

广义 SOM 及其在人脸性别识别中的应用

於东军^{1),3)} 吴小俊²⁾ HANCOCK Edwin R³⁾ 杨静宇¹⁾

¹⁾(南京理工大学计算机科学与技术学院 南京 210094)

²⁾(江南大学物联网工程学院 江苏 无锡 214122)

³⁾(约克大学计算机系 York YO10 5DD 英国)

摘 要 将适用于欧氏空间的经典 SOM 模型,借助于黎曼指数及对数映射,推广至一般黎曼流形空间中,得到了广义 SOM 模型(G-SOM),给出了 G-SOM 模型的序列及批量学习算法. G-SOM 能够在输出空间中保持模式在输入空间中的内蕴拓扑邻域特性,因此,当输入空间为非线性流形时,G-SOM 的性能理论上将优于 SOM. FERET 人脸图像库上的性别鉴定实验支持了理论结果.

关键词 自组织映射;广义自组织映射;黎曼流形;模式识别

中图法分类号 TP391 DOI号: 10.3724/SP.J.1016.2011.01719

Generalized SOM with Application to Facial Gender Identification

YU Dong-Jun^{1),3)} WU Xiao-Jun²⁾ HANCOCK Edwin R³⁾ YANG Jing-Yu¹⁾

¹⁾(School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing 210094)

²⁾(School of IoT Engineering, Jiangnan University, Wuxi, Jiangsu 214122)

³⁾(Department of Computer Science, University of York, York YO10 5DD, UK)

Abstract This paper generalizes the self-organizing map (SOM) in flat Euclidean space with the aid of Riemannian exponential and logarithmic maps and obtains the generalized self-organizing map (G-SOM) on Riemannian manifold. Both sequential and batch learning algorithms for the G-SOM are presented. G-SOM can preserve the intrinsic topological neighborhood structure of the input patterns on output space. Theoretical analysis shows that G-SOM will outperform SOM when the input space is a non-linear manifold. Gender identification experiments on FERET support our theoretical results.

Keywords SOM; generalized SOM; Riemannian manifold; pattern recognition

1 引 言

Kohonen^[1]提出的自组织映射(Self-Organizing Map, SOM),又称为拓扑保持映射,是一种无监督学习模型. SOM 的一个重要特性是可以在没有任何分类信息的条件下,自适应地学习模式在输入空间

中的分布特性,并在 SOM 输出空间中保持输入模式的拓扑邻域信息. 鉴于其学习规则的简单性、学习效果的有效性, SOM 已经在很多领域得到广泛应用,例如:数据挖掘^[2]、人脸识别^[3]、动态系统辨识^[4]等等.

SOM 模型默认模式的输入空间是线性欧氏空间. 亦即, SOM 在输出空间中只能保持模式在输入

收稿日期:2008-09-02;最终修改稿收到日期:2010-09-07. 本课题得到教育部科学技术研究重大项目、国家自然科学基金(60973094)、江苏省自然科学基金(BK2011177)、南京理工大学自主科研专项计划(2011YDXM19)和江苏高校优势学科建设工程项目资助. 於东军,男,1975年生,博士,副教授,主要研究方向为模式识别、神经网络. E-mail: yu.dongjun@yahoo.com.cn. 吴小俊,男,1967年生,博士,教授,博士生导师,研究领域为人脸识别、智能计算、模式识别. HANCOCK Edwin R,男,1956年生,教授,博士生导师,研究领域为计算机视觉、模式识别. 杨静宇,男,1941年生,教授,博士生导师,主要研究领域为计算机视觉、信息融合、模式识别、智能机器人.

空间中的欧氏拓扑邻域关系,而不能有效发现非欧拓拓扑邻域信息^[5].为了能够在 SOM 的输出空间中揭示模式在输入空间中的非欧拓拓扑邻域信息,有学者提出了 Kernel-SOM^[6-7]:使用核函数将输入空间中的模式变换到一个高维(甚至是无限维)的特征空间中;然后将该特征空间作为 SOM 的输入空间.然而,这种方法是核函数依赖的,使用不同的核函数,Kernel-SOM 在输出空间中保持的输入模式的拓扑邻域信息将是不同的.换句话说,Kernel-SOM 在输出空间中保持的依然不是模式在输入空间中内蕴的拓扑邻域结构.

