计算机科学2003Vol. 30№.1

# —种小波域 HMT 模型参数初始化方法\*`

汪西莉<sup>1,2</sup> 刘 芳<sup>3</sup> 焦李成<sup>1</sup>

(西安电子科技大学雷达信号处理国家重点实验室 西安710071)1 (陕西师范大学计算机学院 西安710062)<sup>2</sup> (西安电子科技大学计算机学院 西安710071)<sup>3</sup>

## A Parameter Initialization Method for Wavelet-Based HMT Models

WANG Xi-Li<sup>1,2</sup> LIU Fang<sup>3</sup> JIAO Li-Cheng<sup>1</sup>

(National Key Lab. of Radar Signal Processing, Xidian Univ., Xi'an 710071, China)1

(School of Computer Science, Shannxi Normal Univ., Xi'an 710062, China)<sup>2</sup> (School of Computer Science, Xidian Univ., Xi'an 710071, China)<sup>3</sup>

Abstract Wavelet transformation and hidden Markov model are used in wavelet-based HMT model for analyzing and processing images. Expected Maximization (EM) algorithm used in training model results in slow convergence. The persistence, exponential decay characteristics of wavelet coefficient are analyzed. A model parameter initialization method is proposed. This method provides reasonable initial model value, reduces training time greatly. Its application in image de-noising demonstrates is validity.

Keywords Wavelet, Hidden Markov tree model, Parameter initialization, EM algorithm

# 1 引言

小波理论由于具有坚实的数学基础、快速的变换算法以 及多分辨分析能力,成为信号分析的有力工具。由 M.S. Crouse 等提出的小波域 HMT 模型<sup>[1]</sup>将图像的小波系 数建模为隐 Markov 树模型,该模型考虑了小波系数间的统 计相关性及非高斯性,由于它抓住了小波系数的本质特性及 小波系数间的主要关系,在应用中取得了很好的效果。在图像 处理领域,小波域 HMT 模型用于图像去噪、压缩、检测、识别 等等。

但应用 HMT 模型时,其训练通常采用迭代的期望最大 (EM)算法,虽然每一代复杂度不高,但迭代次数多算法收敛 速度很慢。本文研究了不同尺度间小波系数及其状态转移概 率间的关系,给出了接近实际模型参数的参数初值,由该初值 开始 HMT 模型的训练,大大加快了收敛速度。

### 2 小波域 HMT 模型

u

#### 2.1 图像的离散小波变换

二维离散小波变换(DWT)用三个平移和伸缩的小波基 函数 Ψ<sup>LH</sup>、Ψ<sup>HL</sup>、Ψ<sup>HH</sup>以及一个尺度基函数 Φ<sup>LL</sup>来描述大小为 N×N 的图像  $x(s) \in L^2(\mathbb{R}^2)$ ,令 B={LH,HL,HH},b\in B,  $Ψ_{J,K}^{\delta}(s) = 2^{J} Ψ^{\delta}(2^{J}s - K), Φ_{J,K}^{LL}(s) = 2^{J} Φ^{LL}(2^{J}s - K)$ 构成  $L^{2}$ (R<sup>2</sup>)空间的正交基,则 x(s)可分解为:

$$x(s) = \sum_{K \in Z^2} u_{J_0,K} \Phi_{J_0,K}^{f_1}(s) + \sum_{\delta \in B} \sum_{j=J_{0K \in Z^2}} w_{J,K}^{\delta} \Psi_{J,K}^{\delta}(s) \quad (1)$$
  
$$u_{J_0,K} = \int_{s \in R^2} x(s) \Phi_{J_0,K}^{f_1}(s) ds, w_{J,K}^{\delta} = \int_{s \in R^2} x(s) \Psi_{J,K}^{\delta}(s) ds f ds$$
  
为尺度系数和小波系数。J 代表尺度或分辨率,J 越大分辨率  
越高,K 是平移参数。实际图像具有有限的分辨率,因此式(1)  
中半无穷求和可改为有限个尺度求和,J\_0 < J < J\_1, J\_0, J\_1 < Z.

