

分类号: TS959.9

单位代码:

密 级: 公开

学 号: 21521029

浙江大学

硕士学位论文



中文论文题目: 一种基于语义识别的自动配色方案辅助设计工具

英文论文题目: A Semantic Recognition Based Automatic Color palettes Design Tool

申请人姓名: 胡迪

指导教师: 孙守迁

合作导师: 张克俊

专业名称: 设计学

研究方向: 工业设计

所在学院: 计算机科学与技术学院

论文提交日期 2018年4月20日

一种基于语义识别的自动配色方案辅助设计工具



论文作者签名: _____

指导教师签名: _____

论文评阅人 1: _____

评阅人 2: _____

评阅人 3: _____

评阅人 4: _____

评阅人 5: _____

答辩委员会主席: _____

委员 1: _____

委员 2: _____

委员 3: _____

委员 4: _____

委员 5: _____

答辩日期: _____

浙江大学研究生学位论文独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得 浙江大学 或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文作者签名： 签字日期： 年 月 日

学位论文版权使用授权书

本学位论文作者完全了解 浙江大学 有权保留并向国家有关部门或机构送交本论文的复印件和磁盘，允许论文被查阅和借阅。本人授权 浙江大学 可以将学位论文的全部或部分内容编入有关数据库进行检索和传播，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

(保密的学位论文在解密后适用本授权书)

学位论文作者签名： 导师签名：

签字日期： 年 月 日 签字日期： 年 月 日

摘要

计算机辅助艺术设计例如计算机绘画、计算机自动上色等是一个科研人员长期进行探索的问题，因为其与人类多样化的表达方式相关，而人类感受与复杂多变，使其成为一个挑战性强而又持续探索的领域。随着人工智能与机器学习最近几年的发展，计算机艺术辅助系统有了新的突破，例如 GAN 能够自动生产图片，pixel2pixel 能够自动依据已有风格进行色彩填充，PaintsChainer 能在标注颜色的情况下为二次元图片自动上色。

本文提出了一种基于语义智能理解的自动配色和上色系统。该系统主要解决了目前自动上色系统只能通过标注色彩、有限类型进行色彩填充或依赖已有色彩风格进行迁移，从而使得系统更加智能，可以直接对人类的语言进行响应。对于设计师而言，本系统可以将文本的描述，即时转化为配色方案对设计线稿进行填充，大大方便了其使用便捷性。而配色方案来自于艺术作品，有独特的审美性可以帮助到设计师。

本文为解决这一问题使用了 3 个创新点：1. 使用了深度学习模型对语义进行理解；2. 将图片信息与语义层面的隐藏信息进行关联，其图片风格相似性的偏序特性被保持到其语义描述的偏序特性中；3. 本文使用了泛洪填充算法 (Flood Fill Algorithm) 与随机游走 (Random Walk) 的方式，对图像进行自动填充；

文本解决的难点一共有 4 点：1. 使用 word2vect 对输入的文字、艺术品的文本信息进行语义理解，设计带有语义信息的编码形式；2. 输入文字的编码信息与艺术品文本的编码信息对比，需要搜索其中距离最近者；3. 艺术品的色彩特征提取；4. 上色出现的一系列问题：色彩连续性、色彩比例、效果图的光影效果。

此算法在室内设计线稿图纸的实际测试中取得了良好的结果。达到了目前业界的最好效果。

关键词 计算机辅助设计 深度学习 表示学习 自然语言理解

Abstract

Computer-aided art design such as computer painting, computer automatic coloring, etc. has long been a problem that researchers have long explored, because it is related to the diversity of human expression and the complexity of human experience make it a challenging task which was strongly and continuously explorated by scientists. With the development of artificial intelligence and machine learning in recent years, there has been a new breakthrough in computer art aid systems. For example, GAN can automatically produce pictures, pixel2pixel can automatically fill colors according to existing styles, and PaintsChainer can provide The secondary meta picture is automatically colored.

This paper presents an automatic color palettes generator and coloring system based on semantic intelligence understanding. The system mainly solves that the current automatic coloring system can only color by labels or relying on existing color styles migration, thereby making the system more intelligent and responding directly to human language. For designers, this system can instantly convert the description of the text into color palettes and color the sketch, which greatly facilitates its convenience. The color palettes comes from artworks and has a unique aesthetic that can help designers.

This paper uses three innovations to solve this problem: 1. Using a deep learning model to understand the semantics; 2. Correlating picture information with hidden information at the semantic level, and the partiality characteristics of the picture style similarity are preserved to its partial description of semantic description; 3. In this paper, Flood Fill Algorithm and Random Walk are used to automatically fill in the image.

There are four difficulties in solving the text: 1. To Use word2vect to semantically understand the text and artwork of the input text, and to design the coding format with semantic information. 2. To compare the coding information of the input text and the coded information of the artworks, and to search for the artwork with the closest distance; 3. Extraction of color features of artworks; 4. A series of problems in coloring: color continuity, color ratio, and light and shadow effects of the effect map.

This algorithm has achieved good results in the actual testing of interior design linear draw-

ings, and has achieved the best results in the industry.

Keywords computer aided design nlp representation learning deep learning

目 录

摘要	I
Abstract	II
目录	
插图	VII
表格	IX
缩写、符号清单、术语表	X
1 绪论	1
1.1 课题背景	1
1.2 本文研究目标和内容	3
1.3 本文结构安排	4
2 词向量的构建	6
2.1 词向量在本课题中的研究背景	6
2.2 单词向量化必要性	6
2.3 单词向量化的基本原理	7
2.4 word2vec 单词向量化	10
2.5 本课题单词向量化算法的创新点与具体实现	13
2.6 本章单词向量化方法的结果分析	17
2.7 本章小节	18
3 句向量以及句向量的构建	19
3.1 句子向量化在本课题中的作用	19
3.2 句子向量化的基本原理	19
3.3 本课题所提出的单词向量化方法	19
3.3.1 已经存在的句子向量化的问题	19
3.3.2 LDA 主题自动提取	19
3.3.3 主题权重加权	19
3.3.4 加权的原理	19
3.4 本课题提出的句子向量化方法的实验结果	19

3.4.1 句子向量的相似性判断分析	19
3.4.2 句子向量的聚类分析	19
3.5 本章单词向量化方法的结果分析	19
3.6 本章小节	19
4 技术路线、现有方法分析及背景支持	20
4.1 技术路线分析	20
4.2 国内外研究现状	21
4.2.1 计算机自动图像生成	22
4.2.2 表示学习与 Embedding	26
4.2.3 序列化生成：PixelRNN/CNN	26
4.3 数据库资源	28
4.3.1 艺术品交易数据库	28
4.3.2 维基百科全文	29
5 基于语义识别的自动配色方案辅助设计工具	30
5.1 系统结构	30
5.2 语义编码的实现	32
5.2.1 系统结构之语义搜索	32
5.2.2 文本的向量化	32
5.2.3 本系统的单词向量化	36
5.2.4 本系统的句向量建模	39
5.2.5 搜索算法	40
5.3 色彩提取	42
5.3.1 系统结构之色彩提取	42
5.3.2 色彩模式	43
5.3.3 色彩提取	45
5.4 上色算法	47
5.4.1 系统结构之自动上色	47
5.4.2 泛洪填充算法	47
5.4.3 一些问题	49
5.5 本章小结	50

6 实验结果	53
6.1 输入语句的测试案例	53
6.2 效果图类型的测试案例	58
6.3 本章小结	59
参考文献	60
致谢	63
附录 A 附录：实验结果	64
A.1 文本搜索结果	64
A.1.1 输入：温柔	64
A.1.2 输入：我想要温柔的柔和的感觉	64
A.1.3 输入：温柔的女性化的感觉	68
A.1.4 输入：冷静的男性	71
A.1.5 输入：冷静的男人	74
A.1.6 输入：快乐的儿童	74

插 图

1.1 ImageNet 错误率的下降	2
1.2 从艺术品中借鉴色彩方案的设计作品	3
2.1 word2vec 的语言相似性实例	12
2.2 word2vec 向量聚类可视化	13
2.3 数据库结构	14
2.4 艺术品描述信息	15
2.5 模型计算图	17
2.6 词向量模型 loss 下降过程	17
4.1 设计师使用配色表进行配色	21
4.2 非监督式学习主要用来找到影藏的关系	22
4.3 自编码模型可以从缺少的信息中重建图片	23
4.4 生成模型可以直接生成新的信息而不基于输入的部分信息	23
4.5 GAN 从文字生成真实的图片	24
4.6 从大量图片中学习色彩和谐程度	24
4.7 pixel2pixel 的若干应用	25
4.8 Manga Colorization 在标注颜色的情况下为二次元图片自动上色	25
4.9 Word Embedding 对单词词义的线性保持	26
4.10 pixel network 通过序列化的像素生成图像	27
4.11 pixel network 通过序列化的像素生成图像	27
4.12 数据库结构	28
5.1 内部实现逻辑	30
5.2 用户逻辑	31
5.3 第一步：语义搜索	32
5.4 CBOW model	36
5.5 Skip-gram model	37

5.6 使用哈夫曼树的 CBOW model.....	38
5.7 匹配范围	41
5.8 余弦相似度	42
5.9 第二步：提取色彩	43
5.10 备选方案	44
5.11 第三步：自动上色	47
5.12 泛洪填充	48
5.13 填充过程	48
5.14 漏掉上色的问题.....	50
5.15 色彩连续性的问题	51
5.16 色彩连续性的处理	52
6.1 输入“我想要温柔的柔和的感觉”	54
6.2 输入“温柔的女性化的感觉”	55
6.3 输入“温柔的女性化的感觉”	55
6.4 输入“冷静的男人”	56
6.5 输入“快乐的儿童”	57
6.6 输入“快乐的儿童”	57
6.7 色彩方案：“温柔”对室内设计和服装设计效果图的预览	58

表 格

2.1	one-hot 向量	8
2.2	词频向量	8
2.3	tfidf 向量	9
5.1	one-hot 向量	33
5.2	词频向量	33
5.3	tfidf 向量	34
5.4	上下文向量	35
5.5	输入	35
5.6	输出	35

缩写、符号清单、术语表

NLP	自然语言处理 (Natural Language Processing)
word2vec	神经网络词向量模型
CBOW	词袋模型 (Bag-of-words model)
BP	反向传播 (Back Propagation)
API	应用程序编程接口
CNN	卷积神经网络 (Convolutional Neural Networks)
GAN	对抗神经网络算法 (Generative adversarial networks)
RNN	递归神经网络 (RNN)
tf-idf	term frequency-inverse document frequency
W	权重
V	文本中单词数量
N	词向量的维度
t	单词 (term)
d	文档 (document)
$frequency$	词频
$weight$	权重

1 绪论

1.1 课题背景

如何使计算机能够像人类一样思考、决策、行动这是人工智能科学家长久以来一直思考的问题。^[1]从 20 世纪六十年代，科研工作者普遍关心的是如何利用计算机解决听觉、视觉、语言理解、自动推理等问题。但是这些问题，尤其是计算机视觉和语音识别问题，在相当长的一段时间内发展缓慢。这是因为，基于之前以分析为主的方法，例如基于边缘检测等人工主观判断的特征进行图像识别，或者利用语言学的专家知识进行语义理解和语音识别。自从 1987 年，卡内基梅隆大学李开复博士使用统计概率模型使得当时的语言识别技术大大提升，使得人们开始注重利用统计学习模型进行人工智能相关问题的解决。^{[2][3][4]}

1986 年，时任加州大学圣地亚哥认知科学实验室的 Geoffery Hinton 在 Nature 杂志上发表了一种利用反向传播（Back Propagation）自动优化神经网络模型的方法，利用神经网络进行机器学习第二次成为机器学习研究的重点。当时确实解决了一些复杂的问题，例如 1997 年 YaLeCun 提出的卷积神经网络（Convolutional Neural Networks）能够识别手写的阿拉伯数字，这在计算机尤其是人工智能领域是很重要的进步。但是，神经网络的参数多，为了使得其参数收敛至稳定值，需要大量的训练样本而且其巨大的矩阵运算是一般的计算机所不能承受的，神经网络虽然有了一些优秀的表现，但是并没有产生突破性的表现。

2012 年 ImageNet 计算机视觉识别比赛中，在加拿大蒙特利尔大学担任教授的 Geoffery Hinton 带领团队使得 ImageNet 图像识别的错误率由 30% 以上下降到 15% 一下，这引起了轰动。^[5]

从 2012 年之后，深度学习这一机器学习范式，被用来解决诸多问题。例如此前的 ImageNet 下进行的计算机视觉比赛，在 2015 年其识别的错误率已经低于人类的平均值。2017 年 12 月 22 日，腾讯 DPDAC NLP 实验室使用深度学习神经网络训练的自然语言理解模型，在机器阅读理解的比赛上的正确率已经达到 81.790%，人类在这项测试中的评价准确率为 %82.304。除计算机视觉和自然语言理解领域，深度学习在决策领域也表现出了惊人的进步，2017 年 10 月，Google Deep Mind 的 AlphaGo 战胜了战胜了围棋世界排名第一的柯洁，这表明深度学习在复杂决策上也是具有非常大的潜力。

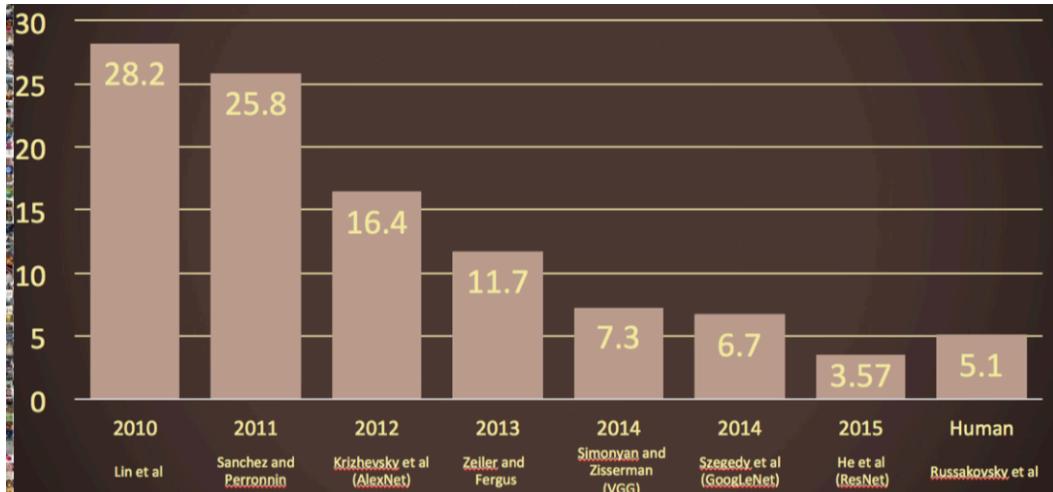


图 1.1 ImageNet 错误率的下降

由于深度学习在计算机视觉、自然语言理解、决策等方面的表现突出。科研工作者希望利用深度学习能够解决更加类似于人类的工作，例如—艺术创作。艺术创作长期以来被认为是人类独有的能力。经过科学家们对于艺术创作过程以及艺术创作心理学的分析发现，人类的艺术创作除了精神与心理学等生理原因，其后天影响，周围环境影响对艺术创作也是具有非常大的影响。或者说，创意，创造力和后天学习具有密不可分的关系。^{[6][7][8]}

因此，计算机辅助艺术设计例如计算机绘画、计算机自动上色等，甚至计算机生成艺术作品在深度学习的逐渐应用的现在，被科学家们寄予希望。计算机辅助设计工具非常常见，各类基于线条信息表示的计算机图像，常常运用在设计图纸中，如建筑图纸、平面图、产品工程图、地形图、流程线框图等。基于明暗的真实感图形，也常常运用在插画、游戏、产品效果图等中，例如 3d 渲染、数字插图等。而从事数字绘画、动画、数字雕塑的艺术家，更是完全的通过计算机进行创作。设计师和艺术家们往往运用软件辅助进行设计和创作，软件是一个执行工具，设计师和艺术家们仍然是设计和创作的主体。

而自动生成图像的工作涉及人类的感受、逻辑、创意。对于人类自身而言，设计是探讨解决问题的方式，常常依赖于设计师敏锐的感知力、知识和经验、逻辑思维、共情能力；艺术更是人类艺术家感受、思考的结晶。对于计算机来说，一方面生成真实感图像的工作，受制于真实感图像兼顾模糊性和逻辑性，故而并不简单；另一方面生成风格化的作品也由于人类感受细腻微妙且复杂多变，人类的表达方式也非常多样化，无法拥有独属于人类的创造力。即便计算机能够将 3d 模型渲染出合理的光影，可是要计算机得到设计和创造的能力实在是困难之极。这一切都使自动生成图像成为一个挑战性极强的领域。而人类持续探索这一领域，也正是因为这些特点带来的魅力，机器是否也可以拥有人类的感受力、理解力、逻辑思维和创造力，这个问题正是所有计算机科学家的终极追求，使得科学家们在

这一领域持续探索。

1.2 本文研究目标和内容

作为一个设计师，在每个项目中，选取合适的配色方案是一个重要的工作。人类能够分辨某些色彩组合是和谐的而另外一些是不和谐的，某些色彩让人感觉兴奋而另一些让人感觉冷静，在色彩理论上，根据人类的生理、文化等等研究的色彩理论非常多，但是却很难讲清楚色彩和感受的关系，没有一个数学模型可以完整的解答这个答案。

设计师对作品采用某种配色都是有目的的，往往需要和环境一起为使用者带来一定的感受。比如儿童乐园的配色、玩具的配色往往鲜艳而给人带来活泼快乐的感受；医院的配色、医疗设施的配色需要给人带来温和舒适并且干净可靠的感受。我们在设计中往往首先使用语言来描述这些满足需求的感受，然后设计师会根据这些需求配色，然而自动化这一程序也同样是很困难的。人类语言描述的需求和色彩给人的感受之间的对应关系也是难以明确的。

事实上，设计色彩方案并制作色卡对于很多设计师也是有挑战性的，设计师常常需要使用线上的配色器，或者借鉴摄影作品、艺术作品等。可以说，艺术作品是设计灵感重要的产生源泉之一。如图 1.2，图中的产品设计和平面设计的配色都来自于莫奈的油画作品图 1.2a。



图 1.2 从艺术品中借鉴色彩方案的设计作品

本文提出了一种基于语义智能理解的自动配色和上色系统。目前自动上色系统通常通过人工标注色彩、有限类型进行色彩填充或依赖已有色彩风格进行迁移，本系统通过从

巨大的数据库中的艺术品学习，从艺术作品中提取色彩方案。依托艺术品图像和文本（包括介绍和题目）对应关系，使得系统直接智能地对人类的语言进行响应。实现根据文本输入来输出设计效果图的配色。

本文的实现方式如下：

1. 本系统依托艺术品交易数据库的内容进行机器学习和训练神经网络，对于数据库中的每一件艺术品有下列信息：艺术品的题目、类型、艺术品的文字描述、艺术品的图片（即绘画作品的内容），作者等。
2. 输入需求的文本描述，通过语义搜索的方式可以找出与之关联度高的艺术品文字及介绍，依托艺术品数据库中艺术品的图像和文字介绍的匹配对应，即找出与输入文本关联度高的图片。
3. 从图片中提取色彩方案，包括绘画作品的色彩和配色比例。
4. 将提取的色彩方案运用到设计效果图上预览。上色使用泛洪填充算法 (Flood Fill Algorithm) 与随机游走 (Random Walk) 的方式，并且使用图层混合算法完成效果图的预览。

对于设计师而言，本系统可以将需求的文本描述即时转化为配色方案对设计线稿进行填充，方便的批量生成多个色彩方案以供选择，大大方便了其使用便捷性。

本文为了实现这个系统使用了三个创新点：1. 使用了深度学习模型对语义进行理解；2. 将图片信息与语义层面的隐藏信息进行关联，其图片风格相似性的偏序特性被保持到其语义描述的偏序特性中；3. 本文使用了泛洪填充算法和随机游走的方式，对图像进行自动填充。

本文解决的难点一共有 4 点：1. 使用 word2vec 对输入的文字、艺术品信息进行语义理解，设计带有语义信息的编码形式；2. 输入文字编码信息与艺术品文本的编码信息对比，在数万条艺术品文本信息中搜索其中距离最近者；3. 艺术品的色彩特征提取；4 上色出现的一系列问题：色彩连续性、色彩比例等。

1.3 本文结构安排

本文为了实现基于语义智能理解的自动配色和上色，首先探索若干类似的课题：

之后本文详细的描述了系统的设计和实现过程。系统的整体结构从算法实现结构以及用户逻辑结构两方面设计。而系统的实现主要涉及自然语言处理，语义编码和搜索，色彩提取，上色四个方面。

最后展示该系统的的结果，通过输入语句案例进行测试，各类效果图测试案例测试该系统的理解能力、准确性、可用性、可拓展性。

2 词向量的构建

2.1 词向量在本课题中的研究背景

如第一章分析，要能够实现基于人类做的自然语言到色彩的自动转化，这首先是要进行文本涵义的表征。需要将日常的文字转换成计算机能够读懂的数字化或者向量化的表征方式。对于一句话的表征，即前文中所叙述的 sentence embedding – 句向量嵌入 – 是目前基于自然语言应用的重要步骤。句子的嵌入化操作，能够使得一句话变成计算机能够直接进行计算的向量。而要进行句子的向量化，目前常见的方法是基于“单词”进行，因为“单词”是能够表征文字信息的最小单位。我们先对单词进行向量化，这种向量化为之后的句子的向量化提供了基础，从而为句子的“语义”级别的理解提供了可能。

由于单词的向量化是长久以来一直研究的问题，其方法众多，而且单词的向量化在 2014 年之后，由于 Google 科学家 Micolove 想提出的向量化方法，众多的科研工作者们在此领域有了长足的发展。本章首先阐述单词向量化的必要性，以及解释目前单词向化不同方法的优劣，从而提出一种适用于本课题的单词向量化方法。

2.2 单词向量化的必要性

要使得计算机具有能够识别语言的能力，计算机机需要能够在某种层面上计算语言的相似性，即，输入两个语义单元，计算机需要能够判断其相似关系。例如，输入“喜欢”和“爱”，由计算机给出的相似性应该要小于“喜欢”和“讨厌”之间的距离。即，计算机应该具有某种函数 $Sim()$ ，该函数接受两个单词的输入 R_{w1}, R_{w2} ，输出正则化的 $-1 \sim 1$ 用以表示两个单词的意思从“很不相思” / “相反” 到“很相似”。

$$Sim(R_{w1}, R_{w2}) \in [-1, +1]$$

如果计算机能够计算这种相似性，那么接下来的句子、段落等更高层面的语义相似性计算机便成为了可能。而要让计算机具有这种能力，首先需要考虑的就是如何表示单词 $w1, w2$ 。

我们可以给单词标号，假设目前有 $|V|$ 个单词，那么依照某种顺序，例如“字典序”

对每个单词进行标号，然后依据字典或其他某种方式定义出来每个单词的近反义词，然后计算单词相似度的时候，检查任意两个单词是否为近义词或者反义词即可。

```
““ vocabulary = 0: 'word': '爱恋', 'syn': '喜爱', '喜欢', '喜爱', 'ant': '讨厌', '憎恨',
1: 'word': '苹果', 'syn': '水果', 'apple', 'ant': ” ””
```

以上的表示方式是长久以来计算机表示单词的方式，在这种实现方式下，能够较为“基础”的实现单词的近反义词的判断。但是这种表示存在 3 个问题：

1. **人工整理成本高**: 该种表示方法，首先需要人工进行同反义词的输入，而人工整理输入就面临着成本高，容易出错等问题；
2. **不能动态适配不同的语料**: 该种表示方法，对于不同的语料库，不能把握其近反义词。例如，对于一个科技类资料库的文本库，“苹果”的近义词应该是“IBM”，“微软”，但是对于生物类的文本库，“苹果”的近义词应该是“梨子”。
3. **不能精确衡量相似度**: 该种表示方法，对于任意两个单词，其内在逻辑是判断这两个单词是否为“近义词”或者“反义词”，输出为一个离散值 $Output \in Syn, Ant \square None$ ，但是用这种离散的方式表示单词关系是不能蕴含词汇之间丰富的关系的。例如，“美国”和“奥巴马”之间，这两个单词之间既不是“相似”单词，也不是“不相似”单词，但是更不是“没有关系”的单词。

由于以上存在的问题，是的我们必须考虑使用新的单词表示方法，该单词的表示方法需要满足以下条件：

1. **计算机能够计算**: 该表示首先需要让计算机能够计算，让计算机能够计算，则需要将其表示成数字（标量或者向量）的形式，我们在这里可以将标量看成是一维的向量，所以下文统称为 **向量**；
2. **能够计算单词之间的相关性**: 利用单词的向量表征，能够计算不同单词之间的相关性，该相关性应该为从 -1 到 1 的连续值。
3. **能够自动化适配于不同的语料库**: 该向量化方式应该能动态适配与不同的语料库，尽可能少的需要人工参与。

2.3 单词向量化的基本原理

单词(或者文本)的向量化在 2014 年之前主要使用 TF-IDF 这种依赖单词概率分布的向量形式进行表示；而对于单个单词主要使用 one-hot 编码进行。

one-hot 编码 one-hot 编码基于上文提到的标号编码的一种向量表示，这两种其实本质是一样的，都是将每个单词表示为唯一的类型 (categorical)。如表 5.1 所表示^[20] 我，例如一句话：“我喜欢吃苹果，而他喜欢吃梨子。”这个句子长度为 7 个不同的单词。我们可以用这样的向量给每个单词编码，每个单词就可以表现为一个 7 维的向量，如表 5.1，每个单词的向量为单词下的一列，就可以将一段文本的信息完整的保存进一个矩阵。

表 2.1 one-hot 向量

	我	喜欢	吃	苹果	而	他	喜欢	吃	梨子
我	1	0	0	0	0	0	0	0	0
喜欢	0	1	0	0	0	0	0	0	0
吃	0	0	1	0	0	0	0	0	0
...									

