

边缘保持的核回归图像插值方法

张峥嵘^a, 刘红毅^a, 韦志辉^b

(南京理工大学 a. 理学院; b. 计算机科学与技术学院, 南京 210094)

摘要: 研究核回归图像插值问题, 基于各向异性距离方法, 提出一种边缘保持的核回归图像插值方法。分析不同参数对该方法以及 Steering 核回归方法的影响。实验结果表明, 边缘保持的核回归图像插值方法对于参数变化具有较强的稳定性, 算法鲁棒性较高, 插值图像的边缘信息也得到更好的保持。

关键词: 图像插值; 核回归; 各向异性; Steering 核; 边缘保持

Edge-preserved Kernel Regression Image Interpolation Method

ZHANG Zheng-rong^a, LIU Hong-yi^a, WEI Zhi-hui^b

(a. School of Science; b. School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing 210094, China)

【Abstract】 This paper studies the kernel regression for image interpolation. Basing on anisotropic distance method, it proposes an edge-preserved kernel regression image interpolation method. It analyzes the influence of parameters to the proposed method and the Steering method. Experimental results show that the proposed method has strong stability and high robustness to variation of parameters, and very successful in preserving edges.

【Key words】 image interpolation; kernel regression; anisotropic; Steering kernel; edge-preserved

DOI: 10.3969/j.issn.1000-3428.2011.19.064

1 概述

图像插值包含了图像处理中的很多内容, 包括图像放大、图像去噪、图像恢复、图像配准、超分辨率重建等。其基本思想是根据周围邻域像素点的值得到邻域中心点的像素值, 或从邻域点的值估计丢失的像素值。图像插值方法可以分为线性和非线性两大类。常用的线性插值方法有最近邻、双线性、双三次等, 这类方法运行效率很高, 但缺点是对于边缘等细节会产生震铃和阶梯效应。非线性插值考虑到了图像的几何结构, 特别是边缘信息。文献[1]提出基于加权 ENO 的图像放大方法。文献[2]提出双边滤波, 不仅利用到了像素点的距离, 还考虑到灰度值的关系。文献[3]利用梯度的协方差矩阵来构造向量的各向异性距离, 提出了基于 Steering 核的核回归图像恢复模型。文献[4]在多视角的图像合成中构造了方向插值滤波器, 其中核函数选取用高斯函数, 距离测量用各向异性距离, 不同于文献[3]的是角度的选取。文献[5]提出边缘驱动的各向异性图像插值方法, 明确提出利用 T 范数代替 2 范数, 实现了各向异性的距离, 其中 T 为结构张量矩阵, 基本思想是通过估计边缘方向来设定适当的权函数以调整各像素点的权值。文献[6]利用结构张量矩阵构造了各向异性的距离, 应用于核回归的图像恢复。

从上述的方法来看, 目前在图像插值领域中存在的主要问题还是权函数和范数(距离)的选取如何自适应于边缘等图像细节。范数的选取可以得到不同的距离, 常用的为 2 范数, 但不具有各向异性, 具有各向异性的 T 范数可以通过取不同的 T 得到不同的衡量距离。权函数反映点与点之间的相关性, 最直接的就是与距离成反比。本文提出一种边缘保持的核回归图像插值方法, 在核函数中用向量的 T 范数代替 2 范数来衡量距离。

2 核回归图像插值问题

核回归是建立在统计学上的一种非参数回归。考虑 $y_i =$

$z(x_i) + \varepsilon_i, i = 1, 2, \dots, p$, 其中, $E(\varepsilon_i) = 0, D(\varepsilon_i) = \sigma^2 < \infty$ 。当 $i \neq j$ 时, ε_i 和 ε_j 是不相关的。如果假定变量 X, Y 都是随机的, 它们有联合分布 $f(x, y)$, 而 X 的边缘分布为 $f(x)$, 一般说来, $z(x)$ 可认为是 Y 在给定了 $X=x$ 之后的条件期望, 即:

$$z(x) = E(Y | X = x) = \int y f(y | x) dy = \int y f(x, y) dy / f(x) \quad (1)$$

