

基于方向场正则化的线描画生成算法

李晶晶¹ 许建楼² 熊静¹ 张选德¹

摘要 图像风格转化在计算机视觉领域广受关注, 其研究目标在于将输入图像利用计算机转化为具有某种特定艺术风格的图像. 线描画作为一种古老的画种, 它通过简单的线条勾勒物体的轮廓, 具有简约、抽象的风格. 本文提出一种基于方向场正则化的线描画生成算法, 该算法由 4 部分构成: 1) 采用非局部平均滤波对输入图像进行预处理; 2) 计算输入图像的方向场, 并基于自表示的思想对方向场进行 Tikhonov 正则化, 为了提高运算速度, 采用 Sherman-Morrison-Woodbury 公式来对正则化算法进行加速; 3) 以正则方向场作为引导, 对预处理图像作高斯差分滤波; 4) 根据人类视觉系统的非线性特点, 设计感知阈值 (Perceptual thresholding) 算法来对高斯差分滤波的结果进行阈值处理, 得到二值化的线描画图像. 仿真实验表明, 该算法可将输入图像转化为线条流畅且能有效表达输入图像主要信息的线描画图像.

关键词 线描画, 方向场, Tikhonov 正则化, 高斯差分滤波, 感知阈值

引用格式 李晶晶, 许建楼, 熊静, 张选德. 基于方向场正则化的线描画生成算法. 自动化学报, 2021, 47(3): 685–694

DOI 10.16383/j.aas.c190393

Line Drawing Generation Algorithm Based on Direction Field Regularization

LI Jing-Jing¹ XU Jian-Lou² XIONG Jing¹ ZHANG Xuan-De¹

Abstract Image style transfer has attracted much attention in the field of computer vision. Its research goal is to transfer the input image into an image with a specific artistic style by computer. Line drawing as an ancient painting, it outlines the object through simple lines, with a simple, abstract style. This paper proposes a line drawing generation algorithm based on direction field regularization, which consists of four parts: 1) Non-local means filtering is used to preprocess the input image; 2) Calculate the direction field of the input image, and the Tikhonov regularization of the direction field based on the self-representation idea, in order to improve the speed of operation, the Sherman-Morrison-Woodbury formula is used to accelerate the regularization algorithm; 3) Using a regular direction field as a guide, the pre-processed image is filtered by difference of Gaussian filter; 4) According to the non-linear characteristics of human visual system, the perceptual thresholding algorithm is designed to process the results of the Gaussian difference of filtering and obtain binary line drawing image. The simulation results show that the algorithm can transfer the input image into a line drawing image with smooth lines and can effectively express the main information of the input image.

Key words Line drawing, direction field, Tikhonov regularization, difference of Gaussian filter, perceptual thresholding

Citation Li Jing-Jing, Xu Jian-Lou, Xiong Jing, Zhang Xuan-De. Line drawing generation algorithm based on direction field regularization. *Acta Automatica Sinica*, 2021, 47(3): 685–694

一直以来, 人们对绘画作品比较感兴趣, 它反映、传播和塑造了人类的文化^[1]. 艺术家们手工绘制铅笔画、卡通画、油画、水彩画等不同风格的作品, 需耗费大量的时间和精力, 几乎每一部经典作品都

凝聚着艺术家的心血. 近年来, 利用算法对输入图像进行快速的风格转化引发了计算机视觉领域广泛的研究兴趣. 图像风格转化 (Image style transfer, IST) 算法可视为一个系统, 系统的输入是利用相机拍摄的自然图像 (Photorealistic image), 输出是具有某种特定艺术风格的图像 (Non-photorealistic image). IST 在娱乐产业和消费电子中有着广泛的应用, 例如智能手机的图片编辑功能中通常提供多个风格转化选项; 在电影和游戏的制作过程中, 常采用 IST 算法来快速生成各种风格的场景.

过去几十年来, 计算机视觉领域对 IST 问题进行了广泛的研究. 1996 年, Decaudin 等^[2]对卡通风格转化进行了明确的定义, 并研究了如何从静态或

收稿日期 2019-05-20 录用日期 2019-06-27
Manuscript received May 20, 2019; accepted June 27, 2019
国家自然科学基金 (61871260, 61603234) 资助
Supported by National Natural Science Foundation of China (61871260, 61603234)
本文责任编辑 黄庆明
Recommended by Associate Editor HUANG Qing-Ming
1. 陕西科技大学电子信息与人工智能学院 西安 710021 2. 河南科技大学数学与统计学院 洛阳 471023
1. School of Electronic Information and Artificial Intelligence, Shaanxi University of Science and Technology, Xi'an 710021 2. School of Mathematics and Statistics, Henan University of Science and Technology, Luoyang 471023

动态 3D 场景中生成卡通风格的 2D 图像或者视频. 2002 年, DeCarlo 等^[3]利用眼动仪 (Eye tracker) 来记录当人观测一幅图像时的眼球运动数据, 并基于眼球运动数据和视觉感知模型来确定图像中的视觉显著元素 (Meaningful elements), 然后基于图像的多尺度表示构造了一种风格转化算法, 风格化后的图像能保持并凸显原图像中的视觉显著元素. 2004 年, Santella 等^[4]借助眼动仪来验证文献 [3] 中构造的算法是否达到了凸显视觉显著元素的目标. 2008 年, Kyprianidis 等^[5]提出一种基于结构自适应滤波的图像风格转化算法. 该算法计算图像在每一像素点处的结构张量, 并以结构张量确定的梯度方向和切线方向作为引导, 实现对图像的局部自适应滤波. 2009 年, Kang 等^[6]采用双边滤波对图像的切线场进行平滑, 并利用光滑切线场来引导风格转化算法. 2015 年, Qian 等^[7]提出一种基于各向异性 Kuwahara 滤波和迭代线积分卷积的风格化方法, 其中 Kuwahara 滤波能较好地克服双边滤波及其均值漂移滤波器在边缘保持方面的局限性.

以上方法属于传统方法, 是研究者基于自己的知识或经验, 启发式地构造的风格转化算法, 这种算法不依赖于样本, 也无需进行训练. 近几年来, 随着深度学习方法的“横空出世”, 研究者们对深度学习在图像风格转化中的应用表现出非常强烈的兴趣^[8-11]. 其中 Gatys 等^[8]采用预训练的、用于一般特征表示 (Feature representation) 的卷积神经网络 (Convolutional neural network, CNN) 来分别处理自然图像的内容和风格, 提出一种基于纹理合成的图像风格转化算法. Johnson 等^[9]采用感知损失函数 (Perceptual loss function) 训练前馈网络来实现图像风格转化. Ulyanov 等^[10]首先训练紧凑的前馈卷积网络, 以生成任意大小且纹理相同的多个样本, 并将艺术风格从给定图像转移到任何其他图像. Elad 等^[11]从字典学习、字典表示的角度拓展了 Kwatra 等^[12]提出的纹理合成方法, 得到了具有与 CNN 类似效果的图像风格转化方法. 基于深度学习的方法能够获得非常好的风格转化效果, 但是深度学习方法依赖于样本进行训练, 运算代价较大.

