即该结点与其相

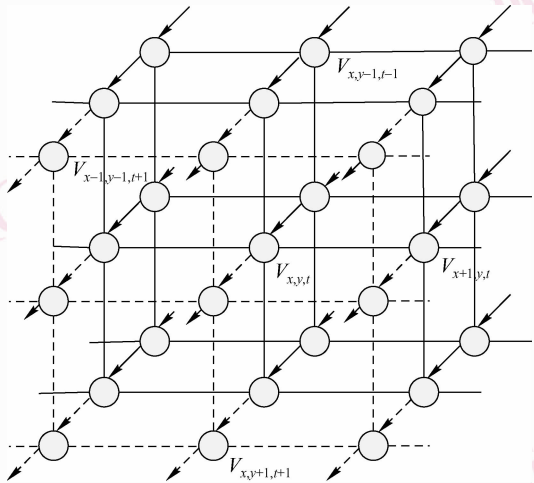


图 2 图像序列的示意图
Fig. 2 Image sequences diagram

邻结点间的关系)来共同决定;而不同时刻下,每一个隐状态都与相邻时刻的各隐状态相互联系。此算法不但可以有效地从复杂变化的背景中分割出运动物体,而且克服了单一模型视频分割算法中的缺点,但文中并未对“未知”的背景变化进行动态建模,尽管文献[34]也曾对此有一定的讨论,但对该问题还缺少系统而有效的算法。

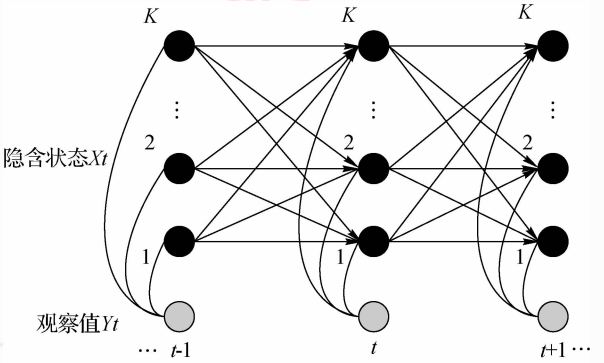


图 3 像素时间序列变化的隐马尔可夫模型描述
Fig. 3 Description of hidden Markov models on time sequence changing of pixel

此外,针对传统结合时间域和空间域视频分割的比重法的不足,文献[35]提出了一种基于 Gibbs 随机场模型的时空分割方法。该方法首先为图像帧的分割模板建立 Gibbs 随机场模型,然后将时间域和空间域分割结果分别作为初始标记场和图像观察场,进一步利用 Gibbs 模型的约束条件将两者进行结合,得到最后的分割标记场。该方法既能利用空间域信息又可以较好地避免比重法在分割时过分依赖于空间域分割精度。但基于 Gibbs 随机场模型方法对不同的视频分割需要不同的参数,如何利用时间域分割结果来指导或自适应的确定参数,将是该方法进一步的改进方向。另外,针对背景中光线变化的规律,文献[36]提出了一种利用贝叶斯进行学习的分割方法,该方法首先在每个像素点处对变化背景进行建模,并计算其对应的颜色直方图,然后利用直方图来表示该像素点处特征向量的概率分布,并利用贝叶斯学习方法来判断其属于前景还是背景。

4.2 PGM 在视频检测和跟踪方面的应用

近年来,随着摄像设备和技术的快速发展,视频的检测和跟踪技术被基于视频的安全导向实际应用所需求,同时也是高层视频语义分析的关键。文献[37]研究了视频序列的不规则数据、视频序列的可疑行为以及视频图像中异常图案的检测和识别方