回归模型 NWE 用下式来估计 $z(x_i)$:

$$\hat{z}(x) = \left[\frac{1}{nh} \sum_{i=1}^n W\left(\frac{x-x_i}{h}\right) y_i \right] / \left[\frac{1}{nh} \sum_{i=1}^n W\left(\frac{x-x_i}{h}\right) \right] \quad (2)$$

函数 $W(\cdot)$ 称为核函数, 它在本质上其实是一个加权平均函数, h 是窗口光滑参数。这种估计 $z(x_i)$ 的方法叫做核估计, 回归模型 NWE 就称为核回归。具体到图像插值, $x_i = (x_{i1} \ x_{i2})^T$ 为采样点, y_i 为采样点的灰度值, $z(x_i)$ 为插值后的图像函数, ε_i 为误差, p 为采样个数。则加权的插值问题可以表达为:

$$\min_z \sum_{i=1}^p \|y_i - z(x_i)\|^2 W_h(x - x_i) \quad (3)$$

进行泰勒展开:

$$z(x_i) \approx z(x) + z'(x)(x_i - x) + \frac{1}{2!} z''(x)(x_i - x)^2 + \dots + \frac{1}{N!} z^{(N)}(x)(x_i - x)^N \quad (4)$$

以上最小二乘问题可以写成矩阵形式:

基金项目: 国家自然科学基金资助项目(61071146); 国家“863”计划基金资助项目(2007AA12E142); 南京理工大学自主科研专项计划基金资助项目(2010ZYT070)

作者简介: 张峥嵘(1977—), 女, 讲师、博士研究生, 主研方向: 图像多尺度分析, 图像恢复; 刘红毅, 讲师、博士研究生; 韦志辉, 教授、博士生导师

收稿日期: 2011-03-28 **E-mail:** zhengrongzhang6@hotmail.com

$$\hat{b} = \arg \min_b \|y - X_x b\|^2 W = \arg \min_b (y - X_x b)^T W (y - X_x b) \quad (5)$$

其中,

$$y = [y_1, y_2, \dots, y_p]^T, b = [z(x), z'(x), \dots, \frac{1}{N!} z^{(N)}(x)]^T \quad (6)$$

$$X_x = \begin{bmatrix} 1 & (x_{11} - x_1) & (x_{12} - x_2) & (x_{11} - x_1)^2 & (x_{11} - x_1)(x_{12} - x_2) & (x_{12} - x_2)^2 & (x_{11} - x_1)^3 & (x_{11} - x_1)^2(x_{12} - x_2) & \dots \\ 1 & (x_{21} - x_1) & (x_{22} - x_2) & (x_{21} - x_1)^2 & (x_{21} - x_1)(x_{22} - x_2) & (x_{22} - x_2)^2 & (x_{21} - x_1)^3 & (x_{21} - x_1)^2(x_{22} - x_2) & \dots \\ \vdots & \vdots \\ 1 & (x_{p1} - x_1) & (x_{p2} - x_2) & (x_{p1} - x_1)^2 & (x_{p1} - x_1)(x_{p2} - x_2) & (x_{p2} - x_2)^2 & (x_{p1} - x_1)^3 & (x_{p1} - x_1)^2(x_{p2} - x_2) & \dots \end{bmatrix} \quad (9)$$

3 各向异性的核函数

在图像插值模型中, 插值方法的核心思想就是如何选择核 W , 核的不同选择就决定了不同的插值方法, 也决定了插值的精确度和计算量。传统的核函数只是考虑到了采样点之间的距离, 双边核不仅考虑到采样点之间的距离, 还有效利用了采样点的灰度值:

$$W_{\text{Bilateral}}(x_i - x, y_i - y) = W_{h_x}(x_i - x) W_{h_y}(y_i - y) \quad (10)$$

