直线特征保持的立体图像缩放

龚永义[1),3)] 李可宏[1)] 姜灵敏[1)] 吴贺丰[1),3)] 罗笑南[2),3)]

[1)](广东外语外贸大学信息学院 广州 510006)

[2)](中山大学信息科学与技术学院 广州 510006)

[3)](国家数字家庭工程技术研究中心(中山大学) 广州 510006)

摘 要 近年来随着 3D 显示设备的普及,内容感知的立体图像缩放问题得到了国内外学者的关注. 与 2D 图像缩放不同,立体图像缩放不仅需要保持并突显图像显著区域,还必须保持立体图像视差的一致性. 文中提出了一种立体直线特征保持的立体图像缩放的方法,通过在基于网格变形的图像缩放的方法中引入立体直线特征保持和内容感知的随机匹配点等约束条件,在保持立体图像视差的一致性的同时,较好地保持立体图像中直线区域的直线特征. 首先,基于立体图像视差图,定义立体图像 Hot-target 图,检测立体图像显著区域,并获取立体图像直线特征;其次,引入内容感知的随机匹配点作为视差一致性约束;最后,构建基于四边域网格的受限能量模型,优化实现立体图像缩放. 文中提供了若干实验数据说明该方法是有效的.

关键词 直线特征保持;内容感知;立体图像缩放;网格变形

中图法分类号 TP391 **DOI 号** 10.11897/SP.J.1016.2016.01307

A Straight-Line Preserved Resizing Method for Stereoscopic Images

GONG Yong-Yi[1),3)] LI Ke-Hong[1)] JIANG Ling-Min[1)] WU He-Feng[1),3)] LUO Xiao-Nan[2),3)]

[1)](*Cisco School of Informatics, Guangdong University of Foreign Studies, Guangzhou* 510006)

[2)](*School of Information Science & Technology, Sun Yat-Sen University, Guangzhou* 510006)

[3)](*National Engineering Research Center of Digital Life, Sun Yat-Sen University, Guangzhou* 510006)

Abstract With the popularity of 3D display equipments, content-aware stereoscopic image resizing problems have attracted great attentions in recent years. Different from 2D image resizing, resizing a pair of stereoscopic images is expected to preserve the disparity consistency of input images, in addition to visual salient contents. A 3D Straight-Line preserved resizing method for stereoscopic images is proposed, which introduces straight lines preservation and applies content-aware random matching to the stereoscopic constraints. Specifically, stereo Hot-target map based on disparity map is firstly calculated to detect visually salient contents and obtain 3D straight-lines of original stereoscopic image. Then, content-aware random matching points are introduced as the stereoscopic constraints. After that, each image is represented as a grid mesh and model an energy distortion function. Each quad mesh is scaled to the target size by minimizing total energy function. Finally, the target stereoscopic image is obtained by interpolation and mapping. Experimental results demonstrate its effectiveness in this paper.

Keywords straight-line preserved; content-aware; retargeting for stereoscopic images; mesh warping

---

收稿日期:2014-08-23;在线出版日期:2015-07-22. 本课题得到国家自然科学基金(61370160,61320106008,61402120)、广东省自然科学基金(2014A030310348)、广东省高等学校科技创新项目(2012KJCX0048)、广州市科技计划项目(2014J4100032)资助. **龚永义**,男,1970 年生,博士,主要研究方向为图像配准、图像分割、图像/视频缩放. E-mail: gongyongyi@gdufs.edu.cn. **李可宏**,男,1990 年生,硕士研究生,主要研究方向为 2D-3D 图像/视频转换、3D 图像缩放. **姜灵敏**,男,1956 年生,博士,教授,主要研究领域为图像处理. **吴贺丰**,男,1985 年生,博士,主要研究方向为计算机视觉、图像处理. **罗笑南**,男,1963 年生,博士,教授,博士生导师,主要研究领域包括数字家庭、移动计算、计算机图形学与 CAD、图像处理.

1 引 言

由于显示设备分辨率的多样性，图像缩放处理已成为图像处理领域的重要研究内容. 典型的图像缩放可视为插值问题，属于需要加入合适的约束条件才能解决的病态问题. 比较常见的约束条件包括保持缩放前后图像的主要特性，避免图像扭曲和不连续，突出显著区域等. 根据被处理图像的维度，图像缩放可分为 2D 图像缩放和 3D(立体)图像缩放.

最早引起国内外研究人员关注的是 2D 图像的缩放，现已涌现出一批重要的研究成果[1-9]. 1990 年至 2000 年前后，该类研究主要集中在图像的等比例缩放领域. 图像等比例缩放的研究大致分为基于插值[1]、基于重建[1-2]、基于学习[1,3]这三类图像缩放算法.

2000 年以来，随着不同规格高清设备、网络电视等可视设备的普及，2D 图像缩放的研究重点转移到图像非等比例缩放，这对图像缩放效果提出了新的更实际的要求. 内容感知的图像非等比例缩放已成为近十几年来图像缩放研究的热点课题. 此类研究的重点是准确检测图像中显著区域，避免图像扭曲、断裂，保持图像重要内容的主要特征和平滑性. 它们可大致分为离散和连续这两类缩放方法. 离散的方法主要包括基于 seam carving[4]、分块处理[5]以及智能裁剪[6]的图像缩放方法. 其不足之处是此类方法可能删除图像的有效信息，破坏其连续性. 连续的方法主要是基于网格变形的图像缩放方法. 首先构造覆盖原图像的网格平面，然后在特定约束下进行网格变形处理，最后将原图像映射到变形后的网格，得到缩放后的图像[7-9]. 约束条件的选择与图像显著区域、图像特征保持等有关. 该类方法的缺陷主要体现在：(1) 若网格变形控制不恰当，则可能出现网格叠加，破坏图像的连续性；(2) 可能引起图像扭曲. 为此，Wang 等人[8]提出了将梯度图和显著图融合的 Hot-Target 检测方法，通过 Hot-Target 约束重要区域的物体尽量不扭曲，同时增强边缘信息. Liang 等人[9]进一步提出了基于 Hot-Target 图和特征边缘保持的图像收缩方法，在图像重要区域检测中综合考虑显著区域和边缘信息，在图像缩放过程中加入直线特征约束，能够更加有效地保持图像的直线特征.

近年来，随着 3D 显示设备的普及，内容感知的立体图像缩放开始得到国内外研究人员的关注[10-22]. 但遗憾的是，人们发现：2D 图像缩放方面的研究成果并不能直接应用于立体图像缩放.

根据视觉成像原理，在双目立体视觉中，人们通过双目视差感知物体的深度信息，因此，保持左右视图视差的一致性非常关键. 直接将 2D 图像非等比例缩放应用于双目立体图像，左右视图对应区域扭曲程度的不一致将破坏左右视图原有的对应关系，产生垂直视差，甚至对应缺失. 如图 1 所示，我们采用 Wang 等人的方法[8]直接对立体图像的左右视图(见图 1(a)、(b))分别进行缩放处理，实验表明由此得到的立体图像效果不佳，给用户带来不舒适的观看体验(见图 1(d)、(e)). 进一步采用 SGM(Semi-Global Matching)半全局匹配方法[12]分析，可以发现，导致这种现象发生的主要原因是：缩放后的立体图像的视差图表明部分区域因左右视图没有找到匹配点而形成大片镂空现象，并产生垂直视差(见图 1(f))，破坏了原立体图像的视差一致性(见图 1(c)).