法,针对传统的诸如图像注意、视频注意、可疑行为及异常物体的识别等计算机视觉问题通常需要分开处理的情况,提出了一个统一的建模架构予以融合并且还提出了一个新的基于图的贝叶斯推理算法,该算法允许在多时空尺度中来探测大的块集,即图像中的空间集合及视频中的时空集合,具体模型如图 4 所示^[37],其中白色结点表示观测变量,灰色结点表示隐含变量,有向弧则表示贝叶斯依赖关系。同样针对上述问题,文献[38]使用 iHMM (infinite HMM) 进行了建模,首先使用 ISA (invariant subspace analysis) 进行视频特征提取,然后通过已训练好了的 iHMM 对这些特征的时间属性进行建模,最后通过训练结果对其事件的正常与否进行判断。这里, iHMM 保留了所有模型参数的后验密度函数,包括基本的 HMM 状态数,同时还应用了分层的狄利克雷框架, iHMM 的后验分布估计是通过基于马尔可夫链的蒙特卡罗和变化的贝叶斯公式获得的,该方法具有非常好的检测精度。

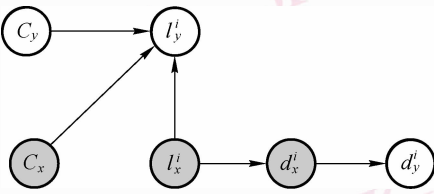


图 4 用于探测异常行为的概率图模型
Fig. 4 Probabilistic graphical model for detecting abnormal behaviors

另外,文献[39]提出一种基于 Laplacianfaces 和 HMM 的视频人脸识别方法,该方法的观测值采用了测试视频中映射到拉普拉斯空间的降维后的人脸图像特征,然后用每一训练视频的隐马尔可夫模型去分析视频数据的统计特性和时间动态性,最后把能够产生该序列的最大概率所对应的模型确定为所匹配类别。文献[40]利用概率图模型对热红外和光谱视频进行融合实现对人物对象的跟踪。该模型依照产生数据的过程,通过隐含变量来对数据进行描述,通过这种方式,可以分别捕获和利用热红外和光谱视频数据的统计结构以及它们的依赖性。这里,模型参数是通过 EM 演绎算法进行学习的,然后,利用观测数据进行对象位置的贝叶斯推理来实现对象的自动跟踪。该模型对噪声及复杂环境具有很好的鲁棒性。文献[41]采用隐马尔可夫模型建模的方式对遥感图像中目标轨迹的动态特性进行刻

画,在模型中,观测输出信号是轨迹信息,而隐状态则表达了连续帧轨迹坐标的转移特性。在识别过程中,将获得的运动目标的轨迹数据送入预处理模块,对完整且有效的轨迹进行 K-均值自动聚类,接着对轨迹样本数据进行预处理得到样本库,从而训练 HMM 模型得到一系列参数,通过这些参数计算并求出产生给定轨迹坐标数据的概率,最大概率值所对应的模型便为轨迹所属的模式类,即为识别结果。该方法具有较高的平均识别率。

4.3 PGM 在视频挖掘和检索中的应用

视频挖掘作为数据挖掘和视频信息处理的交叉研究领域,目前有很多问题未得到很好的解决,比如对大量语义概念的理解和解释,及如何利用它们来探测精确的多重关系,等等。近年来,通过概率图模型表示描述各种多语义概念之间关系的学习方法,为解决该问题提供了一种有效的途径。在文献[42]中通过统一的概率图模型对描述几种针对挖掘视频概念间关系进行表示,其中使用了多种图模型进行处理,包括不同数目的有向图和无向图。最后在同等条件下与多种相关方法进行比较研究,展示了利用无向图去学习概念间关系的潜力和有效性。另外,文献[43]针对体育视频挖掘提出了一个新的多通道分段隐马尔可夫模型,此模型吸收了 CHMMs (coupled HMMs), SHMMs (segmental HMMs), HHMMs (hierarchical HMMs) 的思想,较文献[42]中的模型更具有可变性和多功能特性。为了同时优化模型的结构和参数,在模型训练的过程中应用了熵优先和参数消失 (entropic prior and parameter extinction) 策略。该概率图模型具有 2 个优势:一个是分层和平行结构的一体化,能更灵活地获取多个 MRF 链间的关系;另一个是分段 HMM 的使用,体现了观察值的可变长特性。接下来可以通过提高模型的训练过程、纳入更多通道或分层语义结构、推理高层语义概念等 3 方面进行工作。具体模型参见图 5^[43],图中 S 代表 2 个通道的状态序列; O 代表 2 个通道的观测值; F 代表第 2 层的状态序列,在这里每个状态由下一层的 2 个状态组成,第 1 层包括 2 个 SHMM,第 2 层则捕获了第 1 层 2 个通道的交互信息。

