

## 基于概率随机裁剪的图像缩放算法

郭迎春<sup>1</sup>, 侯骏腾<sup>1</sup>, 于明<sup>1</sup>, 王睿俐<sup>2</sup>

(1. 河北工业大学计算机科学与软件学院, 天津 300401; 2. 梅西大学自然与数学科学系, 新西兰 奥克兰 4442)

**摘要:** 为提高图像缩放的速度, 提出一种结合阈值学习与依概率随机裁剪的快速内容感知图像缩放算法, 通过计算图像的重要度图, 利用径向基函数 (RBF, radial basis function) 神经网络进行阈值学习求出图像的重要度阈值, 根据阈值将图像分成保护区域和非保护区域, 并按缩放要求为其分配不同的缩放比, 分别进行依概率随机裁剪。在 MSRA 图像数据库上与目前流行的内容感知缩放方法进行对比, 实验结果表明, 所提方法的缩放时间明显低于其他算法, 而且在缩放效果上有明显的优势。

**关键词:** 阈值学习; 径向基函数; 依概率随机裁剪; 快速内容感知图像缩放

中图分类号: TN911.73

文献标识码: A

## Content-aware image resizing based on random-carving with probability

GUO Ying-chun<sup>1</sup>, HOU Jun-teng<sup>1</sup>, YU Ming<sup>1</sup>, WANG Rui-li<sup>2</sup>

(1. School of Computer Science and Engineering, Hebei University of Technology, Tianjin 300401, China;

2. Institute of Natural and Mathematical Sciences, Massey University, Auckland 4442, New Zealand)

**Abstract:** To improve the running speed of image resizing, a fast content-aware image resizing algorithm was proposed based on the threshold learning and random-carving with probability. Firstly the important map was calculated by combining the graph-based visual saliency map and gradient map. Then the image threshold value was obtained by radial basis function (RBF) neural network learning. And by the threshold, the original image was separated into the protected part and the unprotected part which was corresponding to the important part and the unimportant part of the original image individually. Finally, the two parts were allocated different resizing scales and the random-carving with probability was applied to them respectively. Experiments results show that the proposed algorithm has lower time cost comparing to the state-of-arts algorithms in MSRA image database, and has a better visual perception on image resizing.

**Key words:** threshold learning, radial basis function, random-carving with probability, rapid content-aware image resizing

### 1 引言

随着数字视频技术的发展, 不同分辨率及尺寸的显示设备种类激增, 如何使图像及视频能完整、适当地显示在不同的设备上, 成为亟待解决的问题。内容感知的图像缩放方法根据图像的内容尽量使图像变形发生在非重要区域, 从而获得更好的视觉效果, 因而成为图像缩放领域的一个研究热点。

在 2007 年, Avidan 等<sup>[1]</sup>首先提出了内容感知的图像缩放方法, 将图像缩放的过程分成图像的重要度图求解与依重要度图的缩放这 2 个部分。该方法经过改进, 大致分为 3 类: 基于线裁剪 (seam carving) 缩放方法、基于图像变形缩放方法和多操作图像缩放方法。

文献[1]首先提出了线裁剪算法, 每次从图像中删除一条能量值最小的八连通曲线, 直到图像满足

收稿日期: 2016-12-20; 修回日期: 2017-03-24

基金项目: 国家自然科学基金资助项目 (No.60302018); 天津市科技计划基金资助项目 (No.14RCGFGX00846, No.15ZCZDNC00130); 河北省自然科学基金资助项目 (No.F2015202239)

Foundation Items: The National Natural Science Foundation of China (No.60302018), Sci-tech Planning Projects Foundation of Tianjin (No.14RCGFGX00846, No. 15ZCZDNC00130), The Natural Science Foundation of Hebei Province (No.F2015202239)

条件为止。该方法由于没有考虑到裁剪线被删除后其左右两侧的像素点相邻后产生的新能量而被称为 Backward 方法。2008 年, Rubinstein 等<sup>[2]</sup>提出 Forward 线裁剪算法, 该算法将删除裁剪线后其左右两侧像素点相邻后产生的能量记入变形能量中, 提高了视觉效果, 但没有考虑到缩放过度对图像结构的破坏。此后, Shi 等<sup>[3]</sup>在 Forward 线裁剪算法基础上进行了改进, 避免了双向线裁剪过程中的图像结构变形。文献[4,5]进一步改进了重要度能量函数。文献[6]对图像的 DCT 域的马尔可夫特征进行扩展, 使该特征能反映更大范围元素间的相关性。文献[7,8]将线裁剪方法引入到立体图像的缩放当中, 扩大了线裁剪方法的应用范围。

基于图像变形缩放方法先用矩形或三角形的网格把图像分成小块, 然后根据图像的重要度对小块进行处理, 尽量使变形发生在重要度低的小块上, 重要度高的小块不改变或进行等长宽比的统一缩放, 该方法首先由 Wang 等<sup>[9]</sup>在 2008 年的 Siggraph Asia 会议上提出, 即优化的尺度变形图像缩放技术 (OSS, optimized scale-and-stretch), 由于这是一种连续的缩放方法, 所以其缩放结果不会因过度缩放而产生锯齿, 当图像缩放的尺度很大时, 也不会引起太大的视觉误差。随后, 雷励星<sup>[10]</sup>改进了图像变形缩放的优化函数, 以避免重要区域的扭曲旋转。Du 等<sup>[11]</sup>根据梯度、显著度和颜色信息得到图像的可伸展性, 把图像分为可伸展块与不可伸展块, 并赋予可伸展块不同系数来进行缩放。Zhang 等<sup>[12]</sup>通过鱼眼算法找出重要区域, 并通过多焦点共存来保护重要区域实现缩放。Lin 等<sup>[13]</sup>在缩放过程中将显著度图检测方法 with 图像分割结合起来, 从而改进了重要区域的能量函数。Liang 等<sup>[14]</sup>将直线特征约束加入到图像变形缩放过程中, 能够有效保持图像中的直线特征。但上述这些方法的共同缺点是效率低。