因此，传统的内容感知立体图像缩放方法的关键在于保持并突显图像显著区域，保持立体图像视差的一致性. 已有的研究成果可分为离散型和连续型两类方法[13]. 离散型方法主要是基于 seam carving 的方法. Utsugi 等人[14]最早将 seam carving 算法应用于双目立体图像，该方法得到左图待删除的 seam 点之后，在右图中找到左图 seam 点的所有对应点，进行相同的删除操作. “遮挡”区域的存在使得左右视图中并非每个点都能找到对应点，因此错误的匹配使得简单的 seam carving 立体图像缩放出现场景扭曲. 基于此，Basha 等人[15-16]提出了几何一致的 seam carving 立体图像缩放算法，该算法严格遵循两个约束，即同时保留或删除左右眼视图的对应点，以及不显示被遮挡点(Occluded Points). Yue 等人[17]将图像分割的思想引入立体图像缩放，首先基于 Panum 融合区域[18]精确分割出图像中的重要物体，进行立体 seam carving 操作时避免删除重要物体的内容，完全保持图像中的显著特征. 离散型方法的主要不足是：删除了图像的有效信息，容易破坏物体结构和连续性. 现有的立体图像 seam carving 方法仅能处理垂直 seam 线，存在一定的局限[15-16]. 连续型方法主要是基于网格变形的方法，此类方法能够弥补离散型方法破坏图像连续性的缺陷[19-22]. Chang 等人[19]提出了内容感知的立体图像编辑方法，利用 SIFT 尺度不变特征变换(Scale-invariant Feature Transform)方法找到左右图的稀疏对应点，并在网格变形的约束能量中加入了稀疏点的对齐

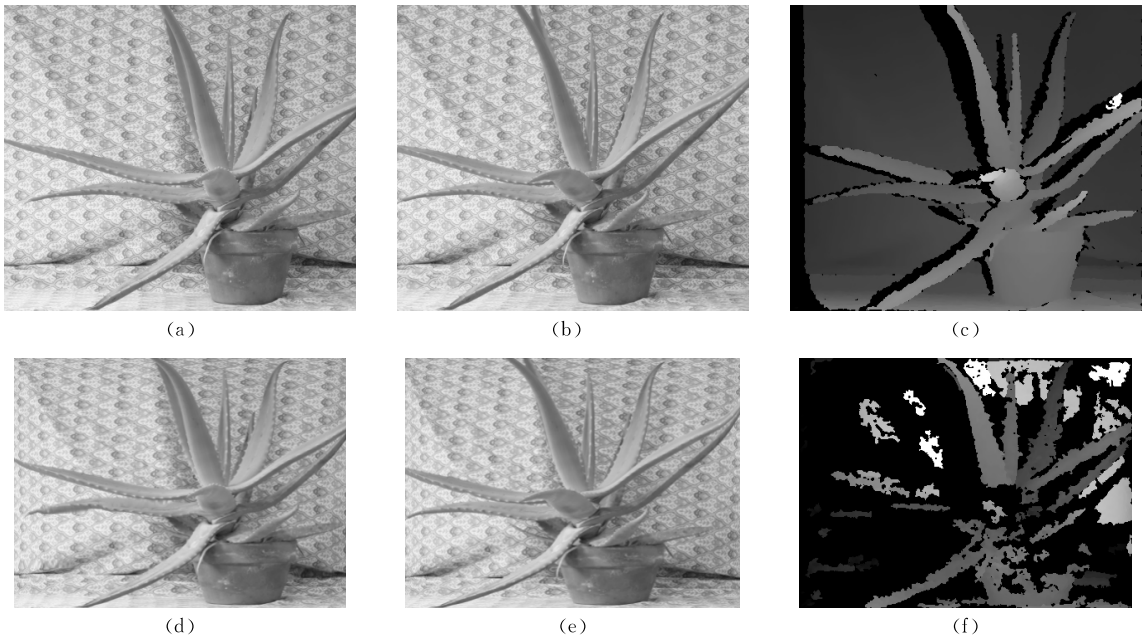


图 1　2D 图像非等比例缩放直接应用于双目立体图像的实验效果((a)、(b)、(c)分别是源立体图像左视图、右视图、对应的视差图.(d)、(e)、(f)分别是缩放后立体图像的左视图、右视图、对应的视差图.对比视差图(c)和(f)表明,直接将 2D 图像缩放算法用于立体图像缩放处理,可能会严重破坏立体图像的视差一致性)

(Alignment)约束和视差一致(Disparity Consistency)约束,较好地保留了图像视差信息.Cho 和 Kang 等人[20]提出了基于立体信息的显著区域检测方法,采用网格变形进行立体图像缩放更好地避免立体图像目标变形与减少视觉不适性.Lee 等人[21]提出 scene warping 立体图像缩放方法,该方法根据深度和色彩信息将立体图像对分解为多个层,对覆盖每一图像层的网格进行不同程度的缩放以达到突出重要物体,避免图像扭曲的目的.Lin 等人[22]提出了目标一致变形的立体图像缩放方法,对图像内容进行分割并对每个目标赋予显著值,将像素级别网格变形方法扩展到目标级别,减少图像扭曲.但是,现有的基于网格变形的方法对处理包含瘦长直线场景的立体图像效果不佳,稀疏对应关系局限于图像的场景结构.由于人眼对直线物体较为敏感,直线的扭曲导致人眼察觉图像失真,造成视觉不适.

在本文中,我们提出了立体直线特征保持的双目立体图像缩放方法.主要贡献包括:(1)我们在网格变形过程中引入了立体直线特征边缘约束,能够保持图像中的特征边缘,如瘦长直线目标不扭曲变形;(2)我们采用内容感知的随机匹配关系代替稀疏匹配点,能够更简单有效地保持视差一致性.另外,我们还采用梯度图和显著图检测图像重要区域以增强图像中的物体边缘信息,更好地保留重要物体的特征.

本文第 2 节详细描述了立体直线特征保持的双目立体图像缩放具体流程,包括立体 Hot-target、立体直线特征的定义,基于四边域网格的受限能量模型的构建与优化等;第 3 节从连续性、重要目标、立体直线特征、视差一致性保留等方面分析实验结果;第 4 节是结论.

2　直线特征保持的立体图像缩放

我们的方法主要由两部分组成:(1)立体图像显著特征的选取;(2)基于四边域网格的受限能量模型及优化.方法基本流程如图 2:首先,对输入的左右视图进行显著特征选取,包括立体 Hot-Target 图,立体直线特征等;然后,构造覆盖左右视图的四边域网格,建立基于网格的受限网格变形模型,其中模型约束条件由图像显著特征,边长约束、边界约束、垂直视差约束和视差一致性等组成;最后,优化模型并采用双线性插值的方法映射得到缩放后的立体图像.

2.1　立体图像显著特征选取

为了使图像中的重要目标和直线物体在缩放中尽可能地保持原来特征,我们选用立体 Hot-target 图、立体直线对为网格模型的主要图像特征约束,用于约束网格变形.其中,立体 Hot-target 图给出立体图像的显著区域信息,立体直线特征给出立体图像中较为显著的立体直线段.

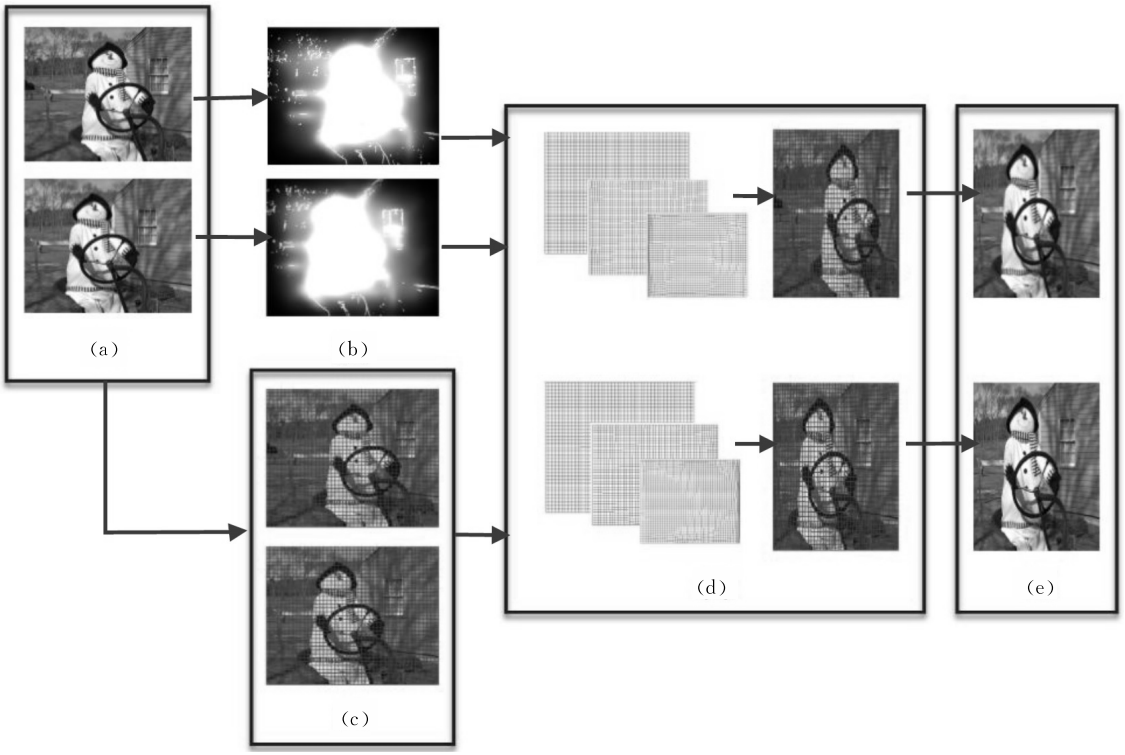


图 2 基本流程图((a)输入立体图像左、右视图;(b)对应的 Hot-Target 图;(c)构建覆盖图像的四边域网格;(d)网格变形过程与最终网格;(e)缩放后的图像)