J1=log2N。图像在各尺度分为三个通道 LH、HL、HH,在第 J

层每个通道有4<sup>1-1</sup>个小波系数,其中任一个小波系数所代表的 信息在第J+1层是由四个小波系数来表示的,故图像的三个 高频通道各自形成一个四叉树结构,使小波系数可以方便地 用离散小波树来描述。

## 2.2 图像的 HMT 模型

一般地,小波系数被建模为联合高斯模型或独立非高斯 模型,它们可以较好地表征小波变换的多数性质,如局部性、 多分辨性、边界检测等,但也有局限性,如联合高斯模型无法 表达压缩性,独立非高斯模型将小波系数建模为统计独立的, 这与实际不完全相符:小波变换虽然具有去相关性,但并没有 使小波系数彼此完全无关。

M.S. Crouse 等提出了图像的小波域隐 Markov 模型,又 称为隐 Markov 树(HMT)模型。HMT 模型将小波系数建模 为非高斯的,用HMT 来描述小波系数间还保留的相关性,因 而可以准确地描述小波系数的性质,这也正是它在实际应用 中表现出色的原因。

HMT 将每一个小波系数的边缘概率密度函数建模成具 有隐状态的混合高斯分布,这里考虑两状态、零均值的高斯分 布(这是因为两状态、零均值的高斯分布已可较好地描述小波 系数的分布,且在计算上具有明显优势),状态 S,的取值为 {S,L},S对应于零均值、小方差。6的高斯分布,L对应于零 均值、大方差 di的高斯分布, di < cd。因此任一小波系数 wi 的 边缘概率密度函数为:

$$f_{W_{i}}(w_{i}) = \sum_{m \in \{S,L\}} p_{S_{i}}(m) f_{W_{i}|S_{i}}(w_{i}|S_{i}=m)$$
(2)  
其中

 $f_{W_{i}|S_{i}}(w_{i}|S_{i}=m) = g(w_{i};0,\sigma_{i,m}^{2}) = \frac{1}{\sqrt{2\pi}\sigma_{i,m}} \exp(-\frac{w_{i}^{2}}{2\sigma_{i,m}^{2}})$ 

为 wi 的条件概率密度函数, ps1(m)表示 wi 是小或大的概率, ps,(L)=1-ps,(S)。尽管当状态变量 S,确定后系数 w,是条 件高斯的,但由于 S, 的随机性,从总体上看 w, 服从非高斯分

• 85 •

<sup>\*)</sup>国家自然科学基金(60073053)和教育部博士点基金资助项目.汪西莉 博士生,讲师,主要研究方向为模式识别,图像处理.刘 芳 副教 授,主要研究方向为人工智能,模式识别. 焦李成 教授,博导,主要研究方向为智能信息处理,模式识别。

布。

HMT 模型还描述了小波系数间的相关性。要建模所有 小波系数间的相关性既复杂又没有必要。小波系数的延续性 表明沿着尺度方向的小波系数(父子系数间)具有较强的相关 性,对每一个通道,用父系数的状态变量和其子节点系数的状 态变量相连的概率树来描述其相关性,和小波系数一样,每个 通道得到一个四叉树,这也是 HMT 模型名称的由来。具体地 说,父状态到子状态的连接关系用状态转移矩阵 A 表示:

 $\mathbf{A}_{i} = \begin{bmatrix} \mathbf{c}_{i,\mathbf{p}(i)}^{\mathbf{c},i} \mathbf{c}_{i,\mathbf{p}(i)}^{\mathbf{c},i} \\ \mathbf{c}_{i,\mathbf{p}(i)}^{\mathbf{c},i} \mathbf{c}_{i,\mathbf{p}(i)}^{\mathbf{c},i} \end{bmatrix}$ (3)

其中 $i,\rho(i)$ 分别为子节点、父节点, $\epsilon_{i,j0}^{ci}$ 表示小波系数w的 父系数状态较小的情况下w、状态也较小的概率, $\epsilon_{i,j0}^{ci}$ =1- $\epsilon_{i,j0}^{ci}$ 。

图像的 HMT 模型 M =  $(\theta^{LH}, \theta^{HL}, \theta^{HL}), \theta^{i} = \{p_{s_1}(S), \mu, \dots, \sigma^{i}, \dots, \varepsilon^{i}, \mu^{i}, \dots\},$ 其中  $p_{s_1}$ 为根节点的状态概率函数,  $\varepsilon^{i}, \mu^{i}, \dots$ 是已 知父节点状态为 n 时子节点状态为 m 的条件概率,  $\mu, \dots$ 和  $\sigma^{i}, \dots$ 是已知系数 w, 的状态为 m 时其均值和方差,  $m, n \in \{S, L\},$ 在零均值时  $\mu, \dots = 0$ 。