多操作图像缩放方法一般结合传统缩放、线裁剪和直接裁剪等多种方法, 该方案比单独使用其中一种方法效果好, 但是时间开销较多, 文献[15]通过比较各种缩放操作的结果选择最佳的方法进行缩放, 文献[16]使用的是将线裁剪与传统裁剪或缩放相结合的方法, 其效果在线裁剪的方法上有一定程度的提高。

在内容感知图像缩放的 3 类方法中, 线裁剪方法明显快于其他 2 种方法, 因此, 很多学者开始通

过改进线裁剪方法来提高运算速度。文献[17]提出一种流裁剪 (stream carving) 的方法, 通过将多条像素宽度为 1 的裁剪线合并成宽度更大的流, 这样, 运算一次就能删除多条裁剪线来减少运算时间。赵旦峰等<sup>[18]</sup>提出了一种基于随机置乱的内容感知图像缩放算法 (CARP, content-aware image resizing algorithm based on random permutation), 该方法简化了缩放的过程, 在求出图像的重要度图后, 计算重要度图每行/列的像素值和, 以此作为行/列能量向量, 将能量向量进行随机重排, 并删除图像中能量向量低于人工阈值的行/列以达到目标要求的高/宽度。该过程方法简单, 缩放迅速, 但是会因过度删除非保护区域而造成图像的局部失真, 并且当目标图像尺寸小于保护区域时没有给出缩放方案致使缩放失败。王会千等<sup>[19]</sup>提出了重要度扩散和自适应采样的缩放方法, 该方法综合考虑了图像的整体结构和重要区域, 根据重要性直接对图像各行和各列进行权值采样。相比于线裁剪方法, 以上 2 种方法将曲线缩放改成直线缩放, 易于操作, 并且有效地提高了图像缩放过程的效率。

鉴于以上方法的优缺点, 本文提出一种基于阈值学习与依概率随机裁剪的快速缩放算法, 以视觉显著度图与梯度图作为重要度图, 得到能量向量, 利用 RBF 神经网络进行阈值学习, 获得自适应能量阈值, 并根据该阈值将图像分为保护区域与非保护区域, 保护区域尽可能多地覆盖到包括视觉显著物体等内容的重要区域, 而非保护区域尽可能多地覆盖原图中重要度低的区域。在删除图像行/列的过程中, 本文采用依概率随机裁剪, 对保护区域与非保护区域采用不同缩放比例进行缩放。本文在有效提高缩放效果的同时, 能保证较少的时间开销。

## 2 依概率随机裁剪算法

内容感知的图像缩放算法分为 2 个部分: 重要度图和依重要度图的缩放。重要度图是根据图像得到标有重要部分的灰度图, 重要度越高, 灰度值越大; 根据重要度图进行图像缩放是内容感知图像缩放算法的核心部分, 当图像大小或比例发生改变时, 尽量使这些变化扩散到相对不重要的图像部分, 从而降低图像变形引起的视觉误差。

### 2.1 重要度图提取

本文将梯度图与视觉显著度图 (GBVS, graph-based visual saliency)<sup>[20]</sup>融合作为重要度图。因为

图像缩放中保护区域与非保护区域的分界处一般在显著物体的边界处，图像的边界信息是重要度图求解过程中的重要内容，梯度图能很好地反映图像的边界信息，而且图像中的主体区域也是图像中的重要部分，GBVS 算法能够准确地识别出图像中的主体区域并加大主体显著度与背景区域显著度的差异，采用 GBVS 与梯度图融合的方法既能保护图像中的结构信息，又能保护图像中的主体区域。其中，梯度图是先将彩色图像转化成灰度图像，并用式(1)进行求解，重要度图  $I_E$  的求解如式(2)所示。

$$I_{\text{Grad}} = \left| \frac{\partial}{\partial x} I_{\text{gray}} \right| + \left| \frac{\partial}{\partial y} I_{\text{gray}} \right| \quad (1)$$

$$I_E = I_{\text{GBVS}} + I_{\text{Grad}} \quad (2)$$

其中， $I_{\text{Grad}}$  为图像  $I$  的梯度图， $I_{\text{gray}}$  为图像  $I$  的灰度图， $I_{\text{GBVS}}$  为图像  $I$  的视觉显著度图。

### 2.2 缩放算法流程

本文方法是从行和列分别进行缩放的双方向缩放方法，以单方向缩放为例，首先用行抽取/复制或列抽取/复制的方法将图像缩放至与目标图像的比例相等，然后进行等比例缩放至目标图像大小。以下的算法流程通过删除图像列为例进行说明。如果图像需要缩放行，只需将图像进行转置并用处理图像列的方法对图像进行处理，完成后再转置回去即可。