本文采用传统方法研究图像风格的转化, 聚焦于“线描画”风格, 提出了一种基于方向场正则化的线描画生成算法 (Direction field regularization based line drawing generation, DFR-LDG). 该算法由以下几部分构成: 1) 为了减轻噪声的影响, 采用非局部平均 (Non-local means, NLM) 滤波对输入图像进行预处理; 2) 计算输入图像的方向场, 并对方向场进行 Tikhonov 正则化, 为了提高运算速度,

采用 Sherman-Morrison-Woodbury 公式来对正则化算法进行加速; 3) 以正则方向场作为引导, 对预处理图像作高斯差分 (Difference of Gaussian, DoG) 滤波; 4) 根据人类视觉系统 (Human visual system, HVS) 的非线性特点, 设计感知阈值 (Perceptual thresholding) 算法来对高斯差分滤波的结果进行阈值处理, 得到二值的线描画图像. 数值实验表明, 提出的算法可将输入图像转化为线条流畅的线描画图像.

本文结构安排如下: 第 1 节首先介绍线描画, 然后分析线描画与边缘检测的关系, 最后探讨边缘检测算子作为线描画生成算法的不足; 第 2 节详细介绍基于方向场正则化的线描画生成算法; 第 3 节进行数值实验; 第 4 节对本文工作进行总结.

1 线描画与边缘检测

1.1 线描画

人类视觉系统具有非常强大的抽象能力, 能够将观测到的对象抽象成线条. 汉字中的象形字便充分地体现了这种抽象能力. 此外, 图像信息是通过灰度值的变化来呈现的, 如果一幅图像的灰度值在空域所有像素点处都一致, 即没有任何灰度变化, 则这幅图像不包含任何信息. 灰度值剧烈变化的区域通常对应语义对象的边缘, 而语义对象的边缘可用线条来勾勒. 因而, 线条自然成为美术构图的最重要元素, 而线描画也成为最古老、最基础的绘画种类. 远古时期, 先民们就开始本能地在岩壁上绘制线描画, 图 1 (a) 所呈现的是贺兰山岩画. 美术教育中, 也通常将线描画作为基础的必修画种 (图 1 (b)). 美术史上, 毕加索和丰子恺都是描述画大师 (图 1 (c) 和图 1 (d)). 毕加索采用更简约流畅的线条, 作品趋于夸张和抽象风格, 而丰子恺的作品则趋于写实和漫画风格.

从图像处理的角度考察, 线描画图像具有以下特点: 1) 线描画属于二值图像, 线条所在像素值为 0, 而背景区域的像素值为 1; 2) 线描画采用流畅的线条勾勒语义对象的轮廓, 比原图像更加简约, 但能够表达原图像中的主要信息. 从而, 线描画与边缘检测 (Edge detection) 密切相关, 而边缘检测算子也可以作为线描画生成算法.

1.2 边缘检测

图像处理中, 边缘检测 (Edge detection) 指确定 (Identify) 图像中具有剧烈灰度变化的、不连续的像素点的方法. 常用的边缘检测算子包括 Canny



图 1 线描画图
Fig.1 Line drawing

算子、Sobel 算子、Prewitt 算子、Roberts 算子等. 其中 Sobel 算子、Prewitt 算子和 Roberts 算子三种方法的基本思想类似, 都是利用模板卷积来逼近图像在水平和垂直方向的方向导数, 然后对梯度模进行阈值得到二值的边缘图像 (Edge image). 不同之处在于采用模板不同, 如图 2 所示. 根据采用的模板分析, Sobel 算子和 Prewitt 算子都可以度量水平和垂直两个方向的变化, 但 Sobel 算子较 Prewitt 算子更强调中心位置像素的作用. Roberts 算子度量主对角和副对角方向的变化. 这三个算子都未对图像作预处理, 由于导数运算对于噪声非常敏感, 这使得这三个算子对噪声的鲁棒性较差. Canny 算子可视为上述三个算子的改进, 其主要步骤包括: 1) 采用高斯滤波来抑制噪声, 提高方向导数计算的鲁棒性; 2) 计算图像的梯度模; 3) 通过非局部极大抑制 (Non-maximum suppression) 来消除虚假边缘; 4) 利用双阈值来确定潜在的边缘; 5) 消除弱的、或者与强的边缘没有连接的孤立的检测点, 得到最终的边缘图像. Canny 算子有严格的理论作支撑, 边缘检测效果相比其他几个算子都要好.

图 3 和图 4 呈现了几种边缘检测算子在测试图像 Baboon 和 Lena 上的检测效果. 从边缘检测的角度分析, 利用 Sobel 算子、Prewitt 算子、Roberts 算子得到的边缘 (分别参考图 3(b) ~ (d) 和图 4(b) ~ (d)) 都存在漏检边缘、边缘不够连续的问题, 而 Canny 算子 (参考图 3 (e) 和图 4 (e)) 能够检测出所有主要的边缘, 而且检测出的边缘也比较连续. 但是, 将以上 4 种边缘检测算子视作线描画生成算

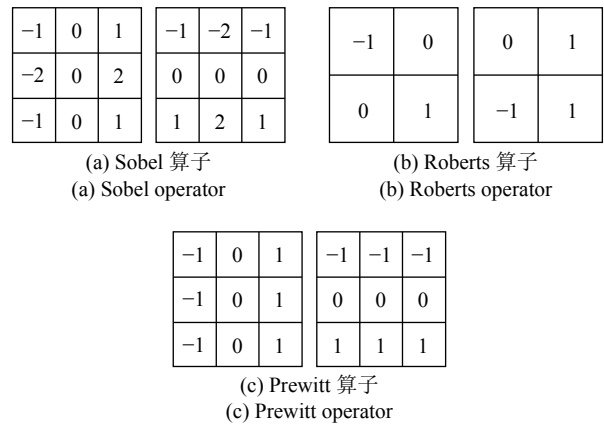


图 2 几种边缘检测算子采用的模板
Fig.2 Templates for several edge detection operators

法, 从线描画风格转化的角度来考察, 则不难发现这几个算子的转化效果都不能令人满意, 其中最显著的问题在于得到的线条不够流畅. 图 3 (f) 和图 4 (f) 还呈现了本文算法得到的结果, 对比可以看出, 本文算法生成的二值图像, 能表达原图像的主要信息, 而且线条更加流畅, 更接近线描画风格.