此外,PGM 还被应用于从数据库中对视频进行检索。Sergio Benini 等人^[44]利用隐马尔可夫模型来表示故事单元,采用马尔可夫熵率来表示复杂索引以便估计结构相似性,取得了很好的效果。

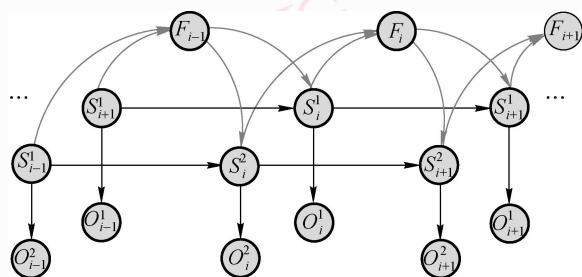


图 5 基于多通道分段隐马尔可夫模型描述

Fig. 5 Description of multi-channel segmental hidden Markov models

5 展 望

概率图模型及其应用是一个比较前沿的研究领域,对解决不确定性问题具有非常好的应用前景,尽管目前已经出现了一些概率图模型及相应的实例,但由于该领域还处于刚刚起步阶段,很多问题还有待于进一步研究和完善,在进一步的研究中如下几个问题值得关注:

(1) 有关 PGM 完整而系统的理论架构还有待于进一步研究,目前的很多应用系统大都使用 PGM 的某一种方法,对不同的应用领域,算法的效率还不是很,还缺少更多适应于各种应用领域的 PGM 模型方法。

(2) 如何将现有的一些图像和视频智能信息工具和方法,诸如图像和视频的多尺度分析方法、视频的运动估计方法和图像、视频的形态学方法等,结合到 PGM 方法中,构建基于多工具的 PGM 模型同样是值得关注的问题。

(3) 目前针对图像和视频智能信息处理的应用范围还不是很广,方法还不是很,探索基于 PGM 的更多种图像和视频智能信息处理的方法是一件很有意义的工作。

(4) PMG 自身的特点使得基于 PMG 的算法通常具有较高的时间复杂度,如何针对具体的应用领域对 PGM 进行优化,包括建模过程、学习过程等,开发出高效、高扩展型的 PGM 实用算法是值得进一步研究的问题。

(5) 不同于其他的操作工具,目前构建的 PGM 对应用领域的依赖性很强,这在一定程度上可以提高模型的效率,但不利于工具包的建立,如何总结领域规律,构建统一而高效的 PGM 模型,形成相对独立的工具平台也是未来值得关注的问题。

6 结 语

在实际应用中,对图像及视频智能信息处理中大量的不确定性问题的研究成为一项重要而又具有挑战性的工作,概率图模型为解决这一问题提供了一种重要途径。本文首先介绍了 PGM 的 3 种表现形式和对应的特性,然后对图像及视频智能信息处理中 PGM 的应用技术进行了综述性研究,最后对该领域进一步的研究进行了展望。

参考文献 (References)