单方向缩放方法的流程如图 1 所示。

假设给定的原图像  $I$  分辨率大小为  $m \times n$ ，缩放后的目标图像  $I_0$  分辨率大小为  $m' \times n'$ 。

单方向缩放方法的流程如下。

首先，利用视觉显著度计算模型 GBVS 计算得到图像的视觉显著度图  $I_{\text{GBVS}}$ ，并用式(1)计算梯度图  $I_{\text{Grad}}$ ，融合得到重要度图  $I_E$ 。根据目标大小判断缩放方式，当  $m < m'$  或  $n < n'$  时，需要删除操作。如果  $\frac{m}{n} > \frac{m'}{n'}$ ，则需要删除图像的列；如果  $\frac{m}{n} < \frac{m'}{n'}$ ，则需要删除图像的行，此时，将图像进行转置变成删除图像列的情况（即  $m、n$  值互换），在图像处理完成后再转置回来；当  $m > m'$  且  $n > n'$  时，需要复制操作：如果  $\frac{m}{n} > \frac{m'}{n'}$ ，则需要复制图像的行；如果  $\frac{m}{n} < \frac{m'}{n'}$ ，则需要复制图像的列，此时，将图像进行转置变成复制图像行的情况（即  $m、n$  值互换），在图像处理完成后再转置回来。对重要度图  $I_E$  的每一列求和得到能量向量，即能量向量为  $\mathbf{v} = \text{sum}(I_E)$ ，其分辨率大小为  $1 \times n$ 。

然后，用 RBF 神经网络求图像的阈值  $T_m$ ，并根据阈值  $T_m$  将图像分成保护区域与非保护区域。设  $v_i$  为向量  $\mathbf{v}$  中的第  $i$  个值，如果  $v_i \geq T_m$ ，则原图像的第  $i$  列属于保护区域；如果  $v_i < T_m$ ，则原图像的第  $i$  列属于非保护区域。如果需要删除/复制的图像列数超过图像非保护区域图像列数的一半，则分别求出保护区域与非保护区域的缩放率，并分别依概率随机裁剪这 2 个区域至与目标图像的长宽比相等。如果需要删除/复制的图像列数没有超过图像非

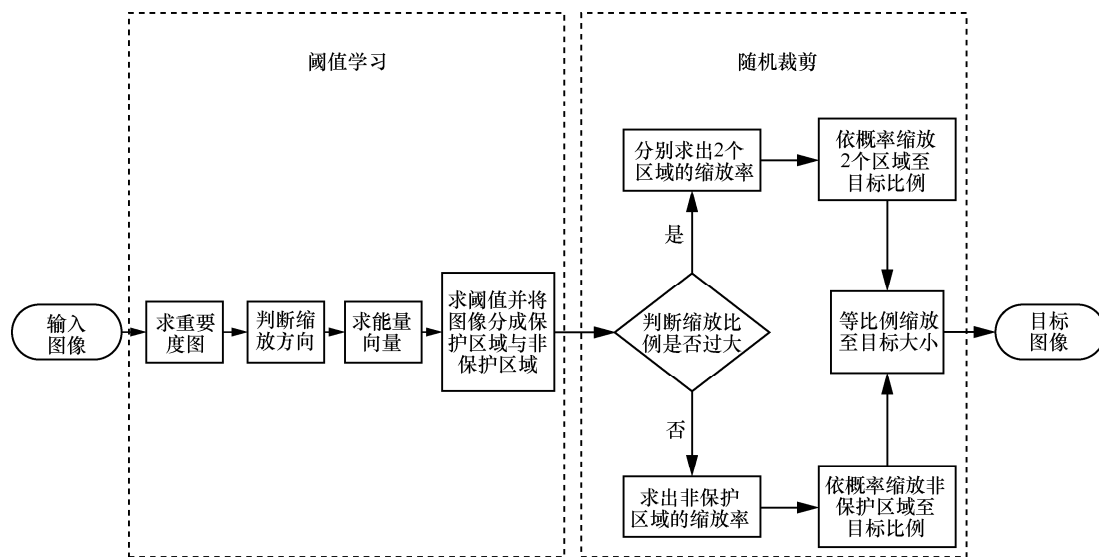


图 1 单方向缩放方法的流程

保护区域图像列数的一半, 则只依概率随机裁剪非保护区域至与目标图像的长宽比相等。

最后, 等比例缩放至目标图像大小。

双方向缩放算法即在  $x$  轴方向和  $y$  轴方向均进行缩放。大多数缩放的算法都会从单一方向扩展到 2 个方向, 本文双方向缩放方法也是从上述的单方向缩放算法扩展过来。

本文所使用的双方向缩放算法是先按照单方向缩放算法流程对原图像的  $x$  轴方向进行缩放至目标大小宽度, 再按照同样的方法对原图像的  $y$  轴方向进行缩放至目标大小高度。但是如果目标图像要求原始图像成倍地减小, 即缩放的比例较大时, 会导致图像列/行的过度缩放, 所以本文设置了一个最大的缩放阈值  $T_{MAX}$ , 当进行删除操作时, 如果图像的  $x$  轴方向与  $y$  轴方向对图像行/列的删除量均超过  $T_{MAX}$ , 则在缩放比例小的方向将图像缩放为原始图像的  $(1-T_{MAX})$  倍, 在缩放比例大的方向将图像缩放