2.1.1 立体 Hot-target 图

我们按以下方式定义立体 Hot-Target 图. 首先,在 Liang 等人[9]的基础上,定义 2D 图像 Hot-Target 图,

$$I_{\text{Hottarget}}=(\alpha\times(k\times I_{\text{Saliency}})+(1-\alpha)\times I_{\text{Edge}})\times I_{\text{Saliency}} \tag{1}$$

Hot-Target 图是融合视觉显著信息与图像边缘信息的像素级重要程度表示图. 其中,I_{Saliency} 为显著图,表示 2D 图像中每个像素的重要程度,我们采用文献[23]基于图论的视觉显著检测算法得到显著图,该方法模拟人体视觉注意机制,根据像素点在颜色、亮度、方向等方面与周边背景的对比生成显著图;I_{Edge} 为梯度图,目的是为了保留重要目标的边缘信息,因为人眼对目标边缘信息较敏感,而显著图中并不包含边缘信息;α 是调节显著图和图像边缘信息比例的权重,本文我们的实验将其设为 0.3;k 是显著图的增强因子,设为 10.

然后,分别计算立体图像左、右视图的 Hot-Target 图 $I_{\text{Hottarget,L}}$,$I_{\text{Hottarget,R}}$;采用 SGM 半全局匹配方法[12],计算左右视图的视差图矩阵 $\boldsymbol{M}$,获取左右视图像素的对应关系:

$$(y_{\text{R}},x_{\text{R}})=(y_{\text{L}},x_{\text{L}}+\boldsymbol{M}(y_{\text{L}},x_{\text{L}})),$$

其中,(y,x)表示像素 P 的行、列坐标;像素 P_{R} 是左视图像素 P_{L} 在右视图中的对应点.

我们定义立体 Hot-Target I_{SHT} 为

$$\begin{cases} I_{\text{SHT,L}}(P_{\text{L}})=I_{\text{Hottarget,L}} \\ I_{\text{SHT,R}}(P_{\text{R}})=I_{\text{Hottarget,R}} \end{cases},\quad 若\ M(P_{\text{L}})=0$$
$$\begin{cases} I_{\text{SHT,L}}(P_{\text{L}})=s \\ I_{\text{SHT,R}}(P_{\text{R}})=s \end{cases},\quad 若\ M(P_{\text{L}})\neq 0$$

其中,$s=\max(I_{\text{Hottarget,L}}(P_{\text{L}}),I_{\text{Hottarget,R}}(P_{\text{R}}))$.

2.1.2 立体直线

由于人眼对图像中的直线特征如垂直线、水平线、斜线比较敏感,当图像中直线发生扭曲变形时,观众感知的图像失真程度非常强烈. 基于此,我们在网格变形过程中加入保持立体直线特征的约束.

我们通过构造左、右视图的梯度图,然后根据 Hough 变换从梯度图中提取直线段,得到立体图像中的初始立体直线. 对于直线特征明显且数目较多的图像,增加直线的检测数量. 对于直线特征不明显的图像则减少直线特征检测数量.

由于算法的误差,我们需要对初始立体直线进行修正,从而保证立体直线的视差一致性. 假设 $P_1^{\text{L}}P_2^{\text{L}}$,$P_1^{\text{R}}P_2^{\text{R}}$ 为立体直线的左、右视图线段,我们依据视差图矩阵 $\boldsymbol{M}$,按以下方式修正右视图线段,

$$x_1^{\text{R}}=x_1^{\text{L}}+M(P_1^{\text{L}}),\ y_1^{\text{R}}=y_1^{\text{L}},$$
$$x_2^{\text{R}}=x_2^{\text{L}}+M(P_2^{\text{L}}),\ y_2^{\text{R}}=y_2^{\text{L}},$$

其中，(y,x)表示像素 P 的行、列坐标.

为了方便起见，本文将修正后所得的立体直线归为 3 类：HL、VL 和 OL 分别为水平立体直线特征集、垂直立体直线特征集和斜立体直线特征集.

2.2　基于四边域网格的受限能量模型

立体图像缩放方法的关键在于保持并突显图像显著区域和保持立体图像视差的一致性. 为此，我们构建基于四边域网格的受限能量模型驱动网格变形，实现立体图像的缩放.

如图 2(c)，假设 $M=(V,E,F)$是覆盖立体图像左(右)视图的四边域网格，$V=[v_1,v_2,\cdots,v_n]$，$v_i\in R^2$为网格顶点集，E 为网格边集，F 为四边域网格集. 假设 $M'=(V',E',F')$是网格变形后得到的目标网格，$V'=[v_1',v_2',\cdots,v_n']$为网格顶点集. 我们引入 5 个能量约束条件，构建基于四边域网格的受限能量模型，

$$D(M')=\min_{M}(\alpha\cdot D_d+\lambda\cdot D_l+\delta\cdot D_b+\gamma\cdot D_v+\eta\cdot D_c) \tag{2}$$

其中：D_d 为网格扭曲能量约束；D_l 为边长失真能量约束；D_b 为边界失真能量约束；D_v 为垂直视差能量约束；D_c 为视差一致性失真能量；$\alpha,\lambda,\delta,\gamma,\eta$ 为权重系数. 模型中各能量约束的定义如下.

2.2.1　网格扭曲能量

网格的非均匀变形将导致扭曲失真，每个网格的变形程度取决于网格的重要程度，重要程度越高的网格非均匀变形的程度越小. 我们采用 D_d 衡量网格扭曲失真程度，$D_d=D_d^{\mathrm{L}}+D_d^{\mathrm{R}}$，$D_d^{\mathrm{L}}$和 D_d^{R}分别为左图和右图的网格扭曲能量，

$$D_d^{\mathrm{L}}=\sum_{(i,j)\in E^{\mathrm{L}}(f)}\omega_{HT}^{\mathrm{L}}\left\|s_{(f)}^{\mathrm{L}}(v_i-v_j)-(v_i'-v_j')\right\|^2,$$

$$D_d^{\mathrm{R}}=\sum_{(i,j)\in E^{\mathrm{R}}(f)}\omega_{HT}^{\mathrm{R}}\left\|s_{(f)}^{\mathrm{R}}(v_i-v_j)-(v_i'-v_j')\right\|^2 \tag{3}$$

其中，ω_{HT}^{L}和 ω_{HT}^{R}分别为左右图每个四边形网格扭曲失真的权重因子，其等于四边形内像素点重要程度的均值. 我们采用 Hot-Target 图表示图像中各像素点的重要程度. 设定式(3)右边的二次项是网格理想缩放与实际缩放的差异，表示每个四边形网格面的失真. $S_{(f)}$ 为四边形网格的缩放比例，将网格缩放分解为 X 方向和 Y 方向的缩放，$S_{(f)}=\{S_x,S_y\}$，其中，

$$S_x=\frac{|v_{x+1,y}'-v_{x,y}'|+|v_{x+1,y+1}'-v_{x,y+1}'|}{|v_{x+1,y}-v_{x,y}|+|v_{x+1,y+1}-v_{x,y+1}|},$$

$$S_y=\frac{|v_{x,y+1}'-v_{x,y}'|+|v_{x+1,y+1}'-v_{x+1,y}'|}{|v_{x,y+1}-v_{x,y}|+|v_{x+1,y+1}-v_{x+1,y}|} \tag{4}$$

保持图像重要特征的非均匀网格缩放会导致非重要区域网格扭曲，使部分网格线段弯曲，过度的图像扭曲将给观众带来不舒适的视觉感受. 我们引入边长失真约束 D_l 和边界失真约束 D_b 控制网格变形，避免图像过度扭曲.