2 基于方向场正则化的线描画生成算法

本文提出一种基于方向场正则化的线描画生成算法, 该算法由以下 4 个部分构成: 1) 为了在消除噪声的同时尽可能保持图像的结构, 采用非局部平均算法对输入图像进行预处理; 2) 考虑到图像在局部具有各向异性结构, 分别利用水平、垂直、主对角、副对角共 4 个方向的模板来计算方向场, 为了能够提取连续流畅的线条, 对方向场进行 Tikhonov 正则化, 为了提高运算速度, 采用 Sherman-Morrison-Woodbury 公式来对正则化算法进行加速; 3) 以正则方向场为引导, 对预处理图像逐点进行一维 DoG 滤波, 即在每一像素点处, 沿变化最大的方向进行 DoG 滤波; 4) 对 DoG 滤波的结果进行非线性阈值处理, 得到二值的线描画图像. 整个算法流程如图 5 所示, 下面对每一步骤逐一阐释.

2.1 预处理

Canny 算子采用高斯滤波来滤除噪声, 但高斯核是各向同性的, 不能自适应于图像的局部结构, 这使得高斯滤波在滤除噪声的同时在一定程度上模糊图像的边缘. 这里采用 NLM^[13] 滤波来对输入图像进行预处理, NLM 利用邻域像素的加权平均来对图像进行逐点估计, 而权重通过相似度来计算.

记输入图像 $V = \{V(i) | i \in \Omega\}$, NLM 滤波后的图像 $U = \{U(i) | i \in \Omega\}$, 其中, Ω 为图像区域, i 为

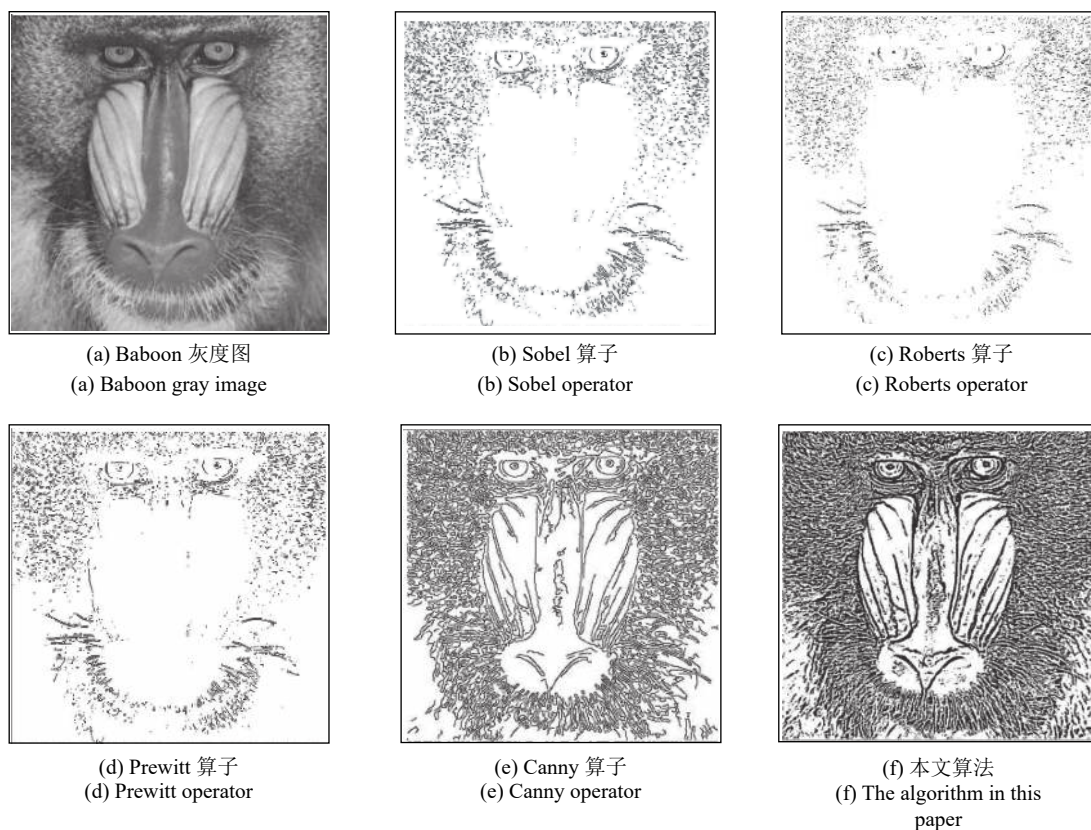


图 3 几种边缘检测算子和本文算法在 Baboon 上的效果对比

Fig. 3 Comparison of several edge detection operators and the algorithm of this paper on Baboon

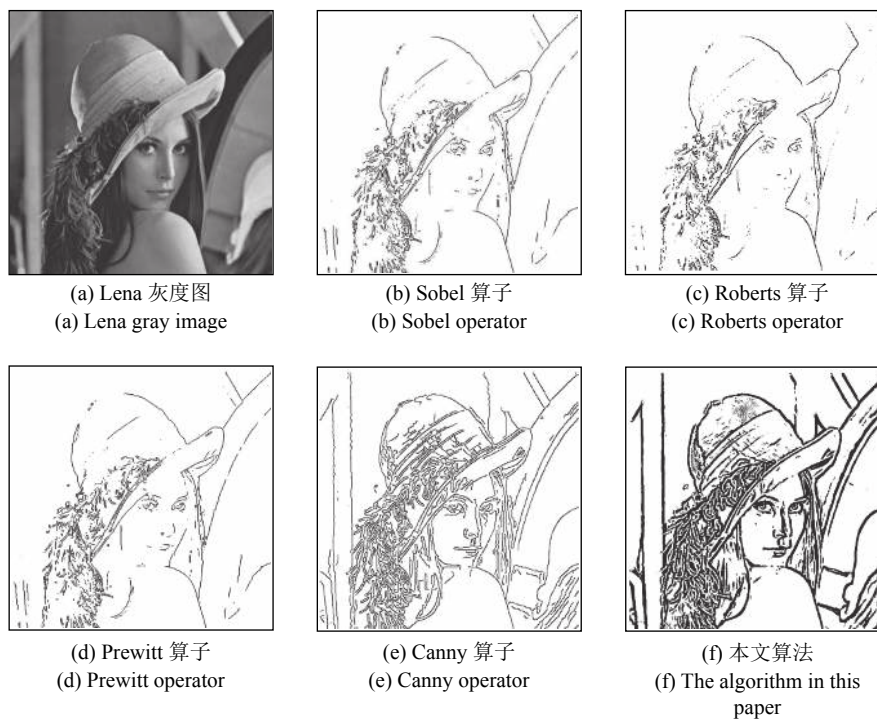


图 4 几种边缘检测算子和本文算法在 Lena 上的效果对比

Fig. 4 Comparison of several edge detection operators and the algorithm of this paper on Lena