为原始图像的  $(1-T_{MAX}) \times \min \left( \frac{m'}{n'}, \frac{n'}{m'} \right)$  倍, 然后将

图像等比例缩放至目标图像大小。当进行复制操作时, 如果图像的  $x$  轴方向与  $y$  轴方向对图像行/列的复制量均超过  $T_{MAX}$ , 则在缩放比例小的方向将图像缩放为原始图像的  $(1+T_{MAX})$  倍, 在缩放比例大的方向将

图像缩放为原始图像的  $(1+T_{MAX}) \times \max \left( \frac{m'}{n'}, \frac{n'}{m'} \right)$

倍, 然后将图像等比例缩放至目标图像大小。其中,  $T_{MAX}$  值是一个经验系数, 本文中  $T_{MAX}=0.3$ 。

### 2.3 阈值学习

对于阈值  $T_m$  的确定, 一般是设置一个固定的阈值, 或是根据图像直接计算一个阈值, 如文献[18]计算得

$$T_{MAX} = \frac{1}{2} \max(\mathbf{v}) \quad (3)$$

但是, 由于图像种类各异, 情况不一, 采用固定的阈值或直接使用一个简单的计算式所计算出阈值很难准确地接近每幅图像的最优阈值  $T_m$ 。RBF 神经网络被广泛地应用于逼近非线性函数, 可以处理系统内难以解析的规律性, 具有良好的泛化能力, 并有很快的学习收敛速度。这里, 利用 RBF 进行机器学习, 为每一幅图计算最优的阈值。本文

以图像的能量向量  $\mathbf{v}$  的分布直方图作为原始图像的特征进行机器学习。其分布直方图的求解过程如下。

1) 求出能量向量  $\mathbf{v}$  的最大值:  $v_{\max} = \max(\mathbf{v})$ 。

2) 能量向量的归一化处理:  $J_i = \frac{v_i}{v_{\max}}$ ,  $J_i$  的取值范围是  $[0,1]$ ,  $i$  的取值范围是  $\{1, \dots, n\}$ 。

3) 将区间  $[0,1]$  平均分成  $N$  份, 每一份的大小是  $\frac{1}{N}$ ,  $k$  是其中任意一份,  $k$  的取值范围是  $\{1, \dots, N\}$ 。

4) 对于每一个给定的  $k$  值, 计算满足  $J_i \leq \frac{k}{N}$  的能量向量值  $J_i$  的个数  $m_k$ 。

5) 计算  $m_k$  在总列数中所占的比例数  $h_k = \frac{m_k}{M}$ ,  $M = \sum_{k=1}^N m_k$ 。

根据  $h_k$  得到分布直方图  $\mathbf{H} = [h_1, h_2, h_3, \dots, h_N]$ 。

不同图像的分布直方图不同, 且分布直方图不受能量向量排列顺序的影响。

本文的最优阈值通过训练阶段得到最优阈值模型, 在测试时利用最优阈值模型和图像的分布直方图就可以得到当前图像的最优阈值。在训练阶段, 首先采用图像分布直方图与真实标注图进行手工标注的方法得到最优阈值, 手动调整阈值, 在图像中能量值大于阈值的部分以覆盖真实标注图中标出的主要物体时的阈值为最优阈值, 然后利用 RBF 神经网络对直方图和得到的阈值进行训练。测试时, 对所测试的图像利用 RBF 神经网络的最优阈值模型和当前图像的分布直方图即可得到图像的最优阈值。

### 2.4 依保护度确定缩放比

根据 2.3 节求出的最优阈值, 将图像分成保护区域与非保护区域 2 个部分。能量值大于阈值的图像列属于保护区域, 能量值小于阈值的图像列属于非保护区域。

一般情况下, 图像中的保护区域在整个图像中所占比例不大, 只对非保护区域进行缩放即可完成缩放目标, 那么可以根据阈值将图像分成保护区域与非保护区域 2 个部分, 删除/复制非保护区域的图像列直到完成缩放目标。

当一幅图像中的保护区域占整个图像的比例过大或要求缩放图像的长宽比与原图像的长宽比差距较大时, 如果只对非保护区域进行缩放, 很容易因需要删除/复制的图像列数目不足而导致缩放

失败，即使能够完成缩放，非重要区域也会因删除/复制的列数过多而变形严重。如图 2 所示，其允许删除/复制的图像列只有 15%，不能进行长宽比差距较大的缩放。

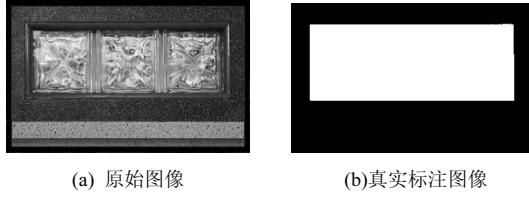


图 2 允许删除图像列较少的原始图像及其对应的真实标注图像

为依保护度确定缩放比，保证图像缩放的整体视觉效果，本文根据算法 1 计算保护区域与非保护区域的删除量。在算法 1 中， $r$  是目标缩放比例， $n_p$  为保护区域的图像列数， $n_{up}$  为非保护区域的图像列数， $n_{pd}$  为保护区域需要删除的图像列数， $n_{upd}$  为非保护区域需要删除的图像列数，本文中阈值  $T_{del}$  设为 0.5， $T_m$  是由径向基函数机器学习得到的。