因此,当模式所在的输入空间是非线性流形空间时,SOM 以及 Kernel-SOM 模型均不能很好地在输出空间中保持输入模式的内蕴拓扑邻域结构.文献[8]提出了基于测地距离的 SOM(GDBSOM),在学习阶段寻找 SOM 的最佳匹配单元(BMU)时,使用测地距离而不是欧氏距离,以期在输出空间保持输入模式的内蕴邻域信息.然而 GDBSOM 在进行输出节点的原型向量更新时,并不能保证更新后的原型向量依旧位于输入空间中,从而降低了在输出空间中保持模式内蕴邻域信息的能力.

本文借助于黎曼对数及指数映射,将 SOM 从线性欧氏空间推广到一般的非线性黎曼流形空间,得到广义 SOM(G-SOM). G-SOM 使用黎曼距离度量模式在输入空间中的相似程度,并在权值更新阶段保证更新后的原型向量依旧位于输入空间中,从而最大程度上在输出空间中保持输入模式的内蕴邻域结构信息.

我们在 FERET 人脸图像库上进行了人脸性别鉴定实验:使用 PGASFS^[9]技术从 2D 人脸图像重构 3D Needle-map,重构出的 Needle-map 位于黎曼流形 $S^2(n)$ 上;使用 G-SOM 对 Needle-map 进行学习及分类.实验结果表明,当模式所在的输入空间为非线性流形时,G-SOM 的性能优于 SOM.

2 自组织映射 SOM

SOM 由两层构成:第 1 层为输入层,第 2 层为输出层.通常,输出神经元在输出层以一维或是二维规则网格结构排列,并且输入神经元和输出神经元是全互联的.SOM 能够将输入空间中(欧氏空间)的模式以非线性的方式映射至输出空间,并在输出空间中保持输入模式的拓扑邻域特性.

设 SOM 有 d 个输入神经元, K 个输出神经元,称输出神经元 $i(1 \leq i \leq K)$ 的权向量 $\mathbf{w}_i \in R^d$ 为原型

向量.SOM 模型可以使用序列学习算法或是批量学习算法进行学习^[1].

2.1 序列学习算法

给定训练数据集 $\{\mathbf{x}_i | \mathbf{x}_i \in R^d\}_{i=1}^N$,SOM 序列学习算法的第 t 次学习步骤如下^[1]:

(1) 从训练数据集中随机选择一个样本作为当前学习样本 $\mathbf{x}(t)$.

(2) 从输出神经元中找到输入为 $\mathbf{x}(t)$ 时的最佳匹配单元(BMU),记为 $c(\mathbf{x}(t))$:

$$c(\mathbf{x}(t)) = \arg \min_{1 \leq i \leq K} \|\mathbf{x}(t) - \mathbf{w}_i(t)\| \quad (1)$$

即在所有输出神经元中,BMU 的原型向量与当前学习样本最为接近.

(3) 更新 BMU 及其邻域内输出神经元的原型向量:

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + \alpha(t) \cdot h_{c(\mathbf{x}(t)),i}(t) \cdot (\mathbf{x}(t) - \mathbf{w}_i(t)) \quad (2)$$

其中 $\alpha(t)$ 为学习效率,邻域函数 $h_{c(\mathbf{x}(t)),i}(t)$ 可以使用高斯型:

$$h_{c(\mathbf{x}(t)),i}(t) = \exp\left(-\frac{\|\mathbf{r}_i - \mathbf{r}_{c(\mathbf{x}(t))}\|^2}{2\sigma^2(t)}\right) \quad (3)$$

式(3)中的 \mathbf{r}_i 和 $\mathbf{r}_{c(\mathbf{x}(t))}$ 分别是输出神经元 i 和 $c(\mathbf{x}(t))$ 在输出层的坐标位置.

2.2 批量学习算法

序列学习算法中,每一步学习只将训练集中的一个样本输入 SOM,然后更新 BMU 及其邻域中输出神经元的原型向量;而批量学习算法每次学习将训练集中所有样本一次性输入 SOM,然后对所有输出神经元的原型向量更新.一轮批量学习的基本步骤简要描述如下^[1]:

(1) 划分训练数据集.依据输出神经元的原型向量集 $\{\mathbf{w}_j\}_{j=1}^K$ 构建 Voronoi 图 $\{V_j\}_{j=1}^K$;若训练样本 $\mathbf{x}_i(1 \leq i \leq N)$ 的 BMU 为输出神经元 j ,则将 \mathbf{x}_i 划分到 Voronoi 区域 V_j .

