

基于 MSD 的图像感兴趣区域自动提取方法研究

刘红霞¹, 谭璐², 吴翊²

(1. 烟台大学数学与信息科学系, 烟台 264005; 2. 国防科技大学数学与系统科学系, 长沙 410073)

摘要: 将单幅图像数据进行分割, 获得高维化后的数据集, 再依据图像数据的最优分解来提取不同图像块之间的数字关联, 利用多维尺度分析(MDS)方法来获取单幅图像数据不同块之间的低维表示。通过对此低维表示的自动分析, 便可获得图像感兴趣区域的精确位置。实例验证了方法的可行性、有效性。

关键词: 图像感兴趣区域; 高维化; 最优分解; 多维尺度分析

Research on Region-of-interest Image Auto-selection Based on MSD

LIU Hongxia¹, TAN Lu², WU Yi²

(1. Dept. of Mathematic and Information Science, Yantai University, Yantai 264005;

2. Dept. of Mathematic and System Science, National University of Defence Technology, Changsha 410073)

【Abstract】 A new method on the region-of-interest image is presented, the image is parted. By the greatest expression of the partition, the relations of the different partitions are gained. At last, according to the MDS, the low-dimension is achieved. Through analyzing the low-dimension data, the precision location of the region-of-interest can be denoted. The examples validate the feasibility and validity.

【Key words】 Region-of-interest; High-dimension; Greatest expression; Multidimensional scaling

感兴趣区域提取作为数据压缩、目标识别、图像分类等方面的一个预处理过程, 在实际数据处理中有着重要的应用。但在实际处理过程中, 由于图像区域“感兴趣”的程度难以度量, 因此图像感兴趣区域的获得经常是依赖于人工判读。这必然造成了判读过程中具有很大程度的主观性, 并且使得结果强烈地依赖于人们的经验认识。

通常“感兴趣区域”提取的出发点是将图像作为三维空间中的一个二维曲面, 提取的方式是点点判断、点点提取^[1,2], 这必然造成直接度量每点感兴趣程度的困难, 由于一点的信息量总是十分稀少的, 若还考虑噪声的影响, 将使得问题难上加难。即使人们可以考虑图像中每点与它的邻域之间的关联, 并通过点的邻域来判断它的感兴趣程度^[3], 但仍然没有摆脱点点考虑的范畴, 从而没有摆脱度量每点感兴趣程度的困难。为了较好地避免这一困难的出现, 不妨将考虑的对象从图像中的点转变为图像中的块, 那么此时便可将需要考虑的点间的关联转化为不同图像块之间的关联, 使得问题变得更加自然(由于我们提取的是图像中感兴趣的区域, 即感兴趣块, 而非感兴趣点)。同时, 基于块与块之间关联的考虑, 还避免了直接度量图像中点的感兴趣程度的问题, 因为只要通过不同块之间的关联, 找到最特别的块即可。

1 图像数据的最优分解

由于图像块的特性与图像的特性是一致的, 因此在不引起误解的情形下, 本节不区分图像块和图像。通常的图像处理技术都是将图像作为二元函数, 利用周围的点来表示此点或先使用二元函数基去逼近原图, 如傅里叶变换、小波变换等, 然后通过函数基的自然扩展来抓住原图的特性。通常这些基都具有可分离的形式, 图像中特征的变化体现在函数系数的改变上。但对于具体的图像数据, 这种事先选取固定基的方法有一定的局限性, 一方面处理效果不一定好, 另一方

面一般需要选取较多的项。因此有大量的文献在研究既能有效地表示图像特征, 又能具有较少项的图像数据表示新方法。这里针对图像特征的多样性, 给出了一种图像最优分解方法^[5], 它是一种特征向量方法, 基本思想是将二维图像自适应地分解为一维向量的有效综合。由于自适应性和较少参数表示, 使其可以获得较好的分解效果。事实上若选取同样个数的基向量, 那么在均方误差意义下, 这种特征向量表示是最优的, 并且从计算上来讲, 它对应于寻求半正定矩阵前最大的 k ($k = rank(X)$, X 代表图像)个特征向量和特征值的问题。并可证明在连续条件下, 当采样无限加细时这一结果可收敛到相应连续的情形^[5,6]。利用这一分解过程, 可以获得图像的基特征描述, 然后利用不同图像基特征的差异来体现它们之间的关联。为简单起见, 本文只选择具有最大分解系数或能量的基。

1.1 处理模型

图像数据进行最优分解的数学模型如下:

设 $F(x_i, y_j)$ (代表原图的离散采样) 为 $F(x, y)$ (代表原图的连续情形) 的采样点, $i=1,2,\dots,M$ $j=1,2,\dots,N$, 寻求 $f(x_i), g(y_j)$ 使得

$$L(f, g) = \min \left\{ \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (F(x_i, y_j) - f(x_i)g(y_j))^2 \right\}$$