使用上述方法进行缩放，对于任意的目标缩放比例都能完成缩放目标，且当目标缩放比例小时，只缩放非保护区域，使保护区域不受任何破坏，尽量减少图像的损失和变形。

计算保护区域与非保护区域删除量的算法表示如下。

**算法 1** 保护区域与非保护区域删除量

输入 测试图像  $I$ ，阈值  $T_m$ ， $T_{del}$ ，图像列数  $n$

输出  $n_{pd}$ ， $n_{upd}$

初始化  $n_p = 0, n_{up} = 0, i = 0$

根据 2.3 节求能量向量  $v$ ;

while  $i < n$

if  $v_i \geq T_m$

$n_p = n_p + 1$ ;

else

$n_{up} = n_{up} + 1$ ;

if  $(1-r) \frac{n}{n_{up}} \leq T_{del}$

计算非保护区域删除量:  $n_{upd} = (1-r)n$ ;

else

计算保护区域缩放比例:  $x = \frac{[n \cdot r - (1 - T_{del})n_{up}]}{n}$ ;

计算非保护区域删除量:  $n_{upd} = (x + T_{del})n_{up}$ ;

计算保护区域删除量:  $n_p = x \cdot n_p$ ;

**2.5 依概率随机裁剪**

根据 2.3 节的最优阈值将图像分成保护区域与非保护区域并确定不同的缩放比例后，需要对保护区域与非保护区域分别进行缩放。本文依概率随机裁剪的方法思路是：根据能量向量  $v$  得到一个关于图像中第  $i$  列的概率函数  $P_i$ ，如果第  $i$  列所对应的能量值  $v_i$  大，则其相应的概率函数  $P_i$  的值就小，即此图像列被删除的概率小；如果第  $i$  列所对应的能量值  $v_i$  小，则对应的概率函数  $P_i$  的值就大，即此图像列被删除的概率大。根据这一思想，得到一组概率模型

$$P_i = 1 - \frac{v_i}{v_{max}} \tag{4}$$

$$P_i = 1 - \frac{v_i - v_{min}}{v_{max} - v_{min}} \tag{5}$$

$$P_i = 1 - \left[ \frac{v_i - v_{min}}{v_{max} - v_{min}} \right]^2 \tag{6}$$

$$P_i = 1 - \left[ \frac{v_i - v_{min}}{v_{max} - v_{min}} \right]^{\frac{1}{2}} \tag{7}$$

$$P_i = 1 - \left[ \frac{v_i - v_{min}}{v_{max} - v_{min}} \right]^{\frac{3}{10}} \tag{8}$$

其中， $v_{min}$  是能量向量  $v$  中的最小值， $v_{max}$  是能量向量  $v$  中的最大值。经实验验证，式(8)的效果最好，本文选用式(8)作为概率函数。

依概率随机裁剪的过程为对于图像的第  $i$  列计算其概率函数值  $P_i$ ，得到一个在 [0,1] 内的随机值  $R_i$ ，比较  $P_i$  与  $R_i$  的大小，如果  $P_i > R_i$ ，则删除所对应的第  $i$  列，如果  $P_i < R_i$ ，则不对第  $i$  列进行处理。按以上方法循环删除直到图像满足目标图像的比例为止。

依概率随机裁剪还有一个优于随机裁剪的特点。在随机裁剪方法<sup>[18]</sup>中，阈值分割处的保护区域一侧对图像不进行任何处理，而在非保护区域一侧对图像进行一定比例的缩放，这样会导致阈值分割处的图像过渡不平滑。而依概率随机裁剪方法中，在阈值分割处附近的概率函数值很小，即此处的图像缩放比例很低，与保护区域一侧对图像进行的零删除率接近，所以，使图像在阈值分割处过渡平滑。

本文提出的依概率随机裁剪的方法既能避免现有的裁剪方法对相邻多列重要度接近的区域过度缩放而导致视觉失真，又能利用重要度图删除图像中不

重要的图像列而保存图像的相对重要的区域。

### 3 实验与分析

本文综合考虑阈值学习、区域保护、依概率随机裁剪缩放和内容感知多种因素使图像缩放达到了很好的效果。验证本文方法有效性的图像库为 MSRA 图像数据集，该图像库包括 1 000 幅原始图像及其真实标注图像，选择这一图像库的原因有以下几点。首先，该图像库包括的图片种类很多，涉及人物、动物、自然景物、几何结构和纹理结构等，并且前景与背景的差别大小不一，有很好的说明性；其次，图中显著物体比较明显，且图像库包含了所有图像的真实标注图像，可以根据真实标注图像得到训练图像的阈值，避免了人工寻找阈值时出现的误差；最后，根据真实标注图像得到训练图像的阈值速度快。本文将这 1 000 幅图像分成 2 个部分，随机选取其中 800 幅作为训练图像，200 幅作为测试图像。该图像库包括的图片种类很多，并不是只针对某一类或几类图像，所以，本文使用 MSRA 图像数据集的 200 幅图像对本文方法与其他图像缩放方法进行对比实验。

