

引文格式:

张开翼, 程鲁豫, 汪云海. 拼贴画自动生成技术综述 [J]. 集成技术, 2021, 10(6): 20-33.

Zhang KY, Cheng LY, Wang YH. A survey on collage auto generation algorithm [J]. Journal of Integration Technology, 2021, 10(6): 20-33.

拼贴画自动生成技术综述

张开翼 程鲁豫 汪云海*

(山东大学 青岛 266237)

摘 要 拼贴画作为常见的设计类型与视觉艺术表现形式, 被广泛应用于海报设计、Logo 制作等领域。但传统人工制作过程需要设计师大量地重复尝试, 且很难依靠人工达到最优布局。该文从自顶向下与自底向上两种思路归纳总结计算机用于拼贴画自动生成的研究进展, 对其中不同的算法进行了分析与对比, 并对该技术的未来发展趋势进行了总结和展望。

关键词 拼贴画; 马赛克艺术; 2D 图形; 可视化

中图分类号 TP 391.7 文献标志码 A doi: 10.12146/j.issn.2095-3135.20210311001

A Survey on Collage Auto Generation Algorithm

ZHANG Kaiyi CHENG Luyu WANG Yunhai*

(Shandong University, Qingdao 266237, China)

*Corresponding Author: cloudseawang@gmail.com

Abstract As a common design type and visual art form, collage is widely used in poster design, logo production, and other applications. However, traditional manual process requires a lot of repetitive attempts and it is difficult for designers to achieve the optimal layout. When using computers to solve problems, there are many different ways. This paper summarizes different collage ideas which are roughly concluded as two ideas of top-down and bottom-up, and their algorithm routes are analyzed and compared. Finally, future development trend of the problem is presented.

Keywords collage; mosaics; 2D shape; visualization

Funding This work is supported by National Natural Science Foundation of China (61772315, 61861136012)

收稿日期: 2021-03-11 修回日期: 2021-06-11

基金项目: 国家自然科学基金 (61772315, 61861136012)

作者简介: 张开翼, 学士, 研究方向为数据可视化; 程鲁豫, 学士, 研究方向为数据可视化; 汪云海(通讯作者), 博士, 教授, 研究方向为数据可视化与可视分析, E-mail: cloudseawang@gmail.com。

1 引言

拼贴画(Collage)是一种经典的视觉艺术形式,通过将图片、文字、剪报、几何图形等元素拼接到画布上,从而产生具有美感的艺术品。目前,较为流行的表现形式有3种:(1)图片马赛克,通过拼接图片或彩色多边形来产生不同的结果;(2)词云(Word Cloud)^[1],作为一种强大的文本可视化工具,通过拼接不同大小(代表权重)的单词来产生美感;(3)传统拼贴画,通过规则和紧密地分布主题较为统一的图形来产生美感。拼贴画强大的艺术表现力具有美观与信息传递的双重优势,能使受众感受到美学的同时获取足够的信息。但由于其组合、角度、位置的多样性,输入特征的复杂性,以及设计目标的差异性,使人工设计过程复杂、设计步骤繁琐。设计师需要反复地试错多种可能的拼接方式,才能制作出一个相对优秀的作品,这个过程需要耗费大量的时间和精力。即便如此,人工制作的拼贴画也很难达到最优的效果。

拼贴画的生成属于可视化与图形学的交叉领域,其领域学者就拼贴画问题提出不同思路。图形学领域学者将该问题视作一个打包问题:先利用形状描述符^[2]对输入的图形进行相似度分析,然后将全局或局部轮廓相似的图形进行拼接;抑或是使用轮廓细分与匹配^[3]的方法,将目标轮廓按一定规则细分成格子后,将与格子相似的图形填充到格子中。可视化领域学者侧重如何传递信息,即如何将输入数据中包含的信息与拼贴画元素进行映射:将输入的图形抽象为特征向量^[4],从而与最终可视化的位置进行映射,更好地展现图形特征数据;抑或是在可视化文本中^[1],将带有词频的文字视作图形输入,通过算法将填充文字与目标轮廓进行映射,从而生成信息量与美观度俱佳的结果。

为更有逻辑地梳理不同学者在拼贴画自动

生成领域提出的思路,本文将其粗略地分为自顶向下的生成技术和自底向上的生成技术,并按两大类四小类的分类方式,从中具体介绍具有代表性或开创性的方法。旨在帮助相关研究人员快速了解本领域演进过程与代表性工作。并将本文中提到的常见算法整理为表1方便查看与对比。

2 形式化定义

如图1所示为常见的3种拼贴画形式,首先对拼贴画自动生成技术进行形式化定义——拼贴画生成的输入集主要分为两部分:

第一个输入集为目标外轮廓约束,一般分为两种:一种为强外轮廓约束,如复杂图形、照片等,填充时需要充分考虑外轮廓的约束,使填充图形的边缘尽可能贴合强外轮廓约束的边缘,这属于较复杂问题;另一种约束为弱外轮廓约束,如矩形、圆形、简单字母等,填充时外轮廓的约束较小,只需考虑边界问题即可。

第二个输入集为填充图形,主要分成两类。第一类是核心图形,即填充中主要考虑的图形。核心图形分为规则图形(如矩形、圆形等)和非规则图形(如照片、文字等)。规则图形主要通过其颜色、大小、密度等参数表示输入的外轮廓,且图形间的拼接相对容易,可简化拼接算法。非规则图形的难度主要在于非规则图形边缘的匹配以及对接,该部分的搜索空间会急剧膨胀,因此如何有效减少搜索空间成为加速和优化匹配速度的核心问题。第二类填充图形为辅助图形,此概念来自设计师作品中为了减少拼接时间、提高填充率的技巧,如图1(c)中小方片与小五边形等。在人工拼接时,依靠穷举达到最佳结果是非常困难的,所以设计师会引入较小且不引人注目的辅助图形,来提高最终结果的填充率和非最优拼接结果的视觉效果。