若令

基金项目: 国家自然科学基金资助项目(60003013); 高等学校博士学科点专项科研基金资助项目(20049998008)

作者简介: 刘红霞(1974-), 女, 讲师、硕士, 主研方向: 图论, 高维数据处理, 数据融合; 谭璐, 博士生; 吴翊, 教授、博导

收稿日期: 2005-12-14 **E-mail:** mqy7174@sina.com

$$F = (F(x_i, y_j))_{M \times N}, f = (f(x_1), f(x_2), \dots, f(x_N)),$$

$$g = (g(y_1), g(y_2), \dots, g(y_N))$$

为方便起见,不妨对 f, g 进行规范化(指 $\|f\| = \|g\| = \sqrt{N}$),规范化后的结果仍记为 f, g 。则问题转化为寻求规范化的 f, g , 及 λ 使得

$$L(f, g, \lambda) = \min \left\{ \left\| F - \lambda f g^T \right\|^2 \right\} \quad (1)$$

$$\text{其中 } \|F\|^2 = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N F(x_i, y_j)^2, (F \in R^{M \times N})$$

称满足式(1)的 f, g, λ 为图像 F 的一步最优分解结果。 λ 为一步最优分解系数, f, g 为一步最优分解因子。

1.2 图像数据的最优分解

定理 1^[5] 给定图像 $F_{M \times N}$, 则存在相应的最优分解 $\{\lambda_i\}$ 、

$$\{f_i\}、\{g_i\}, \text{使得 } F = \sum_{i=1}^k \lambda_i f_i g_i^T, \text{其中 } \{\lambda_i\} \text{ 为 } i \text{ 步分解系数,}$$

$\{f_i\} \{g_i\}$ 为相应的 i 步分解因子 ($i = 1, 2, \dots, k$), $k = \text{Rank}(F)$ $\text{Rank}(F)$, 则有下述结论成立:

(1) $\{f_i\}$ 为 FF^T 的特征向量;

(2) $\{g_i\}$ 为 $F^T F$ 的特征向量;

(3) $\{\lambda_i\}_{i=1}^k$ 单降, 且对 $\forall 1 \leq i \leq k$ $MN\lambda_i^2$ 为 $F^T F$ 的非零特征值;

(4) $\{f_i g_i^T\}_1^k$ 之间为标准正交的, 且 $\langle F, f_i g_i^T \rangle = \lambda_i$, ($i = 1, 2, \dots, k$);

(5) 第 p 步近似误差为

$$L_k = \left\| F - \sum_{i=1}^k \lambda_i f_i g_i^T \right\|^2 = \|F\|^2 - \sum_{i=1}^k \lambda_i^2 \quad (p \leq k)$$

其中在 $R^{M \times N}$ 中定义:

$$\langle A, B \rangle = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N A_{ij} B_{ij}, \quad \|A\|^2 = \langle A, A \rangle, \quad (\forall A, B \in R^{M \times N})$$

定理 1 给出了图像数据的最优分解基的结果, 以及相应的求解方法。可以看出图像的最优分解基 $\{f_i, g_i\}$, $i = 1, 2, \dots, k$, 一定程度上展现了图像数据的本征特性, 从而不同 $\{f_i, g_i\}$ 之间的关联反映了不同图像之间的关联程度, 即 X, Y 为两幅图像, 它们的最优基为 $\{f_i^X, g_i^X\}$ 和 $\{f_i^Y, g_i^Y\}$, 那么 X 和 Y 之间的关联程度可通过

$$R(X, Y) = \sum_{i=k_1}^{k_2} \left(\left| f_i^X \cdot f_i^Y \right| + \left| g_i^X \cdot g_i^Y \right| \right), \quad k_1, k_2 \text{ 为给定的}$$

正整数来表示, 为简单起见, 在后面的探讨中选择 $k_1 = 1$ 或 2 , $k_1 = k_2$, 即

$$R(X, Y) = \left| f_1^X \cdot f_1^Y \right| + \left| g_1^X \cdot g_1^Y \right|, \text{ 或}$$

$$R(X, Y) = \left| f_2^X \cdot f_2^Y \right| + \left| g_2^X \cdot g_2^Y \right|$$

2 关联性的低维表示和可视化

依据第 1 节的探讨, 关联矩阵 $R = (R(X_i, X_j))$ 可看作是数据集 $\{X_i\}$ 的一种信息结构。从而依据关联性来探求特别点的过程, 可通过降维方式自然地获得。即首先探求数据集信息结构的合适低维(如二维)表示, 然后将其表示在低维欧氏空间之中, 此时明显地远离中心的点是特别的; 另外若大多数的点都落于某一低维流形之上, 那么远离数据流形的点是特别的点。当然对“特别的点”的看法不同, 所获得的结果可能是很不一样的。由于本文的探讨旨在说明方法的可行性、有效性, 因此为了方便起见, 本文将远离数据中心的点(指的是图像块)认为是特别的点, 即感兴趣区域。