2.2.2　边长失真能量与边界失真能量

用边长失真约束能量函数来计算网格边长的理想变化与实际变化之间的差异，$D_l=D_l^{\mathrm{L}}+D_l^{\mathrm{R}}$，$D_d^{\mathrm{L}}$和 D_d^{R}分别为左图和右图的网格边长失真能量，

$$D_l^{\mathrm{L}}=\sum_{(i,j)\in E^{\mathrm{L}}(f)}\left\|l_e^{\mathrm{L}}\times(v_i-v_j)-(v_i'-v_j')\right\|,$$

$$D_l^{\mathrm{R}}=\sum_{(i,j)\in E^{\mathrm{R}}(f)}\left\|l_e^{\mathrm{R}}\times(v_i-v_j)-(v_i'-v_j')\right\| \tag{5}$$

边长失真约束可以分解为 X 方向和 Y 方向进行处理，原网格和目标网格的边长比 l_e的构造与缩放比 $S_{(f)}$类似，$l_e=\{l_x,l_y\}$，其中 l_x 和 l_y 分别表示变化前后的网格横向边长比和变化前后的网格纵向边长比，

$$l_x=\frac{|v_{x+1,y}'-v_{x,y}'|}{|v_{x+1,y}-v_{x,y}|},$$

$$l_y=\frac{|v_{x,y+1}'-v_{x,y}'|}{|v_{x,y+1}-v_{x,y}|} \tag{6}$$

网格缩放过程中出现越界而丢失边界信息的现象属于边界失真. 为了保持图像的完整性，图像缩放应保证目标图像边界顶点为原图像边界顶点，假定输入图像的分辨率为 $m\times n$，目标图像的分辨率为 $m'\times n'$，令 $v_{i,x}$，$v_{i,y}$，$v_{i,x}'$，$v_{i,y}'$分别表示原网格顶点 X 坐标、Y 坐标以及目标网格顶点 X 坐标、Y 坐标，并满足以下边界描述：

$$v_{i,x}'=\begin{cases}0, & 若\ v_{i,x}=0\\ n', & 若\ v_{i,x}=n\end{cases},$$

$$v_{i,y}'=\begin{cases}0, & 若\ v_{i,y}=0\\ m', & 若\ v_{i,y}=m\end{cases} \tag{7}$$

边界失真约束能量 D_b 为图像缩放的硬约束，$D_b=D_b^{\mathrm{L}}+D_b^{\mathrm{R}}$，为了保持图像网格不越界，严格满足边界失真约束能量 $D_b=0$. 令 lb,rb,tb,bb 分别表示图像的左边界、右边界、上边界和下边界的网格顶点数. 网格的边界失真约束能量如式(8)所示：

$$D_b^{\mathrm{L}}=\sum_{i=1,i\in E^{\mathrm{L}}(f)}^{lb}(v_{i,x}'-v_{i,x})^2+\sum_{i=1,i\in E^{\mathrm{L}}(f)}^{rb}(v_{i,x}'-v_{i,x})^2+\sum_{i=1,i\in E^{\mathrm{L}}(f)}^{tb}(v_{i,y}'-v_{i,y})^2+\sum_{i=1,i\in E^{\mathrm{L}}(f)}^{bb}(v_{i,y}'-v_{i,y})^2,$$

$$D_b^{\mathrm{R}}=\sum_{i=1,i\in E^{\mathrm{R}}(f)}^{lb}(v_{i,x}'-v_{i,x})^2+\sum_{i=1,i\in E^{\mathrm{R}}(f)}^{rb}(v_{i,x}'-v_{i,x})^2+\sum_{i=1,i\in E^{\mathrm{R}}(f)}^{tb}(v_{i,y}'-v_{i,y})^2+\sum_{i=1,i\in E^{\mathrm{R}}(f)}^{bb}(v_{i,y}'-v_{i,y})^2 \tag{8}$$

2.2.3 垂直视差能量与视差一致性失真能量

不同于内容感知的 2D 图像缩放，我们在立体图像缩放过程中引入了图像立体约束，其中，垂直视差能量 D_v 保持缩放后的左右图对应点水平对齐，避免垂直视差. 视差一致性失真能量 D_c 保持缩放后左右图对应点的视差信息.

我们采用对应点策略，作为垂直视差能量与视差一致性失真能量的构建基础. Chang 等人[10]采用 SIFT 匹配点[24]进行能量约束保持缩放前后左右图对应一致；但是，SIFT 匹配点的选择取决于图像结构，其可能集中在某些特征区域，在 SIFT 匹配点集中区域视差一致效果较好，能量约束限制性较强，SIFT 匹配点稀疏或缺失的位置较难保持缩放前后的视差一致. 基于此，我们采用内容感知的随机匹配方法，构建垂直视差能量与视差一致性失真能量.

我们利用 SGM 半全局匹配方法获取立体图像视差图. 利用均匀随机算法从左图中选取 10% 的像素点 $M^{\mathrm{L}}=[m_1^{\mathrm{L}},m_2^{\mathrm{L}},\cdots,m_i^{\mathrm{L}}]$，并根据视差图计算出左图随机点在右图中的对应位置，得到右图对应像素点 $M^{\mathrm{R}}=[m_1^{\mathrm{R}},m_2^{\mathrm{R}},\cdots,m_i^{\mathrm{R}}]$. 保存随机匹配点 $M=\{M^{\mathrm{L}},M^{\mathrm{R}}\}$.

用垂直视差能量约束来保持左右图对应物体在图像变形前后垂直对齐，以避免左右图不对称而产生垂直视差. 我们考虑左右图全局随机点的对应关系，对这些随机匹配点加以对齐失真约束. 假设选取 n 对全局随机匹配点，能量函数定义如式(9)所示：

$$D_v=\frac{1}{n}\sum_{i=1}^{n}(c_i^{\mathrm{L}'}(y)-c_i^{\mathrm{R}'}(y)) \tag{9}$$

符号 $\boldsymbol{v}(y)$ 和 $\boldsymbol{v}(x)$ 表示向量 $\boldsymbol{v}$ 的 Y 分量和 X 分量. 由于随机匹配点不一定落在网格顶点处，因此，我们需要将随机匹配点映射到网格点坐标.

假设点 m_i 位于图 3 中所示的四边形网格内部，该网格顶点分别为 v_1,v_2,v_3 和 v_4，如图 3 虚线所示，两条垂直线穿过点 m_i 将网格划分为 4 个四边形区域，令每个区域对应的面积分别为 S_1,S_2,S_3 和 S_4，则

$$c=\sum_{i=1}^{4}(\beta_i+0.1\times\omega_{HT})\times v_i,$$

$$\beta_i=S_{5-i}\Big/\sum_{i=1}^{4}S_i.$$

由于距离匹配点越近的网格权重越大，为了使重要物体更能保持原来的视差，我们在权值中加入度量重要程度的权重因子. 同理，

$$c'=\sum_{i=1}^{4}(\beta_i+0.1\times\omega_{HT})\times v_i',$$

v_i' 为缩放后的网格顶点.

D_v 反映了缩放后随机匹配点 Y 分量的差异，左右对应点 Y 坐标的差异导致物体不对齐，产生垂直视差，给观众带来不舒适的观赏体验. 最小化对齐失真约束能量能够有效地避免垂直视差.

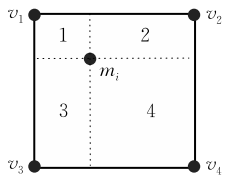


图 3 匹配点和网格顶点的关系

网格变形会导致图像中的物体发生不同程度的扭曲，部分物体的视差将发生变化，导致观众观看立体图像时感知的深度信息与原图像不一致；如果同一物体不同部位的扭曲程度不同，就会导致同一物体存在不同视差. 在本文中，我们考虑缩放前后视差不变的情况，此时视差一致性约束能量函数定义为

$$D_c=\frac{1}{n}\sum_{i=1}^{n}(d_i-d_i')^2 \tag{10}$$

其中，$d_i=c_i^{\mathrm{R}}(x)-c_i^{\mathrm{L}}(x)$，$d_i'=c_i^{\mathrm{R}'}(x)-c_i^{\mathrm{L}'}(x)$ 分别表示缩放前后随机匹配点的视差. 最小化视差一致约束能量旨在保持图像缩放前后的视差信息，避免视差不一致的情况.