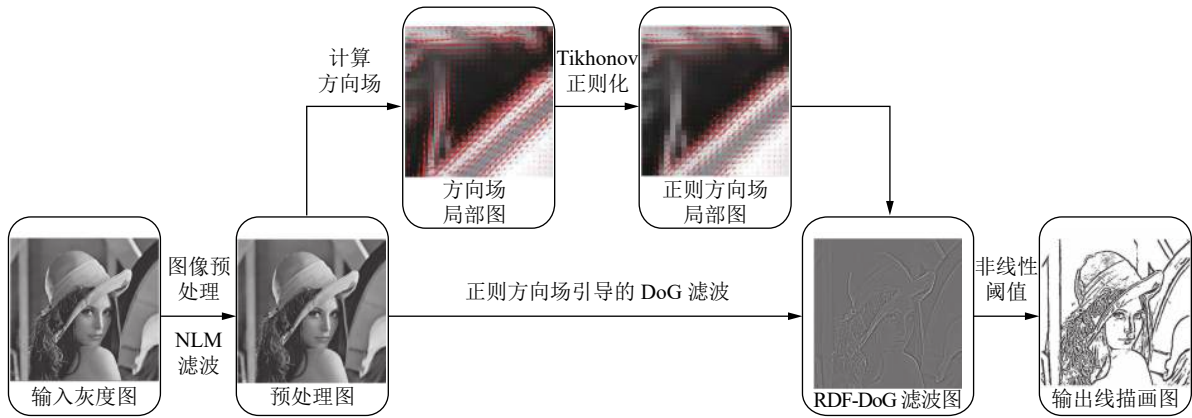


图 5 基于方向场正则化的线描画生成算法框图 (注: 为了使正则化的效果可视, 对方向场进行降维处理且只在 Lena 局部区域上显示)

Fig.5 Block diagram of line drawing generation algorithm based on direction field regularization (Note: In order to make the regularization effect visible, the direction field is dimension-reduced and displayed only on the local area of Lena)

像素索引. 则 $U(i)$ 可表示为

$$U(i) = \frac{\sum_{j \in \Omega_i} w(i, j)V(j)}{\sum_{j \in \Omega_i} w(i, j)} \quad (1)$$

其中, Ω_i 表示以像素 i 为中心, 半径为 t 的方邻域, $w(i, j)$ 为赋予 $V(j)$ 的权值, 分母 $\sum_{j \in \Omega_i} w(i, j)$ 为归一化因子, 保证权值之和为 1. 通常取

$$w(i, j) = \exp\left(-\frac{d(i, j)}{h^2}\right) \quad (2)$$

其中

$$d(i, j) = \|N(i) - N(j)\|_{2, \alpha}^2 \quad (3)$$

其中, $N(i)$, $N(j)$ 分别表示以像素 i , j 为中心, 半径为 f 的图像块; $\|\cdot\|_{2, \alpha}^2$ 表示高斯半范, h 为滤波参数.

NLM 算法中, 滤波的光滑程度由参数 h 和 t 控制. h 决定相似性度量 $d(i, j)$ 对于权重 $w(i, j)$ 的影响程度, h 越大, $d(i, j)$ 对于 $w(i, j)$ 的影响越小, 从而滤波的光滑程度越大; 反之, h 越小, 滤波的光滑程度也越小. t 决定参与加权的像素点的个数, 滤波的光滑程度随着 t 的增大而增大. 图像风格转化问题中, 输入图像通常只包含弱噪声, 因此 h 和 t 的取值都较小. 这里取 $h = 8$, $t = 5$.

2.2 方向场正则化

2.2.1 方向场的计算

经典的 Sobel、Roberts、Prewitt 等一阶梯度算子只能度量水平和垂直两个方向的变化, 但图像在空域具有丰富的多方向结构. 为此, 这里考虑 4 个方向的变化, 在每一像素点计算水平、45° 方向、垂直方向、135° 方向共 4 个方向的方向导数. 方向导数通过模板卷积来计算, 所采用的模板如图 6 所示.

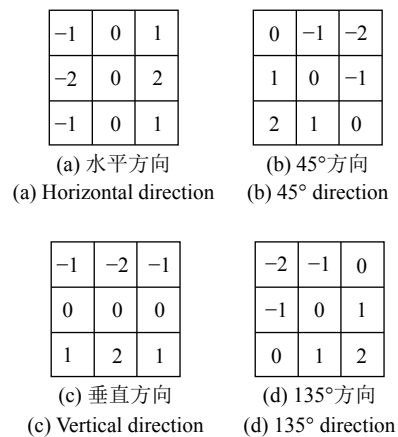


图 6 计算方向导数采用的模板

Fig.6 Template for calculating directional derivatives

以 M_1, M_2, M_3, M_4 分别表示图 6 所示的水平、45° 方向、垂直方向、135° 方向的模板, 则 U 的方向导数为 $G^d = U \otimes M_d$, $d = 1, 2, 3, 4$, 其中 “ \otimes ” 表示卷积运算. 以 $G^d(i), i \in \Omega$ 表示 U 在像素点 i 处、 d 方向上的方向导数, 则 $g(i) = [G^1(i), G^2(i), G^3(i), G^4(i)]^T$ 构成像素点 i 处的方向矢量. 方向矢量刻画了图像的局部结构, 而图像在所有像素点处的方向矢量总体 $\{g(i)|i \in \Omega\}$ 就构成了图像的方向场.