按照关联矩阵来寻找特别点的过程如下:

- (1) 高维化图像, 获得高维空间中的数据集合, 见 2.1 节;
- (2) 计算数据集的信息结构, 即关联矩阵;
- (3) 使用 MDS 方法来获得数据集信息结构的低维表示;
- (4) 寻求低维表示中远离中心的点作为图像的感兴趣区域。

2.1 图像数据的高维化

对图像 $(M \times N)$ 进行分割, 记为 $P[C, L_h \times L_l]$ 。它将原图等分为 C 块, 每块的大小为 $L_h \times L_l$, 可获得 $L_h \times L_l$ 维空间中容量为 C 的数据集合。显然分割过于稠密或过于稀疏都是不可取的, 过于稀疏使获得的高维数据点过于稀少, 由于维数灾难的原因使得降维过程没有意义; 同样过于稠密, 使得数据的信息结构对数据的微小改变非常敏感, 从而难于提取准确的本征结构。故可以想象根据图像感兴趣区域的大小, 存在一个最优分割 P^* , 使得在这一分割之下可以获得最优的效果。但就实际问题而言, 人们经常是在准确性和计算简单性之间进行折中, 即根据感兴趣区域的先验信息来确定分割 $L_h \times L_l$ 的大小。在第 3 节的实例之中, 我们根据感兴趣区域的大小选取 $L_h = L_l = 20$ 和 $L_h = L_l = 40$ 。

2.2 多维尺度分析(MDS)

多维尺度分析(Multidimensional Scaling, MDS)^[7,8]是一种传统的寻求保持数据点之间差异性(或相似性)的降维方法, 它使得在原数据集合中相近的点仍然靠在一起, 远离的点仍然远离。MDS的可保持数据点之间差异的特性, 使得其可广泛应用于社会学、经济学、心理学、教育学等研究领域。在感兴趣区域提取中, 我们便是利用了MDS的这一特性。

MDS 的基本出发点是数据点间的差异性(或相似性)描述。记两点 x_i 与 x_j 之间的差异度(或相似度)为 $d(x_i, x_j)$, 简记为 d_{ij} , $i, j = 1, 2, \dots, N$ 。若实际中我们感兴趣的数据特性是它们之间有什么不同, 那么 $N \times N$ 差异矩阵(如距离矩阵) (d_{ij}) 便完整地概括了我们希望从数据集合中所获得的感兴趣信息, 如图 1 所示。

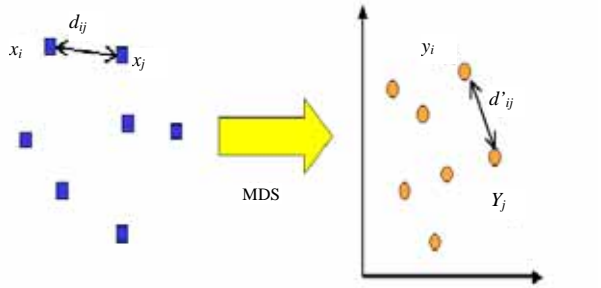
MDS 方法通过 Stress 函数,

$$\text{Stress}(\{y_i\}_{i=1}^N, f) = \sqrt{\frac{\sum_{i,j=1}^N (f(d_{ij}) - d_{ij}^*)^2}{\text{尺度因子}}},$$

其中, 典型的尺度因子为 $\sum_{i,j=1}^N d_{ij}^2$, f 为距离 d_{ij} 的量化函数,

$d_{ij} = d(x_i, x_j)$, $d_{ij}^* = d(y_i, y_j)$, $\{y_i\}_{i=1}^N$ 为 X 在 d 维空间中

的投影,来量化低维数据集对原数据集的“忠实”的程度,不同的 Stress 函数引出了不同的 MDS 方法,本文为方便起见,选取 f 为恒等变换。当 Stress 函数确定后,就可以寻求最优的低维投影,即能够最小化 Stress 函数的 $\{y_i\}_{i=1}^N$, 一般通过最小化 Stress 函数的平方来获得 $\{y_i\}_{i=1}^N$ 。这个集合可以通过标准的梯度下降法求解,即先给出 $\{y_i\}_{i=1}^N$ 的初始值,然后沿着 Stress 函数下降最快的方向去调整,直至 $\{y_i\}_{i=1}^N$ 不再明显改变为止。