2.3 立体直线特征保持的迭代求解

我们采用线性最小二乘法求解受限能量模型式(2). 首先根据能量约束构造 $Ax=b$ 形式的线性方程组，并将均匀缩放的顶点位置赋为迭代初始值，按预先设定的迭代次数对网格顶点的 $\boldsymbol{x}$ 向量和 $\boldsymbol{y}$ 向量进行迭代求解. 必须说明的是：为了简化求解过程，我们并没有将立体直线特征约束直接加入到受限能量模型式(2)中，而是在迭代求解中引入并根据立体直线 HL、VL 和 OL 约束关系调整网格迭代结果，保持立体图像的直线特征. 最后，将得到的左右图最终目标网格通过网格贴图映射获取最终目标立体图像.

在迭代过程中，HL、VL 和 OL 三种立体直线特征约束按以下方式构造.

HL 中每一条水平线经过的网格顶点的 Y 坐标均相同，VL 中每一条垂直线经过的网格顶点 X 坐标均相同. 而 OL 中每一条斜线经过的网格顶点的 X、Y 坐标满足直线函数关系：

$$\begin{cases}HL=\{(\bigcup_{HE}P_{i,j})\mid i\in[1,hn],\ j\in[1,hm]\}\\ VL=\{(\bigcup_{VE}P_{i,j})\mid i\in[1,vn],\ j\in[1,vm]\}\\ OL=\{(\bigcup_{OE}P_{i,j})\mid i\in[1,on],\ j\in[1,om]\}\end{cases} \tag{11}$$

$P_{i,j}$ 表示第 i 条直线特征边缘的顶点集. hn、vn、

on 分别表示水平特征边缘、垂直特征边缘以及斜线特征边缘的条数，hm、vm、om 则表示对应的顶点数. 缩放后直线的特征边缘定义如下：

$$\begin{cases} HL'=\{(\bigcup_{HE} P'_{i,j})\mid i\in[1,hn],\ j\in[1,hm]\} \\ VL'=\{(\bigcup_{VE} P'_{i,j})\mid i\in[1,vn],\ j\in[1,vm]\} \\ OL'=\{(\bigcup_{OE} P'_{i,j})\mid i\in[1,on],\ j\in[1,om]\} \end{cases} \tag{12}$$

$P'_{i,j}$ 表示缩放后的第 i 条直线特征边缘的顶点集. 为了保证缩放前后直线特征边缘不发生扭曲，保持平直状态，HL、VL 和 OL 约束定义如下：

HL 约束关系：缩放前后每条水平特征边缘顶点的纵坐标相等. 即

$$P_{i,x}(y)=P_{i,x+1}(y)=\cdots=P_{i,x+n}(y),$$
$$P'_{i,x}(y)=P'_{i,x+1}(y)=\cdots=P'_{i,x+n}(y) \tag{13}$$

VL 约束关系：缩放前后每条垂直特征边缘顶点的横坐标相等. 即

$$P_{i,y}(x)=P_{i,y+1}(x)=\cdots=P_{i,y+n}(x),$$
$$P'_{i,y}(x)=P'_{i,y+1}(x)=\cdots=P'_{i,y+n}(x) \tag{14}$$

OL 约束关系要求缩放前斜线特征边缘顶点横坐标 $P_{i,j}(x)$ 和纵坐标 $P_{i,j}(y)$ 满足直线函数关系，即 $P_{i,j}(y)=k\times P_{i,j}(x)+b$，$k\in(0,\infty)$，缩放后，斜线特征边缘斜率可能发生变化，需要根据斜线两端顶点的新位置重新计算变形后的斜率 k' 和偏移量 b'，缩放后斜线特征边缘顶点的横纵坐标满足表达式 $P'_{i,j}(y)=k'\times P'_{i,j}(x)+b'$，$k'\in(0,\infty)$. 将原图像的斜线特征边缘上的顶点带入新直线表达式，求出新的顶点位置，更新网格顶点.

3　实验结果分析

我们在 CPU 为 Intel(R) Core(TM) i5-3470 3.20 GHz 和 4 GB 内存的联想启天 M6410/19PC 平台上利用 Matlab 软件实现本文方法. 我们对图像进行缩放操作，网格顶点采样间隔为 10 个像素时，不同的分辨率变化与所需的运行时间如表 1 所示.

表 1　分辨率变化与运行时间关系表

分辨率变化(宽×高)/像素	运行时间/s
300×200 → 200×200	5.74
300×200 → 300×300	5.88
300×200 → 400×400	6.02
420×370 → 200×200	18.57
420×370 → 300×300	18.90
420×370 → 400×400	18.59
500×400 → 300×300	29.14
500×400 → 400×400	29.37
500×400 → 500×500	29.62

以表 1 中图像分辨率从 420×370 缩放至 300×300 的实验为例，改变网格顶点采样间隔并计算运行时间，两者的关系如表 2 所示.

表 2　网格顶点采样间隔与运行时间关系表

网格顶点采样间隔/像素	运行时间/s
10	18.90
12	13.19
15	11.43
20	10.73

上述数据表明该方法的运行时间与立体图像的分辨率、网格密度有关，当处理分辨率越大的图像时，覆盖在图像上的网格数量会增加，因此进行网格缩放所花费的时间会更长，而运行时间受图像分辨率变化的影响相对较小，对相同数量的网格进行处理所需的时间差异不明显. 对同一图像，网格顶点采样间隔越大，网格密度越小，所需的运行时间越短. 由于该方法包括立体 Hot-target 图生成、立体直线检测、网格变形处理以及贴图映射处理几个部分，因此随着网格密度减小，网格变形处理的时间缩短，占整个方法运行时间的比重减少，运行时间的减少呈现变缓的趋势.

在实验过程中，我们根据图像的结构特征对受限能量模型式(2)的权重系数进行适当调整，同时对边缘较多的复杂图像增加特征边缘的检测，以达到理想的缩放效果. 在图 4 的实验中，我们设置受限能量模型(2)的权重系数 $\alpha,\lambda,\delta,\gamma,\eta$ 分别为 50，50，8000，500，2500 可以取得图 4(b)所示的较好的缩放效果. 当权重系数发生变化时，如图 4(c)所示，增加网格扭曲能量 D_d 的权重 α 至 500，更能突显图像的重要区域. 图 4(d)减小边长能量权重 λ 至 5，雪人脸部出现变形的现象表明边长能量权重 λ 的减少导致图像中的物体发生一定的扭曲变形. 图 4(e)减小边界约束 δ 至 80，图中屋顶和天空发生过度扭曲的现象表明边界约束 δ 的减小导致网格出现越界的情况. 垂直视差能量与视差一致能量用于保持立体信息，取值相对较大. 下文的实验分析进一步表明我们的方法能够有效地突出图像的重要区域，保持图像中物体的连续性、直线特征边缘以及视差一致性.

我们采用连续型图像缩放方法，与文献[16]中离散型的立体 seam carving 算法相比，避免了因图像连续性被破坏而导致的图像失真. 立体 seam carving 算法在遮挡检测、视差一致性保持方面取得了较好的效果，但由于移除的 seam 线可能穿过物体的特征边缘，因而导致图像中的物体不连续. 图 5

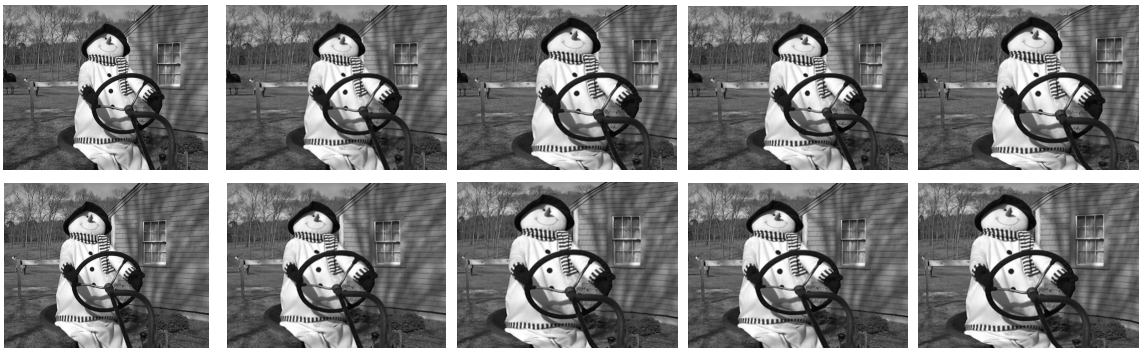


图 4 将图像分辨率从 370×300 缩放至 400×300，预先设定受限能量模型(2)参数 $\alpha,\lambda,\delta,\gamma,\eta$ 分别为 50，50，8000，500，2500，本文方法调整参数对实验结果的影响((a)输入立体图像对，上下图分别为左视图与右视图；(b)预设参数下的缩放结果；(c)其他参数不变 α 调整为 500 的缩放结果；(d)其他参数不变 λ 调整为 5 的缩放结果；(e)其他参数不变 δ 调整为 80 的缩放结果)

用黑色框标注两种方法的实验结果的对应区域，如图 5(b)中黑色框内的花盆形状发生扭曲，部分茎干不连续. 图 5(e)中人物左右两边的跑道线出现不连续的现象. 图 5(c)、(f)采用我们的方法通过网格变形进行图像缩放，能够有效地保持图像的连续性.