2.2.2 方向场的 Tikhonov 正则化

尽管对输入图像的预处理可以一定程度上提高方向导数对于噪声的鲁棒性, 但直接从预处理图像计算得到的方向场不够正则, 尤其边缘处的方向显得比较杂乱, 如图 7 (a) 所示. 为了提取流畅的线条, 这里基于自表示的思想, 利用 Tikhonov 方法对方向场进行正则化处理. 像素点 i 处的方向矢量为

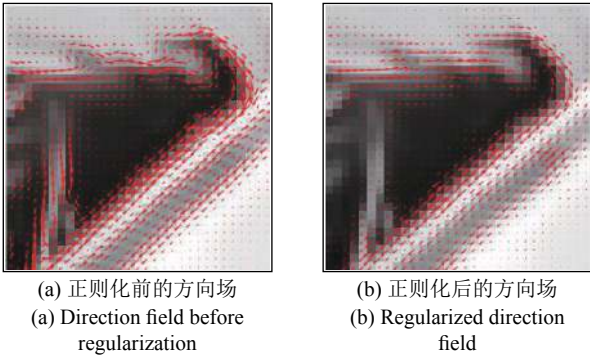


图7 方向场正则化效果图(注:为了使正则化的效果可视,对方向场进行降维处理且只在Lena局部区域上显示)
Fig.7 Directional field regularization effect map (Note: In order to make the regularization effect visible, the direction field is dimension-reduced and displayed only on the local area of Lena)

$\mathbf{g}(i)$, 以 i 为中心、 t 为半径的方邻域 Ω_i 中所有像素点处的方向矢量为 $\mathbf{g}(j)$, $j \in \Omega_i$. 考虑到图像在空域的平稳性及图像中存在大量的自相似结构, $\mathbf{g}(i)$ 可由 $\mathbf{g}(j)$, $j \in \Omega_i$ 线性表示, 即

$$\mathbf{g}(i) = \sum_{j=1}^{(2t+1)^2} \beta_j \mathbf{g}(j) \quad (4)$$

其中, $\beta_j, j = 1, 2, \dots, (2t+1)^2$ 表示系数. 表示成矩阵形式, 有

$$\mathbf{g}(i) = G \times \mathbf{w} \quad (5)$$

其中, G 是以 $\mathbf{g}(j)$, $j \in \Omega_i$ 为列向量的 $4 \times (2t+1)^2$ 的矩阵, $\mathbf{w} = [\beta_1, \beta_2, \dots, \beta_j, \dots, \beta_{(2t+1)^2}]^T$. 注意由于像素点 i 是区域 Ω_i 的中心, 所以 $\mathbf{g}(j), j \in \Omega_i$ 中包含 $\mathbf{g}(i)$, 这使得式 (4) 和式 (5) 存在平凡解 (只需取 $\mathbf{g}(i)$ 的系数为 1, 其余的系数为 0 即可). 为了避免平凡解, 采用如下的 Tikhonov 正则化模型求解表示系数:

$$\hat{\mathbf{w}} = \arg \min_{\mathbf{w}} \|\mathbf{g}(i) - G \times \mathbf{w}\|_2^2 + \lambda \|\mathbf{w}\|_2^2 \quad (6)$$

其中, 第 1 项为数据逼近的忠诚项 (Fitting fidelity), 要求 $\mathbf{g}(i)$ 与 $G \times \mathbf{w}$ 尽可能接近; 第 2 项为正则项, 要求线性表示的系数向量 \mathbf{w} 的 2 范数尽可能小, λ 为平衡两项之间作用的正则参数. 式 (6) 中, 目标函数既是凸的、又是可微的, 因此存在解析解. 对目标函数关于 \mathbf{w} 求导并令导数为零, 可得:

$$\begin{aligned} -G^T \mathbf{g}(i) + G^T G \mathbf{w} + \lambda \mathbf{w} &= 0 \Leftrightarrow \\ (G^T G + \lambda I) \mathbf{w} &= G^T \mathbf{g}(i) \Leftrightarrow \\ \hat{\mathbf{w}} &= (G^T G + \lambda I)^{-1} G^T \mathbf{g}(i) \end{aligned} \quad (7)$$

$\mathbf{g}(i)$ 经正则化处理后记作 $\hat{\mathbf{g}}(i) = [\hat{G}^1(i), \hat{G}^2(i), \hat{G}^3(i),$

$\hat{G}^4(i)]^T$, 则 $\hat{\mathbf{g}}(i) = G \times \hat{\mathbf{w}}$.

2.2.3 正则化算法的加速

第 2.2.2 节中对方向场的正则化需要逐点计算, 且由式 (7) 可见, 每个点上都需要计算一个大小为 $(2t+1)^2 \times (2t+1)^2$ 的矩阵的逆. 当 t 取值较大时, 运算量会非常大. 实验中取 $t = 5$, 那么对于 512×512 像素的图像, 方向场正则化需要计算 262 144 个 121×121 的矩阵的逆, 运算开销很大. 为了减少运算量, 加速算法, 我们引用 Sherman-Morrison-Woodbury^[14] 矩阵恒等式:

$$(A^{-1} + B^T B)^{-1} B^T = A B^T (B A B^T + I)^{-1} \quad (8)$$

令 $A = \lambda^{-1} I$, $B = G$, 代入式 (8) 可得:

$$(G^T G + \lambda I)^{-1} G^T = \lambda^{-1} I G^T (\lambda^{-1} G I G^T + I)^{-1} \quad (9)$$

将式 (9) 代入式 (7) 可得:

$$\hat{\mathbf{w}} = \lambda^{-1} I G^T (\lambda^{-1} G I G^T + I)^{-1} \mathbf{g}(i) \quad (10)$$

对比式 (7) 与式 (10) 不难发现, 利用式 (7), 需要计算 $(2t+1)^2 \times (2t+1)^2$ 的矩阵的逆; 而利用式 (10), 无论 t 如何取值, 都只需计算 4×4 的矩阵的逆; 因此, 当 t 取值较大时, 利用式 (10) 可大大加速算法.

由式 (6) 可见, 当正则参数 $\lambda \rightarrow 0$ 时, Tikhonov 正则化退化为最小二乘法, 这会导致平凡解, 起不到正则化的作用; 由式 (10) 可见, 当 $\lambda \rightarrow \infty$ 时, 线性表示的系数向量 $\hat{\mathbf{w}}$ 中的各个分量趋于相同, 这会使得 Tikhonov 正则化具有类似于均值滤波的效果, 会破坏方向场的固有结构. 为了既能保持方向场固有的大结构, 又能对方向场进行一定的正则化处理. 在数值实现过程中, 取 $\lambda = 25$. 图 7 呈现了方向场正则化的效果图. 从中可见, 经正则化处理后, 方向场更为正则.

2.3 正则方向场引导的高斯差分滤波

自然图像中, 语义对象的边缘对应灰度值变化较显著的区域. 因此, 要勾勒出语义对象的边缘, 首先需要度量图像在空域的灰度值变化. 原则上, 所有的高通滤波都可以度量灰度值的变化. 这里选用一维高斯差分滤波来度量变化, DoG 滤波以下面的函数作为滤波的核函数:

$$K(r) = \mathbf{g}_{\alpha 1}(r) - \mathbf{g}_{\alpha 2}(r) \quad (11)$$

其中

$$\mathbf{g}_{\alpha}(x) = \frac{1}{\sqrt{2\pi\alpha}} \exp\left(-\frac{x^2}{2\alpha^2}\right) \quad (12)$$

可见, $K(r)$ 是两个零均值的高斯函数之差. Marr 等^[15] 指出, DoG 滤波能很好地模拟人类视网膜细胞

对于亮度变化的响应, 并且建议取 $\alpha_2 = 1.6\alpha_1$. 为了提升 DoG 滤波的稳定性, 数值实现过程中使用文献 [16] 的核函数:

$$K(r) = g_{\alpha_1}(r) - (1 - \varepsilon) \cdot g_{\alpha_2}(r) \quad (13)$$

并取 $\varepsilon = 0.001$.