非欧空间中点与点之间的关系 低维欧氏空间中对应的点与点之间的关系

图 1 MDS 保持数据点间距离的过程

2.3 感兴趣区域提取实例

这一节通过单目标图像和多目标图像(脸谱图像)的处理实例,表明基于数据集信息结构的感兴趣区域提取方式,可有效地提取图像数据中感兴趣的信息,获得图像感兴趣区域的精确位置。详见图 2~图 5。

例 1 原图大小为 120×160 , 取 $L_h = L_l = 20$, 共分割为 48 幅图像, 选取 $k_1 = k_2 = 1$ 。



图 2 原图

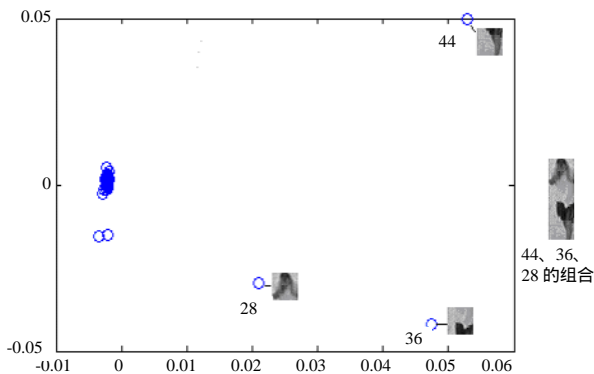


图 3 数据集信息结构的二维表示(44、36、28 为远离中心的 3 点)

例 2 原图大小为 300×250 , 取 $L_h = L_l = 40$, 共分割为 42 幅图像, 选取 $k_1 = k_2 = 2$ 。



图 4 原图

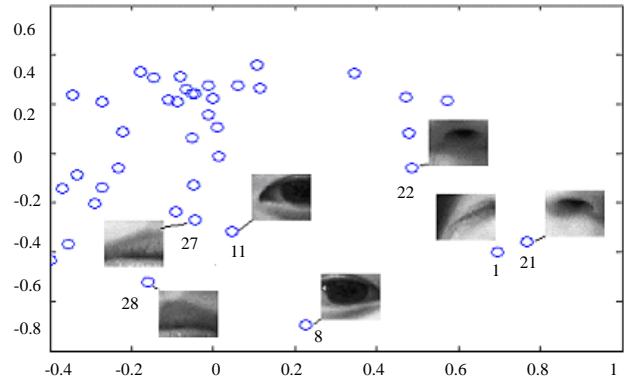


图 5 数据集结构的二维表示

通过例 1、例 2 可以看出,基于数据集信息结构的图像感兴趣区域提取方法,可以较好地提取图像数据的感兴趣区域,并且通过对提取区域的进一步修正(如对边缘进行适当的非线性膨胀、收缩),可获得更加细致的感兴趣区域,从而满足实际数据处理的需求。

3 结论与展望

本文对于图像感兴趣区域提取问题,提出了一种新的解决方案。即将高维数据的非线性降维方法引入到问题中来,利用非线性降维可发现高维数据集信息结构的低维表示这一特性,来获得图像的合理低维表示,从而将逐点计算的感兴趣点的提取问题转化为寻找低维表示中“特别点”(即图像块)的问题。这种考虑与常用方法相比,更加直观,可更好地挖掘不同图像区域之间的本征关联;另外从计算上来讲,它将对每点所作的 $M \times N$ 次计算,转化为对每块 $(L_h \times L_l)$ 的 C 次计算,大大减少了计算量。从而为感兴趣区域的自动提取提供了一种新的解决方案。

尽管我们没有仔细地探讨参数选择对最终结果的影响,但明显地,这种影响是存在的。需要注意的是,这一影响更多的不是来源于方法本身,而是来源于方法的不恰当运用,即由于人们对图像先验认识的不足造成的,从而充分挖掘图像的先验信息在一定程度上可改善最终的结果。

参考文献

- 1 Ueno I, Pearlman W A. Region of Interest Coding in Volumetric Images with Shape-adaptive Wavelet Transform[EB/OL]. http://www.cipr.rpi.edu/~pearlman/papers/ei03_5022-129_up.pdf.
- 2 Gokturk S B, Tomasi C. Medical Image Compression Based on Region of Interest with Application to Colon CT Images[EB/OL]. http://robotics.stanford.edu/~gokturkb/papers/paper_2.pdf.
- 3 Liu Lijie, Fan Guoliang. A New Method for JPEG2000 Region-of-Interest Image Coding: Most Significant Bitplanes Shift [EB/OL]. http://www.vcipl.okstate.edu/Publications/liu_mwscas.pdf.
- 4 宋 健. 高维函数和流形在低维可视空间中的最优表达[J]. 科学通报, 2001, 46(12).

(下转第 236 页)