表 1 拼贴画生成算法结构表

Table 1 Collage auto generation algorithm structure table

生成技术	生成算法	论文名称	核心输入集	算法基础
	基于轮廓细分与匹配	Image Mosaics	目标轮廓图及填充图形	小波变换的图形匹配算法
		Paint By Numbers: Abstract Image Representations	目标轮廓图	Voronoi 算法
		Simulating Decorative Mosaics	目标轮廓图	基于重心的泰森多边形图
		Jigsaw Image Mosaics	目标轮廓图及填充图形	最小化能量函数
自顶向下	基于特征提取的方案	Appearance-guided Synthesis of Element Arrangements by Example	拼贴画	结果统计学习模型
		Interactive By-example Design of Artistic Packing Layouts	填充图形	特征降维
		Digital Micrography	目标轮廓及文本	流线生成
		Spatial Text Visualization Using Automatic Typographic Maps	地图信息及文本	体征提取与刚体拟合
自底向上	基于螺旋线	Pyramid of Arclength Descriptor for Generating Collage of Shapes	目标轮廓图及填充图形	局部形状描述符
		Participatory Visualization with Wordle	文本	阿基米德螺旋线
		Morphable Word Clouds for Time-Varying Text Data Visualization	词云	受约束的刚体动力学
		Geo Word Clouds	有文本标签的目标轮廓	聚类算法
		ShapeWordle: Tailoring Wordles using Shape-aware Archimedean Spirals	目标轮廓及文本	形状感知的阿基米德螺旋线



(a) 图片马赛克

(b) 词云^[5]

(c) 传统拼贴画

图 1 常见的拼贴画设计

Fig. 1 Common collage design

3 自顶向下的生成技术

自顶向下的生成技术思路是从生成的目标轮廓分析或是从填充图形的整体分析,先确定填充图形在轮廓的粗略位置,再小范围地改动图形的细节、角度、位置以得到更优的结果。

3.1 基于轮廓细分与匹配算法

基于轮廓细分与匹配算法常由 3 个步骤构

成:首先对目标轮廓进行细分,将轮廓分成多个规则或不规则的单元格;然后给每个单元格匹配最合适的填充图形;最后对每个单元格的填充图形进行调整以增强效果。

3.1.1 相关算法介绍

Finkelstein^[6]在 1998 年提出 Image Mosaics 技术,其针对的问题为:如何将用户输入的一组图片拼接到另一张目标图片中,产生近距离看是

一组图片, 远距离看是目标图片的效果。在算法上, 首先将目标图片分割成规律的矩形, 利用基于小波变换的图形匹配算法^[7]对每个矩形匹配最相似的输入图片; 然后对矩形内图片的亮度、灰度值进行调整, 从而使结果更接近目标图片。但由于其在分割目标图片时, 将图片仅简单地分割为规则的矩形, 而未考虑图片本身的特点, 故被分割图片不能很好地呈现目标图片的轮廓。

Haeberli^[8]引入 Voronoi 算法, 将不规则图形(如油画的笔触或多边形)随机地摆放并填充至目标区域, 以达到转换输入图片风格的目的。其思路是将油画的笔触或者多边形抽象成填充图形, 并在填充至目标图片过程中, 先使用 Voronoi 算法细分目标图片, 再将笔触填充至网格内。在这一过程中, 为了使笔触更接近真实油画效果, Haeberli 引入根据输入图片计算梯度方向的方式来引导笔触的填充方向, 使其更接近手绘效果。但该方法忽略了原始图片中物体的边缘(如人物图片的人物轮廓、物体图片中不同物体的边缘), 没有使填充图形贴合或者表达原图片中的轮廓, 导致生成结果中的轮廓不清晰。而清晰的轮廓可以帮助用户迅速感知图片本体, 所以该方法生成的结果辨识度不佳。

在 Haeberli 方法基础上, Hausner^[9]引入基于重心的泰森多边形图(Centroidal Voronoi Diagram, CVD)^[10], 其针对的目标问题为: 使用多边形填充目标图片, 着重表现输入图片中的边缘与轮廓, 来增强结果的辨识度。传统的 CVD 图使用欧式距离度量, 通过多次迭代使全局欧式距离最小, 导致计算出来的区块趋近于六边形。在 Hausner 的研究中, 通过将度量修改为曼哈顿距离使最终区块接近正方形, 并用于填充。此外, 也可以通过修改度量的方向来影响 CVD 图中格子的方向。为了高效地生成带方向度量的 CVD 图, 该研究借鉴了 Hoff^[11]提出的生成技术。并在 Hoff 生成技术的基础上, 将欧式

距离替换成曼哈顿距离。为了提取原始图片中的方向信息, Hausner 引入了方向场, 从原始图片的轮廓中产生方向场来指导带方向度量的 CVD 图的生成。结果显示, 作为填充图形的小正方形可以很好地表达原始图片中的边缘, 用户可以非常迅速地辨认仅由带色彩的正方形组成的拼贴画。该方法有足够的创新性, 产生了具有强大表现力的拼贴画。但缺陷在于其构成要素仅为正方形, 很难表现更为复杂的图片, 这很大程度上限制了输入图形的类型, 普适性较差。

3.1.2 代表性算法介绍

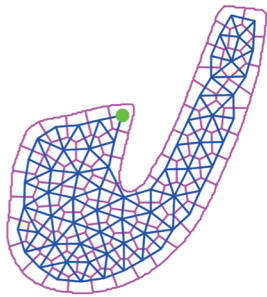
Kim^[3]提出一种基于最小化能量函数的生成技术, 与 Hausner 方法相比, 更具有普适性: 将填充图形从简单的多边形, 扩充到任意的输入图形, 并填充到任意的目标轮廓中。该算法属于基于轮廓细分与匹配算法中较有代表性的算法, 综合了该线路上前人的研究结果。如借鉴 Hausner^[9]的思路, 使用 CVD 图来辅助拼接等, 在此基础上将问题抽象成能量函数最优化的过程, 并结合多种优化方式来加速生成。同时, Kim 将输入集的范围扩充到任意图形, 进一步扩大了算法的适用范围。具体来说, 能量化定义的优势就是可以通过增减或调整各能量项的权重使最终的结果呈现不同的风格, 其能量 E 定义为:

$$E = w_c \cdot E_c + w_G \cdot E_G + w_o \cdot E_o + w_D \cdot E_D \quad (1)$$

其中, 能量 E 由各项的加权和构成; 颜色项 E_c 为图形颜色与当前位置目标容器颜色不符的程度, w_c 为权重; 间隙项 E_G 为最终结果中空白区域的大小, w_G 为权重; 重叠项 E_o 为最终结果中图形之间的重叠程度, w_o 为权重; 变形项 E_D 为图形过度变形的程度, w_D 为权重。