这样的方法，不可避免的有以下问题：

- 存错空间消耗大：如果对于文本长度较大的情况，就需要很大的空间储存这些向量，比如一本上万个单词的书籍，就需要单词的维度上万，里面大多数的位置又都是 0；
- 不同的语料库中，同样的单词表示不同：这种方式对于不同文本来说并没有同一性，对于同一个词，不同的文本中必须有不同的词向量；
- 单词表征无法保持语义的相似性：这些词向量很难表示词的含义，因为，任意两个向量 \vec{v}_1, \vec{v}_2 的单词向量距离都是相等的，都是 $\sqrt{2}$

表 2.2 词频向量

句子	我	喜欢	吃	苹果	而	他	梨子
我喜欢吃苹果而他喜欢吃梨子	1	2	2	1	1	1	1
我喜欢吃苹果，梨子	1	1	1	1	0	0	1
他喜欢吃吃吃	0	1	3	0	0	1	0

tf-idf 词频编码 为了减少向量的维度，增加矩阵的密集程度，出现了基于词频的单词向量化。以我们的例句为例“喜欢”，“吃”这两个单词出现了两次，其他单词出现了一次。词频向量就是基于此构成。如表 5.2，例如一篇文章中有表中的三句话，则每个单

词的向量为单词下的一列，每句的向量为句子右侧的一行。词频向量和 one-hot 向量相比，维度有所下降，使用空间更有效率。词频也会记录一些文本的信息，不同的文本中单词的词频会有区别。然而完全从句子对单词计数，对词序信息完全忽略，同样难以提取文字本身的意义。

在此基础上的 tf-idf 的向量化不仅仅量化一个单词在一段文本中的频率，还考虑到在所有文本中单词出现的情况，这样就可以表征出一个单词和某段文本的关系。^[21] $tfidf = tf(t, d) * idf(t, D)$ ，其中 $tf(t, d)$ (term frequency) 表征单词 t 出现在文本 d 中的频率。它可以有多种形式，比如 $tf(t, d) = f_{td}$, f_{td} 表示单词 t 出现在文本 d 中次数；或者考虑到文本到长度 $tf(t, d) = \frac{f_{td}}{n_{indocment}}$, $n_{indocment}$ 指文章总共的单词数， $tf(t, d)$ 即出现在文本中的频率；甚至也可以是布尔的，当文本 d 中出现词 t 则 $tf(t, d) = 1$ 否则 $tf(t, d) = 0$ ，等等。 idf (inverse document frequency) 定义为 $idf(t, D) = \log \frac{N}{n_{toccurs}}$ ， N 表示总的文档数， $n_{toccurs}$ 表示含有单词 t 的文档数。

以之前的三句话为例：(a) 我喜欢吃苹果而他喜欢吃梨子。(b) 我喜欢吃苹果。(c) 我喜欢吃吃吃。则对于单词“我”，有： $tf("I", a) = \frac{1}{9}$, $tf("I", b) = \frac{1}{4}$ ，那么句子 b 中“我”这个词占整体比例更高。但对于所有文本， $idf("I", D) = \log \frac{3}{2} = 0$ ，每个句子都出现了“我”这个词，所以 $tfidf("I", a) = tfidf("I", b) = 0$ 。处理文档大量时，对于一些非常常见的词，比如“那个”“这个”这样的词，可以认为它们没什么意义，就会出现这种情况。而比如对于单词“苹果”， $tf("apple", a) = \frac{1}{9}$, $tf("apple", b) = \frac{1}{4}$, $idf("apple", D) = \log \frac{3}{2}$ ，就会出现 $tfidf("apple", a) = 0.045$, $tfidf("apple", b) = 0.1$ 。我以这三个句子计算了单词“我”和单词“苹果”的向量如表 5.3：

表 2.3 tfidf 向量

句子	我喜欢吃苹果而他喜欢吃梨子	我喜欢吃苹果	我喜欢吃吃吃
我	0	0	0
苹果	0.045	0.1	0
...			

tf-idf 向量的表示可以从一定程度上衡量不同句子、文本之间的相似性，但是 tfidf 显然还是不能处理同义词和近义词的问题。例如“我早晨去上班”和“我清晨要去打卡”，这两句话的意思其实是很相似的，但是对于 tfidf 向量化来说，对于词表 [“我”，“早上”，“清晨”，“要去”，“去”，“上班”，“打卡”]，第一句话“我早晨去上班”会得到向

量 $\langle a, b, 0, 0, c, d, 0 \rangle$ 而第二句话的向量为 $\langle e, 0, f, g, 0, 0, h \rangle$, 这两句话的向量距离为 $\sqrt{(a-e)^2 + (b-0)^2 + (f-0)^2 + (g-0)^2 + (h-0)^2}$, 这是一个相对比较大的距离。利用 tfidf 表示, 这两句话计算机只能发现“我”和“我”的意思是一样的, 但是对于其他不一样的单词, 虽然是同义词或者近义词, 但是计算机依照这种表示方式, 依然会认为是完全不一样的单词, 从而不能表示其相似性。

2.4 word2vec 单词向量化

如上文所述, 利用 one-hot 或者 tfidf 向量化能够完成使计算机进行计算的问题, 但是存在以下比较显著的问题:

1. 不能识别近义词, 同义词: 不论是 one-hot 还是 tfidf 表示, 都不能表示近反义词等, 而能自动识别近反义词, 同义词其实是语言理解中很重要的一个环节, 因为只有这样, 才有可能完成不基于单纯的“关键词”匹配完成语言分析。
2. 需要很大的存储空间: one-hot 和 tfidf 存储的都是稀疏向量, 需要很大的存储空间;
3. 不能适应性更新 (Adaptive Updating): 该问题值得是, 如何我们已经有了 10000 个单词的 one-hot 词向量编码或者 10000 句话的句子 tfidf 编码, 基于 one-hot 和 tfidf 的原理, 如果我们新增加了 100 个单词或者句子, 我们需要把这 $10000 + 100$ 个句子全部更新。这显然是不合理的。因为如果每次增加新的数据都要对全部的数据进行更新, 那么当已经存储的数据量很大的时候, 每一次增加新词汇、新句子的成本就会很大。

2013 年 Google 的科学家 Mikolov 在《Efficient Estimation of Word Representations in Vector Space》^[23] 中提出的基于神经网络的非监督词向量构建方法, 获得良好的结果, 在 2014 年, 在此基础之上有提出了 Negative Sampling, Hirerachy Softmax 等优方法, 使得该种词向量的计算效率大大提升。

从 2014 年到 2017 年, 基于 word2vec 的词向量表征领域的进步显著, 在 2014 年由 Stanford NLP Group 提出的 GloVe, 2016 年提出的 Fasttext, 2017 年 salesfore 提出的 CoVe (Context Vector), 直到 2018 年提出的 EMLo (Embedding from Language Model) 使得单词的向量表征获得很大的提升。基于词向量的成功, 推动了自然语言理解领域多个子领域的进步, 其中包括: 1. 语言相似判断; 2. 机器阅读理解; 3. 文本匹配; 4. 蕴含识别。

因此, 本课题的单词训练方法也采用 word2vec 的思想, 但是由于本课题所面临数据库的特殊性, 需要在原始的 word2vec 算法上做一定的修改。本小节首先描述 word2vec 的基本原理。

word2vec 是 Mikolov 提出的一种将单词进行向量化的方法。该向量化方法主要基于语言学家 J.R.Firth 在 1957 年提示的理论，“You shall know a word by the company it keeps”，即在语言学中认为，一个单词的意思可以通过其单词周围的单词的来进行判断。在英语考试里边的“完型填空”，我们通过判断周围的单词就可以判断某个缺失的地方应该填入什么单词。例如，如果我们的观察“到了春天，这里的竹笋 [] 得很好”，“这条河两个岸边距离很远，需要建一座很 [] 的桥”。我们可以知道这两句话里边都是要填“长”这个词，但是这两句话里边的“长”字的意思又是不一样的。这表明了一个单词其实依据它周围的词就可以确定出来。

word2vec 基于这种假设，希望能够利用一个单词与其周围单词的关系来确定单词的词向量。其基本思路为，每一个单词都有一个初始的随机向量表示，该向量表示可以是几十到几百维度的向量。我们将单词进行标号，假设其标号为 i ，那么，单词 i 的向量表示为 \vec{w}_i ，如何确定这个单词？

word2vec 提出了两种方法，一种是 skip-gram，一种是 CBOW。其中，skip-gram 设计一个机器学习模型，该模型的目标为输入一个单词的词向量 \vec{w}_i ，对此词向量进行线性变化，之后将该单词进行 softmax 函数变化，使得依据该单词向量，能够求得该单词周围出现的单词。即，skip-gram 希望拟合这样一个模型，公式 ??，该模型通过输入单词的词向量 \vec{w}_i ，能够预测其周围应该分布什么单词。

$$Y(\vec{w}_i; \theta) = w_i, w_i \in context_i \quad (2.1)$$

skip-gram 基于一批训练数据，该训练数据为数据量较大的文本信息。skip-gram 每次指定一个单词，然后训练该模型，使得模型能够预测周围单词的分布，该模型输入一个单词的向量，依据此向量产生的周围单词的分布为 2.2：

$$Pr(\vec{V}_o | \vec{V}_c; \theta) = \frac{e^{\vec{V}_o^t \cdot \vec{V}_c}}{\sum_{i \in V} e^{v_i^t \cdot v_c}} \quad (2.2)$$

获得了周围单词分布的概率估计，就可以使用神经网络的交叉熵 (cross entropy) 方式进行反向传播 (back propagation)，值得指出的是，word2vec 模型（包括 skip-gram 与 CBOW）与普通的神经网络模型并不完全一样。因为，普通的神经网络模型其反向传播之后，利用随机梯度下降 (SGD) 等方式是对参数集 θ 进行更新，对于输入的 X 值，并不更新。但是 word2vec 与其他神经网络模型不同的是，word2vec 的反向传播模型会一直延续到输入值 X ，也就是说，其每一次的输入 V_c 的向量，也是会通过反向传播的方式进行更新的。在每一轮 (epoch) 的更新中，每个单词的向量会得到更新，而这种更新的结果就是，如果两

个单词周围经常出现的单词比较相似，那么这两个单词的词向量也会比较相似。word2vec 模型简单高效容易操作，获得的词向量结果能够良好的保持语言的相似性。例如基于中文维基百科进行 word2vec 的训练，我们可以获得如图 ?? 的近义词关系。

其中图 2.1 一共表现了 word2vec 的四种重要的特性：

1. **同义词近义词发掘:** 例如当输入“数学”的时候，我们可以获得与“数学”单词涵义接近或者类似的单词；
2. **单词线性关系的保持:** 保持线性关系这个特性是 word2vec 模型最为突出的一有点，例如对于单词“女人”和“妈妈”的向量，这两个单词之间的距离最近解决与“男人”和“爸爸”之间的距离。这表明了 word2vec 模型精确得定义了单词之间的语言距离。
3. **同一类型单词的自动识别:** word2vec 因为将具有关联关系的单词进行了自动聚类，所以可以识别出任意的单词是不是属于同一相关类型，例如在图中，能够识别“牛肉”与“黑板，作业，老师”不属于同一个类型。
4. **任意两个单词之间的关系计算:** word2vec 可以计算任意两个单词之间的相关关系。

```

word = '数学'
res = model.wv.most_similar(word, topn = 10)
for item in res:
    print(item[0]+", "+str(item[1]))

微积分, 0.7719005346298218
高等数学, 0.7582617998123169
数学分析, 0.7573831081390381
算术, 0.7523037195205688
数论, 0.7520559232444763
代数, 0.7386804819107056
概率论, 0.7161834239959717
数学家, 0.705559253692627
数理逻辑, 0.7045130729675293
几何学, 0.7014443874359131

model.wv.most_similar(positive=['女人', '妈妈'], negative=['男人'], topn=1)
['爸爸', 0.8544228076934814]

model.wv.doesnt_match("黑板 作业 老师 牛肉".split())
'牛肉'

model.wv.similarity('男人', '工作')
0.0845577918019063

```

图 2.1 word2vec 的语言相似性实例

而且，利用高维向量的降维可视化 TSEN，可以对训练出来的向量进行可视化。TSEN 是一种高维向量的可视化方式，该算法可是使得如果两个向量在高维空间中接近，那么在降维之后两个向量的距离依然相近。尽管我的实际观察发现，其语义相似的单词，经过 word2vec 模型训练得到的最终向量距离确实更相近。如图 2.2 所示，和污染物相关的单词就被聚类到了相近的距离。

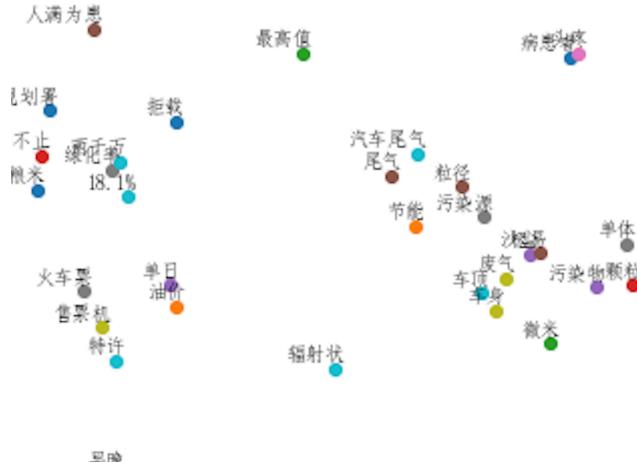


图 2.2 word2vec 向量聚类可视化

除了 skip-gram, word2vec 的方式还有 CBOW(Continuous Bag of Words), 该模型与 skip-gram 类似, 但是不同的时候, skip-gram 输入的是一个单词用以预测其周围的单词, 而 CBOW 模型的输入是一个单词的周围单词, 用来预测这些单词的“中心单词”, 例如, “我今天早上吃了一顿早饭”, skip-gram 输入“早上”, 期望模型能够输出“我”, “今天”, “吃了”, “一顿”, 而 CBOW 的输入是“我”, “今天”, “吃了”, “一顿”, 期望的输出是缺失的单词“早上”。

初次之外, 还有 Stanford NLP Group 提出来的 GloVe 模型, 该模型用来预测两个单词同时出现的概率 (co-occurrence) 的概率; 以及 2017 年和 2018 年提出来的 CoVe(Context Vector) 和 ELMo(Embedding of Language Model), 这两种模型基于机器自动翻译模型的结果, 能够使得每个单词在不同的句子中的单词词向量的表征不一样, 从而更加精确的表征了单词向量。

2.5 本课题单词向量化算法的创新点与具体实现

上文分析了单词向量化的基本原理, 通过该种单词的向量化表示使得我们具有能够表征单词的能力。理论上, 基于 Micolvo 提出来的 word2vec 模型, 可以实现单词的向量化。但是由于本课题使用的数据源的限制, 利用该种 word2vec 模型并不能训练出有效的单词词向量, 本课题需要在已知的词向量方法上, 基于 word2vec 的思想, 设计新的模型算法, 进行艺术相关词汇的词向量的训练。

本课题使用的数据源为雅昌拍卖网艺术品交易数据库 (auction.artron.net), 数据库中关于艺术品以及艺术品交易的信息丰富, 包括流通的艺术品的题目、图片、特征、作者、艺

艺术品类别、交易价格、流通时间等等。本课题所用到的数据库中的信息包括两类如图 4.12，一类是艺术品信息，另一类为艺术品相关的文章。

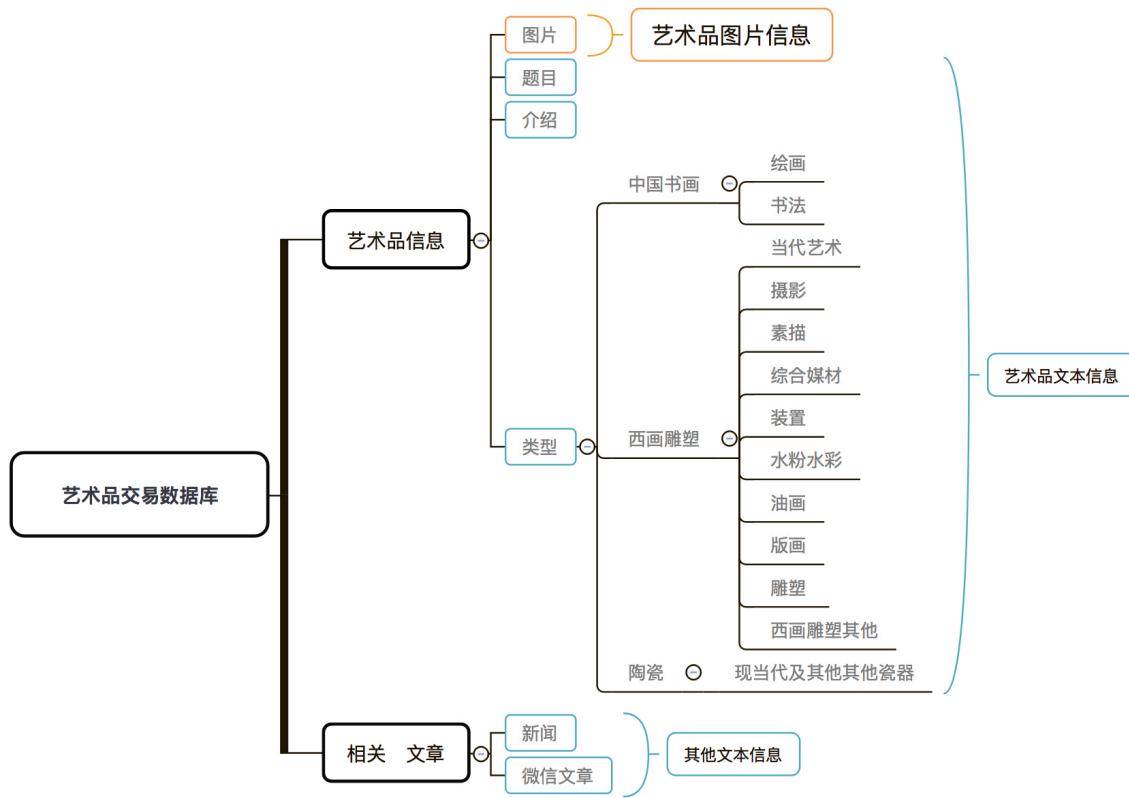


图 2.3 数据库结构

其最重要的艺术品信息的数据样例如图 2.4，该数据库比起 word2vec 原始模型其数据源有以下区别：

- 1. 数据量小：**数据量较小，则会产生模型无法收敛，以至于按照 word2vec 的方式，不能获得有效的单词向量；
- 2. 每句话较短：**由于每句话较短，所以有相当一部分的单词不能具有足够的上下文。所谓上下文 (context) 指的是一个单词的左边和右边的 N 个单词，例如“深度学习模型是一种神经网络模型”，那么，当我们的上下文窗口定义为 3 的时候，“是”的上下文为“深度学习模型”和“一种神经网络”，即，距离在 3(包括 3) 以内的单词都是上下文的。但是由于每句话较短，那么就会之后最中间的几个词汇能够获得足够的上下文，尤其是在这句话的开头和结尾，如果我们把上下文的窗口数量定义为 N，那么我们期望的上下文单词应该是 $2N$ 。这种情况在维基百科等大语料库中是成立的，在这种大语料库中，绝大多数的单词在训练时候都能获得足够的上下文信息。但是

title	oth...	artwor...	artwork_t...	di...	cal...	year	tec...	mo...	display...	calc...	description	image_url
花卉 立轴 设色绢本	NULL	绘画	中国书画	NULL	NULL	暂无	NULL	NULL	104x38cm	3952...	NULL	http://img2.artron.net/auction/old/art4...
书法对联 对联 水墨纸本	NULL	书法	中国书画	NULL	NULL	暂无	NULL	NULL	111x22c...	4884...	NULL	http://img2.artron.net/auction/2010/ar...
行书八言联 对联 洒金笺	NULL	书法	中国书画	NULL	NULL	暂无	NULL	NULL	166x35...	1162...	说明: 上款汪彭年 (1882-19...	http://img3.artron.net/auction/old/art4...
赏花图 镜框 设色纸本	NULL	绘画	中国书画	NULL	NULL	暂无	NULL	NULL	54x49cm	2646...	说明: 2012年2月12日, 此...	http://img1.artron.net/auction/2014/ar...
双吉图 立轴 设色纸本	NULL	绘画	中国书画	NULL	NULL	暂无	NULL	NULL	66x41cm	2706...	NULL	http://img3.artron.net/auction/old/art3...
美人图 设色纸本	NULL	绘画	中国书画	NULL	NULL	暂无	NULL	NULL	65x43cm	2795...	NULL	http://img4.artron.net/auction/2013/ar...
竹子 立轴 设色纸本	NULL	绘画	中国书画	NULL	NULL	暂无	NULL	NULL	79x29cm	2291...	NULL	http://img2.artron.net/auction/2011/ar...
山水 扇片 纸本	NULL	绘画	中国书画	NULL	NULL	暂无	NULL	NULL	--	NULL	NULL	http://img1.artron.net/auction/2013/ar...
1922年作 山水 立轴 水墨纸本	NULL	绘画	中国书画	NULL	1922...	NULL	NULL	NULL	108x54cm	5832...	按: 浩公即赵浩公 王竹虚 (...	http://img5.artron.net/auction/old/art5...
孔雀图 立轴 设色纸本	NULL	绘画	中国书画	NULL	NULL	暂无	NULL	NULL	98x50cm	4900...	NULL	http://img4.artron.net/auction/2011/ar...
1995年作 无限星网 压克力...	NULL	油画	西画雕塑	NULL	NULL	1995...	NULL	NULL	290.5x...	1511...	来源 纽约, Robert Miller画...	http://img4.artron.net/auction/2015/ar...
1970年作 行书 七言诗 立轴...	NULL	书法	中国书画	NULL	NULL	1970...	NULL	NULL	68x30cm	2040...	识文: 昔从姑射见神僊, 绛...	http://img2.artron.net/auction/2011/ar...
书画合璧 册页 (十二开选...	NULL	绘画	中国书画	NULL	NULL	暂无	NULL	NULL	21.5x33...	8514...	徐世昌自藏并题画签 对题: ...	http://img3.artron.net/auction/old/art4...
空林清话 镜心 设色纸本	NULL	绘画	中国书画	NULL	NULL	暂无	NULL	NULL	51.5x64...	3296...	林海钟, 福建永春人, 1990...	http://img4.artron.net/auction/old/art2...
1887年作 游鱼图 立轴 设色...	NULL	绘画	中国书画	NULL	NULL	1887...	NULL	NULL	102.5x4...	5022...	作者简介: 任伯年 (1840~...	http://img5.artron.net/auction/2014/ar...
丁卯 (1747) 年作 椿萱并...	NULL	绘画	中国书画	NULL	NULL	丁卯...	NULL	NULL	120x47...	5700...	说明: 一、高凤翰, 清代画...	http://img2.artron.net/auction/2014/ar...
2003年作 侧面裸体 油彩 布面	NULL	油画	西画雕塑	NULL	NULL	2003...	NULL	NULL	55x50cm	2750...	展览: 郭润文油画作品展, ...	http://img4.artron.net/auction/old/art5...
水果与月季花 油彩 画布·框	NULL	油画	西画雕塑	NULL	NULL	暂无	NULL	NULL	60x50.5...	3030...	备注: 附作品与画家本人的...	http://img3.artron.net/auction/old/art3...
韦陀像 立轴 设色绢本	NULL	绘画	中国书画	NULL	NULL	暂无	NULL	NULL	147.5x8...	1194...	韦陀居四王三十二将之首...	http://img3.artron.net/auction/2011/ar...
溪山访友图	NULL	绘画	中国书画	NULL	NULL	暂无	NULL	NULL	239x70cm	1673...	高术平, 男, 号东齐散人, ...	http://img1.artron.net/auction/2014/ar...
1979年作 丝路花雨 人物形...	NULL	绘画	中国书画	NULL	NULL	1979...	NULL	NULL	27x17.5...	472.50	说明: 1.附本拍品文学台本...	http://img5.artron.net/auction/2013/ar...
1994年作 夏日风情 布面 油彩	NULL	油画	西画雕塑	NULL	NULL	1994...	NULL	NULL	73x91cm	6643...	在当代中国油画家家中, 以风...	http://img2.artron.net/auction/2014/ar...
1971年作 玉手琵琶 立轴 沈	NULL	绘画	中国书画	NULL	NULL	1971...	NULL	NULL	110x69cm	6260...	说明: 1.附本拍品文学台本...	http://img2.artron.net/auction/old/art3...

图 2.4 艺术品描述信息

在本课题的艺术品交易数据库中, 不能获得足够上下文的单词却占大多数。

基于以上两点的分析, 本课题的词向量模型参考 word2vec 的原理, 并且参考 Stanford GloVe 词向量的思想, 设计了新的神经网络模型。该模型输入两个单词, 模型用以计算这两个单词出现在同一个艺术品作品的描述信息中的概率。即, 训练一个模型, 该模型接受两个输入, $\vec{w_1}$ 和 $\vec{w_2}$, 然后在一组参数 θ 的约束之下, 预测该这两个单词出现在同一个艺术品描述信息中的概率 P 。参考前文的 word2vec 原理, 在此模型之下, 对于单词 w_1 和单词 w_2 , 如果两个单词经常出现在同一个作品的描述中, 对于另外一组单词 w_1 和 w_3 也经常同时出现怎么艺术品的描述信息之中。如果 $\langle w_1, w_2 \rangle$ 与 $\langle w_1, w_3 \rangle$ 同时出现的概率类似, 那么 $\vec{w_1}$ 与 $\vec{w_2}$ 输入模型, 与 $\vec{w_1}$ 与 $\vec{w_3}$ 输入模型的输出结果是类似的。从而 $\vec{w_2}$ 与 $\vec{w_3}$ 的单词向量会收敛至相似。而基于上一节所述的 word2vec 的基本思想, 如何两个单词经常出现在类似的位置, 则这两个单词的意思也会相似。所以, 经过以上步骤, 就实现了艺术品交易单词的向量化。

以上分析过程可以形式化的表示为, 我们需要拟合训练一个模型 Y , 该模型的目标函数为 2.3

$$Loss(\vec{w_1}, \vec{w_2}; \theta) = -\log(\sigma(model(\vec{w_1}, \vec{w_2}))) \quad (2.3)$$

数据预处理 在训练之前, 我们首先需要对数据进行预处理, 本课题使用 jieba 分词程序进行自动化分析, 而且使用隐含马尔可夫链进行新词发现。数据预处理的过程如算法 1 所述。

Input: List: A List which contains art work description information

Output: words matrix: a words matrix M, M[i][j] is the probability of $word_i$ 和 $word_j$ co-occurrence

split each string to a words list

$matrix \leftarrow [];$

$total - occurrence \leftarrow 0;$

for one-description \leftarrow List **do**

for $word_i \leftarrow$ one-description **do**

for $word_j \leftarrow$ one-description **do**

if $word_j == word_i$ **then**

 | Continue

end

 ; $matrix[word_i][word_j] += 1;$

$total - occurrence += 1;$

end

end

end

for $word_j \leftarrow matrix$ **do**

for $word_i \leftarrow matrix$ **do**

 | $matrix[word_i][word_j] / = total_occurrence;$

end

end

return $matrix;$

Algorithm 1: co-occurrence 频率预处理

经过算法 1，我们可以获得每个单词对之间出现在同一个艺术品描述信息之中的概率。该概率矩阵为一个 $30,000 * 30,000$ 的矩阵，该矩阵包含了词库中 3 万的词汇的同时出现的概率分布。为之后利用神经网络进行计算提供了基础。

训练神经网络 在定义了 loss 函数之后，本课题建立神经网络模型进行训练。基于之前所定义的 loss 函数，该机器学习模型的输入应该为两个单词的向量， \vec{w}_1 与 \vec{w}_2 ，然后对这两个向量进行线性计算，将该向量映射到一维空间，再对该输出向量进行 sigmoid 函数

的运算，将其限定在 0 1 之间，用以表示概率。

$$Loss(\vec{w}_1, \vec{w}_2; \theta) = \log(sigmoid(\vec{w}_1^t \cdot \vec{w}_2)) \quad (2.4)$$

本课题是神经网络训练环境如下：

1. tensorflow 1.4
2. GPU GTX * 8
3. CPU Intel i5

基于 tensorflow 构建如图 2.6 所示的计算图.