#### 3.1 训练图像的阈值获取

本文采用结合原始图像与真实标注图像进行手工标注的方法得到训练图像的最优阈值。人工标注出的最优阈值是最符合人类视觉审美的，使用该阈值训练得到的最优阈值模型对测试图像进行运算能够得到更加符合人类视觉审美的最优阈值。人工标注的主要过程为将原始图像中能量向量大于阈值  $T_m$  的部分标记出来；与真实标注图像对比，如果被标记的区域没有完全覆盖真实标注图像中标出的主要物体，则减小阈值  $T_m$  直到被标记的区域恰好覆盖主要物体，如果被标记的区域能够完全覆盖主要物体并且有一定的多余部分覆盖到了主要物体之外的区域，则增加阈值  $T_m$  直到被标记的区域恰好覆盖主要物体。

图 3 显示了原始图像、真实标注图像、采用固定阈值区分保护区域与非保护区域的图像 ( $T_m=0.5$ )、人工标注的阈值区分保护区域与非保护区域的图像 ( $T_m=0.3$ )，其中，保护区域加亮显示。

通过以上方法得到 MSRA 图像数据集的 800 幅图像的最优阈值，使用 RBF 神经网络进行训练，并用得到的最优阈值模型对测试图像求解，从而得到测试图像的最优阈值。

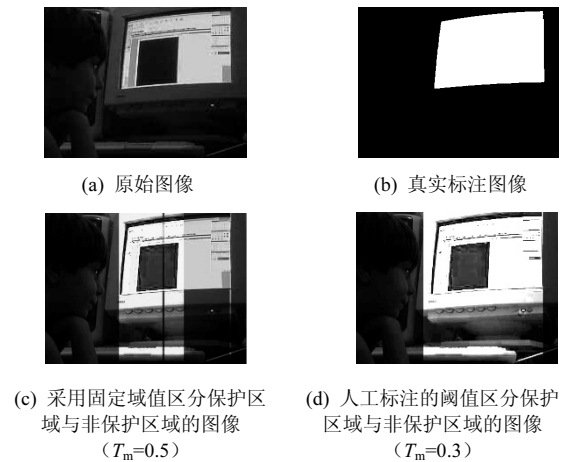


图3 固定阈值与人工阈值的比较

利用每幅训练图像的分布直方图作为目标输入，人工标注的阈值作为目标输出。在测试过程中，输入原始图像的分布直方图即可得到针对当前图像的理想阈值。

#### 3.2 缩放效果对比

本文对比了 CARP<sup>[18]</sup>、线裁剪<sup>[2]</sup>、流裁剪<sup>[17]</sup>、等比例缩放<sup>[21]</sup>和本文方法共 5 种方法的实验结果。其中，等比例缩放方法属于传统缩放方法，其他方法都属于内容感知图像缩放方法中缩放速度较快的方法。本文根据相关文献对以上方法编程实现，并使用这些方法对图像进行高度不变、宽度缩放到原图的 60% 的处理，部分实验结果如图 4 所示，图 4 中包含 4 幅原始图像，每幅原图的第 1 行是不同算法的缩放结构，第 2 行显示不同算法的裁剪线，由于使用 CARP 方法无法缩放第 4 幅图像，这里没有显示出来。

从对图 4 第 1 幅图像的缩放结果可以看出，当缩放后图像的宽度小于原图像中重要物体的宽度时，在 CARP 方法的缩放图中，左侧的伞与右侧的气球都发生了严重的变形；在线裁剪方法中，伞的下部变形也比较明显；在流裁剪方法中，虽然重要物体已经充满了整个图像，但是它也对重要物体的边界上进行了过度的裁剪，从而导致了气球的圆面出现了棱角，伞的下部也出现了变形；等比例缩放方法将重要物体缩到了一个很小的范围；本文方法能在保证整幅图像不发生严重变形的条件下尽量少地裁剪重要物体区域，从而将变形扩散到重要物体的内部，减少因重要物体的局部变形严重而引起的视觉误差。在第 2 幅原始图像中，该图像的真实标注图像中标出的区域是鼻子，但是线裁剪与流裁



图 4 各方法缩放效果及裁剪线

剪方法保护了眼睛，却缩小了鼻子部分，导致重要区域的过度删除，并且流裁剪方法在人脸左右两侧的皮肤与灰色边界处留下了明显的颜色过渡线；CARP、等比例缩放方法与本文方法都较好地保存了鼻子的区域，但是 CARP 方法将人的左眼直接去掉了一半，从而造成了严重的视觉扭曲；而等比例缩放方法相对于本文方法，其鼻子区域缩放太大，没有保护好图像的重要区域。在第 3 幅原始图像中，其重要物体是石碑，而左侧的大树不属于重要物体

的部分，5 种方法中只有 CARP 方法对大树有大量裁剪，从而尽量保护图像中的石碑部分，但是该方法并没有留下完整的重要物体，而是将石碑删除掉了一部分；使用线裁剪与流裁剪方法得到的图像，其石碑的下半部分都发生了严重扭曲，石碑的直线结构遭到严重破坏；本文方法虽然也有裁剪线通过重要物体，使石碑的宽度有所减小，但是本文方法得到的石碑仍然比等比例缩放方法得到的石碑宽，并且石碑中的直线结构没有遭到破坏。在第 4 幅原

始图像中，CARP 方法已经无法对这种重要物体宽度较大的图像完成缩放；由于背景中草的区域图像梯度值较大，线裁剪方法将此区域识别成重要区域，而对真正的重要区域鲜花部分进行了过度删除，效果很差；流裁剪方法有效地保护了鲜花部分的右侧区域，却对鲜花部分的左侧区域进行了过度裁剪，结果只剩半只鲜花，效果也很差；等比例缩放方法得到的图像中鲜花部分所占比例过小；本文方法能将鲜花部分扩展到整幅图像，最大限度地减小了重要物体的变形，同时重要物体的各部分比例也没有发生变化。