拼接时, 是在不对图形进行变形的前提下减少间隙和重叠, 并使图形的颜色与目标容器位置的颜色保持一致, 从而得到图形的初步匹配结果。为使每次拼接操作后, 更新的容器更便于下一次拼接。在每次拼接之前, 借鉴 Hausner^[9]的

思路,按图形的平均大小对容器构建 CVD 图,随机选择邻居数量最少的位置,从而减少拼接后容器产生突出或不连通的问题,便于后续的拼接。如图 2 所示,构建 CVD 邻接图后,图中绿色点为被选中的位置,其仅有 2 个邻居,拼贴在此位置对网格的影响最小。



—: CVD 邻接图 □: CVD 格子 ●: 选中的位置

图 2 使用 CVD 邻接图选择拼接位置

Fig. 2 Use the CVD adjacency graph to select location

当某一步无法找到合适的图形去填充目标容器时,算法需要退回上一能量最小处。为减少这种回溯, Kim^[3]借鉴 Russell^[12]的思想,使用超前技术,惩罚拼接之后无法放置图形的小面积区域。

前文所述,在不对图形变形的情况下进行拼接,尽可能形成少重叠和少间隙填充的目标容器的布局。但其中仍会出现很多空隙与重叠,导致生成结果并不美观。为此,该研究借鉴了 Kass^[13]的思路,使各个图形可以在保持原始形状和最小化能量 E 之间平衡,通过微小的变形来取得更好的结果。在 CVD 图的变形过程中,初始拼接得到的 CVD 结果,会造成较多的重叠与空隙。经过优化过程,重叠与空隙有一定程度的减少。最终,在各个 CVD 格子变化程度较小的前提下,减少了格子之间的重叠与空隙,达到控制变形程度与提高结果填充率、降低重叠率之间的平衡。

因拼贴画的生成过程中需要考虑多种因素,如美观度、空隙等,所以基于最小化能量函数的

生成技术在拼贴画的生成中非常常见。其大致流程为将拼贴画生成中多种因素抽象为能量函数中的参数项,并赋予不同的权重以形成能量函数,之后通过算法最小化该能量函数。Xu^[14]提出的 Calligraphic packing 算法亦是如此,其将拼贴画生成的范围扩展到书法中,将任意文字作为拼贴画生成算法的输入,通过倾斜、扭曲等操作使文字填充到任意给定的图形中,形成有美感的书法艺术作品。并在能量函数中引入风格因素,算法可以根据不同的风格调整能量函数中各项的参数,生成不同风格的输出结果。

3.2 基于特征提取算法

基于特征提取的方法一般分为 3 个步骤:首先对输入图形进行特征提取;然后处理特征信息,并进行初步布局;最后对布局进行优化调整,以减少重叠和间隙。

3.2.1 相关算法介绍

鉴于使用参数控制图形的分布十分困难, Hurtut^[15]提出一种基于结果统计的方式。其针对的问题为:如何从用户输入的拼贴画学习拼贴画结构,并以该结构为基础,生成更大范围的拼贴画结果。算法输入一张包含多种图形的拼贴画,系统使用基于图形外观的分类算法将原始拼贴画中的组成图形根据外观分类,并抽象出布局。系统通过构建统计学习模型,从输入的布局中学习同类别内部以及不同类别之间的布局关系,从而在输入布局的基础上扩充出更大范围的结果。其缺点是不能像传统拼贴画的构造过程,输入足够的图形与轮廓生成指定的结果。并且,由于采用了基于统计的生成方式,因此无法控制生成结果的风格,也很难根据不同的情景生成不同的结果,使用户对结果的控制力大大下降。

在特征提取上, Maharik 等^[16]将特征提取的对象转移到轮廓,其针对的问题为:如何将输入的文本以流线的形式填充到目标轮廓上。该系统首先在目标轮廓中生成矢量场;进而生成符合代

表目标轮廓的流线, 并使其接近正常阅读方式; 再在生成的流线上填充文本, 从而使文本既能够较好地填充目标轮廓, 又具有较强的可读性。其结果在文本对目标轮廓的表现力以及文本的可读性之间取得了较好的平衡。

基于特称提取的算法也被应用在其他类型的可视化中, 如 Afzal 等^[17]提出的 **Typographic Maps**, 其针对的问题为: 如何将文本依照地理信息填充到地图中。该系统以地图中提取的信息和用户的约束作为特征填充文本, 即以输入地理信息作为布局文本的主要依据, 同时接受用户定义的视觉属性和约束条件。从地图中提取的特征为文本布局的位置提供了基础, 而文本又给地图提供了更多的信息, 最终提高了可视化信息传递的效率。

3.2.2 代表性算法介绍

Hurtut^[15]和 Öztireli^[18]都认为利用参数控制图形的分布是非常困难的, 但他们都选择基于统计的方式来布局。而 Reinert^[4]提出一种不基于统计数据的方法, 即从用户的输入中推导布局规则。基于特征提取算法中, 具有代表性的 Reinert 算法摒弃了以往使用参数或基于统计的布局方式, 将拼贴问题形式化为从每个图形的高维特征空间映射到二维的布局问题。同时, 通过在系统中引入交互, 用户可以使用图钉固定一部分图形的位置, 系统从用户确定的子集中计算出用户期望的布局, 并影响图形的最终布局。

具体来说, 系统首先对高维特征空间进行投影, 将图形以一定的规则映射到二维向量上。对每个图形定义特征向量, $\mathbf{f} \in \mathbf{R}^m$ 表示 $m \in \mathbf{N}$ 个不同的特征。特征是由程序自动获取的视觉特征, 如: 大小、形状、亮度、纹理等。通过公式(2), 将高维特征空间降维为二维坐标矩阵 \mathbf{X} :

$$\mathbf{X} = \phi(\mathbf{P}\mathbf{f} + \mathbf{t}) \quad (2)$$

其中, \mathbf{P} 为特征投影矩阵; \mathbf{t} 为转换参数; \mathbf{f} 为特征向量; ϕ 为布局函数; \mathbf{X} 为转换后的坐标矩

阵。 ϕ 可以将结果投影到不同的坐标系统中, \mathbf{t} 可以使参数向量 \mathbf{x} 沿参数的轴作偏移。如在笛卡尔坐标系下, 改变 \mathbf{t} 可以使图形沿坐标轴偏移。