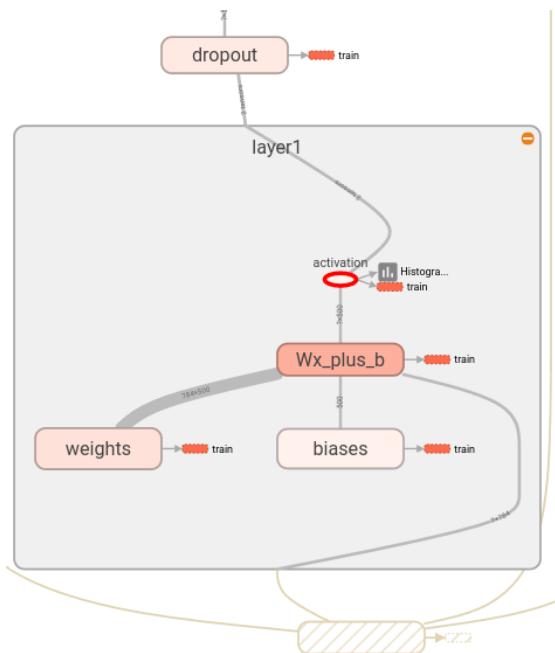


图 2.5 模型计算图

在此基础之上，我们在 GPU 服务器进行模型训练，依据图 ??所示，我们可以发现其 loss 下降很明显。

对于该模型来说，在我们限定的词汇大小上（本课题保留的词汇集大小和 3 万个），平均每个艺术品描述所包含的词汇经过统计为 8 个单词，一共具备 43,200 条描述信息数据。则，对于任意两个单词，判断是否在同一描述信息的准确概率为：

$$baseline = \frac{C\left(\frac{43200}{1}\right)\left(8\right)}{C\left(\frac{43200}{2}\right)} = 0.000185 \quad (2.5)$$

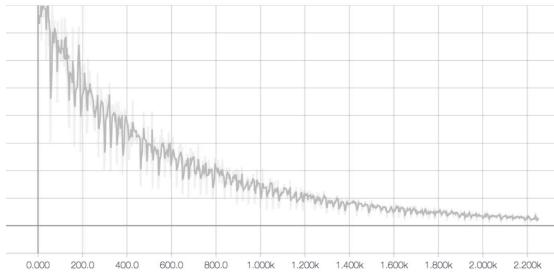


图 2.6 词向量模型 loss 下降过程

基于此基准，则，如果模型是随机进行判断的，那么其 $loss$ 值应该为 $-\log_2(0.000185) = 12$ ，所以如果模型的 $loss$ 值低于 12，则说明该模型是可以训练的。实验证明，如图 2.3 模型最终的 $loss$ 稳定在 **0.1** 左右，说明本课题设计的模型可以进行良好的训练。

2.6 本章单词向量化方法的结果分析

基于以上分析与实验，课题实现了艺术品交易信息的单词向量的训练。

首先，课题获得的词向量保持了单词的语言相似性，如图 XXX 所示；

其次，课题获得的词向量保持了线性计算关系，如图 XXX 所示；

而且，课题获得的词向量可以计算单词是否同属一个类型。

词向量具有以上特点，这表明了单词向量保持语言相似性的良好特性。该向量为之后句子长度、段落长度的语义分析提供了基础。

2.7 本章小节

本章通过分析现有单词词向量的方法，说明了使用简单的 one-hot 与 tfidf 实现词向量的缺点。从而使用了更为先进的，基于深度学习的 word2vec 方法，word2vec 可以良好的保持单词语义级别的相似性。但是，经过分析，word2vec 方法直接应用到本课题的数据是有问题的。因为本课题的每一条艺术品描述信息都比较短，这导致使用 word2vec 模型会缺乏足够的上下文 (context words)，不能良好的保持语言相似性。课题参加 word2vec 的基本思想，还参考了 Stanford NLP Group 提出的 GloVe，设计了一种新的词向量方法。并且在 GPU 服务器与 tensorflow 框架下实现了神经网络结构，实现了词向量的训练。而且本章还展示了训练得出的词向量在：1. 同义词判断；2. 词向量的线性关系；3. 判断语义同属关系。这三个重要特性的表现，说明课题训练的词向量良好的保持了语义相似性，为之后基于更长的自然语义进行语义分析提供了基础。

3 句向量以及句向量的构建

3.1 句子向量化在本课题中的作用

3.2 句子向量化的基本原理

3.3 本课题所提出的单词向量化方法

3.3.1 已经存在的句子向量化的问题

3.3.2 LDA 主题自动提取

3.3.3 主题权重加权

3.3.4 加权的原理

3.4 本课题提出的句子向量化方法的实验结果

3.4.1 句子向量的相似性判断分析

3.4.2 句子向量的聚类分析

3.5 本章单词向量化方法的结果分析

3.6 本章小节

4 技术路线、现有方法分析及背景支持

4.1 技术路线分析

事实上，承上一章所述之背景，设计工作和艺术创作虽然强烈依赖作者的感受力和创造性，但是在某些环境下也是重复的脑力活动。例如在游戏美术、电影场景绘画、高概念电影故事创作以及网络小说的写作等工作中，创作常常符合某种模式，其重复的模式非常多。在此模型下，艺术创作可以抽象为一个制作或者生成过程，如方程(4.1)所示。

$$ArtWork = Generator(f_0, f_1, f_2, \dots, f_N, r_{x \sim p}) \quad (4.1)$$

但是由于传统机器学习的两个特点，使得艺术生成并不适用传统机器学习的方法，第一是：其输出是一个序列性的而非特点长度的标签（Label）或者回归预测值；第二是：艺术的输入特征难以确定，其输入也是一个复杂的序列过程。这两个因素使得计算机进行自动艺术创作在长时间内仍然依赖于人类编写规则，使得计算机艺术创作不能广泛得适用于更加通用的场景。

而深度学习在特征自动学习（Representation Learning）和序列数据特征的提取中均具有良好的效果，所以很多科研工作者开始使用深度学习相关的方法来尝试计算机自动生成艺术作品。

色彩方案设计是在设计领域中普遍的一个工作方式。作为一个设计师，在每个项目中，选取合适的配色方案是一个重要的工作。在设计师设计的过程中，例如设计服饰，软装设计，建筑外观设计等，依据现有的环境特点，应用场景，目标用户，需要制定配色方案，然后将配色方案应用于整体方案上，有时配色方案还会应用到成套的产品中的各个子产品。由于自然语言理解与图像识别等领域目前都有了长足的进步，所以设计作品自动配色这个问题目前有可能使用神经网络和深度学习的内容进行自动化。

而且，由于自然语言理解在最近几年取得的成果，其在语义理解方面可以良好的应用于设计作品自动配色这样领域。目前，计算机辅助设计在这方面的方式是采用一张配色表，使用关键词进行配色，例如图 4.1 这种方式进行。这种方式其实是计算机语言理解的时候的一种折中做法。

这种方式存在两个问题：第一，有限的词汇其实不能表达丰富的含义，设计作品的应



图 4.1 设计师使用配色表进行配色

用涵盖生活的方方面面，其需要表达的语义丰富性是远远高于有限的这些单词的；第二，单个词汇与整体意向组织不具备线性叠加的关系。即，设计师要设计出一个“温柔的春天”这样意境的设计作品，其颜色不能是先选择“温柔”，再选择“春天”，而是需要再次进行加工。将这两种的意向进行重叠。而这种加工是目前的配色表不能完成的。

因此，本课题希望基于目前的自然语言理解的技术与图像自动化处理的方式，实现从作者的直接描述到自动产生配色图的自动化转化。例如，一个设计师接到了室内软装设计的需求，甲方的需求是“洋溢着春天的感觉，要很温柔，然后要像有小孩在草坪上奔跑一样”，本课题要实现的基于语义识别的自动配色系统，只需要将这句话“洋溢着春天的感觉，要很温柔，然后要像有小孩在草坪上奔跑一样”，输入到系统中，就会自动匹配色彩方案，并且可以为直接其线稿进行上色，可以方便的预览，把握整体风格，参考及其方便，甚至批量生成预览图像，为客户展示和预览也节省大量时间。

基于之上的分析，本课题研究一种例如深度学习与表示学习的方式，在此基础上实现从语言理解到色彩搭配的自动转化。之后，本课题在真实数据库上进行了实验，本模型确实能够实现良好的自动化转化，证明了该模型的有效性。

4.2 国内外研究现状

因为本研究课题是机器学习、自然语言处理和图像合成的交叉领域，本小节主要从以下三方面进行国内外现状的分析：

1. 自动图像合成
2. 表示学习 (Representation Learning)
3. 序列化生成

4.2.1 计算机自动图像生成

实现计算机自动图像生成或自动色彩搭配，目前涉及到的方法如下：

- 查表法
- 马尔可夫链
- 机器学习与人工智能方法
- 原细胞有限自动机
- 基于规则的算法 [Wiggins(1991), Nierhaus(2009)]

使用机器学习与人工智能方法主要分为两个类型，一是监督式的 (Supervised)，希望给计算机一定的监督样本 (Samples)，使得计算机能够拟合一个函数，使得该函数能够拟合训练数据到已知样本标签的映射如等式 (4.2)

$$\text{function}(x) \rightarrow y \quad (4.2)$$

监督式的学习方式在图像生成领域应用不多，主要使用因为该种方式主要应用在问题结果确定已知并且有限的情况下，即其目标函数的 $y \in \mathbf{R}^n$ 且 $|y| \leq N$ 。但是，在图像生成领域，目标函数的结果大家的期望并不是从已知的范围内选择一些出来，而是希望能够自动创作。

所以目前机器学习方式主要采用**非监督学习**的方式进行图像生成。非监督的学习方式主要通过从已知的信息中，获得不同类型的信息、数据的隐含关系，例如图 4.2。

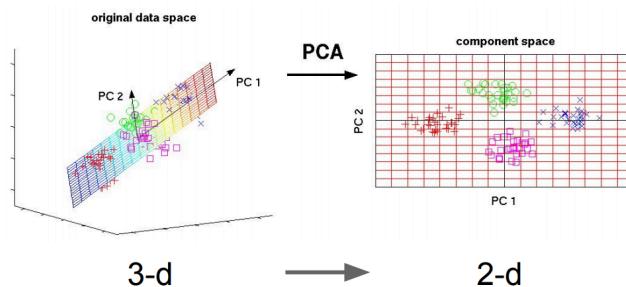


图 4.2 非监督式学习主要用来找到影藏的关系

而基于非监督式的学习，提出的 Autoencoder 自编码模型可以在缺少某些信息的情况下，对图像进行重建。例如图 4.3 所示。

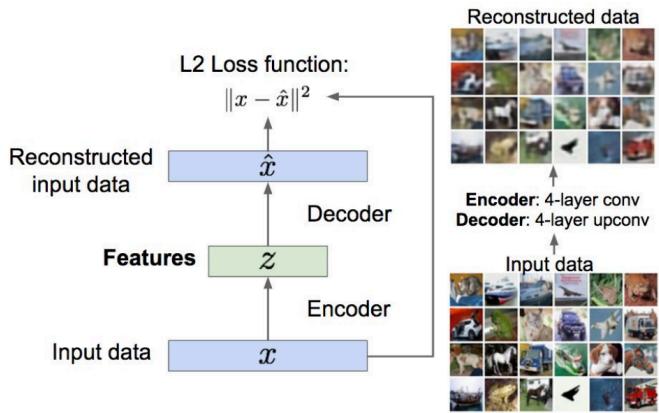


图 4.3 自编码模型可以从缺少的信息中重建图片

由于重建信息的成功，如图 4.3，科研人员考虑如何通过已知的数据，获得其潜在的数据概率分布，从而可以不依靠部分信息而对全部的信息进行重建，以重建其具备与源数据源具有同样数据分布的数据。



图 4.4 生成模型可以直接生成新的信息而不基于输入的部分信息

要完成此任务，目前主流的使用 PixelRNN/CNN, GAN, 马尔科夫连，以及模仿学习 (Imitation Learning)，这些模型都可以模型训练之后，不需要输入部分“初始信息”，而直接获得新的符合原来数据分布的新数据，例如图 4.4 所示，模型可以随机的产生室内的图片，人像，以及自动为线稿配色。

计算机生成图像——GAN 随着理论研究的发展，计算机艺术辅助工具也有了新的突破，通过对抗神经网络算法 (Generative adversarial networks - GAN)，能够自动生产图片。^[9]，GAN 由判别器和生成器实现，判别器对照相机拍的照片和生成器生成的图像进行分类，而生成器从随机噪声中生成尽可能以假乱真的图像，判别器和生成器相互对抗，最后生成判别器无法判别的图像。这种模型，可以生成现实风格的图像，也可以用于将马赛克还原为清晰的照片。使用 GAN 可以根据文字生成图片，如图 4.5。^{[10][11]}

图 4.5 中显示基于 GAN 可以根据一句话，例如“小鸟，胸部和冠羽是粉红色的，黑色的一级和次级飞羽”就可以生成类似照片的图片。使用“胸部和冠羽是粉红色的鸟”和

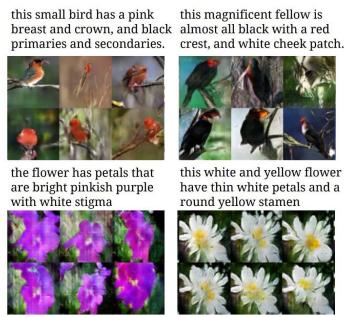


图 4.5 GAN 从文字生成真实的图片

“黑色飞羽的鸟”的照片作为真实照片进行训练，生成器生成尽可能以假乱真的图像，判别器对真实照片和生成器生成的图像进行分类，相互对抗，最后生成判别器无法判别的图像。于是生成器就可以生成满足“小鸟，胸部和冠羽是粉红色的，黑色的一级和次级飞羽”这句话的图片。GAN 生成的图片非常真实，但是每每面对新的一句话，它就需要大量的图片和较长的时间训练 GAN 模型。

自动配色的探索 pixel2pixel 相关的工作中也有关于配色选取的工作。colormind 使用 pix2pix 的方式从输入照片提取生成色彩方案。^{[12][13]} 与此同时也有研究者研究通过大数据探讨关于配色的问题，多伦多大学的 Peter O'Donovan 等人通过 color.adobe.com、www.coluorlover.com 等配色网站上用户对各类配色的打分数据，使用机器学习的方式学习为色彩的和谐程度打分。^[14]

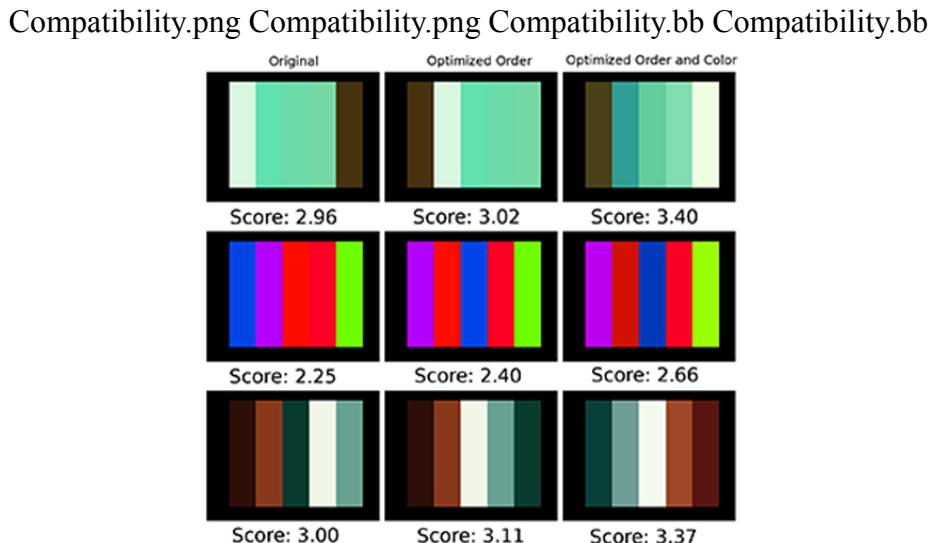


图 4.6 从大量图片中学习色彩和谐程度

自动上色 pixel2pixel 探讨了更多基于对抗神经网络的，从图像生成图像的应用。在这一系列的相关工作中，包括能够为黑白照片自动上现实的合理的颜色，也可以为轮廓填充现实的合理颜色，以及从符号化的色块表达生成现实风格的图片，自动生成遮罩，自动替换照片中的人物。^[15]

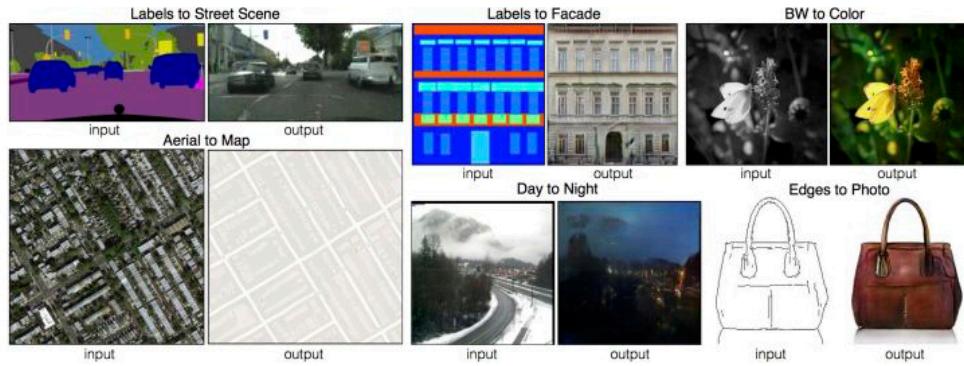


图 4.7 pixel2pixel 的若干应用

有若干的项目都在讨论关于日本式漫画的自动上色问题，例如 PaintsChainer 和 Manga Colorization，如图 4.8。^{[16][17]} 日式漫画风格较为统一，图形基本封闭，模式较为固定，并且对批量上色要求巨大。Manga Colorization 能在标注颜色的情况下为二次元图片自动上色。PaintsChainer 不局限于封闭的图形，并且自动完成上色与阴影模式，但其效果而言，常常出现不可控的色块和理解错误的内容。

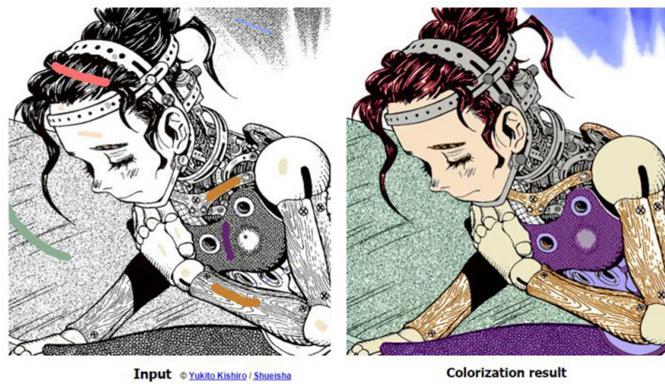


图 4.8 Manga Colorization 在标注颜色的情况下为二次元图片自动上色

4.2.2 表示学习与 Embedding

表示学习是机器学习中重要的研究内容，主要研究如何将数据有效的表示到计算机中。其中，最重要的的方式是如何通过一种方式使得新表现的数据形式能够保持其原始线性关系^[18]，这种行为在计算机及数学中叫做 embedding。Embedding 是指，保持一种偏序关系，是的我们已知的某种关系也保持到新的向量空间中。Embedding 的具体定义如下：

$$\forall x_1, x_2 \in X : x_1 \leq x_2 \Leftrightarrow F(x_1) \leq F(x_2).$$

例如，之前人们对自然语言进行计算的时候，使用的表示方式多是 one-hot 方法表示单词或者利用 tf-idf 等信息表示一句话。但是这种表示并不能表征有效的表征句意。Mikolov 等科学家使用的 word2vec，使得我们能够获得更加深层次的语义^[5]。由于 word2vec 保持了词汇意思的线性关系，使得我们利用已知的单词获得新的，意思解决的单词变成了可能。在 Mikolov 提出了 Skip-Gram 和 CBOW 等 word Embedding 的方法之后，斯坦福大学 NLP 组提出的 Glove，以及 Salesforce 提出的 ContextVec^[19] 都获得更加良好的表现。

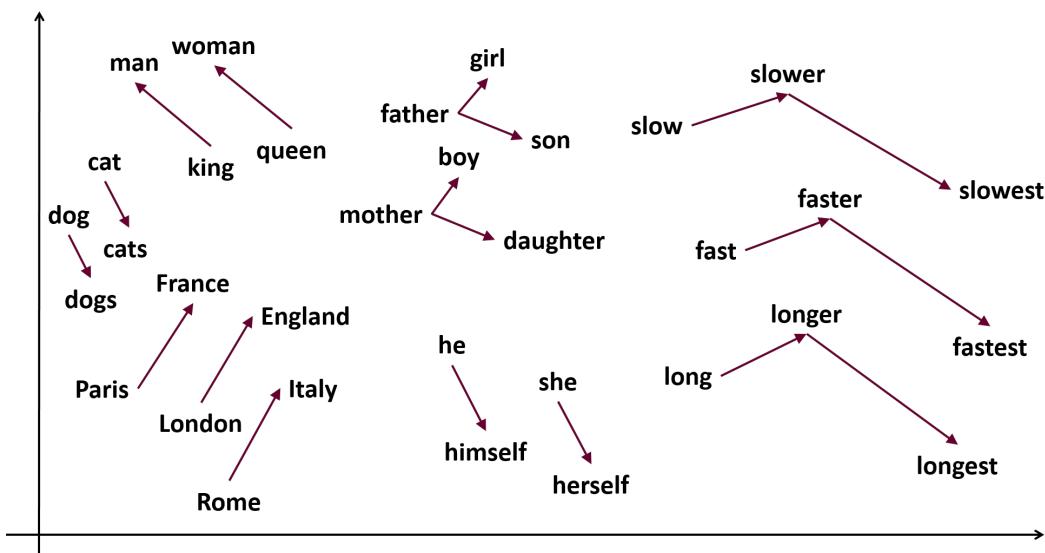


图 4.9 Word Embedding 对单词词义的线性保持

4.2.3 序列化生成：PixelRNN/CNN

$$p(x) = \prod_{i=n}^n p(x_i|x_1, \dots, x_{i-1}) \quad (4.3)$$

该模型基于一个信念网络（Belief network）(4.3), 其中, $p(x)$ 是某图像 x 的似然分布, 而 $pr(x_i|x_1, \dots, x_{i-1})$ 是某个像素的概率在给定该像素之前的所以像素分布之下的条件概率。该信念网络希望通过网络模型获得其条件概率的最大分布。要生成一个像素 pixel 需要获得上一个 pixel 的值, 通过一个 RNN 或者 CNN 网络, 实现序列化的生成, 如图 4.10 所示。

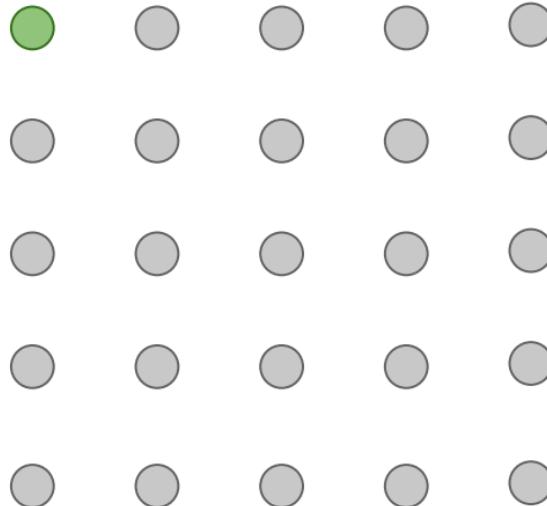


图 4.10 pixel network 通过序列化的像素生成图像

依照此模型, 可以自动生成图像, 例如图 4.11 所示, 该模型在进行 CFAIR10 和 ImageNet 数据库的训练之后, 能够随机生成与之类似的图形.



图 4.11 pixel network 通过序列化的像素生成图像

4.3 数据库资源

4.3.1 艺术品交易数据库

本文使用的数据库资源来自雅昌拍卖网艺术品交易数据库（auction.artron.net）。这是一个艺术品拍卖网站，数据库中关于艺术品以及艺术品交易的信息丰富，包括流通的艺术品的题目、图片、特征、作者、艺术品类别、交易价格、流通时间等等。我所用到的数据
库中的信息包括两类如图 4.12，一类是艺术品信息，另一类为艺术品相关的文章。

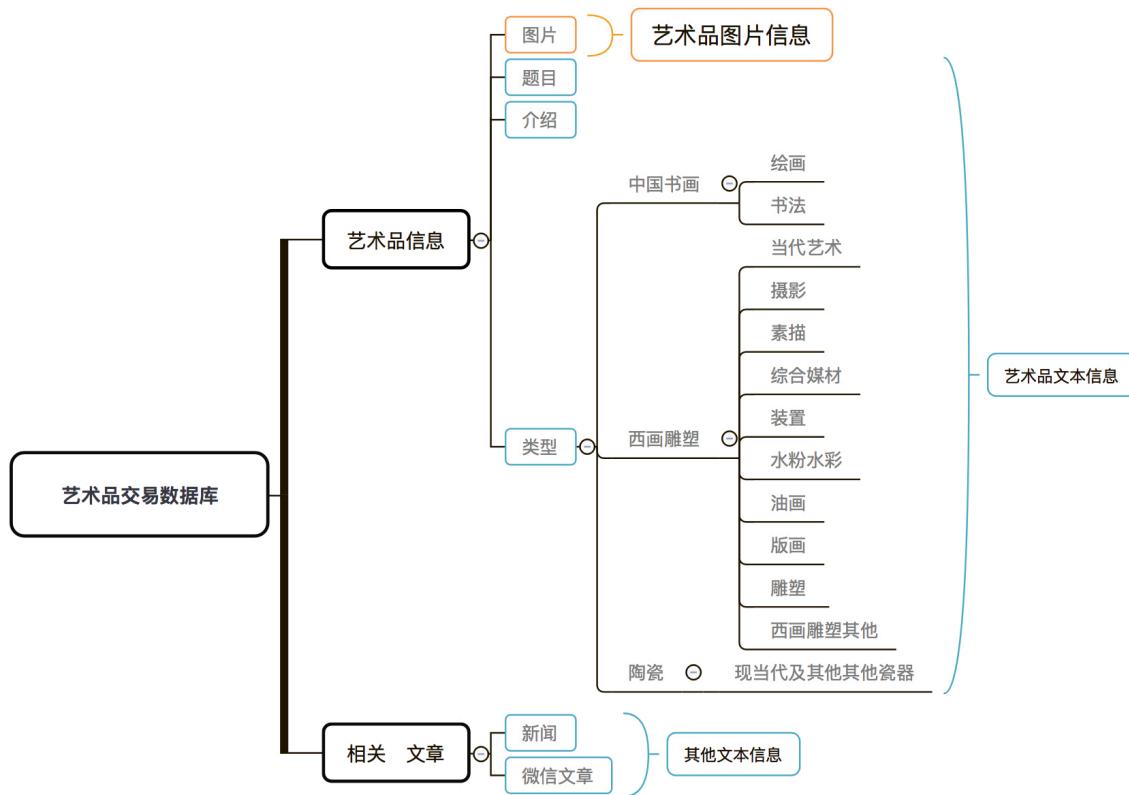


图 4.12 数据库结构

艺术品信息具体是指艺术品的题目、介绍，类型等文字描述（文本信息），以及艺术品的图片所在的 url 地址（可以下载到图像文件）。

每件艺术品都有对应的题目，艺术品的题目就是艺术品的名称，一般由艺术家命名，艺术品的名称和艺术品的内容常常是相匹配的。虽然也有一些情况艺术品的名称和内容表面上完全无关，但是艺术史上叫做“无题”的艺术品也何其之多，但是艺术品的题目作为艺术品的一部分，与艺术品的主题之间的关系紧密是无可辩驳的。