自然图像在空域具有“各向异性”的性质, 灰度值通常在与边界垂直的方向上变化大, 而在与边界平行的方向上变化小. 像素点 i 处的方向矢量为 $\hat{g}(i) = [\hat{G}^1(i), \hat{G}^2(i), \hat{G}^3(i), \hat{G}^4(i)]^T$, 这 4 个分量分别度量了图像在水平、45°、垂直、135° 方向的灰度变化, 如图 8 (a) 所示. 在图 8 (b) 中, $\hat{g}(i)$ 的 4 个分量中绝对值最大者所在的方向为变化最大的方向, 记为 $\max(\hat{g}(i))$, 我们沿着变化最大的方向对预处理图像作一维高斯差分滤波, 而具体实现过程中, 需要对 DoG 滤波的核函数进行离散. 上述 DoG 滤波在每个像素处都需要利用方向矢量 $\hat{g}(i), i \in \Omega$ 来确定滤波的方向, 因而将其称之为“正则方向场引导的 DoG 滤波 (Regularized direction field guided dog filter, RDF-DoG)”. 滤波的结果记作

$$B(i) = DoG(\hat{g}(i), U), \quad i \in \Omega \quad (14)$$

图 8 (c) 中展示了 RDF-DoG 滤波的结果, 从中可见, 滤波结果能较准确地度量图像灰度值的变化.

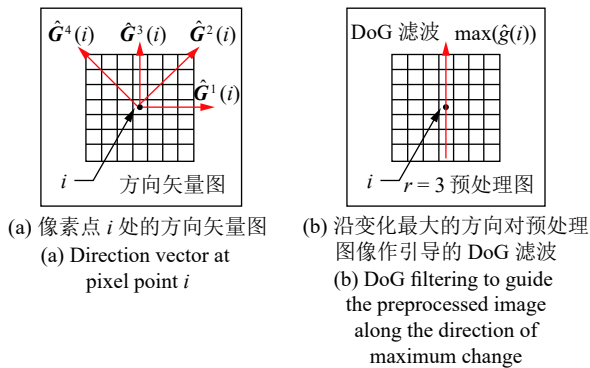


图 8 RDF-DoG 滤波操作示意图

Fig.8 RDF-DoG filtering operation diagram

2.4 感知阈值

类似于边缘检测, 对 RDF-DoG 滤波的结果进行阈值处理就可以勾勒出语义对象的边缘, 这里根据人类视觉系统的非线性特点来设计阈值算法. RDF-DoG 的滤波结果所度量的灰度值变化, 属于客观量 (物理量, 客观刺激), 而 HVS 感知到的变化属于主观量 (心理量, 主观响应), 如图 9 所示. 众所周知, 人类所有的感知系统, 包括听觉系统、视觉系统、触觉系统、味觉系统、嗅觉系统, 都是非线性系统. 生物学和计量心理学中都对人类感知系统的响应规律进行了广泛的研究. 一般认为人类所有的感知系统都具有“双边抑制”效应. 以味觉系统对盐溶液的感知为例, 考察盐浓度 (客观量) 与人感知到的“咸的程度” (主观量) 之间的关系, 人在盐浓度非常小时, 感觉不到咸; 只有盐浓度达到一定程度时, 才开始有咸的感觉; 此时, 随盐浓度的增加, 人会感觉到越来越咸; 但是盐浓度达到一定量时, 再增加盐浓度, 人不会有更咸的感觉 (主观量不会明显增加). HVS 对灰度值变化的感知具有类似的规律, 只有当客观灰度变化达到一定程度时, 才能被 HVS 感觉到. 此时, 随着客观变化的增大, HVS 感知到的变化随之增大, 但是客观变化达到一定量时, HVS 感知到的变化量不再随客观变化的增加而明显增加.

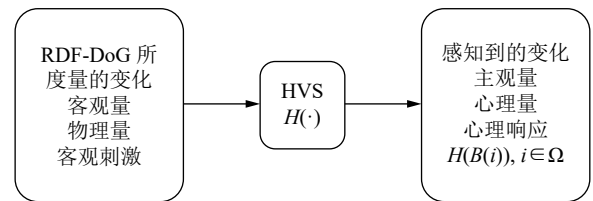


图 9 HVS 对于灰度变化的感知过程

Fig.9 HVS perception process of grayscale changes

人类视觉系统具有的非线性特性与支撑视觉感知过程的复杂生化反应有关, 这种特性是人类在长期的进化中形成的. 这里用双曲正切函数来模拟客观灰度变化 ($B(i), i \in \Omega$) 与 HVS 感知到的变化之间的关系, 即取

$$H(x) = \tanh(x) + 1 = \frac{e^x - e^{-x}}{e^x + e^{-x}} + 1 \quad (15)$$

图 10 展示了 $H(x)$ 的形态, 利用 $H(x)$ 可将客观量映射为主观量.

为了勾勒语义对象的边缘, 这里对主观量 (HVS 感知到的变化) 进行阈值处理, 采用如下形式的阈值函数:

$$\text{threshold}(B(i)) = \begin{cases} 0, & \text{若 } H(B(i)) < \tau \\ 1, & \text{否则} \end{cases}, \quad i \in \Omega \quad (16)$$

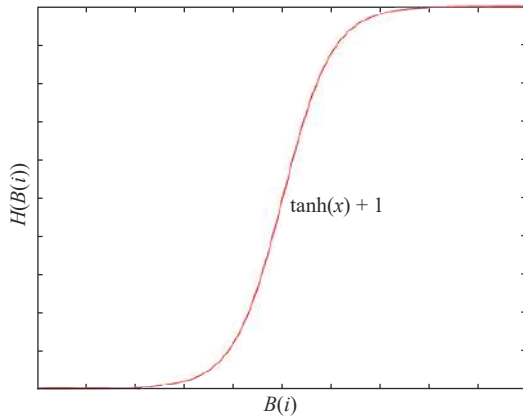


图 10 客观灰度变化与感知到的变化之间的关系
Fig.10 The relationship between objective grayscale changes and perceived changes

其中, τ 表示阈值参数, 正常取值范围 $\tau \in [0, 1]$, 这意味着将 HVS 感知到的变化大到一定程度的部分作为语义对象的边缘. 通常美术构图中需充分考虑 HVS 的感知规律, 因而这一处理更符合美术构图的原则.