其中, W_{h_x} 和 W_{h_y} 取高斯函数; h 为高斯函数的平滑参数。不足的是其中的核并不是自适应的, 距离仍为通常意义下的欧氏距离。文献[5]提出了各向异性距离:

$$\|x - x_i\|_T = (x - x_i)^T T (x - x_i) \quad (11)$$

其中, T 为 2×2 的对称矩阵。文献[5]用协方差矩阵 C 来取代 T , 得到了基于协方差矩阵的各向异性距离:

$$\|x - x_i\|_T = (x - x_i)^T T (x - x_i) = (x - x_i)^T C_i (x - x_i) \quad (12)$$

再将以上距离应用于高斯核, 得到具有自适应的 Steering 核:

$$W_1(x - x_i) = \frac{\sqrt{\det(C_i)}}{2\pi h^2 \mu_i^2} \exp\left\{-\frac{(x - x_i)^T C_i (x - x_i)}{2h^2 \mu_i^2}\right\} \quad (13)$$

图像的局部协方差矩阵 C_i 定义:

$$C_i \approx \begin{bmatrix} \sum_{X_j \in \omega_i} z_{x_1}(X_j) z_{x_1}(X_j) & \sum_{X_j \in \omega_i} z_{x_1}(X_j) z_{x_2}(X_j) \\ \sum_{X_j \in \omega_i} z_{x_1}(X_j) z_{x_2}(X_j) & \sum_{X_j \in \omega_i} z_{x_2}(X_j) z_{x_2}(X_j) \end{bmatrix} \quad (14)$$

其中, ω_i 为所开的局部窗口; $z_{x_1}(\cdot)$ 、 $z_{x_2}(\cdot)$ 分别为沿 x_1, x_2 方向的一阶导数。协方差矩阵 C_i 是对称正定阵, 因此可以正交分解为:

$$C_i = \gamma_i U_{\theta_i} A_i U_{\theta_i}^T \quad (15)$$

$$U_{\theta_i} = \begin{bmatrix} \cos \theta_i & \sin \theta_i \\ -\sin \theta_i & \cos \theta_i \end{bmatrix} \quad (16)$$

$$A_i = \begin{bmatrix} \sigma_i & 0 \\ 0 & \sigma_i^{-1} \end{bmatrix} \quad (17)$$

则协方差矩阵由参数 γ 、 θ 、 σ 决定, 3 个参数的选取依据局部梯度矩阵 C_i , 对局部梯度矩阵进行奇异值分解 (Singular Value Decomposition, SVD) 分解:

$$C_i = \begin{bmatrix} \vdots & \vdots \\ z_{x_1}(X_j) & z_{x_2}(X_j) \\ \vdots & \vdots \end{bmatrix} = U_i S_i V_i^T \quad (18)$$

其中, $S_i = \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix}$ 代表主方向的能量; $V_i = [V_1 \ V_2] = \begin{bmatrix} * & v_1 \\ * & v_2 \end{bmatrix}$ 。

V_2 决定主方向的角度: $\theta_i = \arctan\left(\frac{v_1}{v_2}\right)$, 定义拉长参数 σ_i :

$$\sigma_i = \frac{s_1 + \lambda}{s_2 + \lambda}, \lambda \geq 0, \lambda \text{ 为正则化参数, 用于抑制噪声和限制 } \sigma_i$$

$$W = \text{diag}[W_h(x_1 - x), W_h(x_2 - x), \dots, W_h(x_p - x)] \quad (7)$$

对式(5)求导, 得到:

$$\hat{z}(x) = \hat{\beta}_0 = e_1^T (X_x^T W X_x)^{-1} X_x^T W y \quad (8)$$

其中, e_1^T 为单位向量; X_x 的每个元素为泰勒展开式中 x 的多次项(见式(9))。求出的 β_0 就是 $z(x)$ 的一个估计。

退化 1。定义压缩比例 $\gamma_i = \left(\frac{s_1 \cdot s_2}{M}\right)^\alpha$, M 为所开窗口中的样本数。