系统引入了布局函数 ϕ 可以使布局结果更具多样化。布局函数是一个 $\phi(\mathbf{x}) \in \mathbf{R}^2 \rightarrow \mathbf{R}^2$ 的映射, 可以通过输入不同的布局函数将同样的二维向量映射到不同布局中。如对输入的二维向量的第一个数值映射为 x 轴坐标、第二个数值映射为 y 轴坐标, 则得到在笛卡尔坐标系下的布局。同样, 如果将第一个数值映射到角度而第二个数值映射到半径, 则得到径向布局。

平衡临近图形的距离可以产生良好的最终布局, 但从高维特征空间映射到二维坐标时, 每个图形被看作一个特征向量, 并未考虑图形之间的重叠与间隙。即映射时未考虑图形在空间上的相互关系, 故需在映射之后, 对图形的布局进行优化。

在布局上, 系统也使用了基于 CVD 图的方式, 通过逐步迭代目标函数来达到优化目的。在目标函数中, 使用图形边界与 CVD 区域边界的差值的平方和来衡量布局 \mathbf{X} 偏离平衡的程度:

$$c(\mathbf{X}) = \sum_{i=1}^n \int_{\Omega_i^V} \text{distance}_i(\omega, \mathbf{x}_i)^2 d\omega$$

其中, Ω_i^V 为第 i 个图形的 CVD 区域边界; $\text{distance}_i(\omega, \mathbf{x}_i)$ 为 CVD 边界上的点 ω 和第 i 个图形边界上的点 \mathbf{x}_i 的最短欧几里得距离。为了解决最小化 $c(\mathbf{X})$ 这一问题, 系统使用了一些优化与近似方式来提高计算速度, 并通过 GPU 并行运算的方式实现交互级的流畅度。

交互时, 系统首先生成一个初始结果, 用户可以使用图钉工具在界面中固定一部分图形的位置; 随后, 系统根据用户指定图形的特征向量, 推测用户期望的布局, 并将其应用到全局。

4 自底向上生成技术

自底向上生成技术的思路为从每个图形出

发,依次向已有的图形上拼接最合适的图形,最终逐步填充整个目标轮廓。

4.1 基于形状描述符匹配算法

在自底向上的拼接过程中,由于每当系统将一个图形拼接到拼贴画中,必须不断地对图形进行角度、比例的调整,并评估两者的匹配程度,因此需要高效的局部形状匹配。这个过程涉及巨大的搜索空间——大量的图形,且对于每个图形又有无限个角度、位置、比例的组合。现有的描述符,如基于曲率比例空间(Curvature Scale Space)^[19]、形状上下文(Shape Context)^[20]、三角区域表示(Triangle-area Representation)^[21],可以很好地测量形状之间的整体相似度,但不能直接应用到拼贴画的生成中。拼贴画的生成核心在于两个图形的局部形状的拼接,而并不考虑两个图形的整体相似度,且全局描述符带来的巨大搜索空间会导致拼接效率极速下降。

Kaplan^[22]提出一种利用给定的封闭图形在平面上进行紧密平铺的算法,其使用模拟退火算法优化参数化抽象后的拼接问题,生成“埃舍尔”风格的拼贴画。但其对输入、输出的限制均比较大,只能输入单一的封闭图形,很难限制与自定义输出,这些缺点较大地限制拼贴画的多样性。并且其算法只对图形进行旋转,而不考虑比例缩放等操作,从而影响输出结果的丰富性,很难满足用户多元的需求。

Kwan^[2]提出的算法可以对任意填充元素与任意指定的轮廓进行拼贴画生成,并支持对输入元素进行旋转与比例缩放来优化填充结果,极大地扩展了算法的应用场景。该算法在拼贴画生成领域具有开创性,引入了以局部图形描述符为基础,在填充图形中搜索最适合填充的图形。相比于传统的形状描述符在匹配时需要考虑位置、角度和比例的组合,Kwan提出的PAD(Pyramid of Arclength Descriptor)描述符在匹配时比例和角度不变,这意味着在匹配最相似的图形时,无需计

算角度和比例。通过该描述符确定最优匹配图形以及最优匹配点之后,可使用较小的搜索代价确定该图形的角度和缩放。PAD描述符也可以通过改变参考点的数量,动态调整PAD向量的描述能力和计算复杂度。

PAD描述符的核心是一个基于角度和比例不变域的塔式描述符。首先算法确定一个缩放比例不变域,此处使用的是绝对曲率的积分。给定一段曲线,对 t 、 s 点的绝对曲率积分的定义为:

$$\tau(s:t)=\int_s^t|\kappa(x)|dx \quad (3)$$

其中, $\kappa(x)$ 为 x 点的绝对曲率。研究证明该积分是对缩放比例不变的。

但仅使用绝对曲率的积分并不能准确地表达图形,或者说,不能依靠绝对曲率的积分来进行形状匹配。所以研究中引入了弧长这一形状信息来辅助形状匹配。弧长并不是对缩放比例不变的,但可以通过简单的变形来解决。

首先是对输入图形进行离散化,离散化后的每两个点之间具有相同的绝对曲率的积分变化率($\Delta\tau$)。然后开始构建描述符,对于图形轮廓上的任意点 p ,需要同时考虑 p 点的左侧和右侧。为了更精确描述图形的局部,一个 n 维的向量被引用进来:

$$[l_{n-1} \ l_{n-2} \ \cdots \ l_1 \ l_0 \ r_0 \ r_1 \ \cdots \ r_{n-2} \ r_{n-1}]^T \quad (4)$$

其中, l_i 和 r_i 为 p 点的左侧和右侧在同样的绝对曲率的积分变化率($\Delta\tau$)的间隔弧长; n 可以根据需要进行调整,来获得更高的准确度。则 p 点的描述向量 $m(p)$ 为:

$$\hat{l}_i = \frac{l_i}{2^i l_{n-1}} \quad (5)$$

$$\hat{r}_i = \frac{r_i}{2^i r_{n-1}} \quad (6)$$

$$m(p) = s [\hat{l}_{n-2} \ \cdots \ \hat{l}_1 \ \hat{l}_0 \ \hat{r}_0 \ \hat{r}_1 \ \cdots \ \hat{r}_{n-2}] \quad (7)$$