3 数值实验

为了验证本文提出的线描画提取算法, 我们进行了大量的实验. 实验在 3.6 GHz Intel CPU, 8 GB RAM 的个人计算机上进行, 算法采用 MATLAB 2016a 来实现. 除了式 (16) 中阈值参数 τ 外, 在第 2 节中对算法涉及的所有参数进行了说明, 这里对式 (16) 中阈值参数 τ 选取进行分析, 图 11 呈现了阈值参数

取不同值时算法的输出结果.

从中可见, 阈值参数 τ 的设定对算法的输出有很大的影响, 随着阈值参数的增大, 图像的边缘对噪声更敏感, 边缘涵盖过多的噪声信息直接影响结果的美观性. 而阈值参数偏小时容易遗漏原图像的一些细节信息. 实验中, 我们根据经验取阈值 $\tau = 0.6$.

随后, 我们用本文算法对大量图像进行了测试. 图 12 为任意选取的几幅测试图像, 其实现的线描画效果如图 13 所示. 从中可见, 提取的线描画利用简洁流畅的线描线条勾勒了语义对象的边缘, 能凸显出图像中的主要信息, 视觉上相当美观. 由于该算法直接从自然图像中生成线描画, 与图 1 呈现的艺术家们手工绘制的线描画相对比, 具有“完全写实”的风格.

4 总结与展望

本文构造了一种基于方向场正则化的线描画生成算法, 该算法的显著特点体现在以下两个方面: 1) 基于自表示的思想, 采用 Tikhonov 正则化方法对方向场进行正则化处理, 且对正则化算法进行了加速; 2) 利用 HVS 的非线性特点设计感知阈值算法. 数据实验表明, 本文算法可从任意图像中提取相当美观的线描画. 本文的工作可从以下几个角度进行拓展: 1) 如第 2.3 节中对 RDF-DoG 滤波的结果作进一步处理, 沿着边缘方向作平滑滤波, 有望改善提取的线描画的质量; 2) 本文算法是对灰度图

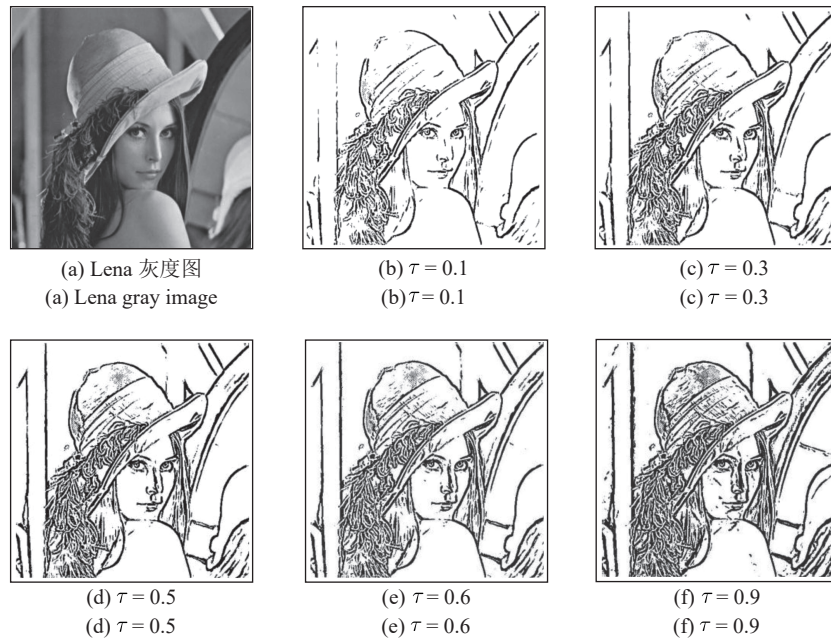


图 11 Lena 图像在不同参数 τ 时的线描画
Fig.11 Line drawing of Lena images at different parameters τ