4 边缘保持的核回归方法

Steering 核回归方法通过协方差矩阵结合了图像的灰度和梯度信息, 可以实现有效的自适应, 恢复图像的边缘也保持的较好。但在实验中发现, 对于图像插值效果中竖性结构的边缘效果却比较差。分析其原因, 高斯函数中采用与距离成反比的权值, 而在图像的竖线形边缘(阶跃), 2 点之间的距离虽然比较小, 却应该赋予较小的权值。因此, 本文提出改进的边缘保持的各向异性权函数的核回归方法。

权函数选取参照文献[5]中的权函数:

$$W(x - x_i) = \left(\frac{R_i - \|x - x_i\|}{R_i \|x - x_i\|}\right)^2 \quad (19)$$

其中, R_i 为常数。将权函数和各向异性距离相结合, 得到新的边缘驱动的各向异性权函数:

$$W_2(x - x_i) = \left(\frac{R_i - \|x - x_i\|_T}{R_i \|x - x_i\|_T}\right)^2 = \left(\frac{R_i - (x - x_i)^T T (x - x_i)}{R_i (x - x_i)^T T (x - x_i)}\right)^2 \quad (20)$$

在此取 $R_i = \max(\|x - x_i\|_T)$ 。

将此权函数带入式(3), 得到了新的改进的边缘保持的图像插值模型:

$$\min_z \sum_{i=1}^p \|y_i - z(x_i)\|^2 W_2(x - x_i) \quad (21)$$

在 $W_2(\cdot)$ 中, 权函数的取值仍然在整体上满足与距离成反比。不同的是, 当两点之间的各向异性距离 $\|x - x_i\|_T$ 比较小时, 此时分式趋于 0 的速度比高斯函数的 e 指数要慢。因此, 在边缘邻域内的点得到的权值稍小于高斯函数的, 不会引起边缘大的波动, 出现伪边缘现象。

5 实验结果及分析

本节对于 Steering 核(W_1 方法)与本文提出的核函数(W_2 方法)进行核回归图像插值实验比较。选用了 256×256 像素大小的 Lena、人工合成图像(Rz)和 Cameraman 图像为实验图像, 如图 1 所示。

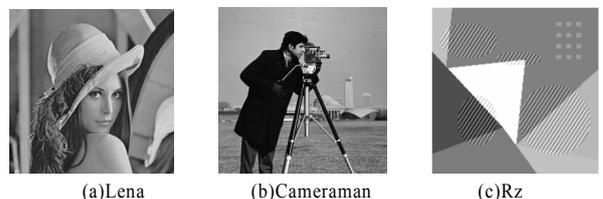


图 1 原始图像

研究图像放大中参数变化对于 2 种方法的影响。对 Lena 进行放大 2 倍的实验。为了更好地进行比较，将原始大小的 Lena 图像降采样到 128×128 像素，再对降采样图像进行放大 2 倍的处理。分别比较插值后图像与原图像的峰值信噪比 (Peak Signal to Noise Ratio, PSNR) 及均方误差 (Root Mean Square Error, RMSE)。

在实验中，参数有窗口大小 ω ，拉长参数 λ ，伸缩参数 α 及高斯的平滑参数 h 。由于在本文方法中没有平滑参数 h ，因此以下研究 2 种方法在协方差矩阵参数 α 、 λ 以及 ω 变化下的情况。

图 2 给出了固定其他参数、变化伸缩参数 α 下 2 种方法的图像放大效果。当 $\alpha < 0.5$ 时，W1 得到的效果较好，当 $\alpha = 0.5$ 时，2 种方法得到了同样的图像插值效果。当 $\alpha > 0.5$ 时，W1 得到的图像在桌脚边以及鼻梁等边缘附近出现了伪边缘，且图像质量整体下降较快。

