第16卷 第2期 自动化学报 Vol. 16, No. 2
1990年3月 ACTA AUTOMATICA SINICA Mar., 1990

基于类别可分离性的遥感图象特征提取方法

谭枫 曾小明
（南京林业大学遥感实验室）

摘 要

本文从类别可分离性出发推导出一种用于遥感图象特征提取的线性变换方法．该方法并不着眼于变换子空间的整体信息保持，而在于使当前待分类别具有最好的分离度．文中给出了变换核的理论推导及计算方法．试验结果表明该方法具有良好处理效果．

关键词——特征提取，线性变换，分离度，本征值和本征向量．

一、采用变换方法进行特征提取的数学模型

1. 一般模型

设图象 X 的取样大小为 $N\times N$，变量维数为 p．由于特征提取仅对变量进行映射变换，因此把图象空间 $X^p_{N\times N}$ 简记为 X．

对于 p 维变量空间任意一组规范正交基 $\{t^{(j)}\}(j=1,2,\cdots,p)$，均可得到 X 的展开式

$$X=\sum_{j=1}^{p} y_j t^{(j)}=TY.$$

其中 $T=(t^{(1)},t^{(2)},\cdots,t^{(p)})$，$t^{(i)}=(t_{1i},t_{2i},\cdots,t_{pi})^T$ 展开式分量

$$y_j=\sum_{k=1}^{p} t_{kj}x_k.$$

只要变换阵 T 是完备的，p 个展开式分量就可以代表原始图象空间中的所有信息．若取其 m 个展开式分量，则得部分和式

$$X(T,m)=\sum_{i=1}^{m} y_i t^{(i)},\ (1\leqslant m\leqslant p).$$

可以用部分和式来表现原始 X 的信息．问题是如何选取 m 个规范正交基

$$t^{(j)}\ (j=1,2,\cdots,m),$$

使得变换空间中的部分分量所包含的原图 X 的信息能够满足后续处理的要求．

本文于1987年3月24日收到．

2. 基于可分离性线性变换的数学推导

对于 p 维随机向量 $\boldsymbol{x}=(x_1, x_2, \cdots, x_p)^T$，其

均值向量 $E(\boldsymbol{x})=(\mu)_{p\times 1}$;

协方差矩阵 $V(\boldsymbol{x})=(\Sigma_x)_{p\times p}$.

令 $\boldsymbol{y}$ 是 $\boldsymbol{x}$ 的线性函数，即有

$$\boldsymbol{y}=a_1x_1+a_2x_2+\cdots+a_px_p=a^T\boldsymbol{x}.$$

其中 $a^T=(a_1,a_2,\cdots,a_P)$.

假设有 k 个总体，第 i 个总体的均值向量为 $\mu_{(i)}$，协方差阵为 $\Sigma_x^{(i)}$. 挑选方向矢量 a 的标准是能够使得各类样本中心在变换空间中分离度尽可能的大. 定义分离度为离差平方和，则选择 a 使得

$$Q=\sum_{j=1}^{k}(a^T\mu_{(j)}-\frac{1}{k}\sum_{i=1}^{k}a^T\mu_{(i)})^2=Q_{\max}.$$

在上式中若不对 a 作任何限制，则 $\|a\|$ 可任意大，求 $Q_{\max}$ 就失去意义. 因此给出约束条件

$$a^T\left[\sum_{i=1}^{k}(\Sigma_x^{(i)})\right]a=1.$$

记 $\sum_{i=1}^{k}(\Sigma_x^{(i)})=\Sigma_k$，则约束条件为 $a^T\Sigma_ka=1$.

令

$$B=u^T\left(I_k-\frac{1}{k}J_k\right)u$$

$$=\sum_{j=1}^{k}\left(\mu_{(j)}-\frac{1}{k}\sum_{i=1}^{k}\mu_{(i)}\right)\left(\mu_{(j)}-\frac{1}{k}\sum_{i=1}^{k}\mu_{(i)}\right)^T.$$