(2) 计算每个 V_j 中样本的均值.用 n_j 表示 Voronoi 区域 V_j 中学习样本的个数,样本的均值记为

$$\bar{\mathbf{x}}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} \mathbf{x}_i \quad (4)$$

(3) 更新输出神经元的原型向量.

$$\mathbf{w}_i(t+1) = \frac{\sum_{j=1}^K h_{ji}(t) \cdot n_j \cdot \bar{\mathbf{x}}_j}{\sum_{j=1}^K h_{ji}(t) \cdot n_j} \quad (5)$$

SOM 由于结构简单及有效性,已经在众多领域得到广泛应用.当数据分布在欧氏空间中时,SOM 可以较好地学习得到模式的拓扑邻域结构信息;然

而,当数据分布在一般的黎曼流形上时,直接应用 SOM 就不再适合.

3 广义 SOM: G-SOM

3.1 黎曼指数映射及黎曼对数映射

设 M 是一个黎曼流形, I 是黎曼流形 M 上的一点, $T_I M$ 表示在该点处的切平面, v 是该切平面上的一个向量 ($v \in T_I M, v \neq \mathbf{0}$), γ_I^v 为通过点 I 并沿着 v 方向的测地曲线. 那么, v 在点 I (基点) 处的黎曼指数映射^[10] $Exp_I(v)$ 是指: 将向量 v 以 I 为基点沿着测地线 γ_I^v 映射到 M 上的一点 I^* , 并且 I 与 I^* 之间的黎曼距离 $d(I, I^*)$ 为 $\|v\|$, 亦即

$$I^* = Exp_I(v) \text{ 且 } d(I, I^*) = \|v\| \quad (6)$$

黎曼对数映射^[10] 是黎曼指数映射的逆映射:

$$v = Log_I(I^*) \quad (7)$$

以二维单位球面流形 S^2 为例: 欧氏空间中的单位向量 $I \in R^3$ 可以看作是位于二维单位球面流形上的点, 亦即 $I \in S^2$. 流形 S^2 上以 I 为基点的黎曼指数映射及黎曼对数映射可以用图 1 表示.

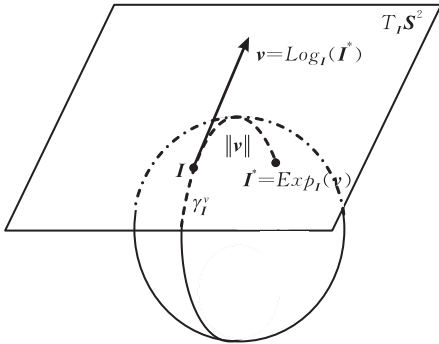


图 1 二维单位球面流形上黎曼指数及对数映射

给定 M 上的 N 个点 $\{x_1, x_2, \dots, x_N\}$, 其内蕴均值定义为

$$\begin{aligned} \bar{x} &= \text{IntrinsicMean}(x_1; x_2; \dots; x_N) \\ &= \arg \min_{x \in M} \sum_{i=1}^N d^2(x, x_i) \end{aligned} \quad (8)$$

内蕴均值可以使用基于梯度下降的迭代法进行求解^[11]:

$$\bar{x}_{t+1} = Exp_{\bar{x}_t} \left(\frac{\tau}{N} \sum_{i=1}^N Log_{\bar{x}_t}(x_i) \right) \quad (9)$$

其中, τ 为迭代步长.

若 w_i 是 x_i 的权重, 带权内蕴均值:

$$\begin{aligned} \bar{x} &= W \text{IntrinsicMean}(w_1, x_1; w_2, x_2; \dots; w_N, x_N) \\ &= \arg \min_{x \in M} \sum_{i=1}^N w_i \cdot d^2(x, x_i) \end{aligned} \quad (10)$$

可以使用下面的迭代式(11)加以求解:

$$\bar{x}_{t+1} = Exp_{\bar{x}_t} \left(\frac{\tau}{N} \sum_{i=1}^N w_i \cdot Log_{\bar{x}_t}(x_i) \right) \quad (11)$$

3.2 G-SOM

3.2.1 序列学习算法

当数据分布于弯曲的黎曼流形 M 上时, SOM 不能正确学习得到输入模式的内蕴拓扑邻域结构信息. 以流形 $M = S^2$ 为例, 若使用 SOM 序列学习算法, 那么迭代式(2)的学习结果是将原型向量 $w_i(t)$ 对应的点沿着方向 $x(t) - w_i(t)$ 移动距离 $\|\Delta\|$ 到达 $w_i'(t+1)$ 所对应的点. 这样存在的问题是: 学习后的原型向量 $w_i'(t+1)$ 所对应的点不能够保证仍然位于流形 S^2 上, 从而导致不能准确学习得到输入模式的内蕴拓扑邻域结构信息. 在这种情形下, 最佳的方法是将原型向量 $w_i(t)$ 对应的点沿着 $x(t)$ 与 $w_i(t)$ 之间的测地线移动距离 $\|\Delta\|$ 到达 $w_i(t+1)$ 所对应的点. 这一目标可以用以下 3 步来实现(如图 2 所示):