- 1 Wang Y, Vassileva J. Bayesian network-based trust model [A]. In: Proceedings of IEEE International Conference on Web Intelligence [C], Halifax, Canada, 2003:372-378.
- 2 Chen J C, Wang Y C, Maa C S, *et al.* Network-side mobile position location using factor graphs [J]. IEEE Transactions on Wireless Communications, 2006, 5(10):2696-2704.
- 3 Whittaker J. Graphical Models in Applied Multivariate Statistics [M]. New York: John Wiley and Sons Inc, 1990.
- 4 Liu Zhen. Probabilistic Graphical Model and Its Application in Video Segmentation [D]. Beijing: Institute of Computing Technology of Chinese Academy Sciences, 2005. [刘震. 概率图模型在视频分割中的应用[D]. 北京:中国科学院计算技术研究所, 2005.]
- 5 Richardson M, Domingos P. Markov logic networks [J]. Machine Learning, 2006, 62(1):107-136.
- 6 Domingos P. Learning, logic and probability: A unified view [A]. In: Proceedings of the 9th Pacific Rim International Conference on Artificial Intelligence: Trends in Artificial Intelligence [C], Guilin, China, 2006, 1:53.
- 7 Pearl J. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference [M]. San Mateo, California, USA: Morgan Kaufmann Publishers Inc, 1988.
- 8 Jia Hai-yang. Research on Some Problems in Learning Bayesian Network [D]. Changchun: Jilin University, 2008. [贾海洋. 贝叶斯网学习若干问题研究[D]. 长春:吉林大学, 2008.]
- 9 Zhao Hai-feng. Study on Mining and Bayesian Network Representation of Association Rules [D]. Chongqing: Chongqing University, 2007. [赵海丰. 关联规则挖掘及贝叶斯网表示研究[D]. 重庆:重庆大学, 2007.]
- 10 Jiang Li-li. Research on Crime Investigation Decision Modelling Based on Bayesian Network [D]. Nanjing: Hohai University, 2007. [蒋莉莉. 基于贝叶斯网络的刑事侦查决策模型研究[D]. 南京:河海大学, 2007.]
- 11 Xiao Ze-lei, Lu Xi-zao. Discussion of Shanghai stock index based on system of Markov chain [J]. Pioneering with Science & Technology Monthly, 2005, 18(9):33-34. [肖泽磊, 卢悉早. 基于马尔可夫链系统的上证指数探讨[J]. 科技创业月刊, 2005, 18(9):33-34.]
- 12 Zhang Rui, Chi Dao-cai, Wang Xiao-yu, *et al.* Application of Markov modified residual error gray and prediction model in disaster predictio [J]. China Rural Water and Hydropower, 2008, (1):7-10. [张瑞, 迟道才, 王晓瑜等. 基于马尔可夫过程的改进残差灰色灾变预测模型研究[J]. 中国农村水利水电, 2008, (1):7-10.]
- 13 Sheng Zhou, Xie Shi-qian, Pan Cheng-yi. Probability and Statistics [M]. Zhejiang: Higher Education Press, 2003. [盛骤, 谢式千, 潘承毅. 概率论与数理统计[M]. 浙江: 高等教育出版社, 2003.]
- 14 Yu Lin-hai. Research of Intelligent Human Face Recognition Technology Base on Hidden Markov Model [D]. Xi'an: Xidian University, 2007. [于林海. 基于隐马尔可夫模型的人工智能人脸识别技术研究[D]. 西安: 西安电子科技大学, 2007.]
- 15 Pearl J. Bayes and Markov Networks: A Comparison of Two Graphical Representations of Probabilistic Knowledge [R]. 860024 (R-46), Los Angeles, California, USA: University of California at Los Angeles, 1986.
- 16 Shi Dong-yu. The Research on Probabilistic Graphical Models for Uncertainty Processing in Distributed Computing Environments and Related Problems [D]. Shanghai: Shanghai Jiaotong University, 2006. [石东昱. 适合分布计算环境不确定性处理的概率图模型若干问题研究[D]. 上海: 上海交通大学, 2006.]
- 17 Hou Chuan-yu. The application of Markov models and hidden Markov models in data mining [J]. Computer Knowledge and Technology, 2008, 1(7):1186-1188. [侯传宇. 马尔可夫及隐马尔可夫模型在数据挖掘中的应用[J]. 电脑知识与技术, 2008, 1(7):1186-1188.]
- 18 Lv Qing-wen. Image Segmentation Based on Mutual Information and Markov Random Field [D]. Guangzhou: Southern Medical University, 2007. [吕庆文. 基于互信息量与马尔可夫随机场的图像分割研究[D]. 广州: 第一军医大学, 2007.]
- 19 Chen Zhong-sheng. Statistical Information Processing Based on Matlab 7.0 [M]. Hunan: Hunan Science & Technology Press, 2005. [陈仲生. 基于 Matlab 7.0 的统计信息处理[M]. 湖南: 湖南科技出版社, 2005.]
- 20 Deng Zhi-chao. Parallel Multiple Sequence Alignment Based on Hidden Markov Model [D]. Changchun: Northeast Normal University, 2007. [邓志超. 基于隐马尔可夫模型的并行多重序列比对[D]. 长春: 东北师范大学, 2007.]
- 21 Liu Wei-qiang. Hidden Markov Model Parameters Estimation for Part-of-Speech Tagging [D]. Dalian: Dalian University of Technology, 2006. [刘伟强. 应用于词性标注的隐马尔可夫模型参数评估[D]. 大连: 大连理工大学, 2006.]
- 22 Li Jie. Research on Hidden Markov Model and its Application to Image Recognition [D]. Beijing: Tsinghua University, 2004. [李杰. 隐马尔可夫模型的研究及其在图像识别中的应用[D]. 北京: 清华大学, 2004.]
- 23 Wang Hong-qun, Peng Jia-xiong, Yu Qiu-ze. A precise eye localization via a dynamic probability distribution HMM model [J]. Journal of Image and Graphics, 2006, 11(1):26-32. [王洪群, 彭嘉雄, 于秋则. 采用动态 HMM 概率分布模型的人眼精确定位[J]. 中国图象图形学报, 2006, 11(1):26-32.]