在突出图像重要区域方面，我们的方法采用立体 Hot-Target 图得到图像的重要区域，将重要程度作为网格扭曲的一个控制指标，通过能量函数最小化进行网格缩放. 相比直接缩放的方法能在一定程度上突出重要物体，避免重要物体过度变形. 从 Hot-Target 图可以看出雪人和飞机处于显著性最大的区域，我们的方法能够有效地保持显著区域的特征，如图 6(d)的雪人和图 6(h)的飞机相比于图 6(c)和图 6(g)更加突出.

由于文献[19]的方法并没有考虑到图像直线特征边缘，对于特征边缘较多的物体可能导致扭曲. 我们采用融合梯度图和显著图的立体 Hot-Target 图定位图像重要区域，考虑了图像的边缘信息，同时在网格变形的迭代计算过程中加入了立体直线特征边缘约束. 图 7(b)为方法检测到的直线特征，基本包含几个英文字母的直线部分. 图 7(c)中上图的 D 的左下角、U 的右部、F 的左部以及下图的 U 右部均

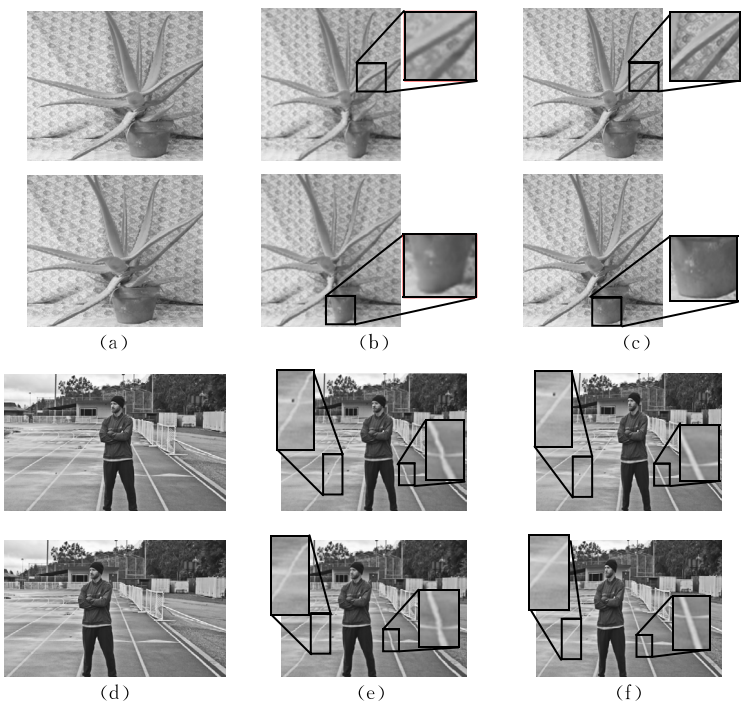


图 5 本文方法保持图像连续性((a)、(d)输入立体图像对；(b)、(e)立体 seam carving 算法[16]的缩放结果；(c)、(f)本文方法的结果)

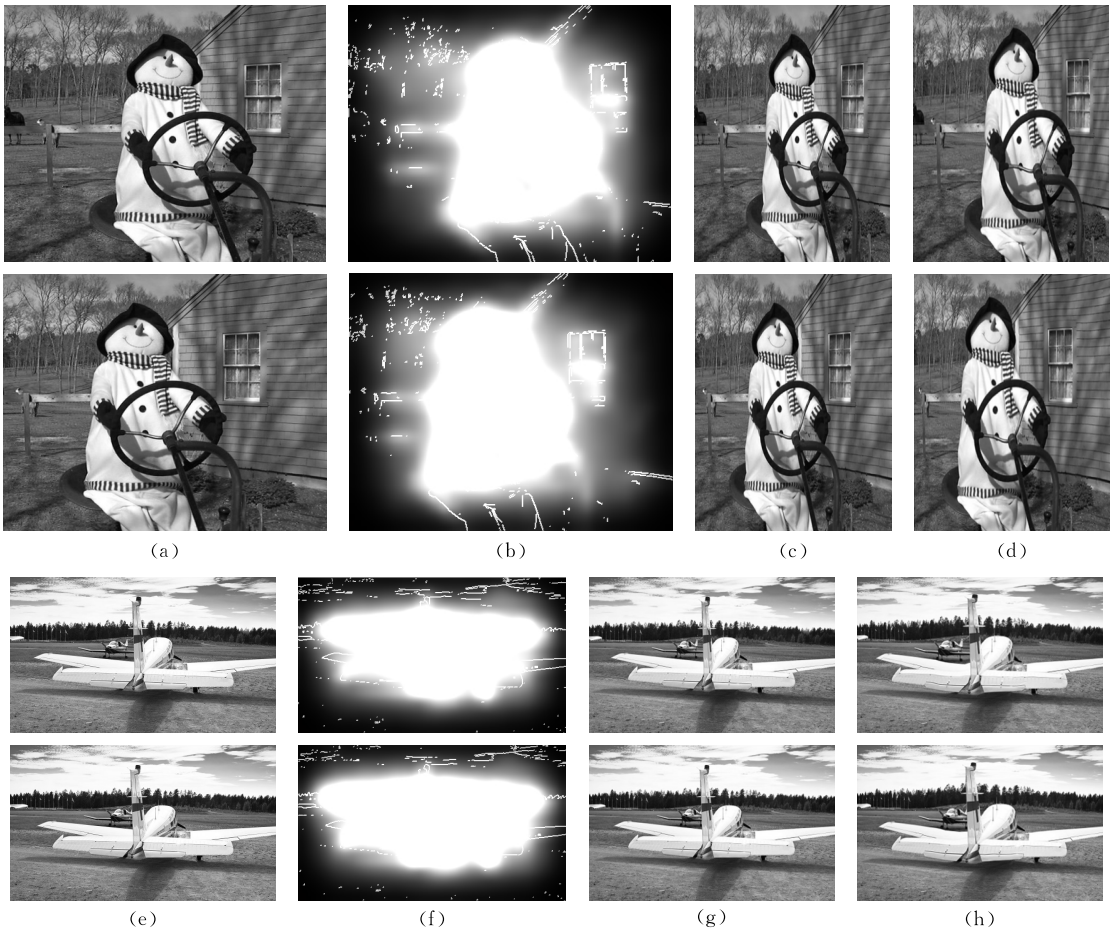


图 6　本文方法突出图像重要区域((a)、(e)输入立体图像对；(b)、(f)对应的 Hot-Target 图；(c)、(g)线性缩放结果；(d)、(h)本文方法的结果)