其中, $s \in \{1, -1\}$ 表示凸(+1)和凹(-1)。

可以通过计算两点的PAD向量之间的距离,来计算两图形轮廓上两点的相似度。其距离

公式为:

$$D(p, q) = \max \{ |m_i(p) - m_i(q)| \} \quad (8)$$

其中, $m_i(p)$ 为 $m(p)$ 中第 i 个元素, 小的 $D(p, q)$ 意味两者更相似, 即使两个图形的缩放比例并不相同。通过描述符确定两个图形的最相似点之后, 算法会在该位置遍历所有可能的角度和缩放比例。因为此时已经确定了相似点, 故遍历的搜索空间并不大, 可以通过并不复杂的计算找到最优的匹配结果。

对图形局部进行匹配之后, 即可进行拼贴画的整体生成。给定一个比例任意的图形库, 其生成是一个迭代的过程。如图 3 所示, 每次迭代都是从一个目标形状开始, 最初的目标形状为种子图形, 由用户指定, 或者随机产生。每次迭代都是从未拼接的图形库中寻找局部轮廓距离最近的 $K\%$ 个作为候选, 并对每个候选的拼接结果通过公式 (9) 进行打分, 选取最终打分高的结果作为最终的拼接图形。

$$Z = \frac{\omega_1 L}{\sqrt{\omega_2 \sum O_i^2 + \omega_3 \sum G_i^2}} \quad (9)$$

其中, L 为两个图形共享的边长; O_i 为第 i 个重

叠部分的面积; G_i 为两图形第 i 个间隙部分的面积; ω_1 、 ω_2 、 ω_3 为权重。由此通过不断的迭代, 将局部相似度高并且适合目标形状的图形进行拼接, 产生最终的结果。

通过弧长与曲率的结合, 对图形局部轮廓进行特征描述, 并计算图形局部的相似度, 可取得较优的拼接效果。但这样可能会陷入仅考虑轮廓的局部而忽略了整体的问题, 从而导致图形局部匹配度较高, 但全局相对较差。该方法通过引入对候选位置打分的方式来选择最优的拼接, 一定程度上缓解了这个问题。但从本质上来说, PAD 的生成过程是贪婪的, 并不能得到全局最优的结果。

4.2 基于螺旋线算法

基于螺旋线算法将从文本中提取的核心词作为主要的输入图形, 文本的大小代表权重, 最终拼接成完整的拼贴画, 其主要用于词云可视化中。Viegas^[1]最初提出使用阿基米德螺旋线来生成词云, 将单词沿着以画布为中心的螺旋线按照权重从大到小的顺序依次拼接, 生成最终结果。词云作为经典的文本可视化工具, 可以迅速生成美观且信息量丰富的可视化结果。但在生成词云的过程

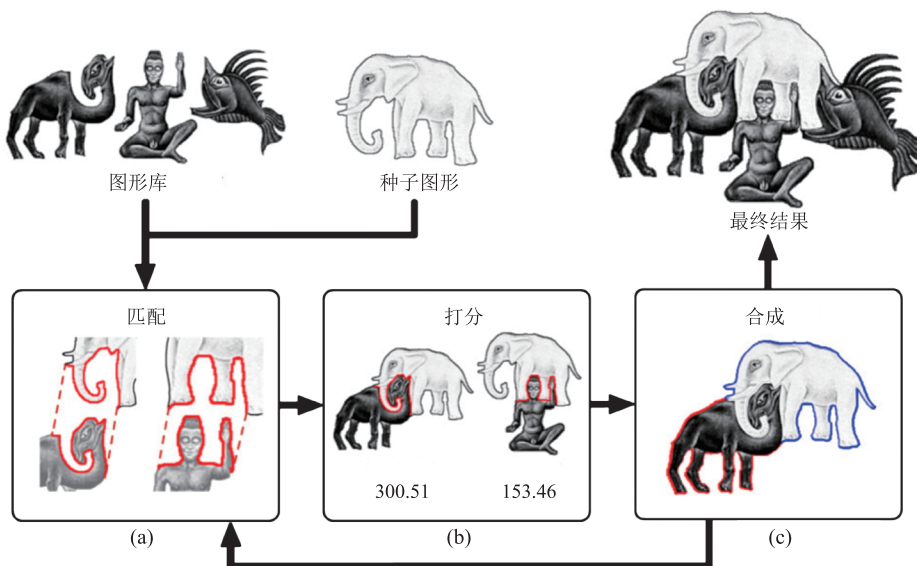


图 3 PAD 拼接示意图

Fig. 3 Schematic diagram of PAD splicing

中,并不会考虑目标轮廓,所以生成的结果不能很好地填充目标轮廓,不适合生成有目标轮廓的拼贴画。故许多研究者在词云算法的基础上进行了诸多改进,使词云可以有效填充目标轮廓。

4.2.1 相关算法介绍

Buchin^[23]提出 Geo Word Clouds 算法,其输入是一个在每个位置均有文本标签的目标轮廓。首先,采用聚类的方式对输入的文本进行分组;然后,根据填充词的频率,在填充目标轮廓时尽可能使大小合适、不重叠的文本靠近。这一程度上解决了词云不能很好地填充目标轮廓的问题,并且取得了相对不错的效果。但其对用户的输入要求比较高,需要目标轮廓的各个位置带有文本标签,不具有普适性。

Chi 等^[24]提出一种不同的思路,其输入不同于传统的带权重文本:将传统螺旋线生成的词云作为输入,使用受约束的刚体动力学牵引单词重新排布为目标形状。这大大降低了用户的输入要求,比较容易使用。但其生成方式的限制在于,如果输入的词云布局与目标轮廓差距较大,则需要多次迭代才能得到较优的布局。而迭代过程需要耗费大量的时间,且其鲁棒性较差,在部分情况下无法保证良好的填充效果,仍需要手动参与。

4.2.2 代表性算法介绍

Wang^[5]提出一种具有轮廓约束的词云算法,降低了用户使用的门槛——仅需要用户输入普通的文本和目标轮廓,而无需额外的信息即可得到填充率高且美观的结果。该算法复杂度较低、生成速度较快,适合大规模的文本可视化生成。

具体来说,系统首先提出了形状感知的阿基米德螺旋线(如图4所示),其核心在于将目标轮廓的距离场和阿基米德螺旋线结合起来。传统的螺旋线是以画布中心为原点,沿着圆的切线方向逐步扩大半径而绘制出来。在这个过程中,螺旋线的方向一直是圆的切线方向,所以绘制结果与