#### 2.3 模型参数的估计

任何基于 HMT 的应用首先都要训练 HMT 模型以得到 模型参数。若小波系数的状态变量 S. 已知,则通过极大似然 (ML)估计可容易地获得模型参数  $\hat{\theta} = \arg \max f(W|\theta)$ 。但实 际情况是 S<sub>i</sub> 未知,无法应用 ML 估计,通常采用 EM 算法<sup>[1,2]</sup> 来估计参数。EM 算法分为 E 步和 M 步两个过程(三个通道 各自处理);

1. 选择初始模型参数 θ°, 置计数器 l=0;

2. E 步 : 计算隐状态变量的联合概率  $p(S|W,\theta)$ ;

3. M 步、更新模型参数  $\theta^{\prime+1} = \arg \max_{\theta} E_s[\ln f(W, S|\theta)]$ 

*W*,θ'];

4. l=l+1,若满足收敛条件则停止,否则转2。

理论证明 EM 算法通过 E 步、M 步的交替迭代最终收敛 于不完全似然函数  $f(W|\theta)$ 的一个局部极大值。

#### 3 模型参数初始化

EM 算法是由不完全训练数据(因为隐状态未知)估计模 型参数的有效且常用的方法,其每一迭代步的计算复杂度为 O(n)(n=N×N),但算法的收敛是耗时的,如果能从合适的 模型参数开始训练,则训练时间会大大缩短。通过分析发现可 以找到比较合适的参数初值。



图1 小波系数的指数衰减性

首先小波系数随着尺度的增加(分辨率增加)具有指数衰减性<sup>(3)</sup>,图1显示了四幅标准灰度图像的小波系数的指数衰减 性。该特性源于图像的整体平滑性,可以把图像看作是由一个 个的平滑区域构成的,其间具有不连续性,由此产生的1/f 谱 特性导致小波系数在尺度间具有指数衰减性。

由于 HMT 将小波系数建模为具有隐状态的混合高斯分 布,混合方差的值也反映了小波系数的值,因此小波系数的指 数衰减性表明方差在尺度间也具有指数衰减性,在初始化时 应将此反映出来。在两状态情况下,令 ஏ.s、ஏ.L表示尺度 J 的 小(状态)方差、大(状态)方差,方差在尺度间的衰减性用式 (4)表示:

 $\sigma_{J,S}^2 = \exp^{C_S - J}, \sigma_{J,L}^2 = \exp^{C_L - J}$ (4)

其中*Cs*、*CL*为常数,对各个尺度要求 *G*.*s* < *G*.*L*,因此 *Cs* < *CL*。 根据式(4)初始化的方差随着尺度的增加满足指数衰减性,和 小波系数的走向吻合。

再来考察小波系数的延续性[4],即较大(或较小)的小波 系数可以在相邻的尺度间传递,这种延续性随着尺度的增加 变得越来越强。小波基函数可以理解为一种局部边界检测算 子,如果在基函数的支持区间内存在边界,相应的小波系数就 较大,而且该边界可能会继续存在于更高分辨率空间的小波 基函数的支持区间内;如果在某一尺度,基函数的支持区间内 不再有边界(边界完全被分离),则随着尺度的增加也不会再 出现边界,这时小的系数产生的孩子其值必定也较小,体现了 小波系数的延续性。状态转移矩阵中的 ɛšɨ̯, ɛɨːšɨ 反映了小 波系数的更新性,一方面小的系数也可能产生大的孩子,因为 在某个小尺度小波基函数的支持区间内如果有多个边界,多 个边界效果相抵致使其小波系数较小,但随着尺度的增加边 界总会分离,使得一个小波基函数的支持区间内只有一个边 界,从而产生大的系数;另一方面大的系数也可能产生小的孩 子,一个边界存在于某个低分辨率空间中,随着分辨率的提 高,由于边界的局部性,边界未必还会存在于所有的高分辨率 子空间中,实际上,边界在任一高分辨率子空间可能还存在也 可能消失,则不存在边界的高分辨率子空间其小波系数将较 小。

根据上述分析得到结论;随着尺度的增加,小系数产生小 系数的概率越来越大,直至1;到了一定的尺度后,边界已完全 被分离开,尺度再增加也不会出现边界,因此当 J→∞时 €<sup>4,3</sup>/<sub>2</sub>, →1。而大系数产生大系数的概率在逐渐减小,直至1/2,即到 达一定的尺度后,在其子空间边界可能存在也可能消失。据此 给出状态转移矩阵的初始化形式;

$$\begin{aligned} & \epsilon_{i,\rho(i)}^{s,\sigma(i)} = 1 - 0.2 * (J - 1), \\ & \epsilon_{i,\rho(i)}^{L,L} = 1 - 0.2 * (J - 1) \\ & \epsilon_{i,\rho(i)}^{S,S} = 1 - \exp^{-J}, \\ & \epsilon_{i,\rho(i)}^{L,L} = \frac{1}{2} + \exp^{-J} \end{aligned}$$
(5)