数据库中艺术品的类型分为中国书画、西画雕塑、陶瓷，在每个类型下再次细分。在

中国书画下分为：绘画、书法。西画雕塑下分为：当代艺术、摄影、素描、综合媒材、装置、水粉水彩、油画、版画、雕塑、西画雕塑其它。陶瓷下分类为：现当代及其他瓷器。

数据库中艺术品的图片包括瓷器、雕塑等三维作品的照片和绘画、书法等二维作品的扫描件，由于数据库的数据缺失，极少量的艺术品并没有对应的艺术品的图片。

数据库中艺术品的介绍是艺术品交易时所附的艺术品介绍，内容包括并不限于对作者背景的介绍、该作品创作背景的介绍、作者风格的描述、作品风格和作品内容的描述等等，由于一些艺术品本身并没有对应的介绍文本以及数据库数据的缺失，一些艺术品没有对应的介绍文本。

艺术品相关的文章包括网站的新闻板块和网站微信公众号的文章，包括艺术品市场关注的新闻事件、热点、拍卖会宣传、艺术家专访、艺术相关杂文、艺术评论等各方面的文章。

4.3.2 维基百科全文

在涉及自然语言处理的语义搜索步骤中，我还使用了维基百科全文作为语料库，这一点将在下一章“语义编码的实现”一节中详细解释。

5 基于语义识别的自动配色方案辅助设计工具

5.1 系统结构

本系统的算法实现结构如图 5.1

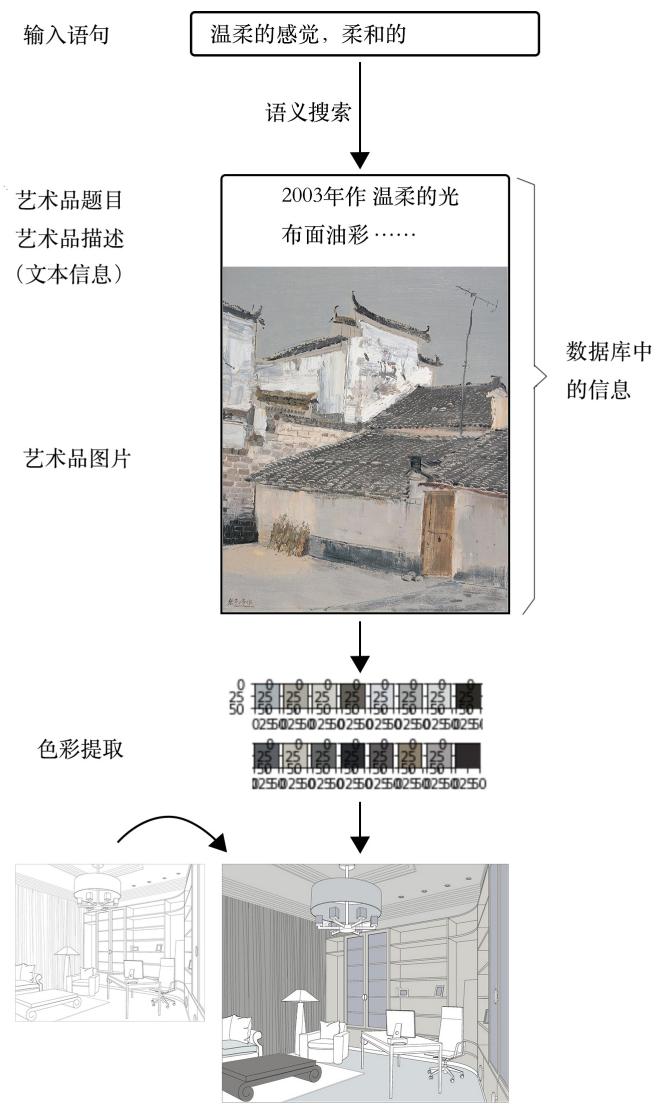


图 5.1 内部实现逻辑

本系统的整体思路为：首先，输入语句是对需求的文本描述，可以是词语，句子或不完整的句子。第二步，针对输入语句对艺术品的文本信息进行语义搜索，找出与之关联度

高的艺术品文本信息。第三步，依据艺术品数据库中艺术品的图片信息和文本信息的隐藏信息关联匹配对应，即可以找出与输入文本关联度高的图片。之后，从图片中提取色彩方案，包括绘画作品的色彩和配色比例。最后，将提取的色彩方案运用到设计效果图上预览。本系统的用户交互框架如图 5.2，对于用户来说，用户的使用逻辑是简单而清晰的，

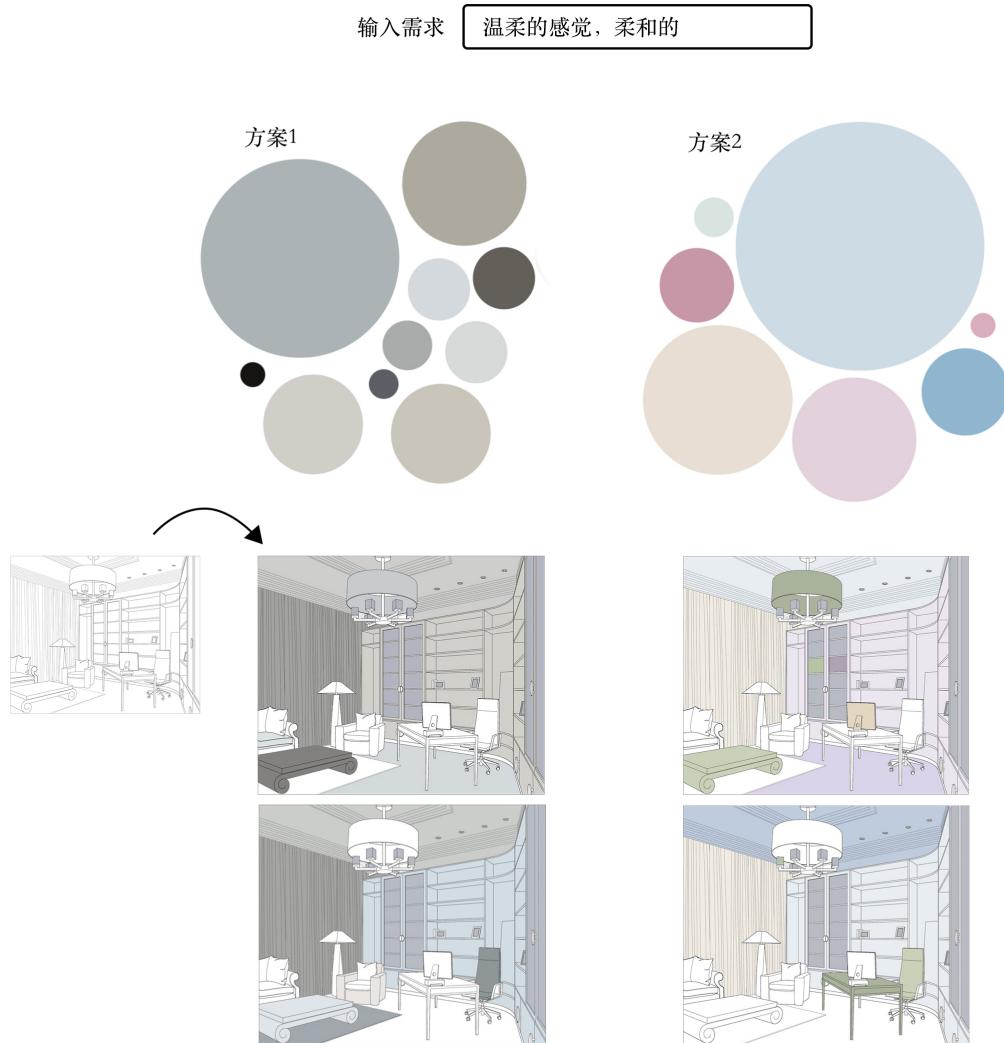


图 5.2 用户逻辑

搜索相匹配的艺术品这一步对用户来说是隐藏的，用户并不能看到相应的艺术品，而只得到色彩方案。这是为了让用户集中精力到他本来的任务中，关心配色方案而不用受到其他影响。输入设计需求的文本描述，就可以得到若干配色方案，并且可以每个配色方案都可以出现若干即时预览图，用户可以对各个配色方案有直接的感受。

5.2 语义编码的实现

5.2.1 系统结构之语义搜索

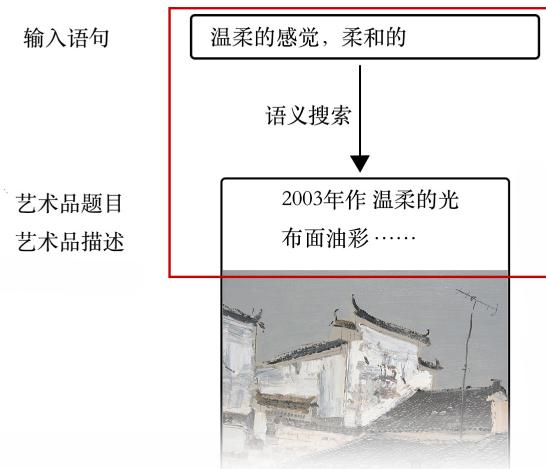


图 5.3 第一步：语义搜索

图 5.1中的第一步如上图 5.3，就是搜索数据库中与用户输入相匹配的内容。要强调的是，我们的需求并非要提取关键词作为匹配，例如搜索含有某个词的字段，而是想做到搜索到词语意义相关的信息，例如输入“沮丧”可以匹配出“伤心，难过，忧伤的人”这样的字段，又例如输入“粉嫩”可以匹配出“少女”这样的字段。这就需要理解用户输入的内容和数据库中的文本内容，在其中作出意义上而非字面上的匹配。

5.2.2 文本的向量化

要使计算机理解文本的含义，从而进行搜索，对文本进行编码就需要满足以下三个要求：1. 编码需要被计算机识别；2. 编码需要包含单词的含义信息；3. 所有的文本编码需要一定的整体同一性。

计算机理解文字的含义是困难的，即使计算机可以轻松的辨别出两句话是否一模一样，一句话中是否出现过某个单词，但是它却无法理解“苹果”和“梨”之间的关系（这对人类来说是常识），它也不能理解作为水果的“苹果”和“苹果”公司有什么不同（对于人类可以清楚的通过上下文理解这一点，不常混淆）。也就是说，文本的含义包括了语义，逻辑，上下文不同带来的多义。要用数字来表达这样复杂的信息，目前最好的解决方式是用向量来表达。

单词向量化是指将单词编码为向量形式，单词的向量化有多种方法，最简单的一种就

是 one-hot 的向量。^[20] 我们用一句话来举例子：“我喜欢吃苹果，而他喜欢吃梨子。”这个句子长度为 7 个不同的单词。我们可以用这样的向量给每个单词编码，每个单词就可以表现为一个 7 维的向量，如表 5.1，每个单词的向量为单词下的一列，就可以将一段文本的信息完整的保存进一个矩阵。

表 5.1 one-hot 向量

	我	喜欢	吃	苹果	而	他	喜欢	吃	梨子
我	1	0	0	0	0	0	0	0	0
喜欢	0	1	0	0	0	0	1	0	0
吃	0	0	1	0	0	0	0	1	0
...									

这样的方法，如果对于文本长度较大的情况，就需要很大的空间储存这些向量，比如一本上万个单词的书籍，就需要单词的维度上万，里面大多数的位置又都是 0。并且，这种方式对于不同文本来说并没有同一性，对于同一个词，不同的文本中必须有不同的词向量。最不好的一点是，这些词向量很难表示词的含义，例如近义词的向量和同类词语的向量，根本没有规律可循。

表 5.2 词频向量

句子	我	喜欢	吃	苹果	而	他	梨子
我喜欢吃苹果而他喜欢吃梨子	1	2	2	1	1	1	1
我喜欢吃苹果，梨子	1	1	1	1	0	0	1
他喜欢吃吃吃	0	1	3	0	0	1	0

为了减少向量的维度，增加矩阵的密集程度，出现了基于词频的单词向量化。以我们的例句为例“喜欢”，“吃”这两个单词出现了两次，其他单词出现了一次。词频向量就是基于此构成。如表 5.2，例如一篇文章中有表中的三句话，则每个单词的向量为单词下的一列，每句的向量为句子右侧的一行。词频向量和 one-hot 向量相比，维度有所下降，使用空间更有效率。词频也会记录一些文本的信息，不同的文本中单词的词频会有区别。然而完全从句子对单词计数，对词序信息完全忽略，同样难以提取文字本身的含义。

在此基础上的 tf-idf 的向量化不仅仅量化一个单词在一段文本中的频率，还考虑到在

所有文本中单词出现的情况，这样就可以表征出一个单词和某段文本的关系。^[21] $tfidf = tf(t, d) * idf(t, D)$ ，其中 $tf(t, d)$ (term frequency) 表征单词 t 出现在文本 d 中的频率。它可以有多种形式，比如 $tf(t, d) = f_{td}$, f_{td} 表示单词 t 出现在文本 d 中次数；或者考虑到文本到长度 $tf(t, d) = \frac{f_{td}}{n_{indocment}}$, $n_{indocment}$ 指文章总共的单词数， $tf(t, d)$ 即出现在文本中的频率；甚至也可以是布尔的，当文本 d 中出现词 t 则 $tf(t, d) = 1$ 否则 $tf(t, d) = 0$ ，等等。 idf (inverse document frequency) 定义为 $idf(t, D) = \log_{\frac{N}{n_{toccurs}}}$ ， N 表示总的文档数， $n_{toccurs}$ 表示含有单词 t 的文档数。

以之前的三句话为例：(a) 我喜欢吃苹果而他喜欢吃梨子。(b) 我喜欢吃苹果。(c) 我喜欢吃吃吃。则对于单词“我”，有： $tf("I", a) = \frac{1}{9}$, $tf("I", b) = \frac{1}{4}$ ，那么句子 b 中“我”这个词占整体比例更高。但对于所有文本， $idf("I", D) = \log_3^3 = 0$ ，每个句子都出现了“我”这个词，所以 $tfidf("I", a) = tfidf("I", b) = 0$ 。处理文档大量时，对于一些非常常见的词，比如“那个”“这个”这样的词，可以认为它们没什么意义，就会出现这种情况。而比如对于单词“苹果”， $tf("apple", a) = \frac{1}{9}$, $tf("apple", b) = \frac{1}{4}$, $idf("apple", D) = \log_2^3 = 1.585$ ，就会出现 $tfidf("apple", a) = 0.045$, $tfidf("apple", b) = 0.1$ 。我以这三个句子计算了单词“我”和单词“苹果”的向量如表 5.3：

表 5.3 tfidf 向量

句子	我喜欢吃苹果而他喜欢吃梨子	我喜欢吃苹果	我喜欢吃吃吃
我	0	0	0
苹果	0.045	0.1	0
...			

只要有足够量的文本，tf-idf 向量可以在文本间做比较，然而还有一个问题没有解决，那就是它仍然不包含单词语义的信息。在构建 tf-idf 向量的过程中，只用到了文本中单词的数量，完全没有使用单词的顺序信息，而文本中单词顺序所含有的信息是很重要的，如何将这些信息提取出来呢。

如果我们考虑上下文构建向量。对于句子“我喜欢吃苹果而他喜欢吃梨子”。将每个词前后对它有影响的单词数量，称为上下文窗口大小 (context window)。比如，对于单词“喜欢”，在窗口大小为 3 时，单词“吃”在它周围出现的次数共 2 次，“我”、“苹果”、“他”、“梨子”、“而”各一次。如表 5.4：

在此基础上，如果我们采用一点统计学和数学的知识，使用神经网络的算法，就可以

表 5.4 上下文向量

句子	我	喜欢	吃	苹果	而	他	梨子
我	0	1	1	1	0	0	0
喜欢	1	0	2	2	2	1	1
吃	1	2	0	1	2	2	1
苹果	1	2	1	0	1	1	0
而	0	2	2	1	0	1	1
他	0	1	2	1	1	0	1
梨子	0	1	1	0	1	1	0

从文本中提取出更多的信息。word2vect 就是这样一种算法，虽然前面所述都是单词向量化方法，但 word2vect 被 Mitolov 等引入到自然语言领域之后，向量前所未有的拥有了密集的信息。对于一个出现在文本中的单词，在知道周围的词语的情况下，就能够预测出这个单词是某个单词的可能性。^[22] 比如“我喜欢吃 (a)”，a 是单词“苹果”的可能性就比 a 是单词“因为”的可能性更高。在非常大量文本的情况下，“苹果”和“梨子”常常出现在相似的环境下，算法就将计算得到它们的向量是相似的，从而被归为相似的词语。^[5]

word2vect 由 CBOW(Continuous bag of words) 和 Skip-gram model 两个部分共同组成。^[19]CBOW 的工作方式为以全文训练神经网络，以求输入神经网络一个单词，输出他周围的一个单词。以 one-hot 向量编码的句子“我喜欢吃苹果而他喜欢吃梨子”作为例子，则全文的矩阵如表 5.1，对于其中“喜欢”这个单词，则具体为假如输入“喜欢”的向量，则期望输出“吃”的向量。^[23]

表 5.5 输入

喜欢	0	1	0	0	0	0	1	0	0
----	---	---	---	---	---	---	---	---	---

表 5.6 输出

吃	0	0	1	0	0	0	0	1	0
---	---	---	---	---	---	---	---	---	---

CBOW 算法结构如图 5.4，假如输入“喜欢”的向量 $(0, 1, 0, 0, 0, 0, 0)$ ，则期望输出“吃”的向量 $(0, 0, 1, 0, 0, 0, 0)$ 。输入为 $1 \times V$ 的矩阵（在这里 V 为 7），隐藏—输入层为

$V \times N$ 的矩阵，隐藏一输出层为 $N \times V$ 。N 就是我们选择的最终得到的词向量的维度。输入为 $V \times 1$ 的矩阵，这个矩阵的每个位置对应的是 one-hot 编码中该位置单词出现的概率。在层与层之间没有激活函数，而是线性函数。使用计算出的输出值和期望的输出值之间的误差作为 loss 函数，反向传播适应得到维度为 N 的权重 \vec{W} 。与此同时，单词“吃”对应的输入还有“苹果”或“梨”，或者在窗口大小更大时的其他单词。多个权重平均的向量就是单词“吃”的向量。

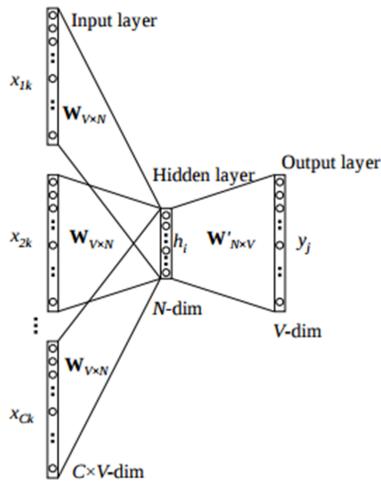


图 5.4 CBOW model

而 Skip-gram model 与之相似，只是输入的单词“喜欢”对应的输出有“我”，“他”，或者在窗口大小更大时的其他单词。对于多个输出单词就有多个误差向量，多个误差向量构成总的误差向量，反向传播适应得到维度为 N 的权重 \vec{W} ，即为单词“喜欢”的向量。

这种充分采用文字顺序的算法拥有很多很好的特性。比如，向量距离接近的单词意义也接近。更神奇的是，在训练文本多且全面的情况下，代表“女人”的向量和代表“妈妈”的向量之间的距离，与代表“男人”的向量和代表“爸爸”的向量之间的距离相等，即“女人” - “妈妈” = “男人” - “爸爸”；“苹果”这个单词的向量，同时和“梨子”的向量以及“IBM”的向量距离相等；由于是从文字的顺序提取信息，所以对于各种语言该模型都适用，也说明了提取的信息是语义上的而不仅仅是文本上的。

5.2.3 本系统的单词向量化

在介绍本系统的单词向量化之前，先要介绍我所使用的训练数据，训练词向量的数据来自两个来源，1. 维基百科全文；2. 艺术品交易数据库中的文本信息。其中艺术品交易数

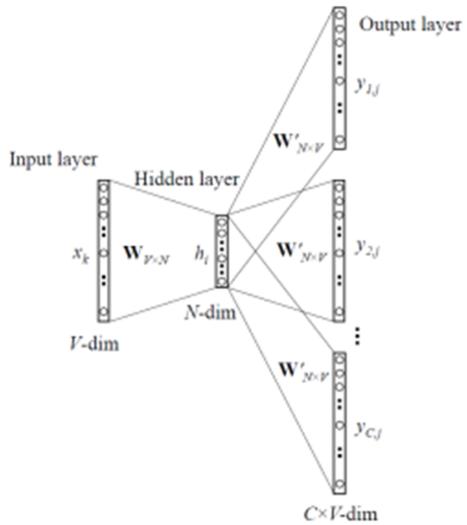


图 5.5 Skip-gram model

据库中的文本信息包含以下几种：1. 艺术品交易数据库中每个艺术品的文本信息（包括艺术品的题目、描述和介绍、类型等）；2. 艺术品交易数据库中的所有文章（包括艺术品相关的新闻报道、微信文章、评论文章、艺术家专访等相关文章）。艺术品交易数据库中的文本信息是对艺术这一话题的加强训练。而采用维基百科的原因是，如果没有维基百科数据库，则当输入艺术品数据库中不存在的日常用语时就无法理解。为了系统能够响应人自然场景下的语言，采用一个普适的文本数据是非常必要的。对文本进行繁简体转化、中文分词之后，就可以作为训练使用的文本。

这样的庞大数据带来了更好的文本语义理解，但是也带来了计算量的问题。使用 CBOW 模型对单词进行 word2vec 向量化的过程中，如果文本有 10000 个单词，那么隐藏—输入层和隐藏—输出层就会是 $10000 \times N$ 和 $N \times 10000$ 的矩阵，需要反向传播调整和适应其中的每个权重计算量将会非常大，而维基百科和艺术品交易数据库带来的词量更是远大于此，对于如此大的计算量需要采用哈夫曼树进行优化，从而可以利用 Hierarchical Softmax 加速。^[24]

哈夫曼树 (Huffman Tree) 又叫最优二叉树，是给定 n 个权值作为 n 个叶子结点，构造一棵二叉树，使得带权路径长度达到最小。哈夫曼树是带权路径长度最短的树，权值越大的结点离根越近。Skip-gram model 的输出层使用哈夫曼树，总的文本中的每个词对应哈夫曼树中的一个叶子节点（每个节点对应一个最后得到的 word2vec 词向量），每个非叶子节点定义一个向量（起到辅助的作用）。对于词典中的任意单词，哈夫曼树中必然存在唯一一条从根节点到该单词对应的路径，这样对于每一个单词，输出层只需要调整哈夫曼树路径上的向量。如果对于有 V 个单词的文本来说，本来隐藏—输入层和隐藏—输出层需要

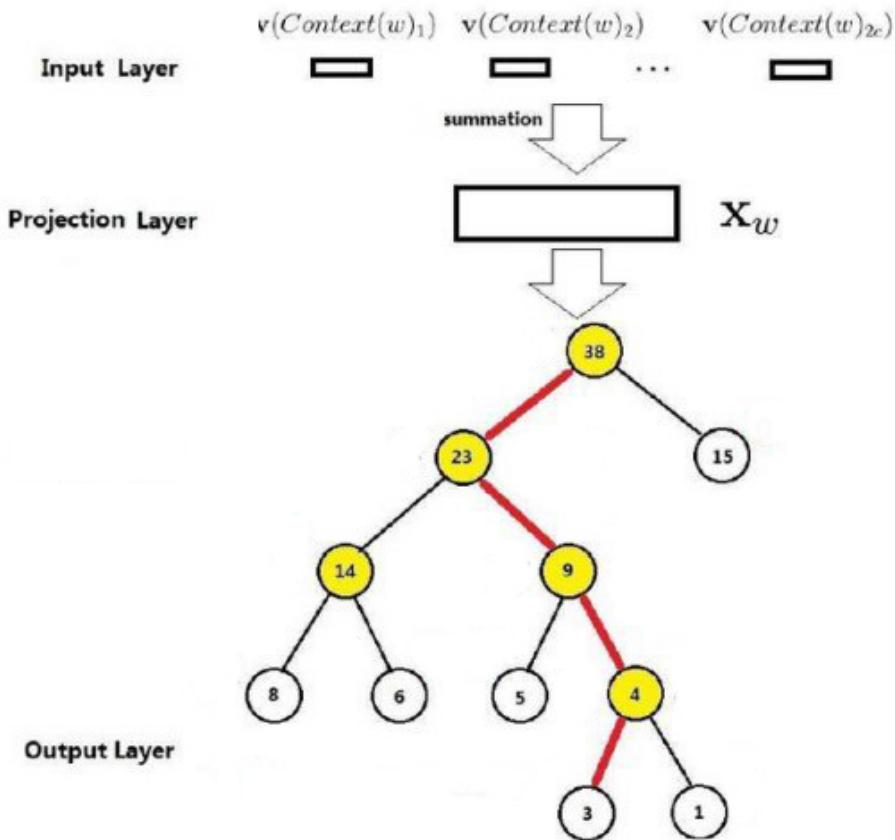


图 5.6 使用哈夫曼树的 CBOW model

反向传播调整 $V \times N$ 和 $N \times V$ 个值，现在在哈夫曼树上每次二分产生一个概率，只需要经过 $\log(V)$ 次运算就得到了单词 a 出现的概率 $p(a|context(a))$ 。

与此同时，本系统的 word2vec 还使用了以下的策略，把常见的词组作为一个新的单词增加到词典中，比如“苹果手机”；少采样常见的词，比如“而”“就”这些出现频率非常高的词其实它的表达的意思很少；采用 Negative Sampling 策略，通过负采样，每次修改其中一小部分 weight，而不是全部。这些策略减少了训练的计算量，而且提升了最后 word vector 的质量。^{[25][5][26]}

CBOW 模型实际上只考虑了单词有哪些上下文，而并未考虑单词的顺序，比如“我喜欢苹果”和“苹果喜欢我”在 CBOW 模型中的特征是一样的。加入了 N-gram 特征将“我-喜欢”和“喜欢-我”也加入单词的特征，也使用 n-gram 特征来将局部词序考虑在内，比如“女-朋友”“陶-瓷”。可以得到表达意义更准确的词向量。

使用上述的文本和模型训练维度为 300 的词向量，词典中包含全部文本中出现的单词。使用此词典构成输入语句和艺术品信息的语句向量。

5.2.4 本系统的句向量建模

在本系统中，需要语义搜索的内容并不是单词。输入语句是使用者输入的，不确定的语句，可能是一个或数个词语、短语、完整的或不完整的句子，总的说来，是一个或多个单词的序列，之后我们将其称为“输入语句”，即和我们平时口语采用的语言是相似的，不拘结构但求表达需求。而数据库中每个艺术品的文本信息由艺术品的题目、介绍，类型三个部分组成，短至一个词语（在文本信息只有一个单词的题目的时候），长至数段话（在文本信息包括大量描述和介绍的时候），总的来说也是一个或多个单词的序列，之后我们将其称为“艺术品语句”。需要从使用者输入的语句定位到某一件艺术品，那么就需要输入语句的含义，和某件艺术品语句的含义相匹配。

要提取一段语言的意义，就需要将这多个表达单词意义的向量转化为一个转化语句意思的向量。在这个将多个词向量加和为语句向量的过程，需要考虑这三点：

1. 词汇的含义，语句由词汇构成，语句中词汇的意思对语句的含义是决定性的，比如一张油画的名字叫做“忧伤的河流”说的就是忧伤，不是喜悦，是河流，不是山川。

2. 词汇的频率，在语句中词汇出现的频率会对语句的意义产生影响，这样的影响并非线性的，有时，频率高的词表达的含义反而少，比如“作者非常非常思念家乡”，中出现了两个“非常”，表达比“非常”更深的程度，但是句子主要要表达的意思仍然是“作者思念家乡”，这句话和“作者很思念家乡”的含义相似，而和“我非常非常喜欢吃苹果”是不同的；而有的时候频率高的单词直接是这段文本中的重要主题和中心。大概率的，总体文本中频率高的词带来的含义有限，这些词比如“有”“这”“那”，并不影响语句的意义，而总体文本中概率一般然而特定文本中频率高的词，比如在某个作品的描述中常常出现“热烈”“欢快”，往往对这一文本的语义的影响很大。