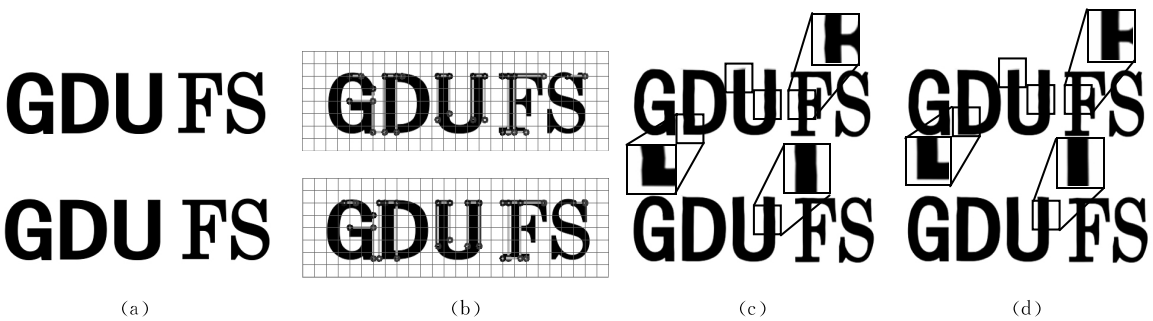


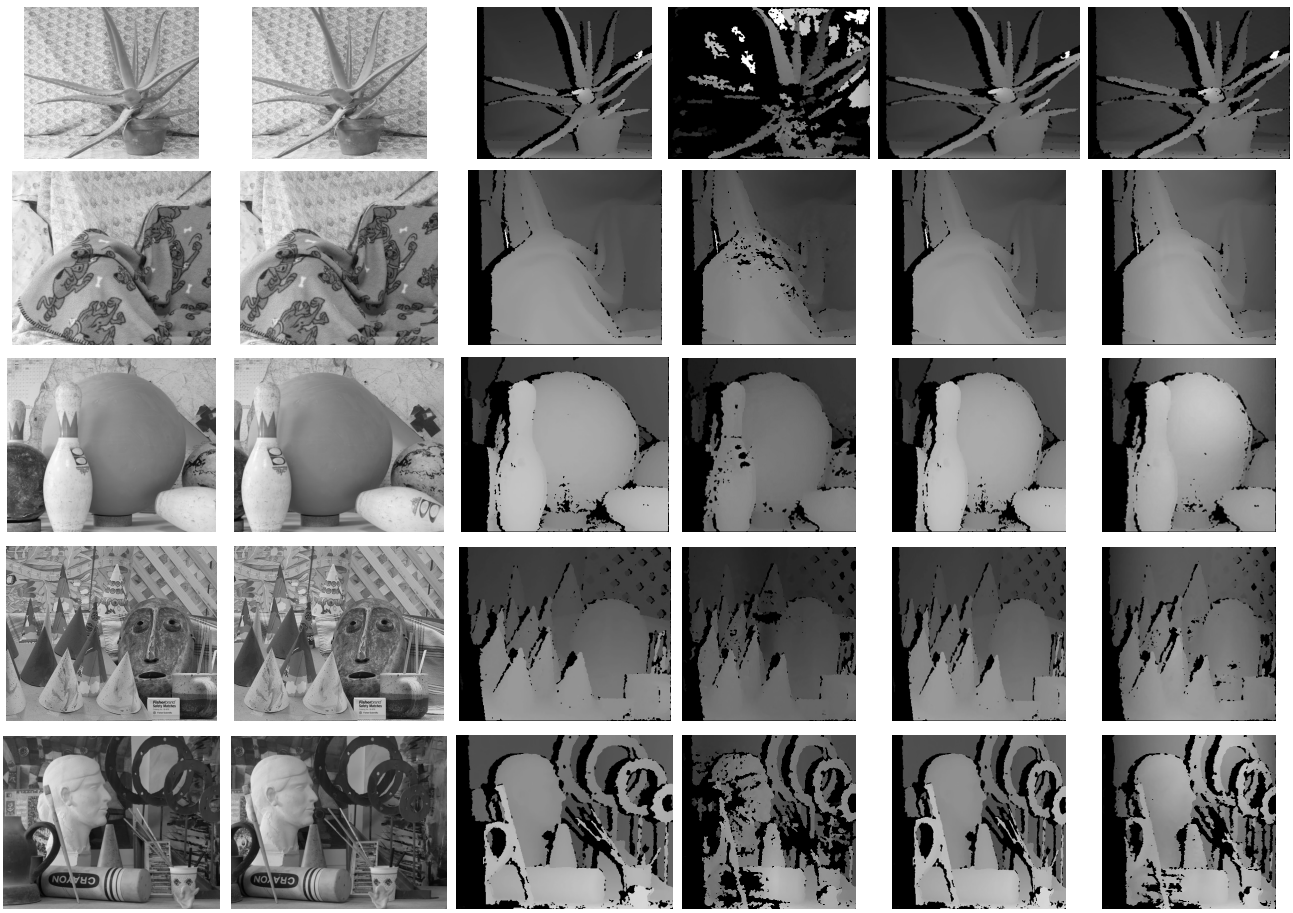
图 7　本文方法保持图像直线特征边缘((a)输入立体图像对；(b)本文方法检测到的直线特征；(c)文献[19]方法的缩放结果；(d)本文方法的缩放结果

出现了扭曲不连续的情况，加入直线特征约束之后，保留检测到的直线边缘，从一定程度上改善了缩放效果，如图 7(d)所示. 图 8(d)的方法没有进行直线特征保持，上下图的旗杆都呈现小幅度的扭曲，使用我们的方法能尽量减少直线物体扭曲的情况.

图 8 显示我们的方法在突出重要区域，避免图像不连续，保持图像直线特征方面取得了较好的效果. 在视差一致性保持方面，我们的方法相比直接使用 2D 图像缩放的方法有明显改善. 采用全局随机匹配点的对应关系将排列约束和视差一致约束作为能量函数，能够有效保留图像的立体信息. 如图 9 所示，我们采用 Middlebury 数据库提供的 5 组立体图像：Aloe，Cloth，Bowling，cones，Art 进行分析. 没有加入此约束之前，图 9(c)显示的左右对应视差图出现大量镂空现象，即缩放后左右图没能找到对应点. 且视差图的灰度值与原始视差图存在一定的差异，视差值发生变化. 图 9(f)为我们的方法求得的效果，采用 10%的随机匹配点，能够覆盖每个网格. 可以得到稠密的对应关系，与原始视差图的灰度比较一致，视差图更平滑有效.

(a) (b) (c) (d) (e)

图 8 不同缩放方法效果比较((a)输入立体图像对;(b)本文方法检测到的直线特征;(c)立体 seam carving 的缩放结果;(d)文献[19]方法的缩放结果;(e)是本文方法的缩放结果)



(a) (b) (c) (d) (e) (f)

图 9 本文方法保持图像立体信息((a)输入立体图像左视图;(b)输入立体图像右视图;(c)输入立体图像的视差图;(d)文献[9]方法缩放结果的视差图;(e)对(c)视差图做与左图相同缩放的视差图;(f)本文方法缩放结果的视差图)

进一步量化分析图 9 实验方法保持视差一致性的效果,我们借鉴 Basha 等人[16]提出的深度扭曲评价方法,构造网格变形方法的视差扭曲评价指标 B. 该指标衡量缩放后的立体图像视差与原始视差之间的差异. 由于缩放后立体图像视差图与原始视差图之间维数并不相同,因此在进行像素级别差异计算时难以做到一一对应,需要对原始视差图进行缩放处理. 本文采用的立体图像视差图是基于左图构造的,反映左视图与右视图之间的水平视差关系. 因此,在进行原始视差图缩放时,采用与左图相同的缩

放方法. 在 Basha 等人[16]提出的评价方法中，将左图 seam carving 的操作直接运用到原始视差图中进行缩放. 与之类似，我们的方法对原始视差图进行基于网格变形的缩放，将左视图网格的缩放操作运用到原始视差图，如图 9(e)所示. 设缩放后的原始视差图为 D_o，缩放后的立体图像计算得到的视差图为 D_a. 视差扭曲评价指标 B 为

$$B=\frac{1}{N}\sum_{(i,j)}T(i,j),$$

$$T(i,j)=\begin{cases}1, & |D_o(i,j)-D_a(i,j)|>1\\ 0, & |D_o(i,j)-D_a(i,j)|\leq 1\end{cases} \quad (15)$$

式中：N 表示缩放后图像的像素总数量；$D(i,j)$ 表示图像 D 在第 i 行第 j 列的值. 该指标计算缩放后的立体图像的视差与原始视差差异大于 1 个像素点的个数占像素总数量的比重，用于评价缩放后立体图像的视差扭曲程度.