- 24 Guo Ping. Study on effective computation of mixture model parameters for bayesian probabilistic image automatic segment [J]. Computer Science, 2002, **29**(8):101-103. [郭平. 贝叶斯概率图像分割中混合模型参数高效计算的研究[J]. 计算机科学, 2002, **29**(8): 101-103.]
- 25 Guo Ping, Lu Han-qing. Research of image automatic segmentation using bayesian probabilistic[J]. Acta Optica Sinica, 2002, **22**(12): 1479-1483. [郭平, 卢汉清. 贝叶斯概率图像自动分割研究[J]. 光学学报, 2002, **22**(12):1479-1483.]
- 26 Guo Song-tao, Sun Qiang, Jiao Li-cheng. Remote-sensing image segmentation based on improved wavelet-domain hidden Markov models[J]. Journal of Electronics & Information Technology, 2005, **27**(2):286-289. [郭松涛, 孙强, 焦李成. 基于改进小波域隐马尔可夫模型的遥感图像分割[J]. 电子与信息报, 2005, **27**(2): 286-289.]
- 27 Lin Ya-zhong. Research of New Approaches in Medical Image Segmentation Based on Gibbs Random Fields[D]. Guangzhou: Southern Medical University, 2004. [林亚忠. 基于 Gibbs 随机场模型的医学图像分割新算法研究[D]. 广州: 第一军医大学, 2004.]
- 28 Guo Hao, Ou Zong-ying. Fingerprint classification based on hidden Markov models[J]. Journal of Chinese Computer Systems, 2006, **27**(9):1718-1721. [郭浩, 欧宗瑛. 基于隐马尔可夫模型的指纹分类研究[J]. 小型微型计算机系统, 2006, **27**(9):1718-1721.]
- 29 Cao Lin, Wang Dong-feng, Zou Mou-yan. Face recognition based on hidden Markov model [J]. Journal of the Graduate School of the Chinese Academy of Sciences, 2006, **23**(1):52-59. [曹林, 王东峰, 邹谋炎. 基于隐马尔可夫模型的人像鉴别算法[J]. 中国科学院研究生院学报, 2006, **23**(1):52-59.]
- 30 Su Cong-yong, Zhuang Yue-ting, Huang Li, *et al.* Probabilistic graphical model for multiple facial feature tracking[J]. Journal of Computer-Aided Design & Computer Graphics, 2004, **16**(11):1523-1527. [苏从勇, 庄越挺, 黄丽等. 基于概率图模型的人脸多特征跟踪[J]. 计算机辅助设计与图形学学报, 2004, **16**(11): 1523-1527.]
- 31 Shen Jie, Wang Zheng-qun, Zou Jun, *et al.* Approach of face recognition based on continuous hidden Markov model [J]. Computer Engineering and Design, 2008, **29**(3):707-709. [沈杰, 王正群, 邹军等. 基于连续隐马尔可夫模型的人脸识别方法[J]. 计算机工程与设计, 2008, **29**(3):707-709.]
- 32 Liu Yi-zhou. Ultrasonic Image Denoise Analysis Based on Markov Random Fields[D]. Wuhan: Wuhan University, 2005. [刘一帆. 基于马尔可夫随机场的超声图像降噪分析[D]. 武汉: 武汉大学, 2005.]
- 33 Zhang Wei, Sui Qing-mei. Image de-noising algorithm using multiple wavelet and universal hidden Markov tree model [J]. Journal of Qingdao University of Science and Technology (Natural Science Edition), 2008, **29**(2):170-174. [张伟, 隋青美. 多小波描述的通用隐马尔可夫树模型图像去噪算法[J]. 青岛科技大学学报, 2008, **29**(2):170-174.]