目标轮廓无关。而引入距离场后,距离场可以很好地反映目标轮廓的信息。在绘制螺旋线时,以距离场的极点为原点,每一点的前进方向并非一直是圆的切线方向,而是根据每一点的距离场信息确定,由距离场引导生成符合目标轮廓的螺旋线。

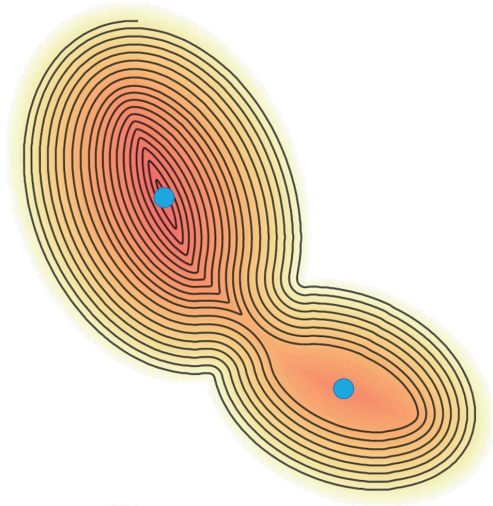


图4 形状感知的阿基米德螺旋线

Fig. 4 Shape-aware Archimedean spirals

为了使算法更具有普适性,系统加入了对复杂图形的处理。为了防止极端情况下,以中心为原点的螺旋线不能很好地对轮廓边缘进行处理,算法引入了图形切割的步骤。在输入图形后系统会首先检测目标轮廓内的所有分区,并为每个分区生成单独的距离场;然后使用一个迭代的梯度下降过程,在每个分区中定位出局部最大值;最后生成单独的螺旋线来排布单词。其效果如图5所示,(a)中图形仅被分为两个区域,填充效果较差,(b)为 ShapeWordle 所采用的图形切割后的分区结果,(c)为在新的分区结果上使用形状感知的螺旋线所产生的效果。结果表明,经过分区后,螺旋线对目标轮廓的贴合能力得到提升,实现了文本填充目标轮廓。但基于螺旋线算法的限制在于,螺旋线的排布适合规则的输入,如矩形文本或常规且相似的图形,而难以对不规则的输入图形进行排布。

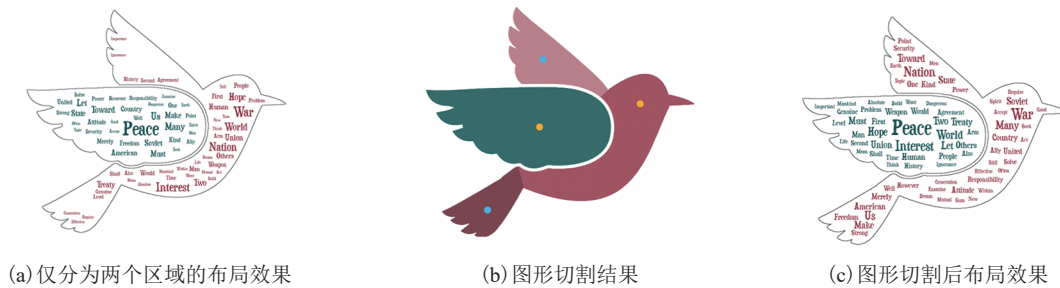


图 5 ShapeWordle 分区方式对比

Fig. 5 ShapeWordle partition mode comparison

5 发展趋势与创新研究

拼贴画的生成算法优势突出——加快设计速度、提高设计质量、降低人工时间,可以迅速应用于工业界。近年来,随着传统电商及社交电商的发展,海报设计、营销设计的需求成几何倍数增长,传统靠人力设计的速度与效率已跟不上时代的需求,自动化海报设计成为大势所趋。如阿里推出的鹿班系统能智能生成大量的海报来应对双十一这类电商节的需求,但其海报生成的结果大多设计简单、内容单调,容易引起大众审美疲劳。现有的拼贴画生成技术非常适合生成风格与众不同的海报设计,鉴于目前学界对两者结合的研究并不多,所以如何针对海报等宣传版面进行拼贴画算法设计将成为未来可能的研究方向。

在生成算法上,自顶向下的生成技术和自底向上的生成技术各自取得了相当优秀的结果,但也存在的一些问题尚待解决。

5.1 自顶向下相关算法

自顶向下的生成技术可以快速地确定物体的初始位置,再对细节进行微调。受限于算力与速度的同时,由于系统确定初始位置时并不会过多考虑图形的轮廓,使得图形轮廓匹配不佳或微调耗时长。基于轮廓细分生成算法思路受限于先将目标轮廓细分,然后将填充图形变形以适应细分后的网格。在网格已经确定的情况下对填充图形进行调整时,若对复杂的填充图形进行较大

的改变易导致其失真,而微调算法也需要较为复杂的运算。且这种变形往往不能取得全局最优的结果。其中一个可行的思路是在进行网格细分之前,先对填充图形的轮廓进行简化,达到计算可以接受的程度。即在网格生成之初,将简化后的图形轮廓与网格细分的形状结合起来,使生成的网格更加接近填充图形的轮廓,以提高匹配度、减少后续微调。

在自顶向下的生成技术中, Hurtut^[15]和 Reinert^[4]将不同的思路引入拼贴画生成中。Hurtut 借助统计模型,将用户生成的拼贴画进行扩展,这种方式非常适合墙纸、瓷砖等需要重复图案生成的应用场景。其问题也相对明显,用户对生成结果的控制力较差,不能有效地对结果进行定制。在该思路的基础上可以尝试拓展可控的参数,增加用户对生成过程的参与度或对结果的可操作性,从而有效提高算法的适用范围。Reinert 的创新之处在于,从用户的输入中推导布局,将拼贴画生成看作每个图形的高维特征向量降维到二维坐标的问题,在降维之后进行局部的微变以优化最终结果。该思路与常规的解决方式有较大区别,可以迅速确定图形的大概位置,但系统中微调的算法只能减少重叠问题,而不能提高图形之间的匹配度,造成结果的填充率较低。该算法的缺点在于,初始降维时图形轮廓完全被忽略。因此,可以在特征向量中进一步引入对图形轮廓的描述符,降维时将图形相似或可拼接的