3. 词汇的位置。词汇的位置有时也代表了词汇含义的重要性。语言是结构性的序列，在结构的某一位置上的词语往往代表更重要的含义。对于词语构成语句的过程，这个位置的重要性很灵活。比如对于语句“作者非常思念家乡”，如果这句话处在一个环境中，上一句话是“谁思念家乡了？”，那么最重要的位置就是主语的“作者”，而如果上一句话是“作者思念什么”，那么最重要的位置就是宾语“家乡”。另外，如果考虑句子构成完整段落的过程，段落的第一句话和最后一句话，由于可能是总起和总结的句子，所以格外重要，这也是考察位置得到的信息。

对单词向量直接加和求得输入语句的向量是最简单的方法，即 $\vec{S} = \sum_{i=0}^n w_i$ 。其中 \vec{S} 指据向量，句子中有 n 个单词， w_i 指第 i 个单词的词向量。^[26] 这种方法简单快捷，但是它只考虑了语句中词汇的含义，并没有考虑到词汇的频率和词汇的位置对句向量的影响。会

导致语句中重复出现的词语对语句影响严重的情况，另外也会让一些本来并不代表意义的单词占比较大，比如对于语句“我要一个柔和一点的色彩方案”，这句话中“柔和”其实是最关注的一个词，但是由于干扰信息太多，如果直接加和，这句话的向量会和“我要一个欢快一点的色彩方案”更近，而与单词“柔和”并不相近。更突出的一个例子是，很多作品的题目中会出现作品的媒介比如“儿童布面油画”，“温柔的风油画”等等，其中，单词“布面”、“油画”等，它们在整体文本中大量出现，但对于每个作品已经有了明确的分类信息，文本信息中更多的是为了匹配作品的内容和色彩风格，它们就并非与作品色彩风格有关的单词。

在这种情况下，我采用了加权加和的方法，即 $\vec{S} = \sum_{i=0}^n \vec{w}_i * weight(w_i)$ ，这样可以通过权重过滤掉干扰信息。^{[19][26]} 我使用单词在总体文本中的词频作为权重分析的来源，在总体文本中出现频率越高的词就说明它是不重要的（如果一个单词出现在每一句话中，那么这个单词当对分辨每句话的意义并没有价值）。如此，则权重 $weight(w_i) = \frac{1}{frequency(w_i)}$ ，句向量的计算方式则为：

$$\vec{S} = normalization(\sum_{i=0}^n \frac{\vec{w}_i}{frequency(w_i)})$$

以词频的倒数为权重，对语句中的单词加权求和，再经过归一化，得到语句向量。对于输入语句如此即可求出其句向量，而对于艺术品语句则还要加上一个内容分类权重，由于艺术品语句的内容分为题目、介绍和类型，三者表达信息的权重是不同的，其中题目与艺术品本身的含义和特点最为联系紧密，而介绍有时详细描述了作品的风格和内容，有时却是和作品本身的内容关系不大的作者背景介绍等，在这样不稳定的情况下，我为系统设置了可调节的权值。

$$\vec{S} = normalization(\sum_{i=0}^n \frac{\vec{w}_i}{frequency(w_i)} \times weight_{texttype})$$

5.2.5 搜索算法

每一件艺术品都有独一无二的文本信息，也就是对应唯一的艺术品语句向量。计算出每件艺术品的艺术品语句向量，将每件艺术品的向量储存在艺术品“词典”，输入语句向量在词典中搜索，找到最匹配的艺术品语句向量，即可定位到最匹配的艺术品。取最符合得到前 10 件艺术品作为备选，可以包容更多的使用情况，这些情况的具体细节在下一小节“色彩提取”中将会详细描述。^[27]

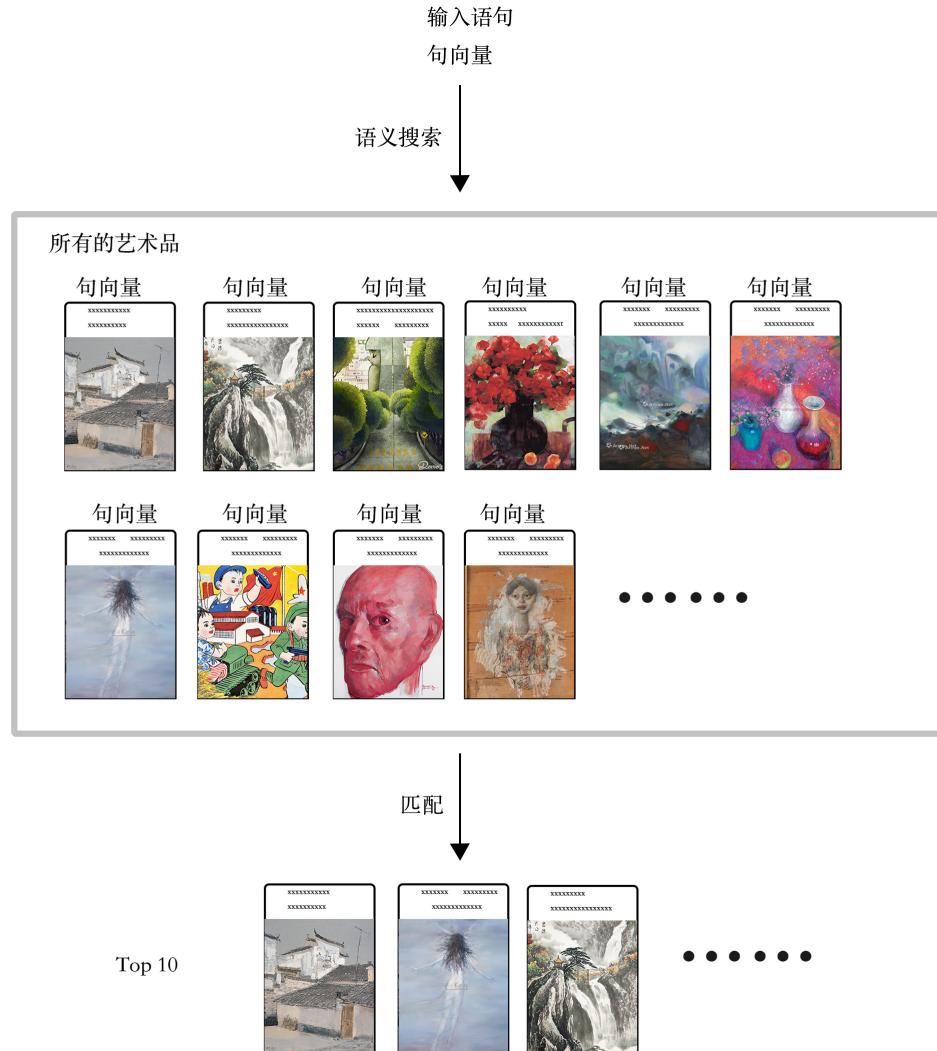


图 5.7 匹配范围

表征向量的相似度使用余弦相似度。词向量和句向量都经过归一化，向量指向同一方向即为匹配，也就是说向量之间的距离大小即为夹角的大小。如图 5.8当两个向量指向相同，余弦为 1，两个向量正交，余弦为 0，指向相反余弦则为-1。两个向量夹角的余弦值可以确定两个向量是否大致指向相同的方向，故而考量输入向量与艺术品向量的距离使用余弦相似度。^[28]

余弦相似度不仅对于二维向量适用，对于任何空间中度向量都适用。

$$\text{similarity} = \cos(\theta) = \frac{\vec{A} \cdot \vec{B}}{\|A\| \cdot \|B\|} = \frac{\sum_{i=0}^n A_i \times B_i}{\sqrt{\sum A_i^2} \times \sqrt{\sum B_i^2}}$$

本系统中的词向量和句向量维度为三百，对于归一化之后的句向量来说，所有的句向量的模都是 1，即 $\|A\| = \|B\| = 1$ ，则

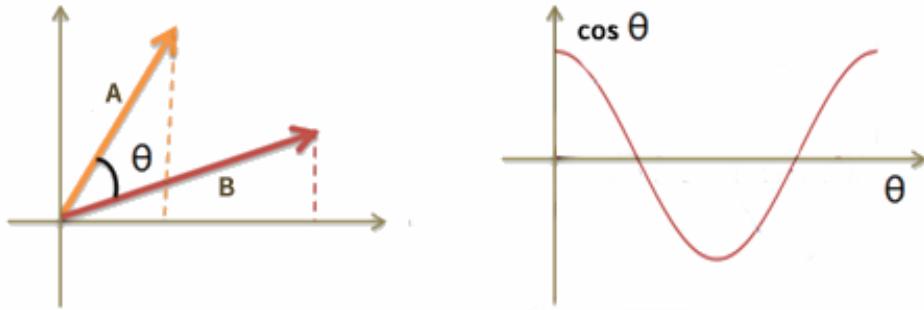


图 5.8 余弦相似度

$$\text{similarity} = \cos(\theta) = \sum_{i=0}^{300} A_i \times B_i = \mathbf{A} \times \mathbf{B}^T$$

对输入语句的句向量与每个艺术品语句的句向量做矩阵乘法，结果最接近 1 的就是最匹配的艺术品语句向量，此艺术品的语句和输入语句在语义上是最接近的，由此决定此艺术品就是与输入的语句最匹配的艺术品。^[29] 接下来就从这个艺术品来提取特征。

在本系统中还为了特定的需求进行一些搜索优化。由于本系统最终是为了从艺术品提取色彩，所以降低色彩很少的作品被搜索到的概率。比如对于书法作品，一般并没有色彩可以用于提取。搜索概率高的艺术类型设置为摄影、综合媒材、水粉水彩、油画、版画，以绘画作品为主。

由于艺术品和文字表达的模糊性，有时文字信息匹配度最高度艺术品事实上并不完全匹配，这种情况的出现往往是由于艺术家为了某些内涵，让艺术品的命名与艺术品内容表面上并不相关，或者艺术品的介绍过于偏重作者的家世、背景，与艺术品的内容关联度很低。为了解决这样的问题，在匹配的过程中，输入语句后，本系统取相似度前 10 的艺术品作为候选，随机选定其中一件作品提取色彩，当用户不满意当前匹配的色彩可以通过重复点击进行切换，用户也可以选择进入页面进行选择。

5.3 色彩提取

5.3.1 系统结构之色彩提取

由于艺术品的图片和文本语义信息的关联，艺术品图像风格的偏序特性保持到语义描述的偏序特性中，当输入语句和艺术品语句匹配成功后，自然就得到了与输入语句匹配的艺术品的图像。从艺术品图像中提取色彩特征，就可以得到相匹配的色彩方案。

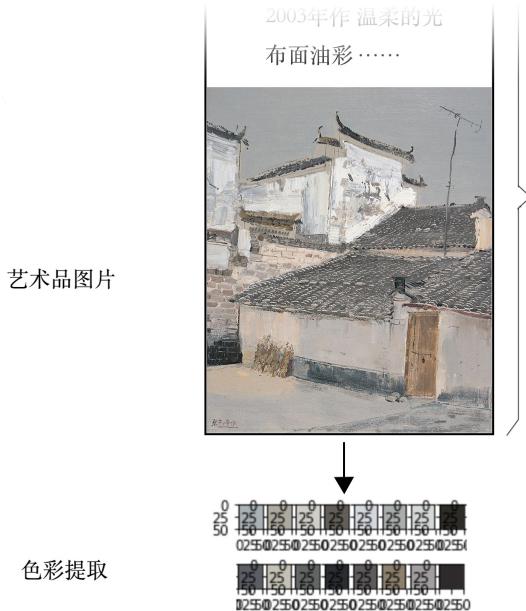


图 5.9 第二步：提取色彩

这一步输出的色彩方案使用 RGB 色彩模式，每个颜色包括 Red-Green-Blue 三个通道的值和该颜色本来在艺术品中占的比例（面积比例）。从艺术品图片中占比例最大的颜色开始，可以自定义提取颜色的个数，在本文的测试中取颜色的个数为 16。

对于用户来说，输入文本并搜索将会出现一组色彩方案，这个色彩方案来自文本相似度与输入文本排前 10 的艺术品之一，用户可以通过不改变输入文本，再次搜索，从相似度前 10 的艺术品中另外随机选择一件，也可以直接打开相似度前 10 的所有色彩方案手动选择。如图 5.10虚线框中为文本相似度与输入文本排前 10 的艺术品，这些内容不展示给用户，用户仅仅关注色彩方案，通过重新搜索，或进入 10 个色彩方案可以选择。

5.3.2 色彩模式

颜色是不同波长的可见光进入人类的眼睛，由视锥细胞感受到不同波长的比例传到大脑所产生的感受。可见光的波长变化 3-5nm，人类就可以分辨为不同的颜色，要对色彩进行数字编码，可以使用多种方法。^[30]

RGB 色彩模式指 Red-Green-Blue 三种色光的混合。人类有三种视锥细胞，分别对红、绿、蓝三种色光最为敏感，所以人类的眼睛可以看到的色彩都是由这三种颜色的光经不同比例构成的。光的混合在牛顿时代就被发现了，使用红、绿、蓝三种色适当混合，可以引起光谱上所有任何颜色的感觉，三原色混合原理广泛地应用于照相、显示技术中。数字化常见的颜色标准也使用红、绿、蓝三个通道混合，称为 RGB 色彩模式。通常情况下，RGB



图 5.10 备选方案

各有 256 级亮度，用数字表示为从 0、1、2... 直到 255。比如 (255,0,0) 指红色，(255,255,0) 则是黄色，(0,0,0) 是三种色光的亮度都为 0，即黑色，(255,255,255) 是三种色光的亮度都为最高，混合起来为白色。在计算机领域 RGB 色彩模式是很常用的，需要注意的是，在编码时有时会使用 BGR 模式，意思并没有区别，但是数字的顺序是反过来的，RGB(255,0,0) 为红色而 BGR(255,0,0) 为蓝色。

常用的色彩模式还有 CMYK，多用于印刷，CMYK 同样将颜色分为数个通道混合，但是通道为 Cyan (青色) -Magenta (品红) -Yellow (黄色) -black (黑色) 四个，这四种颜色的染料可以调配出几乎所有颜色。CMYK 和 RGB 色彩模式可以相互转化，转化的方式如下。

```
def rgb_to_cmyk(r, g, b):
    cmyk_scale = 100
    if (r == 0) and (g == 0) and (b == 0):
        # black
        return 0, 0, 0, cmyk_scale
    # rgb [0, 255] -> cmy [0, 1]
    c = 1 - r / 255.
    m = 1 - g / 255.
```

```

y = 1 - b / 255.

# extract out k [0, 1]
min_cmy = min(c, m, y)
c = (c - min_cmy) / (1 - min_cmy)
m = (m - min_cmy) / (1 - min_cmy)
y = (y - min_cmy) / (1 - min_cmy)
k = min_cmy

# rescale to the range [0, cmyk_scale]
return c*cmyk_scale, m*cmyk_scale, y*cmyk_scale, k*cmyk_scale

```

关于图像等 OpenCV、PIL 等库都能够进行转化。在本系统中，所有的色彩都按照 RGB 模式编码，不同的色彩模式需要进行转化。

5.3.3 色彩提取

色彩提取从图片中占比例最大的颜色开始，可以自定义提取颜色的个数 N_{color} ，输出的每个颜色包括 Red-Green-Blue 三个通道的值和该颜色本来在艺术品中占的比例（面积比例）。

色彩提取的算法采用色彩梯度算法。使用 RGB 色彩模式。首先为了减小计算量通过压缩图像生成一个略缩图。将 R-G-B 各分为步长为 step 的 n 个等级，即总共有 n^3 个颜色等级。对每个颜色等级的像素点个数进行计数，定义色彩 c 的分数为 $score(c) = (saturation(c) + 0.1) \times count(c)$ ，其中 $saturation(c)$ 为色彩的饱和度， $count(c)$ 则为颜色为 x 的像素点个数。非常接近白色的颜色（RGB 三个参数均大于 200 的颜色）和非常接近黑色的颜色（三个参数均小于 50 的颜色）都不参加评分。面积越大、色彩饱和度越高，则色彩得分越高。得分越高的颜色就是图片中主要的颜色，提取前 N_{color} 个得分最高的颜色， $count(c)$ 就代表这个颜色在艺术品中占的比例。

```

def get_dominant_color(image, beside_color, step):

    # 颜色模式转换，以便输出rgb颜色值
    image = image.convert('RGBA')

    # 生成缩略图，减少计算量
    image.thumbnail((200, 200))

    max_score = 0

```

```

dominant_color = 0
proportion = 0

for count, (r, g, b, a) in image.getcolors(image.size[0] * image.size[1]):
    # 跳过透明色
    if a == 0:
        continue
    # 跳过白色
    if ((r>180) & (g>180) & (b>180)):
        continue
    # 跳过黑色
    if ((r<80) & (g<80) & (b<80)):
        continue

    if ((beside_color[0]-step<r<beside_color[0]+step) & (beside_color[1]-step<g<
        beside_color[1]+step) & (beside_color[2]-step<b<beside_color[2]+step)):
        continue

    saturation = colorsys.rgb_to_hsv(r / 255.0, g / 255.0, b / 255.0)[1]
    y = min(abs(r * 2104 + g * 4130 + b * 802 + 4096 + 131072) >> 13, 235)
    y = (y - 16.0) / (235 - 16)

    # 忽略高亮色
    if y > 0.9:
        continue

    score = (saturation + 0.1) * count

    if score > max_score:
        max_score = score
        dominant_color = (r, g, b)
        proportion = count

return dominant_color, proportion

```

这样的方式略过黑色和白色（或者说几乎是黑色和白色的颜色），面积越大、色彩饱和度越高，色彩越排在前面，可以完全忠实于艺术品图片提取色彩。

5.4 上色算法

5.4.1 系统结构之自动上色

从艺术品中提取色彩之后，色彩方案可以自动在效果图线稿上预览。上色预览主要是为了让设计师直观的到色彩方案的风格和潜力，为了满足这个目标有 3 点需求：1. 快速实时；2. 色彩准确和比例较准确；3. 效果图清晰明确。

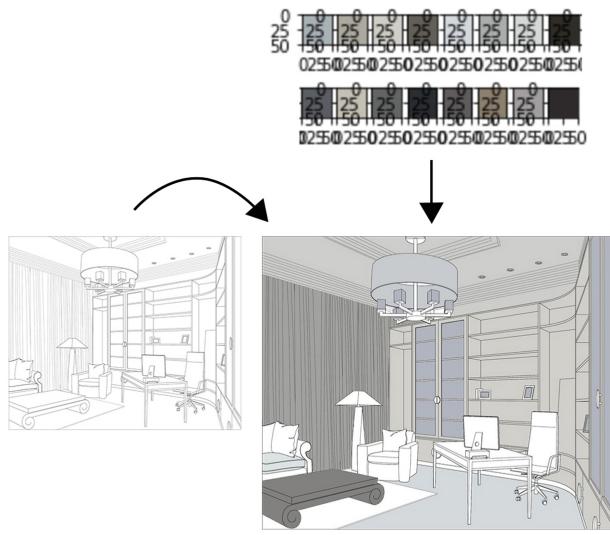


图 5.11 第三步：自动上色

系统准备了室内设计，服装设计各 1 套效果图线稿预览，之后的测试发现，对于上色区域占图片比例较小并且色彩有限的服装、产品效果图本系统也可以使用，但系统效果更适合室内设计效果图。每次生成的预览效果图由于有随机性的参与都不可复制，但可以使用同一色彩方案生成不完全一致但风格同一的预览图。

5.4.2 泛洪填充算法

泛洪填充算法 (Flood Fill Algorithm) 又称漫水填充算法，是在很多图形绘制软件中常用的填充算法。^[31] 泛洪填充算法可以为图像的封闭区域填充颜色，成完整的填充需要的条件有两点：1. 该区域有封闭的边界；2. 该区域的像素点色彩相同。如图 5.12。

泛洪填充算法的原理是，要为某个区域填充色彩 a ，在这个区域中取一个像素点，这个点被称为种子点，种子点原始的颜色为 x ，将种子点的颜色从 x 改为 a ，对于所有相邻且颜色为 x 的点也都染色为 a ，染色后的点相邻且颜色为 x 的点也都染色为 a ，一直传播下去，直到这个区域内所有的点都被填充完为止。如图 5.13。



图 5.12 泛洪填充

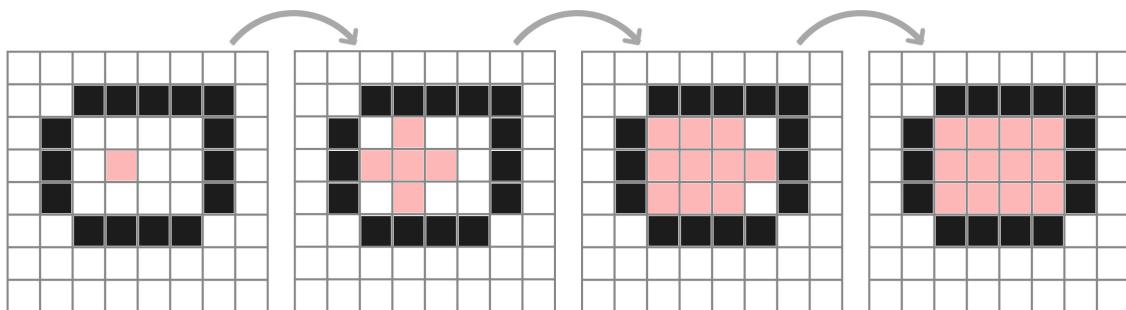


图 5.13 填充过程

实现这样的算法只需要最基础的递归算法或堆栈算法。即

```
def floodfill(x, y, oldColor, newColor):
    # assume surface is a 2D image and surface[x][y] is the color at x, y.
    if surface[x][y] != oldColor: return # the base case
    surface[x][y] = newColor
```

```

floodfill(x + 1, y, oldColor, newColor) # right
floodfill(x - 1, y, oldColor, newColor) # left
floodfill(x, y + 1, oldColor, newColor) # down
floodfill(x, y - 1, oldColor, newColor) # up

```

在实际的使用中，这个简单的算法还需要一些改进。复制制作一个遮罩层，以防图片被直接改变导致返回结果不正常。加入容差的概念，即当染色点原始的颜色为最大负差值不超过 RGB(20,20,20)，最大正差值不超过 RGB(50,50,50)，则可以扩散。

在本系统中，待上色的效果图为室内设计效果图，通过随机选点来为设计效果图上色。每次上色的程序如下：随机选取效果图中像素点，如果这个点为白色则填充当前色彩，同时替换当前色彩为色彩列表下一个色彩；如果这个点非白色，则再次随机选取效果图中的像素点。使用的色彩列表为从艺术品提取的色彩方案，色彩按艺术品中色彩的比例从高到低排序，循环排列，循环多次，直到图片上色完整。

5.4.3 一些问题

漏掉到小面积——随机游走：使用泛洪填充算法可以非常准确工整的为室内设计效果图线稿上色，完全的随机取点会使某些较小的面积被选中的概率太小，所以无法上色的情况严重，如图 5.14(a) 图像中很多地方没有成功上色。

为了解决这个问题，PaintsChainer 所采用的神经网络算法的可以完全的给整张图片上色，如图 5.14(b) 但是对于即时预览色彩方案的效果图来说，不工整的上色、混色、神经网络生成的错误阴影导致预览色彩改变并不适合本系统的需要。生成图像的数十秒延迟也是这个算法不合适的原因。

在原来的方法上面可做一些改进使用随机游走的方式，完全的随机取点会使某些较小的面积被选中的概率太小，无法上色。使用随机游走的方式改进上色算法可以更少出现漏掉面积较小的区域的情况。随机游走的方式倾向于一个区间接一个区间的上色，如图 5.14(c)，循环次数足够多的情况下可以起到更好的效果。

色彩连续性——图片预处理：

使用洪填充算法会出现这样的问题，例如区域 1 和区域 2 都是天花板，一般而言是同种颜色，但是现在的填充算法一点也不懂这件事。

所以将这个效果图进行一点预处理，这样可以基本保证在每个完整的物体会填充同一种色彩。预处理的方式是通过标注的方式将同一物体的信息提取出来，删去物体内部的线

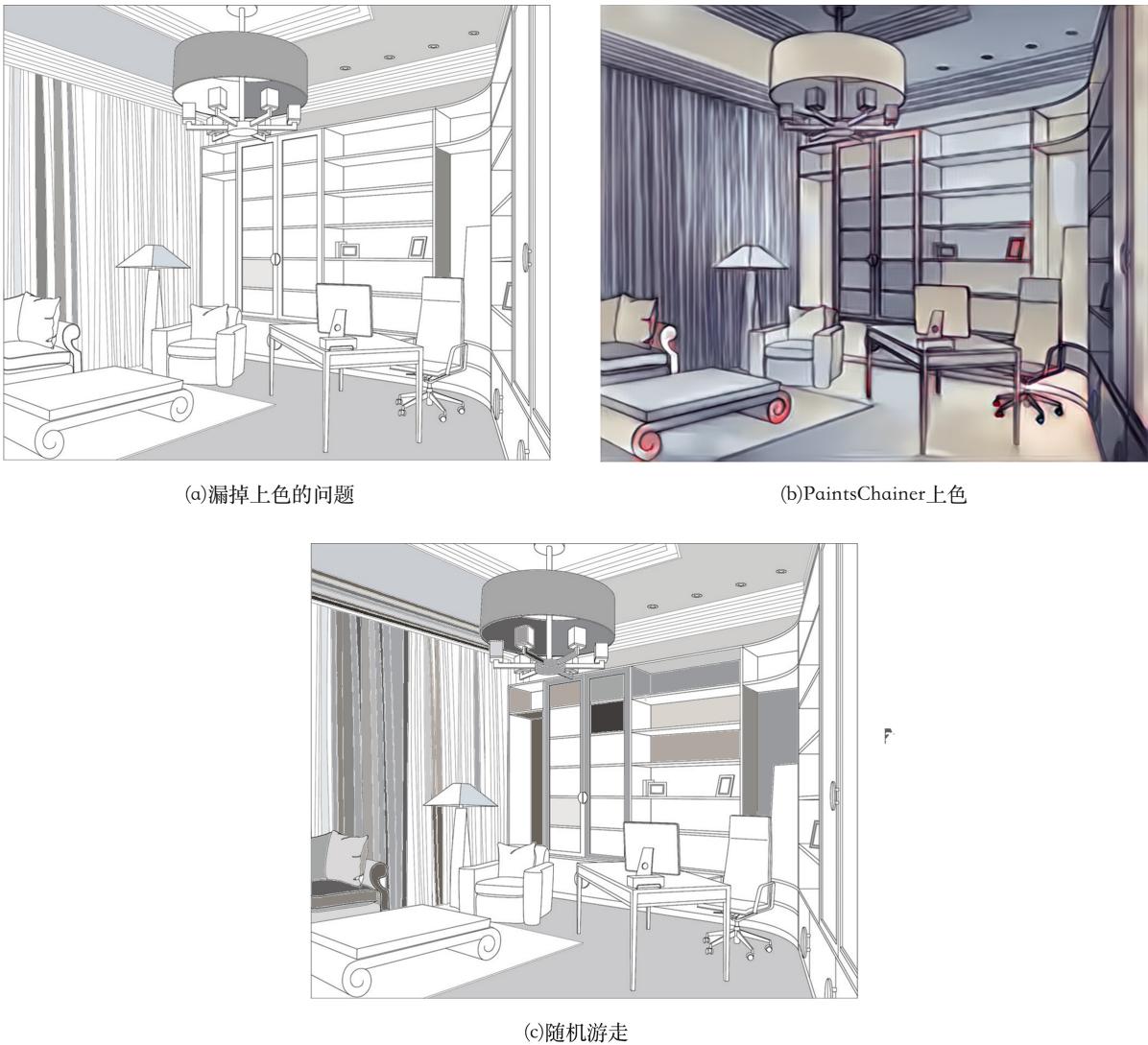


图 5.14 漏掉上色的问题

条，保留物体的外轮廓。与此同时我们还可以做一些其他的标注，比如将透明的物体的色彩标注为某个灰色。之后再进行填充，就可以将同一个物体填充一个颜色。最后，再通过图层叠加的方式将效果图的原线稿叠加在上色稿上，得到最终的效果图。