- 34 Seki M, Wada T, Fujiwara H, *et al.* Background subtraction based on co-occurrence of image variations[A]. In: Proceedings of 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'03) [C], Madison, Wisconsin, USA, 2003, **2**(18-20):65-72.
- 35 Chen Han-feng, Qi Fei-hu. Integration of spatial and temporal segmentation based on Gibbs random field model in video object segmentation [J]. Acta Electronica Sinica, 2004, **32**(1):34-37. [陈韩锋, 戚飞虎. 视频对象分割中基于 Gibbs 随机场模型的时空分割结合方法[J]. 电子学报, 2004, **32**(1):34-37.]
- 36 Wang Lin-bo, Zhao Jie-yu. Video image segmentation based on bayesian learning [J]. Journal of Image and Graphics, 2005, **10**(9):1073-1078. [王林波, 赵杰煜. 基于贝叶斯学习的视频图像分割[J]. 中国图象图形学报, 2005, **10**(9):1073-1078.]
- 37 Boiman Oren, Irani Michal. Detecting irregularities in images and in video[A]. In: Proceedings of the 10th IEEE International Conference on Computer Vision[C], Beijing, China, 2005, **74**(1):17-31.
- 38 Pruteanu-Malinici Iulian, Carin Lawrence. Infinite hidden Markov models for unusual-event detection in video[J]. IEEE Transactions on Image Processing, 2008, **17**(5):811-822.
- 39 Jiang Yan-xia, Zhou Hong-ren, Jing Zhong-liang. Video-based face recognition using Laplacian faces and hidden Markov models [J]. Computer Engineering, 2007, **33**(1):204-206. [江艳霞, 周宏仁, 敬忠良. 基于拉普拉斯脸和隐马尔可夫的视频人脸识别[J]. 计算机工程, 2007, **33**(1):204-206.]
- 40 Chen Si-yue, Zhu Wen-jie, Leung Henry. Thermo-visual video fusion using probabilistic graphical model for human tracking [A]. In: Proceedings of IEEE International Symposium on Circuits and Systems [C], Seattle, Washington, USA, 2008:1926-1929.
- 41 Pan Qi-ming, Cheng Yong-mei. Trajectory recognition of moving objects based on hidden Markov model [J]. Application Research of Computers, 2008, **25**(7):1988-1991. [潘奇明, 程咏梅. 基于隐马尔可夫模型的运动目标轨迹识别[J]. 计算机应用研究, 2008, **25**(7):1988-1991.]
- 42 Yan Rong, Chen Ming-yu, Hauptmann Alexander. Mining relationship between video concepts using probabilistic graphical models[A]. In: Proceedings of IEEE International Conference on Multimedia and Expo[C], Toronto, Ontario, Canada, 2006:301-304.
- 43 Ding Yi, Fan Guo-liang. Multi-channel segmental hidden markov models for sports video mining[A]. In: Proceeding of the 16th ACM International Conference on Multimedia [C], Vancouver, British Columbia, Canada, 2008:697-700.
- 44 Sergio Benini, Pierangelo Migliorati, Riccardo Leonardi. Retrieval of video story units by markov entropy rate[A]. In: Proceedings of the 6th International Workshop on Content-based Multimedia Indexing [C], London, UK, 2008:41-45.