图形尽可能靠近，并在降维后小范围内引入图形描述符来进行局部图形的位置调整。也可以进一步对局部图形的位置进行小范围的再分配，改变其角度与缩放比例，完善局部图形的拼贴效果。

5.2 自底向上相关算法

在自底向上的生成技术中，Kwan^[2]提出的基于局部图形描述符算法在生成时需要对所有图形进行描述与匹配，适用于少数需要精确匹配的场景。对于大规模的图形匹配，则需要借助多块高性能 GPU 进行并行运算，才能将生成时间压缩到一个可以接受的范围。原文使用了 8 块 NVIDIA Tesla K20m 进行并行运算才将生成时间压缩到 3 h 左右。考虑到原文设计的描述符已足够精简，所以在硬件计算能力未达到质的突破情况下，描述符的生成方式并不适合广泛使用。但 PAD 描述符后续被用于小范围的图形识别^[25]说明，局部的描述符不仅可以进行全局、穷举式的匹配，还可以作为其他拼贴算法进行局部优化的补充。通过对小范围图形的局部轮廓进行匹配，将小范围内相似的局部轮廓进行拼接，或作为图形调整角度与缩放的依据，从而优化局部的拼贴效果与填充率。

而对于基于螺旋线算法，目前 Wang^[24]的形状感知的螺旋线在实现将文本填充到目标轮廓上取得了较为不错的效果，且能够移植到其他类矩形的填充图形上。但由于沿螺旋线的初始布局过程中仅考虑了矩形之间简单的重叠避免，不适用于复杂图形的重叠避免与局部轮廓拼接。同时，形状感知的螺旋线在将图形填充到目标轮廓中的价值很大，可以有效地提高填充算法对目标轮廓的感知能力。且其算法复杂度较低，可与其他算法结合，增强其他算法对目标轮廓的感知能力。如先用包围盒算法将图形抽象成易于计算重叠的简单图形，然后使用形状感知的螺旋线将图形拼接到目标轮廓中，以此为基础进行微调，提高局部轮廓的匹配度，从而充分发挥其算法对轮廓的

感知能力。

5.3 智能算法

随着智能算法的发展和应用，越来越多学者尝试将智能算法引入拼贴画领域。根据目的，将应用于拼贴画自动生成的智能算法简单分为两类：第一种是将智能算法应用在对输入图形的预处理中，如提取图片的核心区域、人脸等，或对图形的特征进行提取，并根据提取的特征进行降维，从而确定图形的初始位置；第二种则直接将智能算法应用于确定图形的最终位置，即应用于拼接，如根据统计学习对已有的拼贴画进行扩展。

5.3.1 应用于预处理

对于第一种智能算法的应用，以 Liu^[26]在 2017 年的相关工作为例，其首先使用 VGG-16 特征提取网络对图像进行特征提取，得到对应的特征向量，然后通过 t-SNE 将特征向量维度降至二维平面，并进行聚类。该类智能算法的应用可看作以 Reinert^[4]为代表的特征提取算法进行初步布局的后续工作，通过智能算法可以显著提高根据特征聚类的准确度。但仅针对图形特征预处理的智能算法并不能完善拼贴画生成中的“硬性标准”，如轮廓匹配度、填充率等。

而使用智能算法进行核心区域与人脸的提取，常应用于图片拼贴中，可以有效去除图片中无效的背景区域，使拼贴算法的重心处于图片中的核心区域。

对于该种类型的智能算法应用，核心区域的语义性提取具有较大的研究价值。此类型拼贴画大多为表现同一主题的图片的拼贴结果，提高语义性提取可以增强主题信息表现的力度。目前，在特征提取方向，利用 CNN 进行特征提取可使提取到的特征向量较好地输入集进行聚类。但对用户来说，CNN 的聚类结果是不可解释的，相关智能算法更接近黑盒模式，很难进行语义化的解释，也很难让用户介入参数调整过程。Reinert^[4]的算法说明，通过让用户交互地调整聚

类后的结果, 从调整后的结果中推测用户意图并进行重新聚类, 较为有效地让用户影响聚类结果, 降低用户介入聚类过程的理解成本。

5.3.2 应用于拼接

本文讨论的拼接算法大多为传统算法, 如何将智能算法应用于拼接仍然是一个待探索的问题。Hurtut^[15]基于统计模型学习输入拼贴画的布局, 从而实现对布局的扩展是一个很好的尝试, 但其应用场景较为单一, 并不能广泛用于拼贴画生成中。

Kwan^[2]的工作表现了局部匹配的强大效果, 但由于其忽略整体轮廓, 出现了轮廓重叠现象。在某些输入集中, 局部轮廓的匹配度与整体轮廓的匹配度是相关联的, 能否通过建立模型来构建图形局部与整体之间的关系, 提高局部轮廓计算中对整体轮廓的感知, 从而在使用局部匹配的拼接算法中, 用较低的局部轮廓计算量达到接近整体轮廓的匹配效果。

5.4 交叉领域

拼贴画算法的输入主要为规则的简单图形或不规则的复杂图形, 输出为在少量重叠下高填充率的拼接结果。其布局算法可对其他涉及布局问题的领域有借鉴作用——直接使用算法产生布局结果, 或应用于对布局结果的调优中。如可以借鉴 Reinert^[4]的方法对降维算法结果在二维平面可视化的分布: 在降维后, 对降维结果在二维平面使用 CVD 相关算法进行迭代优化, 降低投影点之间的重叠, 平衡投影点之间的距离, 提高降维结果的易读性。

同时, 拼贴画拼接过程中降低重叠率、提高填充率的目标与其他可视化布局问题的目标基本一致, 有很好的借鉴与启发作用。拼贴画领域针对目标轮廓的相关算法可以提高其他领域布局问题中对目标轮廓的感知能力, 如树图等可视化形式, 可以从 Wang^[24]针对轮廓的螺旋线算法中得到启发, 将树图或其他可视化方式与用户提供的