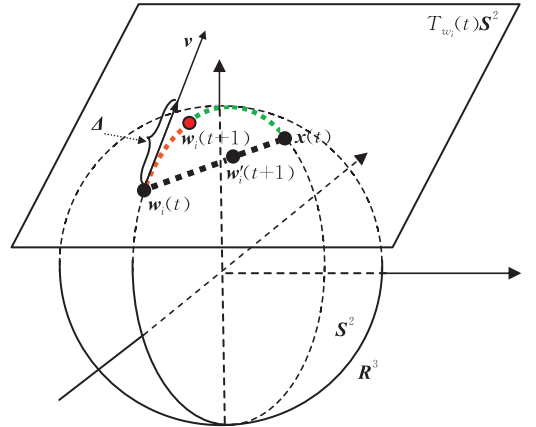


图 2 G-SOM 序列学习示意图

(a) 使用黎曼对数映射将点 $x(t)$ 映射到切平面 $T_{w_i(t)} M$ 上: $v = Log_{w_i(t)}(x(t))$.

(b) 计算在切平面上的调整量: $\Delta = \alpha(t) \cdot h_{c(x(t)), i}(t) \cdot Log_{w_i(t)}(x(t))$.

(c) 再以 $w_i(t)$ 为基点, 使用黎曼指数映射将 Δ 映射回流形 M 上.

上述 3 个步骤可以使用下面的迭代学习公式来统一描述:

$$w_i(t+1) = Exp_{w_i(t)} \left(\underbrace{\alpha(t) \cdot h_{c(x(t)), i}(t) \cdot Log_{w_i(t)}(x(t))}_{(b) \Delta} \right) \quad (12)$$

需要注意, 在确定 $x(t)$ 所对应的 BMU 时, 需要使用黎曼距离而不是欧氏距离, 即

$$c(x(t)) = \arg \min_{1 \leq i \leq k} \|Log_{x(t)}(w_i(t))\| \quad (13)$$

使用 G-SOM 的迭代学习公式(12), 一方面可

以达到将原型向量向当前学习样本移动的目的(SOM的学习目标);另一方面,亦可以保证学习得到的原型向量仍然位于输入流形上.

3.2.2 批量学习算法

当数据分布于弯曲黎曼流形上时,SOM的批量学习算法步骤中的(4)和(5)同样不能保证学习得到的原型向量所对应的点仍然能够位于流形上.为此,G-SOM的批量学习算法在SOM的批量学习算法基础上改进如下:用式(14)来代替式(4)计算每个Voronoi区域 V_j 中所有样本的内蕴均值 \bar{x}_j ;用式(15)来替换式(5)计算 $\bar{x}_j(1 \leq j \leq K)$ 的带权内蕴均值.

$$\bar{x}_j = \text{IntrinsicMean}(\mathbf{x}_1; \mathbf{x}_2; \cdots; \mathbf{x}_{n_j}) \quad (14)$$

$$\mathbf{w}_i(t+1) =$$

$$W \text{IntrinsicMean}(\omega_1, \bar{x}_1; \omega_2, \bar{x}_2; \cdots; \omega_K, \bar{x}_K) \quad (15)$$

其中, ω_j 为 \bar{x}_j 的权重,定义如下:

$$\omega_j = h_{ji}(t) \cdot n_j \quad (16)$$

实际上,G-SOM是SOM的推广;而SOM是G-SOM的一种特殊情形.以序列学习算法为例,当 M 为线性流形时,我们有

$$\text{Log}_{\mathbf{w}_i(t)}(\mathbf{x}(t)) = \mathbf{x}(t) - \mathbf{w}_i(t) \quad (17)$$

$$\text{Exp}_{\mathbf{w}_i(t)}(\Delta) = \mathbf{w}_i(t) + \Delta \quad (18)$$

由此,式(12)的右端为

$$\text{Exp}_{\mathbf{w}_i(t)}(\Delta) = \mathbf{w}_i(t) + \Delta$$

$$= \mathbf{w}_i(t) + \alpha(t) \cdot h_{c(\mathbf{x}(t)),i}(t) \cdot \text{Log}_{\mathbf{w}_i(t)}(\mathbf{x}(t)) \\ = \mathbf{w}_i(t) + \alpha(t) \cdot h_{c(\mathbf{x}(t)),i}(t) \cdot (\mathbf{x}(t) - \mathbf{w}_i(t)) \quad (19)$$

亦即

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + \alpha(t) \cdot h_{c(\mathbf{x}(t)),i}(t) \cdot (\mathbf{x}(t) - \mathbf{w}_i(t)) \quad (20)$$

式(20)完全等同于式(2),亦即,流形 M 退化为线性流形时,G-SOM的序列学习算法就等价于SOM的序列学习算法.批量学习算法亦有类似的结论,不再赘述.