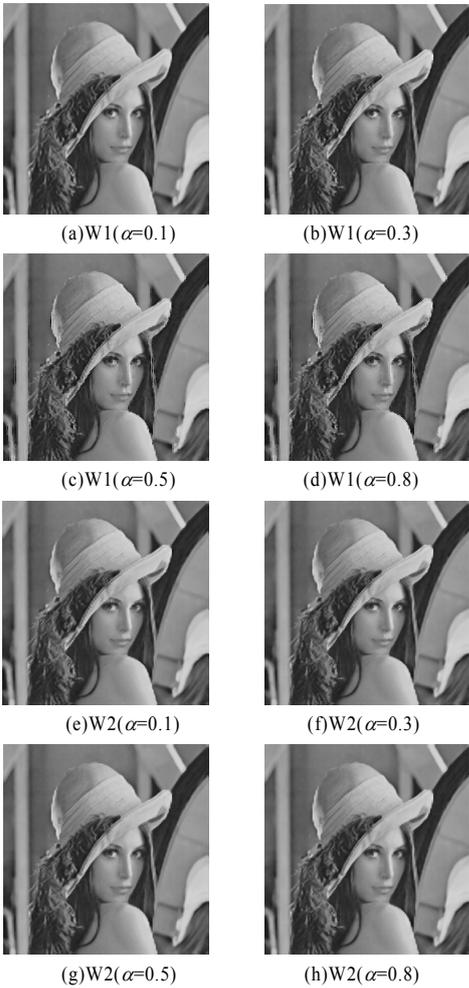


图 2 伸缩参数变化下的插值图像比较

2 种方法在伸缩参数变化下的残差图像如图 3 所示。在 W2 方法的残差图像中几乎没有保留桌脚等边缘的信息，说明插值图像对于这些边缘保持的较好，而 W1 方法的残差图像中很明显的就可以看到头发、帽檐、桌脚等残留的细节， α 越大，W1 方法得到的插值图像的伪边缘越多。

图 4 给出了在固定其他参数、变化拉长参数 λ 的情况下，2 种方法的插值结果及与原图做差后的比较，其中 λ 取值从 0.3~1.5。可以看出，W1 方法在阶跃边缘处仍会产生伪边缘，而 W2 得到的图像很好的保持了边缘及纹理。对于窗口参数 ω ，限于篇幅，本文不再罗列数据。