5.5 本章小结

本系统的整体思路为：1. 输入对需求的文本描述；2. 针对输入语句对艺术品的文本信息进行语义搜索，找出与之关联度高的艺术品文本信息；3 一句艺术品数据库中艺术品的图片信息和文本信息的隐藏存词关联匹配对应，找出与输入文本关联度高的图片；4. 从图片中提取色彩方案，包括绘画作品的色彩和配色比例；5 将所提取的色彩运用到设计效果



图 5.15 色彩连续性的问题

图上预览。

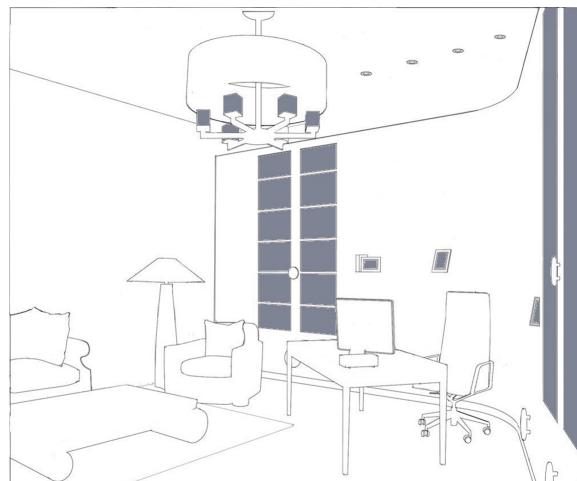
对于用户来说，本系统的使用方式如下：1. 用户输入对需求的文本描述；2. 得到配色方案；3. 当用户不满于当前色彩方案可以重新搜索；4. 使用当前色彩方案为效果图上色；5. 当用户不满于当前效果图可以使用当前色彩方案重新上色。

在实现的过程中，首先需要解决的是自然语言处理的问题。本系统采用 word embedding 的方式实现单词向量化，word embedding 采用 CBOW 模型，使用哈夫曼树、负采样优化、fast text 等策略加快速度以及增加编码的精度，实现对词语语义信息的提取。词向量到句向量的转化采用加权合成的方法，考虑词频、类型权重等实现对句子的精准建模，将句子中重要的语义信息提取出来，略去系统不关心的信息，建立艺术品文本信息句向量词典。使用余弦相似度作为指标，在词典中寻找与输入句子匹配的艺术品。之后需要对图像进行处理，从匹配的艺术品图片中提取色彩。最后对效果图自动上色。

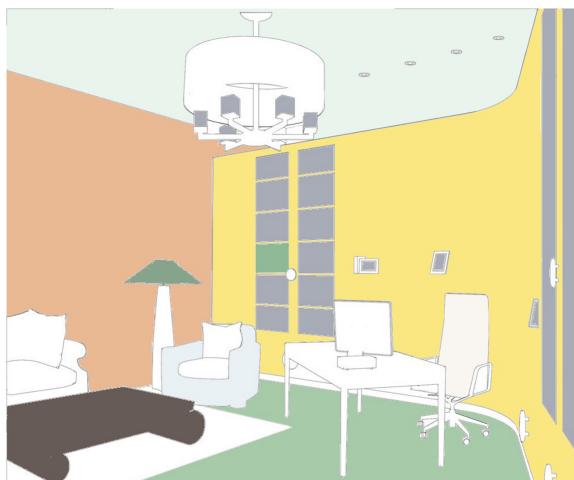
整个系统运行的过程相应迅速，匹配准确，效果图预览清晰有效，可以直接高效的为



(a) 原效果图线稿



(b) 经过标注的线稿



(c) 自动填充



(d) 原效果图线稿图层叠加

图 5.16 色彩连续性的处理

设计师提供快速的参考和辅助。

6 实验结果

6.1 输入语句的测试案例

首先我们聚焦于语义搜索的结果，我们可以看见，搜索“温柔”，得到的前 20 个艺术品为：

“2009 年作你的温柔布面油画”；“2009 年作温柔的坚持布面油画”；“1997 年作温柔的风布面油画”；“1993 年作温柔的光布面油画”；“2007 年作下午温柔的阳光布面油画”；“2003 年作温柔的光布面油彩”；“2003 年作温柔的光布面油彩”；“1989 年作温柔的音乐油彩画布镜框”；“1997 年作温柔的沉思油彩画布”；“1992 年作温柔情人压克力纸本”；“1992 年作温柔情人压克力纸本”；“2013 年作夜色温柔布面油画”；“2010 年作夜色温柔布面油画”；“温柔地杀我油画面布”；“2004 年作温柔地杀我版画”；“2002 年作温柔地杀我布面油画”；“2010 年作优雅的女孩布面油画”；“2009 年作你的柔情布面油彩”；“灿烂的朝阳照在秋日收……”；“2009 年作你的柔情布面油画”。

可以看到，非常明显的，和“温柔”相关的艺术品题目被首先搜索了出来，“xxx 年作”“我的”这样的单词由于在总体文本中高频出现，故而权重低，引起的偏移很低。最后一个题目中带有“温柔”的作品很有意思，叫做“温柔地杀我”，显然，作者在命名这样一个名字的时候使用了对比的手法，“杀”之类的词语的词向量显然和“温柔”方向不同，以至于使得这个题目的句向量偏移到更远的地方，成为了题目中带有“温柔”的作品中排在最后的作品。而更应当注意的是，当题目中带有“温柔”的作品都被搜索出来之后，从第 17 个艺术品开始，我们可以发现题目中不再出现“温柔”这个单词，而出现了“优雅的女孩”，“柔情”等从“温柔”的语义延伸的词语，这更直接的说明 word2vec 发挥了它的优势，在向量化的过程中“温柔”和“柔情”等词的词向量距离很近。

如果加入更多的干扰信息，输入“我想要温柔的柔和的感觉”，得到的前 20 个艺术品为：

“第二年，我去了北京，在那……”；“受马蒂斯的影响，闫平的画……”；“杨飞云的艺术真正打动人的，……”；“这种自觉的留白和色彩的简……”；“居于大地的当代生存者正在描……”；“20 世纪 80 年代以来，他从写实……”；“1993 年作温柔的光布面油画”；“2003 年作温柔的光布面油彩”；“2003 年作温柔的光布面油彩”；“2007 年作下午温柔的阳光布面油画”；“1989 年作温柔的音乐油彩画布镜框”；“刘海粟这个时期的作品有一种……”；“1982 年我毕业后被留校任教，……”；“具象和抽象迭合构成的”古典“作……”；“他的深度绘画表现和塑造的是……”；“毛时安评论说：“俞晓夫是很聪明……”；“2009 年作温柔的坚持布面油画”；“创作于 1995 年的作品《抽象》……”；“《小演员》画面中少女肌肤所充溢……”。

(在这里由于某些文本过长，没有全部显示，可以在附录中查找到实验结果的所有结果。) 受到干扰信息的影响，其中大多数都是带有长长的描述的艺术品文本，在数据库中

出现这样一类的文本，题目为长长的描述，图片的 url 则为空缺。之前我已经进行过数据清理（见第二章），在这里意识到数据清理还需要补充，由于本系统聚焦于色彩方案提取，所以我去除了这些没有图片资料的艺术品。

“1993 年作温柔的光布面油画”；“2003 年作温柔的光布面油彩”；“2003 年作温柔的光布面油彩”；“2007 年作下午温柔的阳光布面油画”；“1989 年作温柔的音乐油彩画布镜框”；“2009 年作温柔的坚持布面油画”。

除了这些本身就没有题目和图片信息的艺术品之外，其他的艺术品文本都是围绕“温柔、柔和”，可以达成使用者的目标。

让我们继续来看的色彩方案和上色结果，输入“我想要温柔的柔和的感觉”，得到的效果图如图 6.1。



图 6.1 输入“我想要温柔的柔和的感觉”

输入“温柔的女性化的感觉”，作为人类的理解，这句话与上一句话的区别为，上一句话主要是“温柔/柔和”，这句话不仅强调“温柔”还强调“女性化”。输入“温柔的女性”，相当于对系统提出了两个要求，要求输出的结果既符合“温柔”又符合“女性”，得到的前 20 个艺术品为：

“2009 年作温柔的坚持布面油画”；“2009 年作你的温柔布面油画”；“新女性布面油画”；“新女性布面油画”；“1993 年作温柔的光布面油画”；“1997 年作温柔的风布面油画”；“2007 年作下午温柔的阳光布面油画”；“2003 年作温柔的光布面油彩”；“2003 年作温柔的光布面油彩”；“1989 年作温柔的音乐油彩画布镜框”；“2006 年作新女性布面油画”；“2003 年作新女性布面油画”；“2003 年作当代女性布面油画”；“1997 年作温柔的沉思油彩画布”；“2002 年作新女性布面油画”；“2011 年作戴银饰的苗族女人布面油画”；“2006 年作现代女性油画面布”；“2007 年作美丽上海—她们的眼睛布面油画”；“2010 年作优雅的女孩布面油画”；“2013 年作优雅女人布面油画”；

可以看到搜索结果其实分为靠近“温柔”的和靠近“女性”的，在“温柔”和“女性”结合的程度上显得不是很好，这样的结果收到数据库的限制，因为数据库中艺术品的文本包括“温柔”和“女性”两个含义的（并非必须这两个词）不常见。这是这个系统的一个

限制，即数据库中的艺术品不能囊括所有的感受，并且即使艺术品中包含的信息和感受很多，艺术品文本包含的信息一定比艺术品本身少，艺术品文本描述的并不一定是艺术品的色彩风格。解决这个问题采用了交互设计的办法，即搜索最为匹配的前 10 个结果，对其中随机的一个提取颜色产生色彩方案。这样用户可以通过不改变输入文本，再次搜索，系统就会另外随机选择一件，直到用户得到满意的方案；用户也可以手动打开 10 个色彩方案手动选择。

对于配色方案我的期望是，比上图有更多的女性特质的方案，也许整体是偏粉红色或有其他符合女性特征的特点。而系统运行的结果如图 6.2，图 6.3，是符合预期的。



图 6.2 输入“温柔的女性化的感觉”

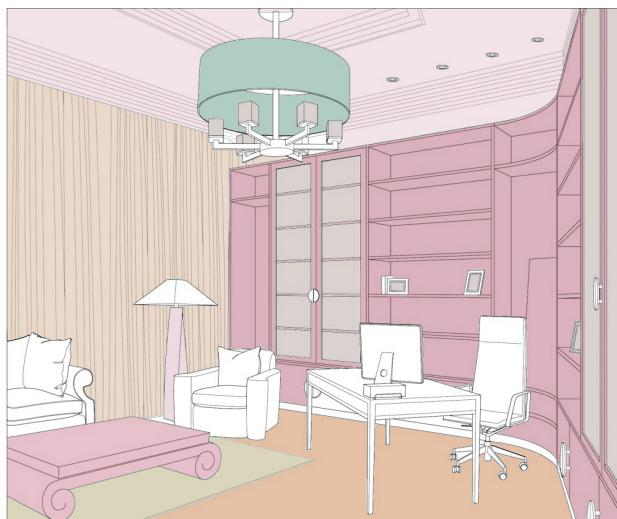


图 6.3 输入“温柔的女性化的感觉”

那么如果输入“冷静的男性”。得到的结果是让人不满意的：

“新女性布面油画”；“新女性布面油画”；“2004年作冷静油彩画布”；“喻红在日常存在中寻找一种诗意……”；“2006年作新女性布面油画”；“陈肇耀历年来的数字合成影像作品……”；“我历年来的数字合成影像作品，皆……”；“刘海粟是要强的人，有强烈的自我……”；“1989年北京的现代艺术大展标志……”；“刘海粟这个时期的作品有一种较……”；“受马蒂斯的影响，间平的画面中，表……”；“2011年作戴银饰的苗族女人布面油画”；“不过站在美术史的角度上来看，从……”；“居于大地的当代生存者正在描摹……”；“2006年作现代女性油画家布”；“最显著的是911事件后，城市、霸权……”；“《男男女女》创作于2006年，是刘……”；“2003年作当代女性布面油画”；“2003年作新女性布面油画”；“1999年作理智与情感油画家布”；

在这里就发生了一个问题，想要男性化的风格得到的搜索结果全是女性。这是由于 word embedding 的词向量本身的性质。由于词向量基于上下文预测，所以最为相似的单词并非语义上最相似，而是用法上最相似的单词，比如“男性”和“女性”往往用在相同的上下文环境中，而“男性”和“男人”相对用在相同上下文环境中的程度更小一些。如果我们适应一下这个系统，输入“冷静的男人”那么就会得到这样的搜索结果：

“2004年作冷静油彩画布”；“1995年作大眼睛男人布面油画”；“2011年作戴银饰的苗族女人布面油画”；“2007年作红色男人布面油画”；“2006年作男人像油画”；“清男人像纸本油画”；“男人像油画家布”；“男人像油画家布”；“2008年作男人布面油画”；“两个男人纸本油画”；“2005年作男人肖像布面油画”；“男人的故事”；“男人布面油画”；“男人布面油画”；“2007年作红色男人油彩布面”；“1999年作男人与马布面油画”；“喻红在日常存在中寻找一种诗意。她……”；“1997年作同志系列：男人布面油画”；“2006年作红色男人像油彩画布”；“2001年作男人·女人布面油画”；

如图 6.4，比起“温柔”来看的确显得更加沉稳厚重。



图 6.4 输入“冷静的男人”

输入“快乐的儿童”，输入“快乐的儿童”搜到的艺术品有：

“1997 年作快乐的儿童第 9 号油画画布”；“2013 年作快乐的孩子们布面油画”；“新中国的儿童（一张）纸本”；“儿童布面油画”；“儿童油画”；“2004 年作儿童时代油画画布”；“快乐 1 油画”；“1997 年作为小孩快乐油画画布”；“2001 年作快乐世界布面油画”；“少女与儿童布面油画”；“2005 年作快乐在一起，2 号油画画布”；“快乐的童年布面油画”；“1999 年作快乐的红色油画画布”；“1992 年作快乐油画画布”；“游戏的儿童布面油画”；“快乐的童年油画画布”；“威尼斯的快乐”；“2002 年作游戏的儿童 9 号油画画布”；“2002 年作游戏的儿童 8 号油画画布”；“2005 年作快乐组合布面油画”。

显示程序完全可以抽象出“儿童”和“小孩”、“孩子”的相似性。得到的图 6.6 的色彩方案，虽然使用的效果图线稿是一个书房，但是也能看出符合方案儿童房的配色。



图 6.5 输入“快乐的儿童”



图 6.6 输入“快乐的儿童”

6.2 效果图类型的测试案例

本系统的色彩方案可以用于室内设计、服装设计、产品设计、平面设计、交互视觉设计等。但是上色预览的算法并不适用于每种工作。系统准备了室内设计，服装设计各 1 套效果图线稿预览，之测试发现，如图 6.7，对于上色区域占图片比例较小并且色彩有限的服装、产品效果图本系统也可以使用，但系统效果更适合室内设计效果图。

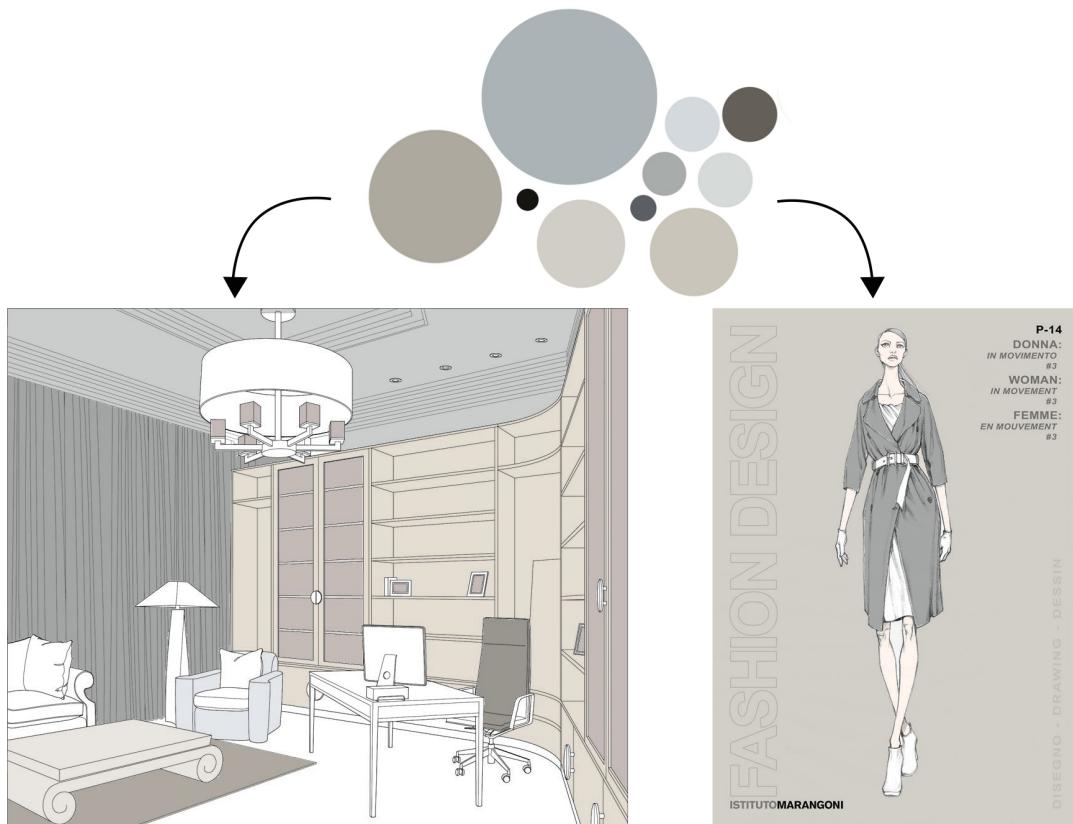


图 6.7 色彩方案：“温柔”对室内设计和服装设计效果图的预览

配色方案对于服装设计是适用的，然而由于服装设计效果图有一些特点，首先服饰在整个图片上的面积较小，填充算法的种子点行走到服装上到几率较小，如果是有花纹的服装，花纹的色彩连续性无法保证，小的配饰更难以填充，所以本系统的上色算法更适合室内设计。

对于其他的效果图，可以使用具体匹配的上色方式。对于产品设计效果图，每个产品上的出现颜色不多，可能有成套的不同颜色的产品，更适合简单的标注上色预览。而对于网站视觉设计等，网页的框架往往预先设定比例和模式，使用堆栈的方式上色是非常适合的，在很多的情况下网站设计适用矢量图片，可以直接标注色彩上色。综上，上色方式并不适用某些效果图，但配色方案对于产品设计、网页设计都是适用的，

6.3 本章小结

经过输入语句测试案例，本系统可以对语义精确的响应。对于“冷静的男性化的”房间、“温柔的女性化的”房间、“快乐的儿童的”房间，配色有明显的差别，符合人类的心理预期。而对于“温柔的”和“温柔的女性化的”，这样信息叠加的语句也有相似而差异的配色方案结果。

对于文本不同而意义相同的语义延伸词语，本系统的 word embedding 模型能够实现同义匹配，并且将其排列在文本完全相同的情况之后搜索到。对于不含有信息的干扰语句，本系统的 word embedding 模型能够实现信息提取，和不含有干扰语句的情况下得到的结果差别很小。对于含有并列信息的输入语句，搜索的结果受到艺术品数据库的限制，不一定存在完全符合所有并列信息的艺术品文本，所以搜索到的文本结果分为偏向各信息的内容，使用交互的方式来解决这各限制带来的偏差。

对于效果图的类型，本系统的色彩可以用于室内设计、服装设计、产品设计、平面设计、交互视觉设计等。但是上色预览的算法并不适用于每种工作，本系统的上色算法更适合室内设计效果图。对于产品设计、网页设计、服装设计等工作，可以使用各自匹配的上色预览算法。

参考文献

- [1] ANON. Artificial Intelligence: A Modern Approach[M]. [S.l.] : Prentice Hall, 2012.
- [2] LUNGARELLA M. LNCS sublibrary: Artificial intelligence : 50 Years of Artificial Intelligence: Essays Dedicated to the 50th Anniversary of Artificial Intelligence[M/OL]. [S.l.] : Springer, 2007.
<https://books.google.com.hk/books?id=vin0j1Z4fZUC>.
- [3] MANNING C, RAGHAVAN P, SCHÜTZE H. Introduction to Information Retrieval[M/OL]. [S.l.] : Cambridge University Press, 2008.
<https://books.google.com.hk/books?id=t1PoSh4uwVcC>.
- [4] ABELSON H, SUSSMAN G, SUSSMAN J. MIT electrical engineering and computer science series : Structure and interpretation of computer programs[M/OL]. [S.l.] : MIT Press, 1985.
<https://books.google.com.hk/books?id=2o9QAAAAMAAJ>.
- [5] MIKOLOV T, CHEN K, CORRADO G, et al. Efficient Estimation of Word Representations in Vector Space[J/OL]. CoRR, 2013, abs/1301.3781.
<http://arxiv.org/abs/1301.3781>.
- [6] 詹建, 高民权. 基于主题词的网络舆情可视化关联分析方法 [J]. 情报资料工作, 2014, 6 : 17–22.
- [7] ADOMAVICIUS G, TUZHILIN A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions[J]. IEEE transactions on knowledge and data engineering, 2005, 17(6) : 734 – 749.
- [8] GIL A, MEDRANO M, MARTORELL I, et al. State of the art on high temperature thermal energy storage for power generation. Part 1—Concepts, materials and modellization[J]. Renewable and Sustainable Energy Reviews, 2010, 14(1) : 31 – 55.
- [9] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks[J]. arXiv preprint arXiv:1511.06434, 2015.
- [10] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[J], 2014 : 2672 – 2680.

- [11] MCALEER S. Generative Adversarial Networks[J]. .
- [12] ISOLA P, ZHU J-Y, ZHOU T, et al. Image-to-Image Translation with Conditional Adversarial Networks[J]. arxiv, 2016.
- [13] CHEN S, ZHANG T. An improved cloud masking algorithm for MODIS ocean colour data processing[J]. Remote Sensing Letters, 2015, 6(3) : 218–227.
- [14] O'DONOVAN P, AGARWALA A, HERTZMANN A. Color Compatibility from Large Datasets[J/OL]. ACM Trans. Graph., 2011, 30(4) : 63:1–63:12.
<http://doi.acm.org/10.1145/2010324.1964958>.
- [15] DENG L, SELTZER M L, YU D, et al. Binary coding of speech spectrograms using a deep auto-encoder[J], 2010.
- [16] QU Y, WONG T-T, HENG P-A. Manga Colorization[J/OL]. ACM Trans. Graph., 2006, 25(3) : 1214–1220.
<http://doi.acm.org/10.1145/1141911.1142017>.
- [17] ZHANG L, JI Y, LIN X. Style Transfer for Anime Sketches with Enhanced Residual U-net and Auxiliary Classifier GAN[J/OL]. CoRR, 2017, abs/1706.03319.
<http://arxiv.org/abs/1706.03319>.
- [18] WIKIPEDIA. Embedding[EB/OL]. 2017.
<https://www.wikiwand.com/en/Embedding>.
- [19] MCCANN B, BRADBURY J, XIONG C, et al. Learned in Translation: Contextualized Word Vectors[J/OL]. CoRR, 2017, abs/1708.00107.
<http://arxiv.org/abs/1708.00107>.
- [20] ATHIER G, FLOQUET P, PIBOULEAU L, et al. Process optimization by simulated annealing and NLP procedures. Application to heat exchanger network synthesis[J]. Computers & chemical engineering, 1997, 21 : S475–S480.
- [21] XIA T, CHAI Y. An improvement to TF-IDF: Term Distribution based Term Weight Algorithm.[J]. JSW, 2011, 6(3) : 413–420.
- [22] BARONI M, DINU G, KRUSZEWSKI G. Don't count, predict! A systematic comparison of context-counting vs. context-predicting semantic vectors[J], 2014, 1 : 238–247.
- [23] MIKOLOV T, SUTSKEVER I, CHEN K, et al. Distributed Representations of Words and Phrases and Their Compositionality[C/OL] // NIPS'13 : Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2. USA : Curran As-

- sociates Inc., 2013 : 3111–3119.
<http://dl.acm.org/citation.cfm?id=2999792.2999959>.
- [24] KAISER L, BENGIO S. Discrete Autoencoders for Sequence Models[J/OL]. CoRR, 2018, abs/1801.09797.
<http://arxiv.org/abs/1801.09797>.
- [25] MIKOLOV T, SUTSKEVER I, CHEN K, et al. Distributed representations of words and phrases and their compositionality[C] // Advances in neural information processing systems. 2013 : 3111–3119.
- [26] LIU P J, SALEH M, POT E, et al. Generating Wikipedia by Summarizing Long Sequences[J/OL]. CoRR, 2018, abs/1801.10198.
<http://arxiv.org/abs/1801.10198>.
- [27] ZHANG Y, LIU Y-D, JI Z, et al. Vector similarity measurement method[J]. Technical Acoustics, 2009, 4(5) : 532–536.
- [28] ANDRECUT M. High-Dimensional Vector Semantics[J/OL]. CoRR, 2018, abs/1802.09914.
<http://arxiv.org/abs/1802.09914>.
- [29] RAVICHANDRAN D, PANTEL P, HOVY E. Randomized algorithms and nlp: using locality sensitive hash function for high speed noun clustering[J], 2005 : 622–629.
- [30] WIKIPEDIA. colormodel[EB/OL]. 2007.
https://en.wikipedia.org/wiki/Color_model#cite_ref-Kandel_1-0.
- [31] WIKIPEDIA. Floodfill[EB/OL]. 2009.
https://en.wikipedia.org/wiki/Flood_fill.