4 实验及分析

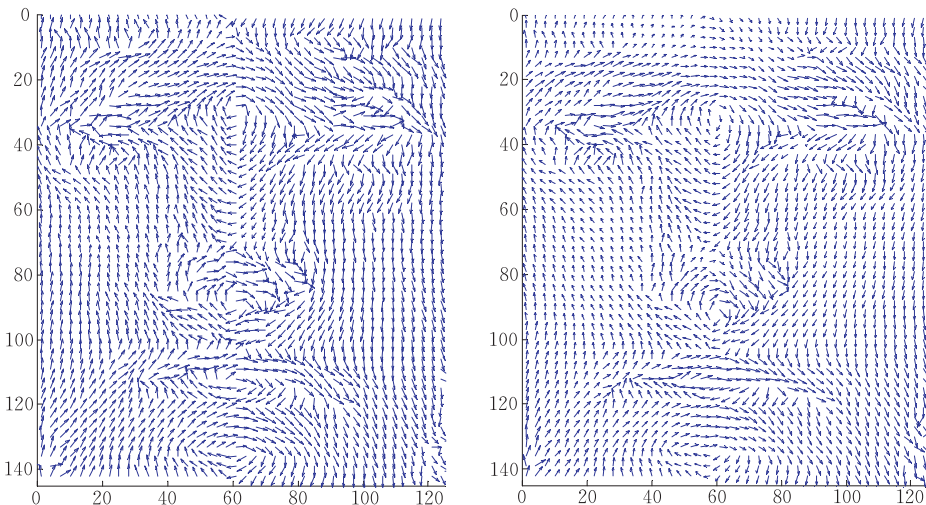
4.1 数据集及预处理

FERET^[12]数据库在人脸识别领域已经成为标准的数据库.FERET数据库中包含有大量的人脸图像分属于不同的人;每个人均有多幅包含不同姿态、视角、表情、照明的人脸图像.此外性别及种族类别信息在该数据库中也得到很好地表达.在下面的实验中,使用FERET的一个子集,该子集包含200个人的200幅图像(103个男性,97个女性,每人一副图像).数据集中的每幅图像均进行预处理:使用剪切、旋转、缩放、对准等几何运算最终得到解析度为 142×124 的图像,并对图像进行直方图均衡化.

图3(a)给出了数据集中的5幅典型的经过预



(a) 数据集中的5幅典型人脸图像



(b) Needle-map,对应图(a)中的第1和第5幅人脸图像

处理后的人脸图像. 图 3(b) 是使用 PGASFS^[9] 人脸形状恢复技术得到的对应于图(a)中第一幅和最后一幅人脸图像的 Needle-map.

4.2 $S^2(n)$ 上的黎曼指数及对数映射

我们已经知道, 给定 2D 人脸图像 $\{x_i | 1 \leq i \leq N\}$, 那么使用 PGASFS^[9] 技术, 可以得到一个对应的 3D Needle-map 集 $\{n_i | 1 \leq i \leq N\}$, 其中 $x_i \in R^{n \times 1}$ 为图像向量, 其对应的 Needle-map 为 $n_i \in R^{n \times 3}$, 并且 n_i 的第 p 行 ($1 \leq p \leq n$, n 是人脸图像像素的个数) $n_i^p \in R^{1 \times 3}$ 是单位向量, 其含义是: 图像 x_i 中第 p 个像素在人脸表面上对应点处的单位法向量.

向量 $n_i^p \in R^{1 \times 3}$ 是位于二维单位球面流形上的点, 亦即 $n_i^p \in S^2$; 因此, $n_i \in R^{n \times 3}$ 是位于流形 $S^2(n)$ 上的点, 即 $n_i \in S^2(n)$, 其中 $S^2(n) = \prod_{i=1}^n S^2$.