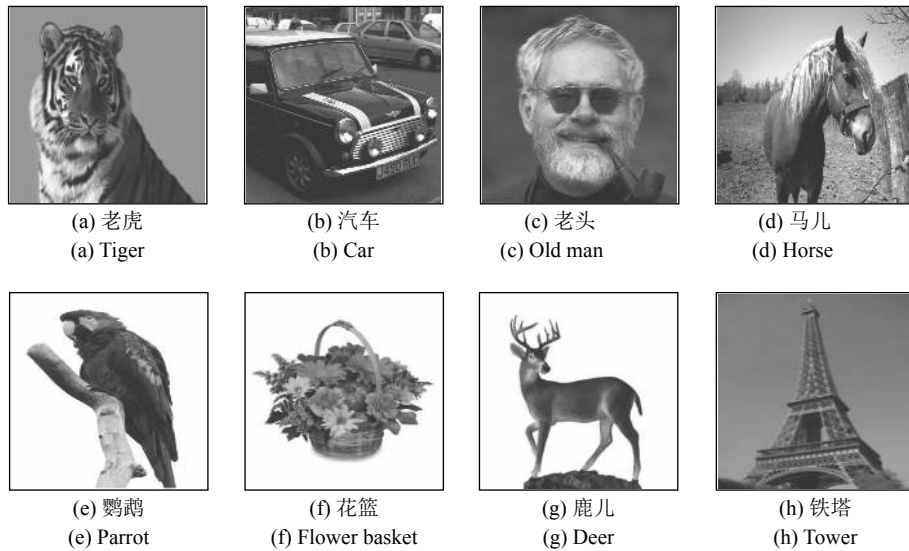


图 12 测试图像
Fig. 12 Test image

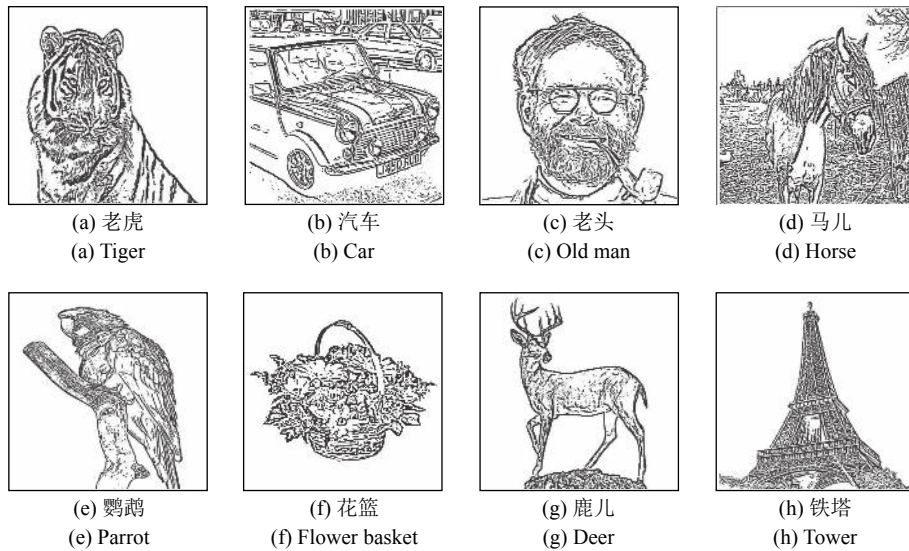


图 13 线描画图
Fig. 13 Line drawing

像设计的, 如何刻画彩色图像的边缘, 从彩色图像中提取线描画, 值得进一步研究; 3) 将本文算法扩展到视频序列, 利用邻近帧的信息来优化算法的性能; 4) 与其他风格转化问题一样, 线描画的风格转化可以利用深度学习方法实现, 当然, 这需要经过一定训练的志愿者根据给定自然图像绘制一定数量的线描画作为训练样本, 工作量很大; 5) 严格来讲, 图像风格转化问题的构建不够明确. 图像的“风格”难以明确定义, 也难以准确量化, 这使得风格转化算法的处理效果只能采用主观评价. 如何客观地评价风格转化算法的性能同样是值得考虑的问题, 对算法性能评价问题的研究会使得风格转化问题的构

建更加明确.

References

- 1 Pandey R K, Karmakar S, Ramakrishnan A G. Computationally efficient approaches for image style transfer. arXiv preprint. arXiv: 1807.05927, 2018.
- 2 Decaudin P. Cartoon-looking rendering of 3D-scenes. *INRIA*, 1996.
- 3 DeCarlo D, Santella A. Stylization and abstraction of photographs. In: *Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques. New York, USA: TOG*, 2002, 21(3): 769-776
- 4 Santella A, DeCarlo D. Visual interest and npr: an evaluation and manifesto. In: *Proceedings of the 3rd International Symposium on Non-Photorealistic Animation and Rendering. New*

York, USA: ACM Press, 2004. 71–78

- 5 Kyprianidis J E, Döllner J. Image abstraction by structure adaptive filtering. In: Proceedings of the 6th Theory and Practice of Computer Graphics Conference. Manchester, UK: TPCG, 2008. 51–58
- 6 Kang H, Lee S, Chui C K. Flow-based image abstraction. *IEEE Transactions on Visualization Computer Graphics*, 2009, **15**(1): 62–76
- 7 Qian W H, Xu D, Yue K, Guan Z. Image abstraction painting of flow-like stylization. *Tehnicky Vjesnik*, 2015, **22**(4): 837–844
- 8 Gatys L A, Ecker A S, Bethge M. Image style transfer using convolutional neural networks. In: Proceedings of the 2016 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016. 2414–2423
- 9 Johnson J, Alahi A, Li F F. Perceptual losses for real-time style transfer and super-resolution. In: Proceedings of the 14th European Conference on Computer Vision. Amsterdam, Netherlands: Springer, 2016. 694–711
- 10 Ulyanov D, Lebedev V, Vedaldi A, Lempitsky V. Texture networks: feed-forward synthesis of textures and stylized images. In: Proceedings of the 33rd International Conference on Machine Learning. New York, USA. 2016. **1**(2): 4
- 11 Elad M, Milanfar P. Style-transfer via texture-synthesis. *IEEE Transactions on Image Processing*, 2017, **26**(5): 2338–2351
- 12 Kwatra V, Essa I, Bobick A. Texture optimization for example-based synthesis. *ACM Transactions on Graphics (ToG)*, *ACM*, 2005, **24**(3): 795–802
- 13 Buades A, Coll B, Morel J M. A non-local algorithm for image denoising. In: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, USA: IEEE, 2005. 60–65
- 14 Horn R A and Johnson C R. *Matrix Analysis*. New York: Cambridge University Press, 2012.
- 15 Marr D, Hildreth E. Theory of edge detection. *Proceedings of the Royal Society B: Biological Sciences*, 1980, **207**(1167): 187–217
- 16 Winnemöller H, Olsen S C, Gooch B. Real-time video abstraction. In: Proceedings of the 2006 ACM Siggraph Papers. New York, USA: TOG, 2006. **25**(3): 1221–1226



李晶晶 陕西科技大学电子信息与人工智能学院硕士研究生. 2017 年获得延安大学西安创新学院物联网工程专业学士学位. 主要研究方向为图像处理, 图像风格转化.

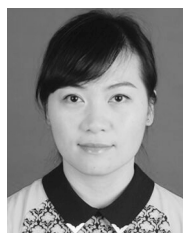
E-mail: li_jing058@163.com

(**LI Jing-Jing** Master student at the School of Electronic Information and Artificial Intelligence, Shaanxi University of Science and Technology. She received her bachelor degree in internet of things engineering from Xi'an Innovation College of Yan'an University in 2017. Her research interest covers image processing and image style transfer.)



许建楼 河南科技大学数学与统计学院副教授. 2013 年获得西安电子科技大学应用数学专业博士学位. 主要研究方向为图像处理变分方法, 稀疏优化. E-mail: xujianlou@126.com

(**XU Jian-Lou** Associate professor at the School of Mathematics and Statistics, Henan University of Science and Technology. He received his Ph.D. degree in applied mathematics from Xidian University in 2013. His research interest covers image processing variational method and sparse optimization.)



熊 静 陕西科技大学电子信息与人工智能学院讲师, 2016 年获得西安电子科技大学计算机应用博士学位. 主要研究方向为视频图像处理, 运动目标检测与跟踪.

E-mail: xiongjing@sust.edu.cn

(**XIONG Jing** Lecturer at the School of Electronic Information and Artificial Intelligence, Shaanxi University of Science and Technology. She received her Ph.D. degree in computer application from Xidian University in 2016. Her research interest covers video image processing, moving target detection and tracking.)



张选德 陕西科技大学电子信息与人工智能学院教授. 2013 年获得西安电子科技大学理学博士学位. 主要研究方向为图像恢复, 图像质量评价, 稀疏表示和低秩逼近理论. 本文通信作者.

E-mail: zhangxuande@sust.edu.cn

(**ZHANG Xuan-De** Professor at the School of Electronic Information and Artificial Intelligence, Shaanxi University of Science and Technology. He received his Ph.D. degree in applied mathematics from Xidian University in 2013. His research interest covers image restoration, image quality evaluation, sparse representation, and low rank approximation theory. Corresponding author of this paper.)