### 3.3 主观评价

本文使用了包括 CARP 方法、线裁剪方法、流裁剪方法、等比例缩放方法和本文方法共 5 种方法对 MSRA 图像数据集的 200 幅图像进行高度不变、宽度缩放到原图 60% 的处理，并对处理结果进行主观评价。为避免随机因素与个人主观差异，由 20 名同学分别对 5 种方法得出的结果进行评价，然后将所得的结果求平均值得出最终的结论。其主观评价的标准如下。

- 1) 是否能成功缩放。
- 2) 重要区域是否有视觉不可接受的明显变形。
- 3) 非重要区域是否有视觉不可接受的明显变形。
- 4) 阈值分割处是否平滑过渡。

本文将 5 种方法的结果图像都显示在同一程序界面中，如图 5 所示。

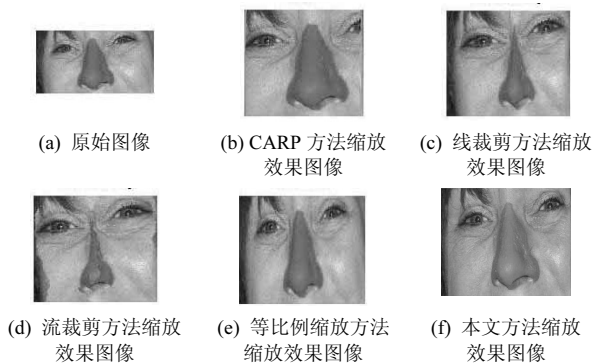


图 5 主观评价对比

为了测试方便，设计了测试程序软件，在软件界面的图片名输入框中输入图片编号后显示图像，程序将会显示包括原图在内的 6 幅图像，评价者根据主观评价标准在 5 幅缩放后的图像中选出 2 幅缩放效果最好的图像。对评价结果进行统计，图片数是评价者选出的缩放效果好的图像总数，百分比是

缩放算法效果好的图像占 200 幅测试图像的比例，统计结果如表 1 所示。

方法	图片数	百分比
CARP 方法	68	34%
线裁剪方法	95	47.5%
流裁剪方法	42	21%
等比例缩放方法	32	16%
本文方法	163	81.5%

由表 1 可知，本文的主观评价结果明显优于其他方法。其中，由于算法的局限性，CARP 方法无法对其中的一部分图片完成缩放，而且即使能完成缩放过程，也会由于区分保护区域与非保护区域时存在误差或非保护区域的过度删除而影响缩放效果。线裁剪方法能有效避开图像中重要度高的区域而尽量减少视觉误差，但是由于其重要度图只用了梯度图，所以当图像中的主体区域比较平滑时，会造成主体区域过度删除。流裁剪方法能删除多个像素宽度的裁剪线，但是当删除裁剪线后，其左右两侧像素点变成相邻像素点，这些相邻像素点差异较大而影响视觉效果。等比例缩放方法由于没有考虑到重要区域而无法得到令人满意的效果。

### 3.4 运行时间对比

本文算法最大的优点是运算速度较快，明显快于线裁剪、流裁剪等方法。为测试时间效率，在同一计算机上对图 4 中的 4 幅原始图像均进行高度不变、宽度变成原来的 60% 的缩放操作。其中，操作系统为：Win7 (64 位)，开发语言为 Matlab R2014a，计算机配置为 Intel(R) Core(TM) i5-4570 CPU，内存为 8 GB。平均运行时间如表 2 所示。

方法	平均运行时间/s
CARP 方法	1.09
线裁剪方法	2.56
流裁剪方法	1.75
等比例缩放方法	0.09
本文方法	0.98

由表 2 可知，等比例缩放方法是最快的一种方法，因为它没有考虑到图像内容的不同重要度，缺少重要物体的识别过程，所以速度很快，但效果不佳。CARP 方法与本文方法的速度都明显快于线裁

剪与流裁剪方法, 本文方法的速度稍微比 CARP 方法快一些。在本文中, 计算重要度图平均用时为 0.97 s, 即缩放过程的用时仅为 0.01 s, 可见本文方法是一种非常快速的图像缩放算法, 只要重要度图的求解速度理想的话, 本文方法完全能应用到视频的实时缩放。

#### 4 结束语

本文提出了一种结合了阈值学习与依概率随机裁剪的快速内容感知图像缩放算法。该算法使用 RBF 神经网络对阈值进行机器学习, 能针对不同的图像选择合适的阈值, 并根据阈值将图像分成保护区域与非保护区域, 有效降低重要区域的误删率, 对保护区域与非保护区域分别进行依概率随机裁剪, 起到了防止过度删除非重要区域、依重要度缩放和阈值边界过渡平滑等作用。本文方法与 CARP 方法、线裁剪方法、流裁剪方法、等比例缩放方法等多种方法进行了实验对比, 实验结果表明本文方法能统筹兼顾保存图像内容、保护图像边缘区域和保护图像结构而提高图像的缩放效果, 并且本文方法是一种快速的图像缩放方法, 其时间效率明显高于其他方法。在重要度图求解比较理想的情况下, 本文方法有望应用于视频的实时缩放中, 这也是后续的研究的重点。