其中 I_k 为 k 阶单位向量，J_k 为元素全为 1 的 k 阶方阵，$u^T=(\mu_{(1)},\mu_{(2)},\cdots,\mu_{(k)})$. 则

$$Q=\sum_{j=1}^{k}\left(a^T\mu_{(j)}-\frac{1}{k}\sum_{i=1}^{k}a^T\mu_{(i)}\right)^2$$

$$=\sum_{j=1}^{k}\left[\left(a^T\mu_{(j)}-\frac{1}{k}\sum_{i=1}^{k}a^T\mu_{(i)}\right)\left(a^T\mu_{(j)}-\frac{1}{k}\sum_{i=1}^{k}a^T\mu_{(i)}\right)^T\right]$$

$$=\sum_{j=1}^{k}a^T\left[\left(\mu_{(j)}-\frac{1}{k}\sum_{i=1}^{k}\mu_{(i)}\right)\left(\mu_{(j)}-\frac{1}{k}\sum_{i=1}^{k}\mu_{(i)}\right)^T\right]a$$

$$=a^TBa.$$

所以问题即为

$$\begin{cases}a^TBa\\ \text{条件 } a^T\Sigma_ka=1\end{cases}\xrightarrow{\text{确定 } a}Q_{\max}.$$

利用 Langalana 乘子法求 $Q_{\max}$:

设 $$\varphi_1=a^TBa-\lambda(a^T\Sigma_ka-1).$$

令
$$\frac{\partial \varphi_1}{\partial a} = 2Ba - 2\lambda\Sigma_k a = 0,$$
则 a 应满足
$$(B - \lambda\Sigma_k)a = 0, \ (\Sigma_k^{-1}B - \lambda I)a = 0.$$
其中 Σ_k^{-1} 是 Σ_k 的逆阵，λ 是 $\Sigma_k^{-1}B$ 的本征值，a 是相应的本征向量.

由于 B, Σ_k 为半正定阵，则 $\Sigma_k^{-1}B$ 的 P 个本征值满足
$$\lambda_1 \geqslant \lambda_2 \geqslant \lambda_3 \geqslant \cdots \geqslant \lambda_p \geqslant 0.$$
相应的 p 个本征向量 $a_{(1)}, a_{(2)}, \cdots, a_{(p)}$. 其中不妨设 $a_{(i)}$ 满足
$$a_{(i)}^T a_{(j)} = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases}$$
定义分离度
$$\Delta(a) = \max_{a^T\Sigma_k a = 1} Q(a).$$
则第一分量分离度
$$\Delta(a_1) = \lambda_1,$$
第一 F 分量
$$y_1 = a_{(1)}^T x = \sum_{i=1}^{p} a_{i1}x_i.$$

设已求得前 $m-1$ 个 F 分量，现欲求第 m 个 F 分量 $(2 \leqslant m \leqslant p)$. 为了尽可能消除数据冗余性，令 $y = a^Tx$ 与前面已求得的 $m-1$ 个分量不相关.

即
$$E\{[y - E(y)][y_i - E(y_i)]^T\} = 0.$$
由于
$$\begin{aligned} E\{[y - E(y)][y_i - E(y_i)]^T\} &= E\{[a^Tx - E(a^Tx)][a_{(i)}^Tx - E(a_{(i)}^Tx)]^T\} \\ &= a^TE\{[x - E(x)][x - E(x)]^T\}a_{(i)} \\ &= a^T\Sigma_x a_{(i)}, \ i = 1, 2, \cdots, m-1. \end{aligned}$$
因为 $\Sigma_x a_{(i)} = \gamma_i a_{(i)}$，其中 γ_i 为常量，

所以
$$E\{[y - E(y)][y_i - E(y_i)]^T\} = 0$$
等价于
$$a^Ta_{(i)} = 0, \ i = 1,2,\cdots,m-1.$$
则有
$$\begin{cases} a^TBa \\ a^T\Sigma_k a = 1 \\ a^Ta_{(i)} = 0 \end{cases} \xrightarrow{\text{确定 } a} Q_{\max}, \ i = 1,2,\cdots,m-1.$$
由 Langalana 乘子法

令
$$\varphi_m = a^TBa - \lambda(a^T\Sigma_k a - 1) - 2\sum_{i=1}^{m-1}\gamma_i a^T a(i),$$
其中 λ, γ_i 是待定系数，

又令
$$\frac{\partial \varphi_m}{\partial a} = 2Ba - 2\lambda\Sigma_k a - 2\sum_{i=1}^{m-1}\gamma_i a_{(i)} = 0,$$