致谢

岁月如梭，转眼间，三年的研究生求学生涯即将结束，站在毕业的门槛上，回首往昔，奋斗和辛劳成为丝丝的记忆，甜美与欢笑也都尘埃落定。浙江大学以其优良的学习风气、严谨的科研氛围教我求学，以其博大包容的情怀胸襟、浪漫充实的校园生活育我成人。值此毕业论文完成之际，我谨向所有关心、爱护、帮助我的人们表示最诚挚的感谢与最美好的祝愿。

本论文是在导师孙守迁教授的悉心指导下完成的。三年来，导师渊博的专业知识，严谨的治学态度，精益求精的工作作风，诲人不倦的高尚师德，朴实无华、平易近人的人格魅力对我影响深远。导师不仅授我以文，而且教我做人，虽历时三载，却赋予我终生受益无穷之道。本论文从选题到完成，几易其稿，每一步都是在导师的指导下完成的，倾注了导师大量的心血，在此我向我的导师孙守迁教授表示深切的谢意与祝福！

本论文的完成也离不开其他各位老师、同学和朋友的关心与帮助。在此也要感谢张克俊、汤永川等各位老师在论文开题、初稿、预答辩期间所提出的宝贵意见，感谢创新设计课题组为本论文提供的数据和建议，还要感谢同门的师兄师妹们，在科研过程中给我以许多鼓励和帮助。回想整个论文的写作过程，虽有不易，却让我除却浮躁，经历了思考和启示，因此倍感珍惜。

还要感谢父母在我求学生涯中给与我无微不至的关怀和照顾，一如既往地支持我、鼓励我。最后要感谢我的男友，在学习期间给我的支持和帮助，也在生活上给予我包容与鼓励。

于求是园 2018年4月16日

胡迪

附录 A 附录：实验结果

A.1 文本搜索结果

A.1.1 输入：温柔

“2009 年作你的温柔布面油画”;
“2009 年作温柔的坚持布面油画”;
“1997 年作温柔的风布面油画”;
“1993 年作温柔的光布面油画”;
“2007 年作下午温柔的阳光布面油画”;
“2003 年作温柔的光布面油彩”;
“2003 年作温柔的光布面油彩”;
“1989 年作温柔的音乐油彩画布镜框”;
“1997 年作温柔的沉思油彩画布”;
“1992 年作温柔情人压克力纸本”;
“1992 年作温柔情人压克力纸本”;
“2013 年作夜色温柔布面油画”;
“2010 年作夜色温柔布面油画”;
“温柔地杀我油画画布”;
“2004 年作温柔地杀我版画”;
“2002 年作温柔地杀我布面油画”;
“2010 年作优雅的女孩布面油画”;
“2009 年作你的柔情布面油彩”;
“灿烂的朝阳照在秋日收获的黄土高原上，辛劳纯朴的陕北女孩儿用心爱的笛子吹起悠扬的音乐。这笛声飘荡在静穆的村庄上空，诉说着古老而遥远的传说。陕北女孩儿的柔情与这秋日辉煌形成鲜明的对比，从而更加突出了女孩儿的纯情和对未来人生的美好憧憬。”;
“2009 年作你的柔情布面油画”;

A.1.2 输入：我想要温柔的柔和的感觉

“第二年，我去了北京，在那里整整住了一个冬天，窗外飘着大雪。味道不佳的室内，暖烘烘的，在这里我画了一批工厂和管道与植物交织的作品，牧歌式的景观被分裂和忧郁的超现实图式所代替，这一个是著名的 85 美术思潮涌动的时候，超现实主义的大师达利、契里柯的追随者在中国的这页美术史中大放异彩，与注重哲学思考和文化批判的理性思考不同的是：我是由纯粹个人经历和感觉自然的过渡到超现实语境的。我喜欢已沉迷于文化风暴而心灵脆弱的内心为表达主题：因为想太多而行太少，因为书读得过多而体会过深，人格失调、精神分裂、失意、堕落或彷徨接踵而来。《离开和留在最后一块草地上的两个人》、《屋外的马窥视的她和被他端视的我们》、《最后的花园》、《听见毕加索马叫的隐者和鸟》、《春天唤醒冬眠者》是这时的作品。虚幻失真的田园憧憬，平庸沉闷的现实生活、虚假的文化表象与躁动

的思潮迷雾之中，是绘画使我解放出来，为我展开一个新的景界。一切静止不动，一成不变的东西都开始颤栗了。不只是被人赋诗的星星、月亮、树林、花草，甚至一团发电厂的烟灰，一张被风刮到街角的报纸，一块不知所以然被蚂蚁从草丛里咬出的树皮，一座锈蚀的、停摆的座钟……所有的一切都显露给我一张自己的脸，它内在的生命，这些沉默的心灵。经常无声胜有声，无论静或动的点或线对我都是这般生动，这就足够使我投入整个生命感觉去体会艺术的可能性和它的生命。自然地将自己的梦境表现在画布上，并且化痛苦为诙谐，化诙谐为抒情，像鱼一样，流连在自己的幻想中，那时，对我而言，绘画几乎是人与动植物和工厂之间的一种混合体。”；

“受马蒂斯的影响，闫平的画面中，表达观念时，依靠的是色彩而不再是传统写实中的光与影，体现光线的感觉是通过色彩的对比与排列来表现的，而不仅仅是描绘形体。从画面的色彩来看，戏剧的服饰能够充分彰显出色彩的艳丽，乍看张扬，细看则处处小心，如同喜剧演员台上台下判若两人，色彩的叙述有一定的私密性能让她尽情地表达内心世界。颜色是外在的，而内心要表达的情绪置于其中则具有保护性、自由性和私密性。在此画中，我们不仅可以看到有后印象派的影响，还能看到对中国传统艺术文化的传承，画面中融会贯通了西方的理性秩序和中国的意向表现，在理好的色系画面上，加上特立独行的焦墨，恰到好处。纳比派的主观表现给了闫平以启迪，加强了用色的浓度和画面的装饰性处理，使色彩相互交织，对比与协调并存。画面中，闫平大胆运用强有力的笔触进行塑造，由薄到厚，由于作品中用笔交错相叠，产生了一些犹如砾石般的肌理，使得画面的表现更加丰富。从画面的构图来看，她把握大局，对于画面的四边四角和主体物之间的关系的处理非常考究。这里面包括：画面的分割，每一个线条的长短，每一个方圆的比较，画面中疏密的安排以及造型方式的扩展等等。”；

“杨飞云的艺术真正打动人的，与其说是他通过学习西方古典艺术的绘画形态而将古典主义移植到我们这块土地，不如说是他将古典的精神通过他独特的灵魂体认而呈现在我们面前。为一个人，他有幸感悟到永恒的生命奥秘，作为一个画家，他有幸以第二双眼睛看到了世界的荣光之美，正像他所说的：“感谢造物主，给我们提供了如此缤纷煌丽深沉广博的自然世界，内里深藏井然有序的规律、外有瞬息万变的勃勃生机，丰丰满满的不缺少什么，供我们探求和寻找想欲的一切。””；

“这种自觉的留白和色彩的简括处理是一种新的绘画性格，它是油画形态发生改变的重要端倪，它必定被后人发现和珍视，这种新的端倪预示了在西方光色和结构两大语言谱系之外，油画语言开放的另一通道。刘海粟通过书写与塑造的重新匹配，油画被另一种民族智慧所把握，补充和开放了油画这一古老艺术样式的“音域”，这就是贡献，对绘画表达和视觉审美经验的双重贡献。这是一个东方少年天才自悟自通的发明。《圣扬乔而夫飞瀑》无论是作品形态和完整度还是语言的成熟度，都已经达到那个时代的最高水平。在一定意义上，这是被逼迫和放纵出来的才情，任性和要强的情绪通过书写的动态挥策而出，如瀑布流泉一样奔涌、流淌出来，笔下的物象已经不是静态的客观物象，更像是一个充满热情和煽动感的视觉宣言，预示着观看习惯和表现陈习的巨变。尽管这个巨变被中断、覆盖，在历史视野中消逝了近八十年，今天让我们看见它，就如蓦然看见了旷代知音，本能般萌生一种把臂欲言的冲动。”；

“居于大地的当代生存者正在描摹一幅高度机器化的是日常生活图景，已经取得的物质成果并没使人们获得更多的幸福和精神自由。人们陷身于自己的创造物中并被其所奴役。丁方为我们提供了新的视觉经验：广阔空无的黑暗，在我们的体内发出回声，一个恍惚而苍白的飞翔人形在浊浪翻滚的昏暗中挣扎飘荡，所有鲜活的生命都被这一片混沌无情遮没着，即沉重工业的压迫之力迫使人们集结突围的勇气和力量，抵抗和承担由恐惧所预感到的否定性（《我逃向何方？》）；一条没有来处和去处的褐色河流挟裹着颓废的建筑残体和无知无觉的肉身，面对这种毫无反抗的挣扎努力的自溺情形，绝望感几乎要漫过我们的头顶（《罪灵们》）。此时，坚信者的呼告之身形迎面而来；精砺的铅色背景上，一个巨大的鸟形幻像，犹如折翼的天使，正飞向黄昏中模糊的地平线，羽翅上青冽的光芒耀亮即将陷入昏暗的荒野（《折羽的天使》、《请退让吧，忧郁的阴影》）。这些画幅突破了艺术的唯美法则的有限而均突出了其思想的锐利；神圣关怀并没有抛离人，只要人愿意，就在此时此刻，他会来与我们相遇，以其苦痛搀扶人的苦痛。光影。古典的光影效果重新返回丁方的画面，光对画面的统领使其显示出蕴于悲剧深处的庄严性。丁方画面中精神性光影的设置，作为神圣俯临此世的标志，它象征着叹息、救扶、怜悯、启明、爱等精神内涵，它不是来自熟常的经验世界，它温抚着画面肌理的紧张断裂，呼应着上升的精神，漫渗到流幻的色彩深处，洞开灵魂的奥秘。它在显示生命的悲剧况味的同时，“给予灵魂以永

恒的召唤”。在这种精神之光的携领下，丁方的画语世界凸现出几个非常重要的意象造型：以黄土高原、荒野自然为原型的大地意象。作为众生灵终其一生、存活于其上的大地，其深邃、厚重的自然存在形态在丁方的画面中取得超拔于日常感觉经验的独语性：奔泻而下的色流形成大地强烈的动势，坚岩般的肌理沉凝滞重，深含着忍耐、重负和强劲的生存精神。”；

“20世纪80年代以来，他从写实到表现，走上了一条更为宽阔的道路。他的表现不是苦涩的，而是歌吟式的，也就是抒情的，这是他的人生观与世界观的体现。他总是希望现实与理想接近一些，试图用自然生命的热烈与蓬勃气象缓解现实生活的苦恼与矛盾，用美好的向往引领人们的精神世界。他的艺术风格兼蓄了许多属于“成熟”境界的因素，例如文艺复兴诸家晚年作品那种浑然的氛围、印象派各位到后来信笔流淌出来的华彩、属于表现主义语言根本的那种急疾而鲜活的“手气”。他也将中国传统美学中崇尚“骨气”，注意综合，以阐发人与自然亲和关系等优点融化到自己的作品之中，由此不断朝向表达更自由、笔力更老辣、抒发更从容的境界迈进。”；

“那年冬天，我经常去美术馆看一个接一个的展览，劳申伯的大型展给我不同以往的震撼。它虽不像其它我所看到的《韩默藏画展》、《法国250年藏画》、《德国表现主义画展》、赵无极、戈雅、柯尔韦尔个展等古典和现代主义那种对自我情节的顶礼膜拜具有令人驻足叹服，品评赏析的细节，但却使我第一次强烈地感受到当代艺术直接切入周遭生活的力量。这与我几年来接受的为艺术而艺术的教育很不相同，这也使我先前即已产生的对“纯美术”意义的质疑更得到证明。1986年朋友们同我一起发起和组织的《西南艺术群体》，使我更直接投身于“新潮美术”的潮流中。我们彻夜地谈论种种宣言和观念，策划在昆明、上海、重庆和北京的展览，在各种报刊上撰写文章，对那些同样激动得面红耳赤的人解释何为“新具像”——那的确是一个冲动的年代，经历过八十年代前期文化洗礼的人们，例如我本人，总是难以忘怀那些热闹的场面。虽然那时的文化存在诸多的夸张，自以为是和种种谬误，但是，那毕竟是一个在文化上有追求的年代。毛旭辉永远紧锁着眉头，在和平村的小屋中挤满了《他的红色体积》和其它许多的作品，他的东西令人压抑、难受。与此相反他却能写一手漂亮的激扬文字。潘德海总是结结巴巴地发明一些惊人的词汇，除了画《杂乱的、杂乱的》形象，还用泥巴做了一大堆黑不溜秋的怪物。张晓刚则一直有点善感和见异思迁；胃病住院就画些痛苦的痉挛不堪的《白色幽灵》；搬到桃花山又显得花花草草地吟诵《遗梦集》……今天，当年年轻气盛的他们都已一个个成为了不起的画家了。那个时候，我在医院里一面等待女儿的出生，一面写《自然意识略述》和《丛林猛兽》。在省图书馆的展览开始时，朋友们在请柬上写满了对未出生女儿的祝福……同当时的许多知识分子一样，我们怀有真实的人文主义理想，虽然今天我的同辈们大多数未必赞成那种理想，然而在情感上却仍又存在深深的怀念。”；

“1993年作温柔的光布面油画”；

“2003年作温柔的光布面油彩”；

“2003年作温柔的光布面油彩”；

“2007年作下午温柔的阳光布面油画”；

“1989年作温柔的音乐油彩画布镜框”；

“刘海粟这个时期的作品有一种较劲和任性的情绪，与日本第一代“准印象派”艺术家的作品相比，他自觉的东方性格已经完整地呈现出来，甚至是近乎偏执的强调这种性格。他的写生作品与日本前辈画家黑田清辉和藤岛武二的相比，更具书写的自由和流畅感，情绪更为敏感、张扬，色彩更为刺激、强烈，他自觉的弱化了色层堆叠和塑造感，以柔性的笔调写出阳刚的气质，他完成的语言难度是巨大的——以柔致刚，以简驭繁，以轻薄而致厚重。在他1930年代的代表作中，我们可以看到他明确而自觉的语言策略。一个优秀的艺术家，策略自觉、成熟的标志就是——他将那些难以结合的矛盾双方确立为自己的表现范畴，将相反者结成紧密的同一，这就是艺术命题的方向，天才往往就是在那些难以逾越鸿沟间自由腾挪的人。从文化背景上来看，刘海粟当年也是承受着双重的文化压力，反者道之动，最对立的矛盾双方往往是触发天才属性的直接动因，只有那些能主动挑战并立志于战胜这些截然对立矛盾的人，才可能创造出历史的奇迹——不可思议的难度是历史奇迹的前提！刘海粟之所以伟大，在于他克服甚至颠覆了中国传统文人的内向、惰性、柔美的审美趣味和文化性格，但又将文人细腻、敏感、神秘的趣味特征糅合进了全新的表现形式之中，结成了相反者之同一；同时，他颠覆了当时国内以及整个东方艺术领域对客观再现的迷恋——他在这两种主流美学趣味的压力之下寻求

突围的通道，以非常态的学习方式和创作风格如一匹黑马腾空而起，成为时代的英雄。他 1930 年代创作巅峰的作品基本致力于转化后期印象派的语言和形态，具有超前的实验精神，比日本梅原隆三郎 1930 年代晚期的中国写生，更有一种从容的气度和灵敏的语言才华。”；

“1982 年我毕业后被留校任教，浪漫的、不着边际的学生生活同平庸乏味的现实生活的差别很快显露出来，尽管此时我被周围的朋友们认为是个“幸运儿”——在毕业前我已是“美协会员”，作品也参加和发表在各种展览、杂志上，抽屉里藏满了入选、获奖证书和来自各地的初学者的来信。但是很快这种小小的满足感就消失的一干二净，首先是一度在全国独领风骚的四川油画，已失去其尖锐的社会抗争性，在受到普遍的赞誉和认同后，业已成为一种日益陈旧和庸俗的样式范本，从思想与风格都同此有距离的我的作品几乎没有起码的参展和发表机会。其次，作为一个为人师表的教师的现实与我内心幻想的那种“波西米亚式”的艺术家生涯相去甚远，同时我居住、工作的美院四周重庆典型的旧工业景观使我倍感压抑。透过发电厂巨大的烟囱冒出的滚滚浓烟，阴霾的天空下，灰色的街道泥烂如河，排放废气的喧哗声有如雷鸣，夜空为之震颤——在蓝天和树荫下长大的我想躲避的就是这个世界！对付这种日子的最好办法是去一个防空洞里开设的小酒馆里喝酒，每天，从黄昏时分起我就开始酝酿情绪，等待那喝酒的时刻终于来临。小酒馆里坐满了疲惫的夜班的巡道工，兴高采烈的白日扛活的搬运工以及就着铁丁下酒的酒鬼和守在角落里的默默无语的失意者。我虽是他们中间的一员，却难以排解难言的孤独，在一次半醉半醒之间的日记中我问自己：“你同周围世界格格不入，你无法和他人苦乐与共，你感到自己和他们隔膜甚深，这时你该怎样生活？”有时，我也只好自我安慰：“我还是不断地听到一种责备，指责自己缺乏现实的感觉，我的确是不尊重现实，我认为，现实最不需要人们充分地去注意，人生活于现实中永远不可能满意，因为现实是一种偶然性，是生命的垃圾，对于这种可怜的现实，我们除了否定它之外，别无选择。””；

“具象和抽象迭合构成的”古典”作品，给人一种包容性的丰富特征，这也就使得解读杨飞云的作品并非那么简单，在他那里既有古典大师的决定性的影响，也不乏塞尚、马蒂斯等现代画家的印迹。不过，难能可贵的是，杨飞云艺术作品的多层次，并没有给人生硬组合之感，读他的作品犹如聆听一曲旋律美妙清纯的小提琴独奏，其境界有追随东方高山流水和西方天堂之音的韵味。”；

“他的深度绘画表现和塑造的是具有当下性的永恒图景，其重要特征展现为一种弥漫于画面中的重返生存大地的勇气与决心。在这块安置人但又不能使人安居的大地上，他感受到表面欢娱下的人的苦难与绝望，并在绝望的深处发现了与人类苦难同在的神圣。无形的灵魂世界运行于涌流的色群之中，将要急掠而去的云，希望与绝望接壤的破晓，世界的恐惧、悬浮的命运、尖锐的力量—每一事象，都隐没于无边无垠的空间之中，没有一条直线的旋律，能在浊存浑重的色群中挣脱出来，一道来自天边的光芒照耀着侍立静听的人，犹如一个遥远而又切近的声音在寂静中传谕启示之语。”；

“毛时安评论说：“俞晓夫是很聪明而不是小聪明的画家。他的聪明使他的作品不会有宏大叙事而赤膊上阵、笨重不堪的毛病，相反总是那么灵动恣肆，神采飞扬，生机盎然。他的图像，看看清晰看看模糊，看看写实看看写意，让你不断交替戴上近视镜和望远镜，视觉总是处在交换的过程中。我们很难用一种‘主义’来概括他的画风。你可以说他写实，但他总是变形，年轻时扎实的造型基础，使他的变形总是那么恰到好处地鲜明准确传神，他的灵魂，时常在一个深邃广阔的背景下深思游荡。平民气质和贵族趣味在他身上和谐地得到了统一。””；

“2009 年作温柔的坚持布面油画”；

“创作于 1995 年的作品《抽象》亦是这一时期善用白色经典之作，同时朱德群采用了他最为钟爱的色调，以多样化的青色搭配蓝绿黑黄。艺术家亦曾说过，蓝色让他联想到海洋的神秘气质、变化万千的形态，时而静默、时而狂啸，也为跃动的画面增添沈潜大气的风范。同时作品笔触婉转，饱含中国书法性的线条，勾勒牵曳，使色彩有了流畅的韵律感，飞舞的色块犹如音乐流淌，轻盈醉人，恰好形成视觉印象里静与动的对比，亦是朱德群创作所表现的主要特色。朱德群并未大片地将青与蓝绿铺排于画面上，而是以大笔刷或长或短地挥动翻飞，表现出青色的空灵径走与蓝绿的沈静卓绝。画中光源仿佛来自未能探知距离的遥远处所，光线行而至此，并未将形与色构筑的空间全然劈裂击碎，而是如空气般隐动漂浮着。奔驰跳跃的深色点线是画中流动的能量，主导着观者视线游移的走向，深色渲染蹲踞在画面角落，让青绿色层的幻化不羁。而隐约透出黄色，借由观众对意象的认知，指引出光的行迹，与透视深远的延伸，并为色彩提供了游走的空间，是诗人寄托在隐喻中的自我观照，也是串联起平仄起伏的诗韵，精炼而不泛滥。”；

“《小演员》画面中少女肌肤所充溢的活力经由特定状态下的纯净观照，从而使她无限丰富性的形体得到了诗化的提升，所表现出来的则是另外一种近观所给出的人物写真，于是瞬间的思绪构成了整幅画面中充盈着的令人感动的美质，画面体现出来的平衡完整、松紧有序的生动结构，以及由大块的黑丝绒和皮肤质感的对比所透露出的气息，表明画家注重当下的感知，追求在古典画风中渗入现实感受的因素。”；

A.1.3 输入：温柔的女性化的感觉

“2009 年作温柔的坚持布面油画”；
“1993 年作温柔的光布面油画”；
“1997 年作温柔的风布面油画”；
“1989 年作温柔的音乐油彩画布镜框”；
“2009 年作你的温柔布面油画”；
“2003 年作温柔的光布面油彩”；
“2003 年作温柔的光布面油彩”；
“2007 年作下午温柔的阳光布面油画”；

“居于大地的当代生存者正在描摹一幅高度机器化的是日常生活图景，已经取得的物质成果并没使人们获得更多的幸福和精神自由。人们陷身于自己的创造物中并被其所奴役。丁方为我们提供了新的视觉经验：广阔空无的黑暗，在我们的体内发出回声，一个恍惚而苍白的飞翔人形在浊浪翻滚的昏暗中挣扎飘荡，所有鲜活的生命都被这一片混沌无情遮没着，即沉重工业的压迫之力逼使人们集结突围的勇气和力量，抵抗和承担由恐惧所预感到的否定性（《我逃向何方？》）；一条没有来处和去处的褐色河流挟裹着颓废的建筑残体和无知无觉的肉身，面对这种毫无反抗的挣扎努力的自溺情形，绝望感几乎要漫过我们的头顶（《罪灵们》）。此时，坚信者的呼告之身形迎面而来；精砾的铅色背景上，一个巨大的鸟形幻像，犹如折翼的天使，正飞向黄昏中模糊的地平线，羽翅上青冽的光芒耀亮即将陷入昏暗的荒野（《折羽的天使》、《请退让吧，忧郁的阴影》）。这些画幅突破了艺术的唯美法则的宥限而均突出了其思想的锐利；神圣关怀并没有抛离人，只要人愿意，就在此时此刻，他会来与我们相遇，以其苦痛搀扶人的苦痛。光影。古典的光影效果重新返回丁方的画面，光对画面的统领使其显示出蕴于悲剧深处的庄严性。丁方画面中精神性光影的设置，作为神圣俯临此世的标志，它象征着叹息、救扶、怜悯、启明、爱等精神内涵，它不是来自熟常的经验世界，它温抚着画面肌理的紧张断裂，呼应着上升的精神，漫渗到流幻的色彩深处，洞开灵魂的奥秘。它在显示生命的悲剧况味的同时，“给予灵魂以永恒的召唤”。在这种精神之光的携领下，丁方的画语世界凸现出几个非常重要的意象造型：以黄土高原、荒野自然为原型的大地意象。作为众生灵终其一生、存活于其上的大地，其深邃、厚重的自然存在形态在丁方的画面中取得超拔于日常感觉经验的独语性：奔泻而下的色流形成大地强烈的动势，坚岩般的肌理沉凝滞重，深含着忍耐、重负和强劲的生存精神。”；

“陈擎耀历年来的数字合成影像作品，看似皆以自我扮演的虚拟行动，搭配对象收集、挪用，或自行制作的过程，以时空跳跃并置与角色扮演错置等手法，用轻松诙谐甚至是无理嬉闹的姿态，拼凑并组合出我等凡人看似熟悉却又被抽离的时空情境，进行一种自我调侃的《自娱娱人》，以颠覆被社会礼教所制约的道德规范。更深入一层剖析，艺术家藉由角色装扮甚至是场景的虚拟建构，彷彿进行着一出全球化消费体系中，不断挪用符号、弱化历史感与刺激消费的某种文化拼图游戏，藉此拼凑出消费体系残存脑海中各式被灌输的想象景象，从中追寻也实践著作者心中的欲望草图；画面中充满荒谬与对立矛盾之角色扮演，宛如一场游戏，一个自我价值感所建构的心理催眠术。什么次文化或后殖民等庞大议题，在整个拍摄行为中已不重要，取而代之的是某种不需负责任、不需投入深刻情感的装腔作势，只要快速将情绪投入当下的扮演角色中，便能获得某种抽离于这个世界的快感。但就算人类历史命运这个巨大的命题对陈擎耀而言，充其量不过是另一种形式的《蒙太奇》，一个比鸿毛还轻的玩世不恭，不过仍可从其作品中读出对台湾社会奇观的另类思考，放大至当下台湾被欧美日文化殖民的处境反思，我们所认知的《他者》(other)，除了是由媒体片面所建构的绮想之外，也含有某种曲解成分与破碎断裂的文化意涵。不可否认地，在台湾目前的浅碟文化现况上，透过新作《天工开物》

的黑色喜剧般戏谑性，除了搬演了外籍劳工来台打工的狂想曲外，也同时点出了在消费廉价劳力的同时，也消遣着在这个过度天工开物（《物》生自天，《工》开于人，以天工为基础，顺应自然而制造出有利用价值的东西，才是人类技术存在的意义）的全球化消费主义时代，所不可避免因着大量生产所产生出的文明残渣，及其所衍生出来的阶级意识价值，应从更为人性的观点切入，实为本系列的言外之意。””；

“刘海粟这个时期的作品有一种较劲和任性的情绪，与日本第一代“准印象派”艺术家的作品相比，他自觉的东方性格已经完整地呈现出来，甚至是近乎偏执的强调这种性格。他的写生作品与日本前辈画家黑田清辉和藤岛武二的相比，更具书写的自由和流畅感，情绪更为敏感、张扬，色彩更为刺激、强烈，他自觉的弱化了色层堆叠和塑造感，以柔性的笔调书写出阳刚的气质，他完成的语言难度是巨大的——以柔致刚，以简驭繁，以轻薄而致厚重。在他 1930 年代的代表作中，我们可以看到他明确而自觉的语言策略。一个优秀的艺术家，策略自觉、成熟的标志就是——他将那些难以结合的矛盾双方确立为自己的表现范畴，将相反者结成紧密的同一，这就是艺术命题的方向，天才往往就是在那些难以逾越鸿沟间自由腾挪的人。从文化背景上来看，刘海粟当年也是承受着双重的文化压力，反者道之动，最对立的矛盾双方往往是触发天才属性的直接动因，只有那些能主动挑战并立志于战胜这些截然对立矛盾的人，才可能创造出历史的奇迹——不可思议的难度是历史奇迹的前提！刘海粟之所以伟大，在于他克服甚至颠覆了中国传统文人的内向、惰性、柔美的审美趣味和文化性格，但又将文人细腻、敏感、神秘的趣味特征糅合进了全新的表现形式之中，结成了相反者之同一；同时，他颠覆了当时国内以及整个东方艺术领域对客观再现的迷恋——他在这两种主流美学趣味的压力之下寻求突围的通道，以非常态的学习方式和创作风格如一匹黑马腾空而起，成为时代的英雄。他 1930 年代创作颠峰的作品基本致力于转化后期印象派的语言和形态，具有超前的实验精神，比日本梅原隆三郎 1930 年代晚期的中国写生，更有一种从容的气度和灵敏的语言才华。”；

“岳敏君所采取的特殊技法，使他的画面看上去通常都很简单、明快，甚至还有点艳俗、有点单调。这些简单的画面让人们把注意力全部集中在作品所表达的深意上，《大烟囱》也是这样。画中的形象很简单，只有一个用渐变的灰蓝色画成的巨大烟囱，占据了画面的大半空间，一些红脸孔的大笑人，一轮金黄色的太阳，画面构图基本上是对称式的，没有太多的节奏和变化，空间深度也十分有限，看上去就像一张平面招贴。但这简单的画面，却具有很强的视觉冲击力，也许正是因为它简单，才能将它的全部力量都集中起来，对人们的视觉和观念造成巨大的冲击。画中的红色、黄色与蓝色尽管都被中性化了，其对比却依然十分强烈。那些显得有些僵硬的形体，恰恰传达出了政治的无情与生硬、人的灵魂的麻木不仁。”；