对于流形 S^2 , 当基点 $I = (0, 0, 1)$ 时, 那么切平面 $T_I S^2$ 上的一个向量可以记为 $v = (v_1, v_2, 0)$. 因此黎曼指数映射为^[11]

$$Exp_I(v) = \left(v_1 \cdot \frac{\sin \|v\|}{\|v\|}, v_2 \cdot \frac{\sin \|v\|}{\|v\|}, \cos \|v\| \right) \quad (21)$$

对于 S^2 上的一点 $I^* = (x, y, z)$, 当基点为 $I = (0, 0, 1)$ 时, 黎曼对数映射为

$$Log_I(I^*) = \left(x \cdot \frac{\theta}{\sin \theta}, y \cdot \frac{\theta}{\sin \theta} \right) \quad (22)$$

其中 $\theta = \arccos(z)$.

若给定的基点 $I = (x, y, z) \neq (0, 0, 1)$, 必定可以通过连续的两步旋转 (先绕 x 轴, 再绕 y 轴, 反之亦可) 运算将 I 的坐标变换为 $(0, 0, 1)$, 然后再利用上述两式得到黎曼指数和黎曼对数映射. 流形 $S^2(n)$ 上的黎曼指数映射等价于 n 个 S^2 流形上黎曼指数映射的直积, 黎曼对数映射类似^[9].

4.3 训练及识别

基于减少 SOM 输入维数及使用人脸局部特征的优势的目的^[3], 在学习阶段, 将每个 Needle-map 分割为 $L = \left\lfloor \frac{h}{h'} \right\rfloor \cdot \left\lfloor \frac{w}{w'} \right\rfloor$ 个互不重叠的子块, 其中 h 和 w 分别是图像的高度和宽度, h' 和 w' 分别是子块的高度和宽度. 每个子块是分布在流形 $S^2(m)$ 上的一个点, $m = h' \cdot w'$. 需要注意的是, 在训练阶段, 是将每个子块作为输入模式, 共有 $N \cdot L$ 个输入模式.

模型学习完毕后, 可以得到 N 个用于训练的 Needle-map 的 BMU 矩阵:

$$B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1L} \\ b_{21} & b_{22} & \cdots & b_{2L} \\ \cdots & \cdots & b_{ij} & \vdots \\ b_{N1} & b_{N2} & \cdots & b_{NL} \end{pmatrix} \quad (23)$$

其中 b_{ij} 表示将第 i 个训练 Needle-map 的第 j 个子块作为输入时 SOM 模型的 BMU.

在识别阶段, 对于一个待识别的 Needle-map, 首先将它划分为 L 个子块; 依次将这些子块输入 SOM 模型, 可得到一个 BMU 向量:

$$B' = (t_1, t_2, \cdots, t_L) \quad (24)$$

其中 t_i 是第 i 个子块作为输入时 SOM 模型的 BMU.

基于 B' 和 B , 使用 Soft-KNN 机制^[3], 可以得到一个信度矩阵, 用于度量待识别的 Needle-map 与所有训练样本在各个子块上的相似程度:

$$C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1L} \\ c_{21} & c_{22} & \cdots & c_{2L} \\ \cdots & \cdots & c_{ij} & \vdots \\ c_{N1} & c_{N2} & \cdots & c_{NL} \end{pmatrix} \quad (25)$$

其中 c_{ij} 是表示待识别的 Needle-map 与第 i 训练样本在第 j 个子块上相似的信度值.

用 M 表示训练样本中男性的集合, F 表示女性的集合, 那么测试样本属于男性或是女性的信度可以用下面两式分别求得:

$$Score_M = \frac{1}{\|M\|} \cdot \sum_{x_i \in M} \sum_{j=1}^L c_{ij} \quad (26)$$

$$Score_F = \frac{1}{\|F\|} \cdot \sum_{x_i \in F} \sum_{j=1}^L c_{ij} \quad (27)$$

依据 $Score_M$ 和 $Score_F$ 值的大小, 可以完成测试样本的性别判定.

4.4 结果及分析

为了便于比较, 在实验中, G-SOM 及 SOM 的输出层神经元统一配置为 30×20 的正六边形网格, 且均使用批量学习算法. 对于每种不同的子块配置 ($h' \times w'$), 实验进行 10 次; 每次实验中, 随机选择 50% 的样本作为学习样本 (为降低学习过程的偏向性, 在随机选择学习样本时, 同时需兼顾男女性样本数的平衡), 识别正确率取 10 次实验结果的平均值.