#### 参考文献:

- [1] AVIDAN S, SHAMIR A. Seam carving for content-aware image resizing[J]. ACM Transactions on Graphics, 2007, 26(3): 10-18.
- [2] RUBINSTEIN M, SHAMIR A, AVIDAN S. Improved seam carving for video retargeting[J]. ACM Transactions on Graphics, 2008, 27(3): 1-9.
- [3] SHI M L, PENG G Q, YANG L, et al. Optimal bi-directional seam carving for content-aware image resizing[C]//The 5th International Conference on E-learning and Games, Edutainment. 2010: 456-467.
- [4] LIU Z, YAN H B, SHEN L Q, et al. Adaptive image retargeting using saliency-based continuous seam carving[J]. Optical Engineering, 2010, 49(1): 017002.
- [5] CHO S H, CHOI H, MATSUSHITA Y, et al. Image retargeting using importance diffusion[C]// The ICIP. 2009: 977-980.
- [6] 盛国瑞, 高铁杠, 范礼, 等. 基于扩展马尔可夫特征的 seam-carving 篡改检测[J]. 通信学报, 2014, 35(6): 39-46.  
SHENG G R, GAO T G, FAN L, et al. Seam-carving forgery detection based on expanded Markov features [J]. Journal on Communications, 2014, 35(6): 39-46.
- [7] BASHA T, MOSES Y, AVIDAN S. Stereo carving a geometrically consistent approach[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(10): 2513-2525.
- [8] YUE B, HOU C, ZHOU Y. Improved seam carving for stereo image resizing[J]. EURASIP Journal on Wireless Communications and Networking, 2013(1): 1-6.
- [9] WANG Y S, TAI C L, SORKING O, et al. Optimized scale-and-stretch for image resizing[J]. ACM Transactions on Graphics, 2008, 27(5): 1-8.
- [10] 雷励星. 基于混合能量的内容敏感图像缩放新方法[J]. 计算机学报, 2010, 33(10): 2015-2021.
- LEI L X. Content-aware image resizing based on hybrid energy[J]. Chinese Journal of Computer, 2010, 33(10): 2015-2021.
- [11] DU H, LIU Z, JIANG J, et al. Stretchability-aware block scaling for image retargeting[J]. Journal of Visual Communication & Image Representation, 2013, 24(4): 499-508.
- [12] ZHANG L, SONG H, OU Z, et al. Image retargeting with multifocus fisheye transformation[J]. Visual Computer, 2012, 29(5): 407-420.
- [13] LIN S S, YE H I C, LIN C H, et al. Patch-based image warping for content-aware retargeting[J]. IEEE Transactions on Multimedia, 2013, 15(2): 359-368.
- [14] LIANG Y, ZHUO S U, LUO X N, et al. Image shrinkage based on hot-target map and featured edge preservation[J]. Journal of Software, 2011, 22(4): 789-800.
- [15] LIANG Y, SU Z, WANG C, et al. Optimised image retargeting using aesthetic-based cropping and scaling[J]. IET Image Processing, 2013, 7(1): 61-69.
- [16] RUBINSTEIN M, SHAMIR A, AVIDAN S. Multi-operator media retargeting[J]. ACM Transactions on Graphics, 2009, 28(3): 23.1-23.11.
- [17] DONG W M, ZHOU N, PAUL J C, et al. Optimized image resizing using seam carving and scaling[J]. ACM Transactions on Graphics, 2009, 28(5): 125.1-125.10.
- [18] 赵旦峰, 王博, 杨大伟. 基于随机置乱的内容感知图像缩放算法[J]. 吉林大学学报(工学版), 2015, 45(4): 1324-1328.  
ZHAO D F, WANG B, YANG D W. Content-aware image resizing based on random permutation[J]. Journal of Jilin University (Engineering and Technology Edition), 2015, 45(4): 1324-1328.
- [19] 王会千, 杨高波, 张兆扬, 等. 基于重要度扩散和自适应采样的图像/视频缩放技术[J]. 通信学报, 2011, 32(8): 165-170.  
WANG H Q, YANG G B, ZHANG Z Y, et al. Image and video resizing algorithm based on importance diffusion and adaptive scaling[J]. Journal on Communications, 2011, 32(8): 165-170.
- [20] HAREL J, KOCH C, PERONA P. Graph-based visual saliency[C]// Advances in Neural Information Processing Systems. 2006: 545-552.
- [21] RUBINSTEIN M, GUTIERREZ D, SORKINE O, et al. A comparative study of image retargeting[J]. ACM Transactions on Graphics, 2010, 29(6): 81-95.

#### 作者简介:



郭迎春 (1970-), 女, 河北张家口人, 河北工业大学副教授, 主要研究方向为数字图像处理、图像显著性检测、图像缩放、图像质量评价和人脸识别表情识别等。

侯骏腾 (1990-), 男, 河北黄骅人, 河北工业大学硕士生, 主要研究方向为图像处理与模式识别。

于明 (1964-), 男, 河北秦皇岛人, 河北工业大学教授, 主要研究方向为语音与图像视觉信息融合的生物特征识别、图像数学变换、图像与视频编码的高效算法和视觉计算及其应用 (视频数据挖掘、人脸识别、笔迹鉴别系统及其应用) 等。

王睿俐 (1965-), 男, 新西兰奥克兰人, 博士, 梅西大学副教授, 主要研究方向为模式识别、机器学习和复杂系统等。