通过分析小波系数和状态转移概率的特点,我们给出了 式(4)、(5)来初始化 HMT 模型中的参数,虽然它不十分精 确,但也足以表征方差和状态转移概率的趋势。

## 4 方法在图像去噪中的应用

采用上述方法对一些标准图像进行了实验。四幅灰度图像(lenna、boat、bridge、aerial)加上均值 μ<sub>n</sub>=0、方差 σ<sub>n</sub><sup>2</sup>=0.1 的白高斯噪声得到各噪声图像。称利用本文的模型参数初始 (下转第77页)

• 86 •

B<sub>1</sub><sup>•</sup> = [1,1]/v<sub>1</sub> + [0.95,0.98]/v<sub>2</sub> + [0.9,0.93]/v<sub>3</sub> + [0.94,0.97]/v<sub>4</sub>+[0.96,1]/v<sub>5</sub> 2)T(A<sup>•</sup>,A<sub>2</sub>)=0.294<α,故放弃该条规则

3)T(A<sup>•</sup>,A<sub>3</sub>)=0.71≥0.5≥α且D(A<sup>•</sup>,A<sub>3</sub>)=0.29>0 由式(7)及式(8)得:

- $B_{3}^{\bullet} = [0.96, 0.99]/v_{1} + [0.98, 1]/v_{2} + [0.96, 0.99]/v_{3} + [1, 1]/v_{4} + [0.97, 0.99]/v_{5}$
- 4) 0. 5  $\geq$  T(A<sup>•</sup>, A<sub>4</sub>) = 0. 352  $\geq$   $\alpha$
- 由式(7)及式(8)得:
- $B_4^{\bullet} = [0.04, 0.18]/v_1 + [0.26, 0.3]/v_2 + [0.32, 0.33]/v_3 \\ + [0.35, 0.35]/v_4 + [0.32, 0.33]/v_5$
- 5)T(A<sup>•</sup>, A<sub>5</sub>)=0.756 $\geq$ 0.5 $\geq$  $\alpha$  <u>H</u> D(A<sup>•</sup>, A<sub>5</sub>)=0.244>
- 由式(7)及式(8)得:

0

- $B_5^* = [0.3, 0.4]/v_1 + [0.35, 0.48]/v_2 + [0.45, 0.55]/v_3 \\ + [0.4, 0.46]/v_4 + [0.36, 0.4]/v_5$
- 最后的推理结果为:
- $B^{\bullet} = B_1^{\bullet} \cup B_3^{\bullet} \cup B_4^{\bullet} \cup B_5^{\bullet}$ 
  - $= [1, 1]/v_1 + [0.98, 1]/v_2 + [0.96, 0.99]/v_3 + [1, 1]/v_4 + [0.97, 1]/v_5$

另一方面,若输入事实 B<sup>\*</sup> = [0.1,0.14]/v<sub>1</sub> + [0.15, 0.17]/v<sub>2</sub> + [0.2,0.25]/v<sub>3</sub> + [0.15,0.2]/v<sub>4</sub> + [0.11,0.16]/ v<sub>5</sub>,假定规则启动阈值 α=0.25,则逆向近似推理应按如下推 理过程进行:

1)T(B\*,B<sub>1</sub>)=0.2<α,故放弃该条规则 2)0.5≥T(B\*,B<sub>2</sub>)=0.45≥α 由式(9)及式(10)得:

- A<sub>2</sub><sup>•</sup>=[0.02,0.04]/u<sub>1</sub>+[0.04,0.05]/u<sub>2</sub>+[0.05,0.07]/ u<sub>3</sub>+[0.03,0.05]/u<sub>4</sub>+[0.02,0.04]/u<sub>5</sub> 3)T(B<sup>•</sup>,B<sub>3</sub>)=0.186<a,故放弃该条规则
- 4)0. 5≥ $T(B^*, B_4) = 0.375 ≥ \alpha$
- 由式(9)及式(10)得:
- $A_4^* = [0.04, 0.06]/u_1 + [0.05, 0.07]/u_2 + [0.07, 0.09]/$