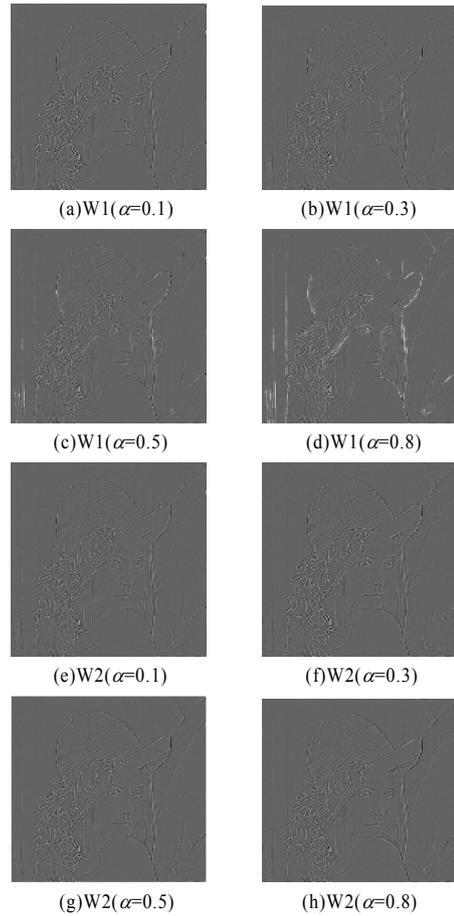


图 3 伸缩参数变化下的残差图像比较

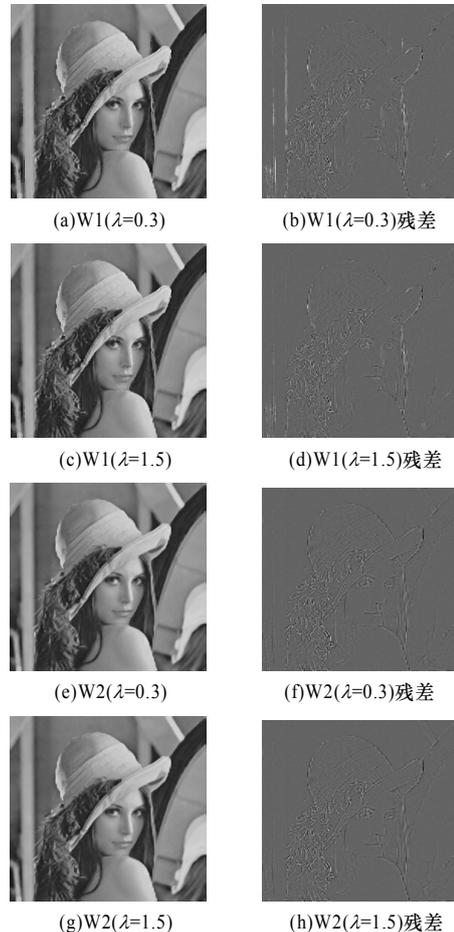


图 4 拉长参数变化下的插值图像与残差图像

选取 Cameraman 和具有几何结构的人工合成图像 Rz 为实验对象, 对于 2 种方法固定参数为 $\alpha=0.5$ 、 $\lambda=1$ 、 $\omega=7$, 得到 2 种方法的图像放大结果如图 5、图 6 所示。在图 5 中, W2 方法对于照相机的支架、人的衣领等部分边缘得到了很好的保持, 在图 6 中, 对于 Rz 图像, W1 方法在图像的左边角处出现了伪边缘, 在白色的三角形的三条边上都有较大的振荡, 而这些在 W2 方法中却都很不明显。



图 5 Cameraman 图像放大结果比较

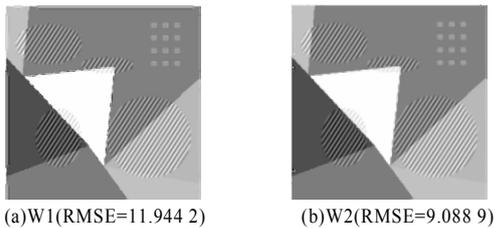


图 6 Rz 图像放大结果比较

6 结束语

本文提出了用向量的加权范数(T 范数)来反映各向异性距离, 同时和新的核函数相结合, 得到了边缘保持的各向异性核回归插值方法。实验证明, 高斯函数的各向异性距离对于参数过于敏感, 恢复图像中的阶跃形边缘会出现伪边缘效

应。而本文提出的边缘驱动的各向异性核, 具有各向异性特点, 通过对边缘的自适应的调节, 很好地保持了阶跃边缘及图像的细节部分, 图像恢复质量得到提升, 同时具有较高且稳定的峰值信噪比以及较低的均方误差。本文方法虽然对于边缘的保持有明显的改善, 但对于图像中的细小纹理还有所丢失, 这是下一步的研究方向。

参考文献

- [1] 姜东焕, 徐光宝. 基于加权 ENO 的图像放大算法[J]. 计算机工程, 2009, 35(6): 222-224.
- [2] Tomasi C, Manduchi R. Bilateral Filtering for Gray and Color Images[C]//Proc. of the 6th International Conf. on Computer Vision. Washington D. C., USA: IEEE Computer Society, 1998.
- [3] Takeda H, Farsiu S, Milanfar P. Kernel Regression for Image Processing and Reconstruction[J]. IEEE Trans. on Image Processing, 2007, 16(2): 349-366.
- [4] Yang Ping, Tong Xin, Zheng Xiaozhen. A Gradient Based Adaptive Interpolation Filter for Multiple View Synthesis[C]//Proc. of the 10th Pacific-rim Conf. on Multimedia. Bangkok, Thailand: [s. n.], 2009.
- [5] Casciola G, Ontefusco L B, Morigi S. Edge-driven Image Interpolation Using Adaptive Anisotropic Radial Basis Function[J]. Journal of Mathematical Imaging Vision, 2010, 36(2): 125-129.
- [6] Liu Hongyi, Wei Zhihui. An Edge-adaptive Structure Tensor Kernel Regression for Image Interpolation[C]//Proc. of the 2nd International Conf. on Future Computer and Communication. Wuhan, China: [s. n.], 2010.