表 1 中给出了 G-SOM 与 SOM 在不同子块配置下识别结果的比较. 图 4 则给出了 G-SOM 在子

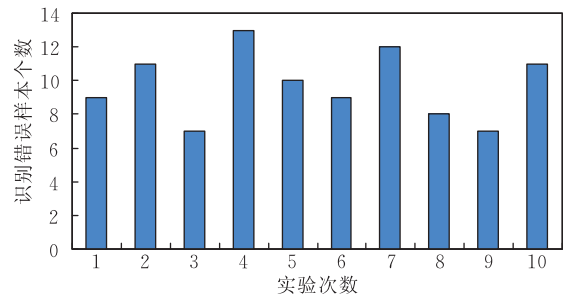


图 4 G-SOM 在子块配置为 4×4 时 10 次实验的错误识别样本个数分布

块配置为 4×4 时 10 次实验的错误识别样本个数分布情况。

表 1 G-SOM 与 SOM 的性能比较

子块大小 $h' \times w'$	G-SOM		SOM	
	错误 识别个数	识别 正确率/%	错误 识别个数	识别 正确率/%
4×4	97	90.3	162	83.8
8×8	124	87.6	173	82.7
16×16	153	84.7	212	78.8
32×32	175	82.5	268	73.2

注:表中错误识别个数是指 10 次实验中累计识别错误的样本个数;识别正确率是指 10 次实验的平均识别正确率。

从表 1 可以看出,在子块配置为 4×4 时,G-SOM 及 SOM 的识别正确率均为最高,此时,G-SOM 的识别正确率为 90.3%,比 SOM 的识别正确率高 6.5 个百分点。同时,可以发现:随着子块规模逐步变大,G-SOM 和 SOM 的识别率都呈下降趋势,但是在相同子块配置下,G-SOM 的识别率总是高于 SOM 的识别率。图 4 则给出了 G-SOM 在子块配置为 4×4 时 10 次实验的错误识别样本个数分布。观察图 4 可以发现,10 次实验中,每次识别错误的样本个数差异不是很大,表明 G-SOM 的学习结果是比较稳定的。

在实验中,模式分布的输入空间是弯曲的黎曼流形。G-SOM 能够在输出空间中保持模式在输入空间中的内蕴拓扑分布特性;而 SOM 总是假设输入空间是线性的欧氏空间,性能劣于 G-SOM 在理论上是可以预见的,而实验结果完全支持我们之前的理论分析。

我们还在同样的数据集(恢复得到的 200 幅 Needle-map)上实现了 PCA+LDA^[13] 及 PGA+LDA^[14] 的性别识别方法。实验同样进行 10 次,每次实验中,随机选择 50% 的样本作为学习样本,识别正确率取 10 次实验结果的平均值。表 2 给出了对比结果。

表 2 G-SOM 与其它方法的性能比较

	错误识别个数	识别正确率/%
SOM	162	83.8
PCA+LDA	143	85.7
G-SOM	97	90.3
PGA+LDA	76	92.4

文献[14]提出的 PGA+LDA 方法是基于人脸形状信息的性别识别的最新研究成果,G-SOM 的识别正确率(90.3%)非常接近 PGA+LDA 方法的正确率(92.4%),并且明显地优于通常的 PCA+LDA 方法。这表明,将 G-SOM 用于人脸性别识别

是一种有效且可行的方法。

5 结 论

本文将经典的适用于线性输入空间的 SOM 算法推广到一般黎曼流形输入空间得到 G-SOM,给出了 G-SOM 序列学习算法及批量学习算法。理论分析表明:G-SOM 是 SOM 的推广;SOM 是 G-SOM 的特殊情形。G-SOM 能够依据模式的实际输入空间特性,在输出空间中保持输入模式的内蕴拓扑分布特性;SOM 则总是假设模式的输入空间是线性流形。因此,当输入模式位于非线性流形上时,SOM 势必劣于 G-SOM。在基于人脸形状的性别识别上的实验结果支持了上述结论。

需要指出的是,对于 G-SOM 的序列学习算法及批量学习算法而言,重要的一个步骤是需要知道输入模式所在流形之上的黎曼指数映射及黎曼对数映射。如果能够准确知道输入模式位于何种流形之上(如文中的 Needle-map 必定位于流形 $S^2(n)$ 上),应用 G-SOM 是自然的。但是,在很多应用中,往往只能知道输入模式位于某个特定的流形上,却并不能准确知道是何种流形;也就是说,我们不能显式给出黎曼指数映射及黎曼对数映射。这时,应用 G-SOM 就会存在困难。如何根据训练模式集找到或是逼近数据所在流形的黎曼指数映射及黎曼对数映射,这是我们需要进一步研究的工作。