对图 9 中的实验效果计算视差扭曲评价指标 B，得到表 3 的结果，其中左列是采用文献[9]方法缩放结果得到的视差扭曲指标 B_s，右列是采用本文方法缩放结果得到的视差扭曲指标 B_o. 右列的视差扭曲指标数值远小于左列的数值，说明我们的方法在避免垂直视差、保持视差一致性方面都取得了一定成效. 由于原始视差图的计算过程中仍存在误差，能量函数各个约束的制约关系使得本文方法也存在一些视差扭曲现象，从图 9(e)、(f)中对比看出本文方法结果视差图与原始视差图差异很小，不影响观看效果.

表 3　缩放方法视差扭曲程度对比

实验图像	视差扭曲指标 B_s/%	视差扭曲指标 B_o/%
Aloe	74.42	4.37
Cloth	9.86	1.68
Bowling	66.57	18.77
Cones	43.91	7.98
Art	43.60	28.82

4　结　论

我们提出直线特征保持的立体图像缩放方法. 该方法将图像缩放视为对覆盖图像的四边形网格进行迭代处理，在立体直线特征边缘约束的条件下，最小化包含网格扭曲约束、边长约束、边界约束、垂直视差约束和视差一致性约束的能量函数. 使用该方法进行图像缩放能够有效地突出图像显著区域和重要目标，保持图像的直线特征，尽量避免图像扭曲和不连续，同时减少了立体图像垂直视差的产生，保持了左右图的视差一致性. 该方法的另一个优点是采用求解线性方程组的方法对网格变形进行求解，提高了运算速度.

我们的方法对于重要物体的检测依赖于显著图的准确求取. 当显著图检测发生错误时，将使得图像中的重要物体发生扭曲. 该方法根据经验值设置网格变形的受限能量模型权重系数，对不同内容结构的图像尚需进行系数调整. 由于网格变形属于联动操作，缩放后的物体很难保持原物体的尺寸特征，依然存在一定的变宽或变窄的现象，如图 7 非自然图像中重要物体占图像空间位置较大，缩放程度增加时容易导致重要物体扭曲变形. 另外，特征边缘的提取和识别在一定程度上也影响了图像缩放效果，特征边缘的错误识别将使得图像出现不必要的扭曲. 以上是我们进一步开展研究需要解决的问题.

致　谢　感谢团队成员对本文成果的辛勤付出，感谢美国西肯塔基大学计算机系李祺博士的指点，感谢中山大学软件学院曾坤博士提供的帮助，感谢评审专家提供的宝贵意见和建议！

参　考　文　献

[1] Keys R. Cubic convolution interpolation for digital image. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1981, 29(6): 1153-1160

[2] Baker S, Kanade T. Limits on super-resolution and how to break them. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 24(9): 2372-2379

[3] Kopf J, Cohen M F, Lischinski D, Uyttendaele M. Joint bilateral upsampling. ACM Transactions on Graphics, 2007, 26(3): 1-5

[4] Avidan S, Shamir A. Seam carving for content-aware image resizing. ACM Transactions on Graphics, 2007, 26(3): 10:1-10:9

[5] Cho T S, Butman M, Avidan S, Freeman W T. The patch transform and its application to image editing//Proceedings of the Computer Vision and Pattern Recognition. Anchorage, USA, 2008: 1-8

[6] Setlur V, Takagi S, Raskar R. Automatic image retargeting //Proceedings of the 4th International Conference on Mobile and Ubiquitous Multimedia. Christchurch, New Zealand, 2005: 59-68

[7] Wolf L, Guttmann M, Cohen-Or D. Non-homogeneous content-driven video-retargeting//Proceedings of the 11th IEEE International Conference on Computer Vision (ICCV 2007). Rio de Janeiro, Brazil, 2007: 1-6

[8] Wang Y S, Tai C L, Sorkine O, Lee T Y. Optimized scale-and-stretch for image resizing. ACM Transactions on Graphics, 2008, 27(5): 1-8

[9] Liang Yun, Su Zhou, Luo Xiao-Nan, Wang Dong. Image shrinkage based on Hot-Target map and featured edge preservation. Journal of Software, 2011, 22(4): 789-800(in Chinese)
(梁云，苏卓，罗笑南，王栋. 基于 Hot-Target 图和特征边缘保持的图像收缩方法. 软件学报，2011，22(4)：789-800)

[10] Chauvier L, Murray K, Parnall S, et al. Does size matter? The impact of screen size on stereoscopic 3DTV//Proceedings of the IBC Conference. Amsterdam, Netherlands, 2010: 1-14

[11] Wei Q. Converting 2D to 3D: A survey//Proceedings of the International Conference. Delft, Netherlands, 2005, 7: 14

[12] Hirschmuller H. Stereo processing by semiglobal matching and mutual information. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(2): 328-341

[13] Rubinstein M, Gutierrez D, Sorkine O, Shamir A. A comparative study of image retargeting. ACM Transactions on Graphics, 2010, 29(6): 160

[14] Utsugi K, Shibahara T, Koike T, et al. Seam carving for stereo images//Proceedings of the 3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video. Tampere, Finland, 2010: 1-4

[15] Basha T, Moses Y, Avidan S. Geometrically consistent stereo seam carving//Proceedings of the 2011 IEEE International Conference on Computer Vision (ICCV). Barcelona, Spain, 2011: 1816-1823

[16] Basha T, Moses Y, Avidan S. Stereo seam carving a geometrically consistent approach. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(10): 2513-2525

[17] Yue B, Hou C, Zhou Y. Improved seam carving for stereo image resizing. EURASIP Journal on Wireless Communications and Networking, 2013, 2013(1): 1-6

[18] Krol J D, van der Grind W A. Rehabilitation of a classical notion of Panum's fusional area. Perception, 1982, 11(5): 615-619

[19] Chang C H, Liang C K, Chuang Y Y. Content-aware display adaptation and interactive editing for stereoscopic images. IEEE Transactions on Multimedia, 2011, 13(4): 589-601

[20] Cho S H, Kang H B. A new visual comfort-based stereoscopic image retargeting method//Bebis G, Boyle R, Parvin B, et al, eds. Advances in Visual Computing. Berlin Heidelberg: Springer, 2013: 290-300

[21] Lee K Y, Chung C D, Chuang Y Y. Scene warping: Layer-based stereoscopic image resizing//Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Providence, USA, 2012: 49-56

[22] Lin S, Lin C, Chang S, et al. Object-Coherence Warping for Stereoscopic Image Retargeting. IEEE Transactions on Circuits and Systems for Video Technology, 2014, 24(5): 759-768

[23] Harel J, Koch C, Perona P. Graph-based visual saliency. Advances in Neural Information Processing Systems, 2007, 19: 545-552

[24] Lowe D G. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 2004, 60(2): 91-110

GONG Yong-Yi, born in 1970, Ph. D. His research interests include image registration, image segmentation and image/video upsampling.

LI Ke-Hong, born in 1990, M. S. candidate. His research interests include 2D-3D image/video conversion and stereo image retargeting.

JIANG Ling-Min, born in 1956, Ph. D., professor. His research interest is image processing.

WU He-Feng, born in 1985, Ph.D. His research interests include computer vision and image processing.

LUO Xiao-Nan, born in 1963, Ph. D., professor, Ph. D. supervisor. His research interests include digital home technology, mobile computing, computer graphics & CAD and image processing.

Background

Content-aware image scaling methods have received great attentions in the past 10 years. Most of them are proposed only for monocular 2D images. With the development of 3D display techniques, it is very desirable to extend these methods on stereoscopic images.

However, directly applying existing monocular image scaling methods to stereoscopic images is problematic as it often introduces vertical disparities and damages the original disparity distribution. The existing stereo methods fail in preserving objects with straight lines. In this paper, we propose a stereoscopic image scaling method based on straight line preservation. Our method belongs to the continuous approaches. We use content-aware random matching instead of sparse matching points that well preserve disparity consistency. Since human eyes are more sensitive to straight objects, we introduce straight line constraint in the process of mesh deformation that avoids twist of the straight lines.

The work was supported by the National Natural Science Foundation Grant of China (Grant Nos. 61370160, 61320106008, 61402120), the Guangdong Natural Science Fund Project 2014A030310348, the Guangdong Colleges and Universities Scientific and Technological Innovation Project 2012KJCX0048, the Guangzhou Scientific and Technological Plan Project 2014J4100032.