“受马蒂斯的影响，闫平的画面中，表达观念时，依靠的是色彩而不是传统写实中的光与影，体现光线的感觉是通过色彩的对比与排列来表现的，而不仅仅是描绘形体。从画面的色彩来看，戏剧的服饰能够充分彰显出色彩的艳丽，乍看张扬，细看则处处小心，如同喜剧演员台上台下判若两人，色彩的叙述有一定的私密性能让她尽情地表达内心世界。颜色是外在的，而内心要表达的情绪置于其中则具有保护性、自由性和私密性。在此画中，我们不仅可以看到有后印象派的影响，还能看到对中国传统艺术文化的传承，画面中融会贯通了西方的理性秩序和中国的意向表现，在理好的色系画面上，加上特立独行的焦墨，恰到好处。纳比派的主观表现给了闫平以启迪，加强了用色的浓度和画面的装饰性处理，使色彩相互交织，对比与协调并存。画面中，闫平大胆运用强有力的笔触进行塑造，由薄到厚，由于作品中用笔交错相叠，产生了一些犹如砾石般的肌理，使得画面的表现更加丰富。从画面的构图来看，她把握大局，对于画面的四边四角和主体物之间的关系的处理非常考究。这里面包括：画面的分割，每一个线条的长短，每一个方圆的比较，画面中疏密的安排以及造型方式的扩展等等。”；

“这种自觉的留白和色彩的简括处理是一种新的绘画性格，它是油画形态发生改变的重要端倪，它必定被后人发现和珍视，这种新的端倪预示了在西方光色和结构两大语言谱系之外，油画语言开放的另一通道。刘海粟通过书写与塑造的重新匹配，油画被另一种民族智慧所把握，补充和开放了油画这一古老艺术样式的“音域”，这就是贡献，对绘画表达和视觉审美经验的双重贡献。这是一个东方少年天才自悟自通的发明。《圣扬乔而夫飞瀑》无论是作品形态和完整度还是语言的成熟度，都已经达到那个时代的最高水平。在一定意义上，这是被逼迫和放纵出来的才情，任性和要强的情绪通过书写的动态挥策而出，如瀑布流泉一样奔涌、流淌出来，笔下的物象已经不是静态的客观物象，更像是一个充满

热情和煽动感的视觉宣言，预示着观看习惯和表现陈习的巨变。尽管这个巨变被中断、覆盖，在历史视野中消逝了近八十年，今天让我们看见它，就如蓦然看见了旷代知音，本能般萌生一种把臂欲言的冲动。”；

“第二年，我去了北京，在那里整整住了一个冬天，窗外飘着大雪。味道不佳的室内，暖烘烘的，在这里我画了一批工厂和管道与植物交织的作品，牧歌式的景观被分裂和忧郁的超现实图式所代替，这一个是著名的 85 美术思潮涌动的时候，超现实主义的大师达利、契里柯的追随者在中国的这页美术史中大放异彩，与注重哲学思考和文化批判的理性思考不同的是：我是由纯粹个人经历和感觉自然的过渡到超现实语境的。我喜欢已沉迷于文化风暴而心灵脆弱的内心为表达主题：因为想太多而行太少，因为书读得过多而体会过深，人格失调、精神分裂、失意、堕落或彷徨接踵而来。《离开和留在最后一块草地上的两个人》、《屋外的马窥视的她和被他端视的我们》、《最后的花园》、《听见毕加索马叫的隐者和鸟》、《春天唤醒冬眠者》是这时的作品。虚幻失真的田园憧憬，平庸沉闷的现实生活、虚假的文化表象与躁动的思潮迷雾之中，是绘画使我解放出来，为我展开一个新的境界。一切静止不动，一成不变的东西都开始颤栗了。不只是被人赋诗的星星、月亮、树林、花草，甚至一团发电厂的烟灰，一张被风刮到街角的报纸，一块不知所以然被蚂蚁从草丛里咬出的树皮，一座锈蚀的、停摆的座钟……所有的一切都显露给我一张自己的脸，它内在的生命，这些沉默的心灵。经常无声胜有声，无论静或动的点或线对我都是这般生动，这就足够使我投入整个生命感觉去体会艺术的可能性和它的生命。自然地将自己的梦境表现在画布上，并且化痛苦为诙谐，化诙谐为抒情，像鱼一样，流连在自己的幻想中，那时，对我而言，绘画几乎是人与动植物和工厂之间的一种混合体。”；

“对世界独特角度的观察、真诚和同情的气质、与自在的无聊无奈的心情构成了更深的矛盾；对普通人生活的人性的感受、对生活常态的表述和对表述本身的痴迷使这个没心没肺的世界变得不可知和神圣起来。”；

“我历年来的数字合成影像作品，皆以自我扮演的虚拟行动，搭配对象收集、挪用，或自行制作衣服、道具的过程，以时空跳跃并置与角色扮演错置等手法，用轻松诙谐甚至是无理嬉闹的姿态，拼凑并组合出一般大众看似熟悉却又被抽离的时空情境，进行一种自我调侃的“自娱娱人”行为，以颠覆被社会礼教所制约的道德规范。我想藉由角色装扮甚至是场景的虚拟建构，进行着一出全球化消费体系中，不断挪用符号、弱化历史感与刺激消费的某种文化拼图游戏，藉此拼凑出消费体系残存脑海中各式被灌输的想象景象，从中追寻也实践着可能潜藏在我心中的欲望草图；在我的作品画面中充满荒谬与对立矛盾之角色扮演，宛如一场游戏，一个自我价值感所建构的心理催眠术。次文化或后殖民等庞大议题，在整个拍摄行为当中已显得不重要，取而代之的是某种不需负责任、不需投入深刻情感的装腔作势，只要快速将情绪投入当下的扮演角色中，便能获得某种抽离于这个世界的快感。人类历史命运这个巨大的命题，充其量不过是另一种形式的“蒙太奇”，一个比鸿毛还轻的玩世不恭。”；

“以后的日子里，新潮那种太多太杂乱的文化负荷使当时的创作状态难以为继。绘画被淹没在文字和哲学的定义和图解中。我那时的一张油画《奔逃者》在创作中，令我困惑的是形象已无法恢复纯净的本来原面目。大而空的文化指标下，绘画——那最高灵性活动的象征，已一文不值冷落在一个无人过问的隅角。对于那时的艺术家，意义的追问已经沦为一种附会、说教或者一种宣泄，属于人性不健全的无节制的快感。面对快速变换的现实，需要繁复的感情和思想，更需要特定的语言表现这个世界。需要感觉体味揉合我所需要的人生一致的真淳；或者悲壮，成为时代的讴歌；或者深邃，成为灵魂的震撼。在它所有的要求中，对于我自身，如今最先满足的，不是前期浪子式情感的挥霍，而是绘画感觉本身。应该有一种形式，那形式与我们面对真实的周遭现实时引起的冲动是合拍的，在那里，语言、文字、推理、心理分析都噤然无声。一些偶然的机会，我开始迷恋上水墨和丙烯。经常地、不断地在纸上涂画，总使我沉浸在这种状态中，它使我渐渐中止对琐碎外表的观察和摹仿，而把我自身投入到同观察本身同样有趣，但又更觉动人的，由全部有趣的事物所给予的印象的综合想象之中去。因此，它所产生的景象是一种和我们命运想象的，一种悲哀与甘美混合在一处的东西。”；

“那年冬天，我经常去美术馆看一个接一个的展览，劳申伯的大型展给我不同与往的震撼。它虽不像其它我所看到的《韩默藏画展》、《法国 250 年藏画》、《德国表现主义画展》、赵无极、戈雅、柯尔韦尔个展等古典和现代主义那种对自我情节的顶礼膜拜具有令人驻足叹服，品评赏析的细节，但却使我第一次强烈地感受到当代艺术直接切入周遭生活的力量。这与我几年来接受的为艺术而艺术的教育很不相同，这也使我先前即已产生的对“纯美术”意义的质疑更得到证

明。1986 年朋友们同我一起发起和组织的《西南艺术群体》，使我更直接投身于“新潮美术”的潮流中。我们彻夜地谈论种种宣言和观念，策划在昆明、上海、重庆和北京的展览，在各种报刊上撰写文章，对那些同样激动得面红耳赤的人解释何为“新具像”——那的确是一个冲动的年代，经历过八十年前期文化洗礼的人们，例如我本人，总是难以忘怀那些热闹的场面。虽然那时的文化存在诸多的夸张，自以为是和种种谬误，但是，那毕竟是一个在文化上有追求的年代。毛旭辉永远紧锁着眉头，在和平村的小屋中挤满了《他的红色体积》和其它许多的作品，他的东西令人压抑、难受。与此相反他却能写一手漂亮的激扬文字。潘德海总是结结巴巴地发明一些惊人的词汇，除了画《杂乱的、杂乱的》形象，还用泥巴做了一大堆黑不溜秋的怪物。张晓刚则一直有点善感和见异思迁；胃病住院就画些痛苦的痉挛不堪的《白色幽灵》；搬到桃花山又显得花花草草地吟诵《遗梦集》……今天，当年年轻气盛的他们都已一个个成为了不起的画家了。那个时候，我在医院里一面等待女儿的出生，一面写《自然意识略述》和《丛林猛兽》。在省图书馆的展览开始时，朋友们在请柬上写满了对未出生女儿的祝福……同当时的许多知识分子一样，我们怀有真实的人文主义理想，虽然今天我的同辈们大多数未必赞成那种理想，然而在情感上却仍又存在深深的怀念。”；

“2011 年作戴银饰的苗族女人布面油画”；

A.1.4 输入：冷静的男性

“新女性布面油画”；

“新女性布面油画”；

“2004 年作冷静油彩画布”；

“喻红在日常存在中寻找到一种诗意。她描述我们每天都曾经历过的简单的喜悦，从和朋友开怀大笑，到静坐沉思，到陷入爱情和操持家务。她关注对个人的演绎可以追溯到在北京中央美术学院油画系读本科时所接受的彻底的写实主义技巧训练。她的早期作品将高度写实的肖像与非现实的环境与色彩相结合，展示出一种来源于现实世界的错位感。渐渐的，生活的安定反映在作品中，她抛弃了其艺术中超现实的策略。同时，用心磨砺观察力，使她在刻画人物面部表情和肢体语言方面表现出了敏锐的才能。喻红大多数作品的题材是女性。在中国美术史中很少见到以理解和有节制的方式描述女性观点的作品。一般来说，早期表现女性日常生活的作品多由男性创作，他们用没完没了的象征为题材镀金。他们即不关心女性本身，也不在意女性观点。现在，中国当代女性艺术家热衷于女性题材，她们表现的形象带有沉重的象征性，因而失去了独立的个体性。喻红尊重人生从儿童到成熟各不同阶段的个体价值（林似竹）”；

“2006 年作新女性布面油画”；

“陈擎耀历年来的数字合成影像作品，看似皆以自我扮演的虚拟行动，搭配对象收集、挪用，或自行制作的过程，以时空跳跃并置与角色扮演错置等手法，用轻松诙谐甚至是无理嬉闹的姿态，拼凑并组合出我等凡人看似熟悉却又被抽离的时空情境，进行一种自我调侃的《自娱娱人》，以颠覆被社会礼教所制约的道德规范。更深入一层剖析，艺术家藉由角色装扮甚至是场景的虚拟建构，彷彿进行着一出全球化消费体系中，不断挪用符号、弱化历史感与刺激消费的某种文化拼图游戏，藉此拼凑出消费体系残存脑海中各式被灌输的想象景象，从中追寻也实践著作者心中的欲望草图；画面中充满荒谬与对立矛盾之角色扮演，宛如一场游戏，一个自我价值感所建构的心理催眠术。什么次文化或后殖民等庞大议题，在整个拍摄行为中已不重要，取而代之的是某种不需负责任、不需投入深刻情感的装腔作势，只要快速将情绪投入当下的扮演角色中，便能获得某种抽离于这个世界的快感。但就算人类历史命运这个巨大的命题对陈擎耀而言，充其量不过是另一种形式的《蒙太奇》，一个比鸿毛还轻的玩世不恭，不过仍可从其作品中读出对台湾社会奇观的另类思考，放大至当下台湾被欧美日文化殖民的处境反思，我们所认知的《他者》(other)，除了是由媒体片面所建构的绮想之外，也含有某种曲解成分与破碎断裂的文化意涵。不可否认地，在台湾目前的浅薄文化现况上，透过新作《天工开物》的黑色喜剧般戏谑性，除了搬演了外籍劳工来台打工的狂想曲外，也同时点出了在消费廉价劳力的同时，也消遣着在这个过度天工开物(《物》生自天，《工》开于人，以天工为基础，顺应自然而制造出有利用价值的东西，才是人类技术存在的意义)的全球化消费主义时代，所不可避免因着大量生产所产生出的文明残渣，及其所衍生出来的阶级意识价值，应从更为人性的观点切入，实为本系列的言外之意。””；

“我历年来的数字合成影像作品，皆以自我扮演的虚拟行动，搭配对象收集、挪用，或自行制作衣服、道具的过程，以时空跳跃并置与角色扮演错置等手法，用轻松诙谐甚至是无理嬉闹的姿态，拼凑并组合出一般大众看似熟悉却又被抽离的时空情境，进行一种自我调侃的“自娱娱人”行为，以颠覆被社会礼教所制约的道德规范。我想藉由角色装扮甚至是场景的虚拟建构，进行着一出全球化消费体系中，不断挪用符号、弱化历史感与刺激消费的某种文化拼图游戏，藉此拼凑出消费体系残存脑海中各式被灌输的想象景象，从中追寻也实践着可能潜藏在我心中的欲望草图；在我的作品画面中充满荒谬与对立矛盾之角色扮演，宛如一场游戏，一个自我价值感所建构的心理催眠术。次文化或后殖民等庞大议题，在整个拍摄行为当中已显得不重要，取而代之的是某种不需负责任、不需投入深刻情感的装腔作势，只要快速将情绪投入当下的扮演角色中，便能获得某种抽离于这个世界的快感。人类历史命运这个巨大的命题，充其量不过是另一种形式的“蒙太奇”，一个比鸿毛还轻的玩世不恭。”；

“刘海粟是要强的人，有强烈的自我意识和表现欲望，他的表现欲之强烈，恐怕在中国第一代现代主义艺术家中是可以考第一名的。如果我们要突破传统的美术嫡传谱系来寻找中国现代艺术的先锋，无疑刘海粟足以荣膺这个桂冠。他富有传奇色彩和非常态的学习和创作经历，以及对大文化的主动影响、导向，为美术史贡献了一种新的性格和艺术致力方向。他颠覆了嫡传谱系的美术史论述，将艺术的天才属性和个性特质上升到了新的历史命题和策略的高度，为此，当代的我们，对未来和未知的好奇与兴奋，远比依靠传统记忆而获得的安全感更具有吸引力，新生的艺术力量因此而显得年轻而有活力。中国的当代艺术性格乃至文化性格的巨大转变并非偶然——尽管被中断、覆盖了很长时间，历史的一维性仍冷峻地证明着，天才属性是艺术的第一属性，这种属性的价值回归将使我们当代中国艺术性格不再温驯、内向、惰性，相反，她已经秉具了进取的勇气和机敏、思辨、开放的品质。（文/漆澜）”；

“1989年北京的现代艺术大展标志着80年代的终结，美术馆一楼诸多的标新立异的实验和哗众取宠的新闻闹剧使我觉得江郎才尽的疲软，肖鲁的一声枪响，更使我所在的二楼东厅的古典现代主义作品暗然失色。然而花样翻新并未迎来文化复兴的时代，相反，它却耗尽了新潮美术最后一点革命的热情和创新的想象力——当我对八十年代后期直到九十年代初期的文化现状进行抨击时，我不仅陷入矛盾的困境，而且陷入一种深深的，近乎感伤的忿恨。八十年代的知识分子主流文化对我这个年龄层的人来说，一直是一个巨大的舞台，我们怀着兴奋的心情观赏，偶尔参与，随后却在努力等待这个伟大神话破产。八十年代后期，大而无当的文化批判解体的实现，其实也为这辈人和后来者在文化上新的冲锋陷阵提供了一个战场。然而，文化的溃败是如此的迅速而彻底，历史没有如期提供一个底盘，而且根本就没有那样一个底盘。我们不得不在批判八十年代神话内容的同时，却又怀念那样一种文化形式——文化的历史机遇和历史实践的方式。八十年代已经颓然死去，留下的不仅是一大堆问题，而且还是一种回忆。”；

“刘海粟这个时期的作品有一种较劲和任性的情绪，与日本第一代“准印象派”艺术家的作品相比，他自觉的东方性格已经完整地呈现出来，甚至是近乎偏执的强调这种性格。他的写生作品与日本前辈画家黑田清辉和藤岛武二的相比，更具书写的自由和流畅感，情绪更为敏感、张扬，色彩更为刺激、强烈，他自觉的弱化了色层堆叠和塑造感，以柔性的笔调写出阳刚的气质，他完成的语言难度是巨大的——以柔致刚，以简驭繁，以轻薄而致厚重。在他1930年代的代表作中，我们可以看到他明确而自觉的语言策略。一个优秀的艺术家，策略自觉、成熟的标志就是——他将那些难以结合的矛盾双方确立为自己的表现范畴，将相反者结成紧密的同一，这就是艺术命题的方向，天才往往就是在那些难以逾越鸿沟间自由腾挪的人。从文化背景上来看，刘海粟当年也是承受着双重的文化压力，反者道之动，最对立的矛盾双方往往是触发天才属性的直接动因，只有那些能主动挑战并立志于战胜这些截然对立矛盾的人，才可能创造出历史的奇迹——不可思议的难度是历史奇迹的前提！刘海粟之所以伟大，在于他克服甚至颠覆了中国传统文人的内向、惰性、柔美的审美趣味和文化性格，但又将文人细腻、敏感、神秘的趣味特征糅合进了全新的表现形式之中，结成了相反者之同一；同时，他颠覆了当时国内以及整个东方艺术领域对客观再现的迷恋——他在这两种主流美学趣味的压力之下寻求突围的通道，以非常态的学习方式和创作风格如一匹黑马腾空而起，成为时代的英雄。他1930年代创作巅峰的作品基本致力于转化后期印象派的语言和形态，具有超前的实验精神，比日本梅原隆三郎1930年代晚期的中国写生，更有一种从容的气度和灵敏的语言才华。”；

“受马蒂斯的影响，闫平的画面中，表达观念时，依靠的是色彩不再是传统写实中的光与影，体现光线的感觉是

通过色彩的对比与排列来表现的，而不仅仅是描绘形体。从画面的色彩来看，戏剧的服饰能够充分彰显出色彩的艳丽，乍看张扬，细看则处处小心，如同喜剧演员台上台下判若两人，色彩的叙述有一定的私密性能让她尽情地表达内心世界。颜色是外在的，而内心要表达的情绪置于其中则具有保护性、自由性和私密性。在此画中，我们不仅可以看到有后印象派的影响，还能看到对中国传统艺术文化的传承，画面中融会贯通了西方的理性秩序和中国的意向表现，在理好的色系画面上，加上特立独行的焦墨，恰到好处。纳比派的主观表现给了闫平以启迪，加强了用色的浓度和画面的装饰性处理，使色彩相互交织，对比与协调并存。画面中，闫平大胆运用强有力的笔触进行塑造，由薄到厚，由于作品中用笔交错相叠，产生了一些犹如砾石般的肌理，使得画面的表现更加丰富。从画面的构图来看，她把握大局，对于画面的四边四角和主体物之间的关系的处理非常考究。这里面包括：画面的分割，每一个线条的长短，每一个方圆的比较，画面中疏密的安排以及造型方式的扩展等等。”；

“2011 年作戴银饰的苗族女人布面油画”；

“不过站在美术史的角度上来看，从文艺复兴到十八世纪末期对油画技法的争论和改革是相当激烈的，因为当时的画家几乎在同类型的绘画模式下创作，也就是说，采取的是古典写实手法。那么油画技术的高低和改革决定了画家作品优劣，以及在有限范围内的个人艺术风格的突现程度。在十九世纪大部分时间里对古典油画技法的革命依然有空间，这点我们可以从新古典主义、浪漫主义、写实主义画家的作品中进行比较而得知。但这些依然是建立在古典写实之上的。直到印象派之后，人们从精神上到绘画手法上都背离古典写实的宗旨，画家们各自在所建立的绘画样式的支持下，对自己的绘画技术进行研究和发展，以便有效地表达思想，体现对时代的关注（当然这里也包括“借用”古典写实手法的画家）。那么，在这样宽松多样化的艺术氛围内，还继续纠缠对古典油画技法接受还是排斥问题，并同现代还是保守同时牵扯到一起去，就没什么意义了。每个艺术家的学术背景不一样，艺术气息也不一样，他们有权力根据个人的情况选择适合的艺术形式和手法。如果把追随时尚风格当作评判一件作品是否具有当代性的标准，无疑是不妥的。”；

“居于大地的当代生存者正在描摹一幅高度机器化的是日常生活图景，已经取得的物质成果并没使人们获得更多的幸福和精神自由。人们陷身于自己的创造物中并被其所奴役。丁方为我们提供了新的视觉经验：广阔空无的黑暗，在我们的体内发出回声，一个恍惚而苍白的飞翔人形在浊浪翻滚的昏暗中挣扎飘荡，所有鲜活的生命都被这一片混沌无情遮没着，即沉重工业的压迫之力迫使人们集结突围的勇气和力量，抵抗和承担由恐惧所预感到的否定性（《我逃向何方？》）；一条没有来处和去处的褐色河流挟裹着颓废的建筑残体和无知无觉的肉身，面对这种毫无反抗的挣扎努力的自溺情形，绝望感几乎要漫过我们的头顶（《罪灵们》）。此时，坚信者的呼告之身形迎面而来；精砺的铅色背景上，一个巨大的鸟形幻像，犹如折翼的天使，正飞向黄昏中模糊的地平线，羽翅上青冽的光芒耀亮即将陷入昏暗的荒野（《折羽的天使》、《请退让吧，忧郁的阴影》）。这些画幅突破了艺术的唯美法则的宥限而均突出了其思想的锐利；神圣关怀并没有抛离人，只要人愿意，就在此时此刻，他会来与我们相遇，以其苦痛搀扶人的苦痛。光影。古典的光影效果重新返回丁方的画面，光对画面的统领使其显示出蕴于悲剧深处的庄严性。丁方画面中精神性光影的设置，作为神圣俯临此世的标志，它象征着叹息、救扶、怜悯、启明、爱等精神内涵，它不是来自熟常的经验世界，它温抚着画面肌理的紧张断裂，呼应着上升的精神，漫渗到流幻的色彩深处，洞开灵魂的奥秘。它在显示生命的悲剧况味的同时，“给予灵魂以永恒的召唤”。在这种精神之光的携领下，丁方的画语世界凸现出几个非常重要的意象造型：以黄土高原、荒野自然为原型的大地意象。作为众生灵终其一生、存活于其上的大地，其深邃、厚重的自然存在形态在丁方的画面中取得超拔于日常感觉经验的独语性：奔泻而下的色流形成大地强烈的动势，坚岩般的肌理沉凝滞重，深含着忍耐、重负和强劲的生存精神。”；

“2006 年作现代女性油画面布”；

“最显著的是 911 事件后，城市、霸权、安全成为陈劭雄作品中的话题。在他参加上海双年展的《花样反恐》（2002 年）和 2003 年个展作品《BBS - 如果他们来了》中，艺术家以一种虚拟的方式调侃着城市人的安全感。他假设大楼是可以躲避飞机的，假设任何风景都是不安全的，同时也以幽默的方式提出自己的“反恐方案”。这些作品的玩笑实在过火，以至于英国著名的费顿出版社在去年挑选了陈劭雄作为过去两年最重要的 100 个艺术家时，故意回避了这件作品。其实陈劭雄试图说明他的作品在政治上，表达的是旁观者的角度，而其本质在于表明在公共空间的城市与政治交锋的极

端手段中，安全感和不安全感其本质都源于“幻想”。可怕的不是不安全感，而是用这种幻想去改变现实。”；

“《男男女女》创作于2006年，是刘小东现实主义题材的又一力作。长达5米的画面横躺着五男三女，他们半裸或全裸，中间以衣物隔开，具有很强的叙事感。画面人物的身份性并不明晰，男性或为下工休息中的民工，女性或为对身体不以为然的小姐，作者注重的不是人物的具体所指，画面整体有一种电影般的感觉和戏剧化的效果。”；

“2003年作当代女性布面油画”；

“2003年作新女性布面油画”；

“1999年作理智与情感油画画布”；

A.1.5 输入：冷静的男人

“2004年作冷静油彩画布”；

“1995年作大眼睛男人布面油画”；

“2011年作戴银饰的苗族女人布面油画”；

“2007年作红色男人布面油画”；

“2006年作男人像油画”；

“清男人像纸本油画”；

“男人像油画画布”；

“男人像油画画布”；

“2008年作男人布面油画”；

“两个男人纸本油画”；

“2005年作男人肖像布面油画”；

“男人的故事”；

“男人布面油画”；

“男人布面油画”；

“2007年作红色男人油彩布面”；

“1999年作男人与马布面油画”；

“喻红在日常存在中寻找一种诗意。她描述我们每天都曾经历过的简单的喜悦，从和朋友开怀大笑，到静坐沉思，到陷入爱情和操持家务。她关注对个人的演绎可以追溯到在北京中央美术学院油画系读本科时所接受的彻底的写实主义技巧训练。她的早期作品将高度写实的肖像与非现实的环境与色彩相结合，展示出一种来源于现实世界的错位感。渐渐的，生活的安定反映在作品中，她抛弃了其艺术中超现实的策略。同时，用心磨砺观察力，使她在刻画人物面部表情和肢体语言方面表现出了敏锐的才能。喻红大多数作品的题材是女性。在中国美术史中很少见到以理解和有节制的方式描述女性观点的作品。一般来说，早期表现女性日常生活的作品多由男性创作，他们用没完没了的象征为题材镀金。他们即不关心女性本身，也不在意女性观点。现在，中国当代女性艺术家热衷于女性题材，她们表现的形象带有沉重的象征性，因而失去了独立的个体性。喻红尊重人生从儿童到成熟各不同阶段的个体价值（林似竹）”；

“1997年作同志系列：男人布面油画”；

“2006年作红色男人像油彩画布”；

“2001年作男人·女人布面油画”；

A.1.6 输入：快乐的儿童

“1997年作快乐的儿童第9号油画画布”；

“2013年作快乐的孩子们布面油画”；

“新中国的儿童（一张）纸本”；
“儿童布面油画”；
“儿童油画”；
“2004 年作儿童时代油画面布”；
“快乐 1 油画”；
“1997 年作为小孩快乐油画面布”；
“2001 年作快乐世界布面油画”；
“少女与儿童布面油画”；
“2005 年作快乐在一起，2 号油画面布”；
“快乐的童年布面油画”；
“1999 年作快乐的红色油画面布”；
“1992 年作快乐油画面布”；
“游戏的儿童布面油画”；
“快乐的童年油画面布”；
“威尼斯的快乐”；
“2002 年作游戏的儿童 9 号油画面布”；
“2002 年作游戏的儿童 8 号油画面布”；
“2005 年作快乐组合布面油画”；