# (上接第86页)

化方式训练 HMT 模型的方法为 F-HMT,模型采用两状态、0 均值。在实验中取  $C_s = 3$ 、 $C_L = 6$ ,  $p_{s_1} = (0.5, 0.5)$ 。实验采用 Daubechies4小波。表1列出了采用 HMT 和 F-HMT 对四幅含 噪图像训练,算法所用的迭代次数(三个通道之和)、噪声图像 和去噪后的图像的信噪比 PSNR,其定义为: PSNR = - $10\log_{10}||y-\hat{y}||^2/n, y, \hat{y}$ 分别为原始图像和去噪图像。PSNR 越 大说明图像越清晰。由实验结果可见,本文的参数初始化方法 较好地反映了参数间的关系,逼近参数的实际值,因而对各图 像训练时间都减少了2/3左右。

	表]	实	验	结	果
--	----	---	---	---	---

图像	F-HMT		HMT		噪声图像
	迭代次数	PSNR	迭代次数	PSNR	PSNR
lenna	28	29. 5685	94	29. 5002	19.9875
boat	21	27.6753	83	27.6823	19.9942
bridge	33	25.1634	80	25. 1821	19.9983
aerial	25	25.684	79	25.6528	20. 0128

 $u_3 + [0.05, 0.07]/u_4 + [0.03, 0.05]/u_5$ 

5)T(B<sup>•</sup>, B<sub>5</sub>)=0.848 $\ge$ 0.5 $\ge \alpha$  <u>H</u> D(B<sup>•</sup>, B<sub>5</sub>)=-1.52<

由式(9)及式(10)得:

0

 $A_{5}^{*} = [0.49, 0.55]/u_{1} + [0.71, 0.77]/u_{2} + [0.55, 0.6]/u_{3} + [0.39, 0.44]/u_{4} + [0.24, 0.29]/u_{5}$ 

最后的推理结果为:

- $A^* = A_2^* \cup A_4^* \cup A_5^*$ 
  - $= [0.49, 0.55]/u_1 + [0.71, 0.77]/u_2 + [0.55, 0.6]/$  $u_3 + [0.39, 0.44]/u_4 + [0.24, 0.29]/u_5$

结论 本文在指出文[1]中提出的相似度量方法的缺陷 的基础上,提出了一种新的相似度量方法,同时提出了 Vague 集间相似方向的概念。在此基础上给出的一个新的基于 Vague 集的双向近似推理方法,由于该方法考虑了 Vague 集 包含信息的精确性以及采用我们提出的相似性度量方法,从 而使得推理结果更加精确和符合实际情况。这为智能系统中 的模糊推理提供了一个十分有用的工具。

# 参考文献

- Chen S M. Measures of Similarity Between Vague Sets. Fuzzy Sets Systems , 1995, 74(2): 217~223
- 2 Chun M-G. A similarity-based Bidirectional approximate reasoning method for decision-making systems. Fuzzy Sets and Systems, 2001,117:269~278
- 3 Gau W L, Buehrer D J. Vague sets. IEEE Trans. Systems Man Cybernetics, 1993, 23 (2): 610~614
- 4 Chen S M, Hsiao W H, Jong W T. Bidirectional approximate reasoning based on interval-valued fuzzy sets. Fuzzy Sets and Systems, 1997, 41: 339~353
- 5 李凡.模糊信息处理系统.北京:北京大学科学出版社,1998
- 6 Zadeh L A. Fuzzy Sets and Their Applications to Cognitive and Decision Processes. L. A. Zadeh, et al. eds. Academic Press, New York, 1975. 1~39

小波域 HMT 模型精确地描述了小波系数间的关系,显示出良好的应用前景。但它在模型参数估计时收敛速度慢,防碍了它在实时或需要快速处理的场合的应用。本文提出了一种有效的模型参数初始化方法,应用于图像去噪的结果说明,该方法大大加快了训练速度,同时得到了良好的去噪结果。

# 参考文献

- 1 Crouse M S, Nowak R D, Baraniuk R G. Wavelet-Based Statistical Signal Processing Using Hidden Markov Models. IEEE Trans. on Signal Processing, 1998, 46(4):886~902
- 2 Dempster A P, Laird N M, Rubin D B. Maximum Likelihood from Incomplete Data via the EM Algorithm. Journal of the Royal Statistical Society, Series B, 1977, 39:1~38
- 3 Mallat S. A Wavelet Tour of Signal Processing. San Diego: Academic Press, 1998
- 4 Mallat S, Zhong S. Characterization of Signals from Multiscale Edges. IEEE Trans. on Pattern Recognition Machine Intelligence, 1992,14(7):710~732