目标轮廓相结合, 增强结果的视觉效果。

5.5 创新趋势

在设计拼贴画生成算法之外, 扩大拼贴画算法的适用范围也是发展该领域的重要方向。如 Xu^[14]将拼贴画的输入扩展到书法领域, 通过应用拼贴画算法的思想将文字填充生成艺术性较强的文字画。基于这一思路, 可将拼贴画的输入扩大到其他元素, 从而产生创意性的结果。将算法的输出载体扩展到其他领域亦是如此, 如 Gal^[27]将拼贴这一概念扩展到三维空间, 使三维的物体拼接到三维的轮廓中, 增强了结果的表现能力与表现空间。在拓展维度方面, 也可以在不同载体上做更多的尝试, 如将载体扩展到 3D 打印、虚拟现实技术与增强现实技术等方向。

在当前算法的基础上进行创意性改良也是一个有趣的研究方向。Huang^[28]将传统的用户给定输入集更改为用户指定输入集的主题, 算法从互联网中匹配适合指定主题的拼贴元素。该改进思路既减少了用户寻找输入集的繁琐工序, 也由于每次互联网图片的随机性, 增加了生成结果的多样性。鉴于拼贴画的结果与艺术有一定的相关性, 研究者可在传统算法基础上, 添加更多创意性、艺术性的改进, 扩展和增强算法结果的艺术表现能力。

6 总 结

本文从自顶向下与自底向上两种思路归纳概述了计算机用于拼贴画自动生成的研究进展, 对其中较有创新性的算法进行了分析, 并对该技术的未来发展趋势进行了总结和展望。

随着社交电商与社交互联网的发展, 工业界对自动生成较为复杂的设计需求将持续增长, 而拼贴画自动生成技术作为结合了图形学与可视化两个领域的自动技术, 可在可视信息表达与图形展示效果之间取得良好平衡。目前, 该领域仍然

存在许多重要的问题和挑战, 需要更多的研究人员继续进行深入研究探索。

参 考 文 献

- [1] Viégas FB, Wattenberg M, Feinberg J. Participatory visualization with wordle [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2009, 15(6): 1137-1144.
- [2] Kwan KC, Sinn LT, Han C, et al. Pyramid of arclength descriptor for generating collage of shapes [J]. *ACM Transactions Graphics*, 2016, 35(6): 229.
- [3] Kim J, Pellacini F. Jigsaw image mosaics [J]. *ACM Transactions on Graphics*, 2002, 21(3): 657-664.
- [4] Reinert B, Ritschel T, Seidel HP. Interactive by-example design of artistic packing layouts [J]. *ACM Transactions on Graphics*, 2013, 32(6): 218.
- [5] Wang YH, Chu XW, Zhang KY, et al. ShapeWordle: tailoring wordles using shape-aware archimedean spirals [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2020, 26(1): 991-1000.
- [6] Finkelstein A, Range M. Image mosaics [C] // *International Conference on Raster Imaging and Digital Typography*, 1998: 11-22.
- [7] Finkelstein A, Jacobs CE, Salesin DH. Multiresolution video [C] // *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, 1996: 281-290.
- [8] Haeberli P. Paint by numbers [C] // *Proceedings of the 17th Annual Conference on Computer Graphics and Interactive Techniques*, 1990: 207-214.
- [9] Hausner A. Simulating decorative mosaics [C] // *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, 2001: 573-580.
- [10] Lloyd S. Least square quantization in PCM [J]. *IEEE Transactions on Information Theory*, 1982, 28(2): 129-137.
- [11] Hoff K, Keyser J, Lin M, et al. Fast computation of generalized voronoi diagrams using graphics hardware [C] // *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques*, 1999: 277-286.
- [12] Russell S, Norvig P. *Artificial Intelligence: A Modern Approach* [M]. Prentice Hall: 1994.
- [13] Kass M, Witkin A, Terzopoulos D. Snakes: active contour models [J]. *International Journal of Computer Vision*, 1987, 1(1): 321-331.
- [14] Xu J, Kaplan CS. Calligraphic packing [C] // *Proceedings of Graphics Interface*, 2007: 43-50.
- [15] Hurtut T, Landes PE, Thollot J, et al. Appearance-guided synthesis of element arrangements by example [C] // *Proceedings of the 7th International Symposium on Non-Photorealistic Animation and Rendering*, 2009: 51-60.
- [16] Maharik R, Bessmeltsev M, Sheffer A, et al. Digital micrography [J]. *ACM Transactions on Graphics*, 2011, 30(4): 1-12.
- [17] Afzal S, Maciejewski R, Jang Y, et al. Spatial text visualization using automatic typographic maps [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2012, 18(12): 2556-2564.
- [18] Öztireli AC, Gross M. Analysis and synthesis of point distributions based on pair correlation [J]. *ACM Transactions on Graphics*, 2012, 31(6): 170.
- [19] Mokhtarian F, Mackworth AK. A theory of multiscale, curvature-based shape representation for planar curves [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1992, 14(8): 789-805.
- [20] Belongie S, Malik J, Puzicha J. Shape matching and object recognition using shape contexts [J]. *IEEE*

- Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(4): 509-522.
- [21] Alajlan N, Elrube I, Kamel MS, et al. Shape retrieval using triangle-area representation and dynamic space warping [J]. Pattern Recognition, 2007, 40(7): 1911-1920.
- [22] Kaplan CS, Salesin DH. Escherization [C] // Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, 2000: 499-510.
- [23] Buchin K, Creemers D, Lazzarotto A, et al. Geo word clouds [C] // 2016 IEEE Pacific Visualization Symposium (PacificVis), 2016: 144-151.
- [24] Chi MT, Lin SS, Chen SY, et al. Morphable word clouds for time-varying text data visualization [J]. IEEE Transactions on Visualization and Computer Graphics, 2015, 21(12): 1415-1426.
- [25] Han C, Liu XT, Sinn LT, et al. TransHist: Occlusion-robust shape detection in cluttered images [J]. Computational Visual Media, 2018, 4(1): 161-172.
- [26] Liu LJ, Zhang HJ, Jing GM, et al. Correlation-preserving photo collage [J]. IEEE Transactions on Visualization and Computer Graphics, 2018, 24(6): 1956-1968.
- [27] Gal R, Sorkine O, Popa T, et al. 3D collage: expressive non-realistic modeling [C] // Proceedings of the 5th International Symposium on Non-Photorealistic Animation and Rendering, 2008: 7-14.
- [28] Huang H, Zhang L, Zhag HC. Arcimboldo-like collage using internet images [C] // Proceedings of the 2011 SIGGRAPH Asia Conference, 2011: 1-8.