编辑 顾姣健

(上接第 190 页)

序损失(ranking loss), 平均精度(average precision)。

表 2 分类后的结果

评价标准	θ -MLkNN	结合方法
hamming loss	0.362 5	0.372 9
one-error	0.450 0	0.312 5
coverage	2.925 0	2.925 0
ranking loss	0.334 6	0.296 0
average-precision	0.694 7	0.747 0

由于会出现预测所得的标记与实际标记不一致数目增多的情况, 因此 hamming loss 的性能会略降低, 但并不影响整体的性能, 所以 average-precision 的值有了较大提高。

one-error 的性能明显提高, 是因为音频文件所具有的隶属度最高的标记也包括这 2 类, 所以会直接降低 one-error 的值。也会降低音频文件对其所属标记的隶属度低于非所属标记的隶属度的可能性, 即降低了 ranking loss 的值。

而在这 2 种方法中, 平均约用 3 个标记类别就可以形容音频文件所具有全部情感类别, 因此 coverage 值不变。

6 结束语

本文针对歌曲情感的多标记分类问题, 提出一种结合了 θ -MLkNN 和 TF-IDF 的方法。音频文件分为 6 类, θ -MLkNN 多标记分类后, 由于 angry/fearful, happy/pleased 这 2 类情感是各类情感中人们感受最明显、最重要的情感, 再利用 TF-IDF 分别计算歌词这两类的情感分数, 得出其所属类别, 从而修正音乐内容分类错误的类别标记, 之后重新进行 θ -MLkNN 分类。通过实验可以看出, 得到了较好的准确

率。但是对歌词进行情感计算, 类别数目还较少, 今后工作需要考虑适合更多情感类别的歌词情感计算方法及歌词与音频的分类结合方法, 从而能够更准确地识别歌曲的情感类别。

参考文献

- [1] Feng Yazhong, Zhuang Yueting, Pan Yunhe. Music Information Retrieval by Detecting Mood via Computational Media Aesthetics[C]//Proc. of IEEE/WC Int'l Conf. on Web Intelligence. Halifax, Canada: [s. n.], 2003.
- [2] Trohidis K. Multi-label Classification of Music into Emotions[C]//Proc. of the 9th Int'l Conf. on Music Information Retrieval. Philadelphia, USA: [s. n.], 2008.
- [3] Yang Yi-Hsuan, Lin Yu-Ching. A Regression Approach to Music Emotion Recognition[J]. IEEE Trans. on Audio, Speech and Language Processing, 2008, 16(2): 448-457.
- [4] Tumbull D. Semantic Annotation and Retrieval of Music and Sound Effects[J]. IEEE Trans. on Audio, Speech, and Language Processing, 2008, 16(2): 467-476.
- [5] Hu Yajie, Chen Xiaou, Yang Deshun. Lyric-based Song Emotion Detection with Affective Lexicon and Fuzzy Clustering Method[C]//Proc. of the 10th Int'l Conf. on Music Information Retrieval. Kobe, Japan: [s. n.], 2009.
- [6] Zhang Minling, Zhou Zhihua. MI-knn: A Lazy Learning Approach to Multi-label Learning[J]. Pattern Recognition, 2007, 40(7): 2038-2048.
- [7] 吕小勇, 石洪波. 基于频繁项集的多标签文本分类算法[J]. 计算机工程, 2010, 36(15): 83-85.

编辑 陈文