左乘 $a_{(l)}^T$ 得

$$a_{(l)}^T B a - \lambda a_{(l)}^T \Sigma_k a - 2\sum_{i=1}^{m-1}\gamma_i a_{(i)}^T a_{(i)} = 0.$$

由于 $a_{(l)}$ 是 $\Sigma_k^{-1} B$ 相应的第 l 个本征值 λ_l 所对应的本征向量，根据同时对角化定理可得

$$\begin{aligned} a_{(l)}^T B a - \lambda a_{(l)}^T \Sigma_k a &= a^T B a_{(l)} - \lambda a^T \Sigma_k a_{(l)} \\ &= a^T \alpha_1 a_{(l)} - \lambda \alpha_2 a^T a_{(l)} = 0. \end{aligned}$$

其中 λ, α_1, α_2 均为系数。

又因为　$a_{(l)}^T a_{(l)} = 1$,

故有　$\gamma_l = 0.$

所以　$Ba - \lambda\Sigma_k a = 0.$

第 m 分量分离度　$\Delta(a_m) = \lambda_m.$

相应的变换系数即为 $\Sigma_k^{-1} B$ 的相应于 λ_m 的本征向量

$$a_{(m)} = (a_{1m}, a_{2m}, \cdots, a_{pm})^T.$$

第 m 个 F 分量　$y_m = a_{(m)}^T \boldsymbol{x}.$

综合以上推导有：

对于 P 元随机向量 $\boldsymbol{x} = (x_1, x_2, \cdots, x_p)^T$，设其均值向量为 μ，协方差阵为 Σ_x；$A = (a_{(1)}, \cdots, a_{(p)})$，$\Lambda = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_p \end{pmatrix}$ 分别为 $\Sigma_k^{-1} B$ 的本征矢量和本征值矩阵。其中

$$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_p \geqslant 0;$$

$Ba_{(j)} - \lambda_j \Sigma_k a_{(j)} = 0$，并且 $Y = (y_1, y_2, \cdots, y_p)$ 的各分量互不相关，第 j 个分量

$$y_j = a_{(j)}^T X$$

在与 $y_1, y_2, \cdots, y_{j-1}$ 不相关的所有线性组合中，各总体的可分离性最好，且分离度

$$\Delta(a_{(j)}) = \lambda_j.$$

定义. 判别总贡献 $\Delta(A) = \sum_{i=1}^{p} \lambda_i$,

第 i 维 F 分量的判别贡献率 $\Delta(a_i) = \lambda_i / \Delta(A)$，前 m 个 F 分量的累计贡献率

$$\Delta(a_{1-m}) = \sum_{i=1}^{m} \lambda_i / \Delta(A).$$

二、试 验 结 果

利用上述的线性变换方法，我们对陆地卫星图象进行处理，结果见表 1。

表 1 根据十大类地物训练样本计算所得的变换矩阵

本征值	46.90191	19.59429	1.07224	0.34297
本征向量	0.16673 0.51968 0.21683 1.08215	−0.33777 −0.56302 −0.02104 0.56543	−0.90532 0.37385 0.10114 −0.06482	0.14632 −0.25118 0.41541 −0.29057
贡献率	69.06%	28.85%	1.58%	0.51%
累计贡献率	69.06%	97.91%	99.49%	100.00%
分类结果	取第 1 个分量	取前 2 个分量	取前 3 个分量	取 4 个分量
总样本 3734	正确 1859	正确 3456	正确 3529	正确 3536

参 考 文 献

[1] Fu, K.S., Min, P. J., and Li, T.J., Feature Selection in Pattern Recognition, *IEEE Trans. Syst. Sci. Cybern.*, SSC-6 (1), 1970.

[2] 程民德,沈燮昌等,图象识别导论,上海科技出版社, 1983.

A FEATURE EXTRACTION ALGORITHM FOR REMOTE SENSING IMAGE CLASSIFICATION BASED ON CLASS SEPARABILITY

TAN FENG ZENG XIAOMING

(*Nanjing Forestry University*)

ABSTRACT

This paper develops a linear transformation algorithm for feature extraction based on class separability. The algorithm takes the separability of the current classes as objective function instead of maintaining perfect information of the original image. The derivation of the transformation kernel function and the computational method are given. Experimental results show the effectiveness of the algorithm presented.

Key words ——Feature extraction; linear transformation; separability; eigenvalue and eigenvector.