参 考 文 献

- [1] Kohonen T. The self-organizing map. Proceedings of IEEE, 1990, 78(1): 1464-1480
- [2] Kohonen T, Kaski S, Lagus K, Salojarvi J, Honkela J, Paatero V, Saarela A. Self organization of a massive document collection. IEEE Transactions on Neural Networks, 2000, 11(3): 574-585
- [3] Tan Xiao-Yan, Chen Song-Can, Zhou Zhi-Hua Zhang Fu-Yan. Recognizing partially occluded, expression variant faces from single training image per person with SOM and soft KNN ensemble. IEEE Transactions on Neural Networks, 2005, 16(4): 875-886
- [4] Barreto Guilherme A, Araujo Aluizio F R. Identification and control of dynamical using the self-organizing map. IEEE Transactions on Neural Networks, 2004, 15(5): 1244-1259
- [5] Andras P. Kernel-Kohonen networks. International Journal of Neural System, 2002, 12(2): 117-135
- [6] Manuel A M, Alberto M. Extending the SOM algorithm to non-Euclidean distances via the kernel trick//Proceedings of

the 11th International Conference on Neural Information Processing, Calcutta, India, 2004: 150-157

- [7] Pan Zhi-Song, Chen Song-Can, Zhang Dao-Qiang. A kernel-based SOM classification in input space. *Chinese Journal of Electronics*, 2004, 32(2): 227-231(in Chinese)
(潘志松, 陈松灿, 张道强. 原空间中的核 SOM 分类器. *电子学报*, 2004, 32(2): 227-231)
- [8] Shi Chun-Qi, Zhang Su-Lan, Zheng Zheng, Shi Zhong-Zhi. Geodesic distance based SOM for image clustering//Proceedings of the International Conference on Sensing, Computing and Automation. Chongqing, China, 2006: 2483-2488
- [9] Smith William, Hancock Edwin R. Facial shape-from-shading and recognition using principal geodesic analysis and robust statistics. *International Journal of Computer Vision*, 2008, 76(1): 71-91
- [10] Pennec X. Probabilities and statistics on Riemannian manifolds; Basic tools for geometric measurements//Proceedings

of the IEEE Workshop on Nonlinear Signal and Image Processing. Antalya, Turkey, 1999: 194-198

- [11] Fletcher P T, Lu Cong-Lin, Pizer S M, Joshi S. Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Transactions on Medical Imaging*, 2004, 23(8): 995-1005
- [12] Phillips P J, Moon H, Rauss P J, Rizvi S. The FERET evaluation methodology for face recognition algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(10): 1090-1104
- [13] Buchala Samarasena, Dacey Neil, Gale Tim M. Dimensionality reduction of face images for gender classification. Department of Computer Science, The University of Hertfordshire, UK; Technical Report 408, 2004
- [14] Wu Jing, Smith William, Hancock Edwin R. Facial gender classification using shape-from-shading. *Image and Vision Computing*, 2010, 28(6): 1039-1048



YU Dong-Jun, born in 1975, Ph. D., associate professor. His research interests include pattern recognition, neural network and fuzzy system.

WU Xiao-Jun, born in 1967, Ph. D., professor, Ph. D. supervisor. His research interests include pattern recogni-

tion, intelligent computation, image processing and artificial intelligence.

HANCOCK Edwin R, born in 1956, professor, Ph. D. supervisor. His research interests include computer vision and pattern recognition.

YANG Jing-Yu, born in 1941, professor, Ph. D. supervisor. His research interests include pattern recognition, robot vision, image processing, data fusion, and artificial intelligence.

Background

The self-organizing map has been widely used in many fields. A defect of the classic SOM is that it supposes the input space always be a Euclidean space thus can not reveal the intrinsic topological neighborhood structure of the input patterns when the input space is a non-linear manifold. This paper aims to generalize the classic SOM to a non-linear manifold input space. The core idea underlying the generalized SOM is to move the codebook vector closer to the learning sample along a geodesic. The obtained G-SOM is the generalization of the classic SOM, and the classic SOM is the spe-

cial case of the G-SOM. Theoretical and experimental results demonstrate that the G-SOM outperforms the classic SOM when data lying on curved manifold space.

The work is supported by the Key grant Project of Chinese Ministry of Education, the National Natural Science Foundation(Grant No. 60973094), the Natural Science of Jiangsu Province (Grant No. BK2011177), the NUST Research Funding (Grant No. 2011YDXM19), and A Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions.