

分类号: _____

学校代码: 11078

学 号: 2111906078



密 级: _____

保密日期: _____

保密期限: _____

广州大学

硕士学位论文

题目: 交互式教育资源可视化预览研究

学 生 类 型: ☐ 学术学位硕士
☒ 专业学位硕士

学 位 申 请 人: 王剑雄

导师姓名及职称: 张景中 教授

学 院: 计算机科学与网络工程学院

学科专业 (领域): 计算机技术

论 文 提 交 日 期: 二〇二二年六月

广州大学学位论文原创性声明

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究作出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

学位论文作者签名：王剑雄

2022年6月1日

广州大学学位论文版权使用授权书

本学位论文作者完全了解广州大学有关保留、使用学位论文的规定，即：研究生在校攻读学位期间论文工作的知识产权单位属广州大学。学校有权保留并向国家主管部门或其指定机构送交论文的电子版和纸质版，允许学位论文被检索、查阅和借阅；学校可以公布学位论文的全部或部分内容，可以允许采用影印、缩印、数字化或其他手段保存、汇编学位论文。（保密论文保密期满后，适用本声明）

本学位论文电子版和纸质版内容完全一致，属于：

☐ 保密，在 _____ 年解密后适用本授权书。

☒ 不保密。

（请在以上相应方框内打“√”）

学位论文作者签名：王剑雄

2022年6月1日

导师签名：张新

2022年6月1日

分类号:
学校代码: 11078
学 号: 2111906078

密 级:
保密日期:
保密期限:

广州大学硕士学位论文

交互式教育资源可视化预览研究

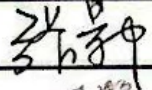
Research on Visual Preview of Interactive Educational Resources

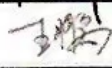
王剑雄

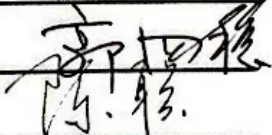
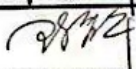
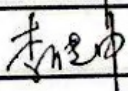
学科专业（领域）： 计算机技术

研 究 方 向： 教育资源预览

论 文 答 辩 日 期： 2022年5月27日

指导教师（签名）： 

答辩委员会主席（签名）： 

答辩委员会委员（签名）： 



摘要

教育对国家的发展起着举足轻重的作用，伴随着互联网技术的飞速发展，大量线上教育资源平台应运而生，使教育资源更易获得。在这些在线教育资源中，交互式教育资源因能与之交互带来的趣味性，倍受用户青睐。然而，与传统资源相比，交互式教育资源具有隐藏的交互信息和动态内容。只有当用户正确地执行交互操作时，才能获得资源所传递的知识，这给知识的获取增加了难度。同时，随着在线教育资源数量日益增加，如何帮助用户理解交互式教育资源的主要内容，并找到感兴趣的资源成为一个挑战。基于此，本文分析了用户交互理论与手动交互过程，提出了一种生成交互式教育资源可视化预览（Interactive Education Resource Visual Preview，简称 IERVP）的方法，并实现了网络画板资源的可视化预览。主要研究工作包括：

（1）提出一种生成交互式教育资源可视化预览（IERVP）的方法。首先，以用户交互过程相关理论为基础，分析用户与资源手动交互过程，提取手动交互过程重要信息。其次，对存储交互资源信息的文件进行分析，获取交互元素信息，提取约束关系，生成操作序列。再次，使用自动化软件工具执行操作序列，模拟用户与资源之间的交互过程，并保存过程中的截图。最后，提取截图的颜色直方图特征与边缘特征并组合这两种特征，使用 K-Means 方法进行聚类，获取截图中的关键图片，生成交互式教育资源的图片预览。

（2）基于 IERVP 方法，实现动态数学平台网络画板资源的可视化预览，并从节约时间比、预览正确性、预览完整性、预览反映内容程度四个维度对可视化预览效果进行测评。实验结果显示，该方法能够有效提取资源中的交互语义信息、模拟交互过程、提取关键图片，生成的图片预览平均节省用户 87.3% 的时间，帮助用户理解 88.6% 的资源内容，对用户理解交互式教育资源的内容有较大帮助。

关键词：交互式教育资源；资源预览；操作序列；网络画板

ABSTRACT

Education plays a pivotal role in the development of a country. With the rapid development of Internet technology, a large number of online educational resource platforms have emerged, making educational resources more accessible. Among these online educational resources, interactive educational resources are favored by users because of the interestingness brought by interaction with them. However, compared with traditional resources, interactive educational resources have hidden interactive information and dynamic content. The knowledge transmitted by the resource can only be obtained when the user performs the interactive operation correctly, which increases the difficulty of knowledge acquisition. At the same time, with the increasing number of online educational resources, how to help users understand the main content of interactive educational resources and find the resources they are interested in has become a challenge. Based on this, this paper analyzes the user interaction theory and manual interaction process, proposes a method to generate Interactive Education Resource Visual Preview (IERVP), and realizes the visual preview of the resources on Netpad. The main research work includes:

(1) A method for generating Interactive Educational Resource Visualization Preview (IERVP) is proposed. Firstly, based on the relevant theory of user interaction process, the manual interaction process between users and resources is analyzed, and the important information of the manual interaction process is extracted. Secondly, analyze the files storing the interactive resource information, obtain the interactive element information, extract the constraint relationship, and generate the operation sequence. Again, use automated software tools to execute the operation sequence, simulate the interaction between the user and the resource, and save screenshots of the process. Finally, extract the color histogram features and edge features of the screenshots and combine these two features, use the K-Means method for clustering, obtain key pictures in the screenshots, and generate image previews of interactive educational resources.

(2) Based on the IERVP method, realize the visual preview of the resources on a dynamic mathematics platform, NetPad, and evaluate the visual preview effect from four dimensions: saving time ratio, accuracy of the preview, completeness of the preview and content reflected

by the preview. The experimental results show that the IERVP method can effectively extract interactive semantic information in resources, simulate the interaction process, and extract key pictures. The generated picture preview can save users 87.3% of the time on average and help users understand 88.6% of the content of resources. The generated picture preview is of great help to the users to understand the content of interactive educational resources.

Keywords: Interactive Educational Resources, Resource Preview, Operate Sequence, NetPad

目 录

摘 要.....	I
ABSTRACT.....	III
第一章 绪论.....	1
1.1 研究背景及意义.....	1
1.2 国内外研究现状.....	3
1.3 研究内容及组织架构.....	6
第二章 用户交互相关研究与分析.....	11
2.1 用户交互过程理论基础分析.....	11
2.2 用户手动交互过程分析.....	13
2.3 本章小结.....	15
第三章 交互式教育资源可视化预览（IERVP）方法.....	17
3.1 交互语义信息提取.....	17
3.2 模拟交互过程.....	26
3.3 关键图片获取.....	28
3.4 关键图片组合.....	33
3.5 本章小结.....	33
第四章 基于 IERVP 方法的网络画板资源可视化预览.....	35
4.1 获取交互语义信息.....	36
4.2 获取交互过程图片.....	43
4.3 可视化预览.....	44
4.4 本章小结.....	45
第五章 实验与分析.....	47
5.1 实验.....	47
5.2 实验结果及分析.....	49
5.3 本章小结.....	53
第六章 总结与展望.....	55
6.1 研究总结.....	55
6.2 存在问题和不足.....	55

6.3 展望	56
参考文献	57
附录 问卷调查	63
攻读硕士学位期间取得的科研成果	67
致 谢	69

第一章 绪论

1.1 研究背景及意义

1.1.1 研究背景

公众普遍认为高质量的知识应该是公开的、容易获得的，教育资源应该是公平的。信息技术飞速发展^[1]，促进了教育与信息技术深度融合，教育信息化应运而生^[2]，促进了优质教育资源流动共享，是提高教育教学质量的重要途径之一。目前，国内外众多教育组织纷纷建设在线教育平台，促进了教育资源的大规模共建共享。然而，各种资源库都有不同的资源建设标准，这些标准并不统一且多数不公开。随着在线教育资源种类和数量的迅速增加，如何高效地浏览和理解资源的主要内容已经成为查找资源的一个巨大挑战，而资源预览能够反映资源的主要内容，因此已成为近几年一个热门的研究领域。

教育资源是网络教育平台的基础构成资源，包括传统教育资源和交互式教育资源。常见的传统教育资源包括 WORD、PDF、PPT、音视频等^[3]，满足了大部分教学需求，但在使用传统教育资源教学过程中，知识的传播方向一般是单向的，从教师方流向学生方，学生缺少参与感，使得知识获取过程不够生动有趣。而交互式教育资源不仅包含文本、图形、音频、视频等常见元素，还包含具有更复杂语义信息的交互元素。一般而言，交互元素包括按钮、滑竿和可拖动的对象等。交互元素可以增加用户的学习兴趣，减少其认知负荷，增强其对知识的理解。用户可以通过多个交互操作与这些元素进行交互，伴随着交互操作，资源能够可视化地展现动态内容，从而呈现出相关的知识和知识形成过程。

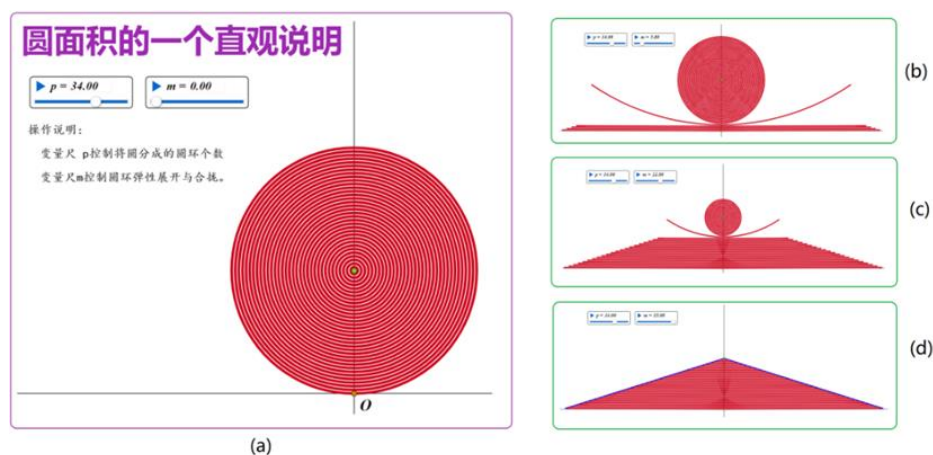


图 1-1 “圆面积的一个直观说明”教学课件

图 1-1 是一个由动态数学软件——网络画板制作的教学课件，课件名为“圆面积的一个直观说明”。在该课件中，点击变量 p 对应的按钮可以改变变量 p 的值，从而改变圆环的数目，当 p 值达到最大值时，圆环最密集，这些圆环可近似组成一个圆，点击变量 m 对应的按钮，可以使得变量 m 的值持续改变，从而动态连续地展示每个圆环拆分一条线、整个圆拆解为三角形的过程。在动态过程中，能够清晰地观察到圆面积等于三角形面积。从这个案例中可以了解到交互式教学案例存在的意义，它能够使得知识传播过程生动形象，能够帮助学生更加直观深刻地获取知识。

1.1.2 研究意义

1.1.2.1 用户意义

在交互式教育资源中，交互元素之间有特定的交互顺序，对交互元素按不同的顺序执行交互操作会产生不一样的动态内容，从而导致不同的知识呈现效果，这一方面会给用户带来交互困惑，导致用户对资源内容的理解存在障碍，影响资源的获取与使用，另一方面也对资源的语义信息提取带来了很大挑战，增加了资源可视化预览的难度。

在缺乏资源预览服务的情况下，用户需手动打开资源，观察资源的文本和图像等特征，还需正确执行操作顺序，才能准确理解资源的内容。从用户的角度来说，用户更希望用一种简洁友好的方式去快速了解资源内容。

因此，资源预览可以帮助用户快速了解资源内容，降低知识获取的门槛。

1.1.2.2 工程意义

交互式教育资源预览有利于开展资源其他相关工作，例如资源检索。资源检索的过程可拆分为特征提取和特征匹配两部分，经过多年的研究，基于相似度度量的特征匹配算法日趋成熟，已经取得很大的进展，因此，特征提取对资源检索的准确性起决定性作用。对于交互式教育资源检索，基于图像的检索方法（即“以图搜图”）是其中一个有效手段。

目前交互式教育资源检索面临的一个难题是：无法获取到能体现交互过程、代表资源关键内容的图像。现有的资源预览方式可以实现对资源的重要内容进行预览，但是这些预览方法若应用在交互式教育资源中，却缺乏对资源的动态和互动过程的预览。若能使资源中的动态内容可视化，生成反映交互过程的图像，从中选取预览图片从而更准确、

更完整地反映资源的内容，消除冗余，那么以预览图像作为特征，应用到资源检索或资源分类工作，将会有效提高交互式资源检索和分类的准确性。

1.2 国内外研究现状

随着教育技术的不断普及，教育软件、在线教育平台的应用得也越来越多^[4]，教育资源数量也越来越多。如何高效地浏览和理解资源的主要内容已经成为查找资源的一个巨大挑战，而资源预览旨在帮助用户在不需点开或下载文件的前提下，快速了解资源的主要内容，能有效节省用户的时间，因此已成为近几年一个热门的研究领域。研究者们已经在传统资源预览方面进行了大量研究，例如，对图像、文本和视频资源的预览。

1.2.1 图像及文本资源预览

图像资源预览通常将原图提取为缩略图，Johannes Kopf 等人^[5]提出内容自适应的图像缩小方法：核函数由分别定义在空间和颜色上的两个高斯核组成，然后不断迭代优化核函数，使核函数能够很好得与图像内容重合。Cengiz 等人^[6]提出了基于感知的图像缩小方法，以人的感知来评估图像的质量，保留了感知上的图像细节，且速度要快于 Kopf 的算法。Nicolas Weber 等人^[7]提出了基于卷积滤波器的算法，该算法认为与邻域像素的颜色差异越大的像素，对图像细节的贡献越大。Gastal 等人^[8]提出能够适当保留高频特征的图像缩小方法。

文本资源在线预览可以使用户在不下载资源的情况下，在浏览器上查看文件。百度网盘等公司将常见的 Word、PDF、Excel 等格式的文件进行转码，转换成 SWF 格式，即 Flash 动画文件格式^[9]。文档预览除了强调预览的表现形式之外，还需要以提取文档中心内容为基础。研究者们为了总结文档主要内容，提出了很多方法，如进化算法^[10]，增强遗传算法^[11]，整数线性规划^[12]，深度学习^[13]，矩阵因子分解^[14]，将传统机器学习算法和深度学习方法应用到文档内容总结领域。

1.2.2 视频资源预览

由于视频资源的持续时间一般较长，用户若要完整观看视频，需付出较多的时间成本。为了让用户能够快速粗略地了解视频内容，出现了一些视频预览方法。Schoeffmann 等人提出一种使用交互式导航摘要进行视频浏览的方法^[15]。这种方法提供了用户自定义级别的内容摘要可视化，因此可以快速地向用户显示视频的内容。针对传统的视频缩略

图生成方法大部分基于视频的视觉特征、没有考虑用户的查询意向、不能提供用户关注的视频预览的情况,Yuan 等人^[16]提出一种基于查询语句的动态视频缩略图的生成方法,并给出了一种新的图形卷积视频缩略图指针(GTP)。GTP 利用图卷积网络对句子—视频语义交互以及视频与句子信息结合的内在关系进行建模。最后,选择多个视频片段中与查询语句高度相关的帧,并将这些帧组合成视频的缩略图。生成的缩略图不仅提供了视频内容的简明预览,而且还根据用户的查询语句,动态关联用户的查询意向。

Zhao 等人提出一种教育视频可视化导航方法^[17]。该方法首先对关键帧进行图像融合,然后结合语音识别与光字符识别技术来获取视频中的文字信息,并将提取到的内容通过系统的可视化组件,直观全面地呈现给教育视频可视化导航系统,最后呈现给用户融合后的图像与关键的文字信息,使得用户可以在不浏览视频的前提下更加准确快速地定位到感兴趣的内容节点。

Zhao 等人设计了一种基于多线索的视频检索结果可视化浏览系统^[18]。图 1-2 展示了系统检索结果界面,该系统首先从后端视频库中检索相关视频,然后根据网格结构布局显示搜索结果。在这个网页中,每个网格单元显示搜索结果中一个视频的内容海报(见图中区域 2)。此外,还显示了视频对应的标题和时长(参见图中区域 4)。在这两个区域之间显示一条窄带,可能携带多个蓝点(见图中区域 3),称此片段为视频的时间线片段,这些蓝点在条带上的空间分布,可视化地显示了与用户输入查询或查询的子部分相关的视频内容片段的时间分布。为了指定与查询相关的单个视频片段,在界面的左上角提供了一系列的复选框(图中区域 1),每个复选框指定在查询中包含一个关键字的要求,通过操作这些复选框,用户可以指定一个视频片段是否必须匹配整个查询,或仅匹配查询的特定子部分,以被认为是相关的查询。



图 1-2 系统检索结果

1.2.3 关键帧提取

关键帧提取是视频预览的前提，因本文涉及到关键图片提取，与关键帧提取的思想类似，都是从一组图片中提取有代表性的部分图片，故对该方面研究状况展开介绍。关键帧由视频中的若干图像组成，能够代表视频的主要内容，关键帧的用途广泛，可用于视频摘要、视频分割和视频检索，能够有效减少视频处理过程中的数据量，降低数据存储的负担。

常见的关键帧提取方法主要包括以下三类：第一类基于光流法^[19]。Wayne Wolf^[20]提出光流法，计算相邻两幅帧中特征像素点的位移即移动向量，将该向量投影到二维平面得到的向量即为该像素点的光流，将所有特征像素点的光流组合在一起即为该图像帧的光流场。这种方法的基本思想是首先提取第一帧中的特征像素点，将该特征点的光流组合得到每帧图像的光流场，得到每帧图像对应的光流图，最后选定具有极小值的帧作为关键帧^[21]。这种方法能够较好地识别运动物体，但是受光线限制较大，且计算量与特征像素点数目成正比，可能耗费时间较久^[22]。

第二类是基于图像内容特征的帧差法。这种方法共同思想为首先提取每幅图像的特征，然后用相邻帧图像间特征的变化程度衡量相邻帧内容差异，差异度较大的即为关键帧。在选取图像特征时，Zhang 等人^[23]提取视频中每帧图像的颜色和纹理特征，当这些特征发生显著变化的时候，表明视频的内容发生了变化。Chen 等人^[24]在 Zhang 的基础上，又融合了颜色矩、梯度特征。基于图像内容特征的帧差法优点是关键帧选取较为灵活，不受给定关键帧数目的限制，缺点是关键帧的提取结果很大程度上依赖阈值的选择，不一定具有代表意义。

第三类是基于聚类的方法^[25]。该方法首先确定聚类数，然后基于相似度度量，对所有视频帧进行聚类，在每类中按特定规则选取关键帧，这种方法获取的关键帧可以代表视频的关键内容。但是，这种方法需要设置聚类中心点的坐标，而且聚类的数量必须事先确定，这使得初始准备过程更为复杂。此外，基于聚类的方法没有考虑到帧的时间序列^[26]。周舟等人^[27]在采用聚类方法后，将图像清晰度作为二次筛选关键帧的指标。

此外，还有一些提取关键帧的其它方法。对于非结构化用户视频，Wang 等人^[28]提出了一种两层组稀疏表示方法，用于分析视频数据的时空相关性。该方法提取用户视频的关键帧，不同于传统的基于聚类的方法，依赖于信号重构的角度来选择关键帧。Chen 等人^[29]认为不同的用户可能对同一视频中的不同片段感兴趣，提出了一种基于用户兴趣

的关键帧提取方法。Zhao 等人^[30]在数学计算中使用黄金分割法提取关键帧。Ma 等人^[31]利用资源获取交互视频后,使用分类算法获取关键帧,使用 Heatmap 算法渲染交互区域。在视频搜索领域,视频缩略图提供了视频内容的预览。Song 等人^[32]考虑了缩略图的审美价值,并结合了视频内容的高相关性和优越的审美品质这两个重要特征。由于传统的视频缩略图只根据视频的视觉特征生成一次,没有考虑用户的搜索意图,一些研究者考虑了查询语句^{[33][34]},结合视频上下文对视频片段进行划分。

现有的资源预览方法已经能够预览传统资源的重要内容,如图像预览、文档预览、视频资源预览,但是这些预览方法若应用在交互式教育资源中,却缺乏对资源的动态和互动过程的预览。

从研究者在资源预览领域所做的相关工作中,可以总结出一个普遍的模式。首先,必须提取资源的核心内容,然后以便于用户接受的形式展现出来。本文将这种模式转移到交互式教育资源预览工作。第一个问题是如何提取资源的核心内容。由于隐藏的交互信息和动态内容既是这类资源的基本特征,又是资源的核心内容,因此需要获取资源的交互信息和动态内容。为了获取这些动态内容,需要展现用户与资源的交互过程,因此可以借助自动化软件工具来自动模拟交互过程。根据资源的动态特性,选取交互动作后一段时间的截图作为资源的重点内容。第二个问题是如何呈现预览。数字图像作为一种重要的多媒体数据类型,已经得到了广泛的应用。另外,与视频相比,图片占用的内存更少,用户可以更快地掌握图片中的内容,所以本文选择图片作为交互式教育资源的预览形式。

1.3 研究内容及组织架构

1.3.1 研究主要内容

要获得交互式教育资源较好的预览效果,就必须充分考虑获取交互式教育资源主要内容的过程。首先,由于交互式教育资源通常需要专业的软件工具或网页来浏览,用户需要下载资源和相应的工具,在本地计算机上打开资源进行浏览,然后判断是否需要这些资源。对用户来说,这个过程比较耗费时间。其次,打开资源后,用户需要手动与交互元素进行交互。然而,交互式教育资源隐藏的交互信息和动态内容,很难从这些元素中准确触发出来。交互元素之间存在一定的操作规则,交互元素的不同操作序列会产生不同的知识表示效果。只有按照适当的顺序操作交互元素,才能触发隐藏的交互信息、

显示动态内容、获取资源所传递的知识。这一方面会给用户在理解资源内容时带来交互困惑和障碍，影响资源的获取和使用。对于那些不熟悉该资源平台的用户来说，学习相关的交互操作知识会大大增加时间和脑力成本。另一方面，也给资源动态内容的提取带来了很大的挑战，增加了资源预览的难度。

本文在分析用户交互理论与手动交互过程的基础上，提出了一种生成交互式教育资源可视化预览（Interactive Education Resource Visual Preview，简称 IERVVP）的方法。具体来说，该方法首先分析存储资源信息的文件，获取交互元素信息，提取约束关系，生成操作序列，然后使用自动化软件工具执行操作序列，模拟用户与资源之间的交互过程，并保存过程中的截图，最后，提取截图中的关键图片，生成交互式教育资源的图片预览。

本文的创新点和主要贡献如下：

（1）提出一种生成交互式教育资源可视化预览（IERVVP）的方法。首先，以用户交互过程相关理论为基础，分析用户与资源手动交互过程，提取手动交互过程重要信息。其次，对存储交互资源信息的文件进行分析，获取交互元素信息，提取约束关系，分析合适的交互顺序，生成操作序列；再次，使用自动化软件工具执行操作序列，模拟用户与资源之间的交互过程，在特定时刻对交互界面截图并保存截图；最后，提取截图的颜色直方图特征与边缘特征，并组合这两种特征，使用 K-Means 方法进行聚类，获取截图中的关键图片，组合为 GIF 动图，生成交互式教育资源的图片预览。图片预览方法可以应用于交互式资源的其它相关工作，如资源的自动分类、通过图片检索资源等。

（2）基于 IERVVP 方法，实现动态数学平台网络画板资源的可视化预览，并从节约时间比、预览正确性、预览完整性、预览反映内容程度四个维度对可视化预览效果进行测评。实验结果显示，该方法能够有效提取资源中的交互语义信息、模拟交互过程、提取关键图片，生成的图片预览平均节省用户 87.3% 的时间，帮助用户理解 88.6% 的资源内容，对用户理解交互式教育资源的内容有较大帮助。

1.3.2 组织架构

本文的章节组织如下：

第一章：绪论。本章介绍了交互式教育资源预览的研究背景及意义，分析了国内外学者在资源预览领域的研究现状，梳理了目前在交互式教育资源预览方面存在的问题，从国内外学者在其它资源预览方面的工作中总结普遍模式，最后介绍了本文的研究内容、创新点及组织架构。

第二章：用户交互相关研究与分析。本章从认知心理学的视角，对手动交互过程的理论储备基础进行了分析，然后具体分析了用户与交互式教育资源进行的手动交互过程，对此过程对用户造成的时间负担和脑力认知负荷等问题展开说明，从而证明进行交互式教育资源预览的必要性。从手动交互过程中分析出信息要素，为交互式教育资源预览提供指导方向。

第三章：交互式教育资源可视化预览（IERVP）方法。本章首先介绍了在提取资源交互语义信息过程中用到的一些方法，包括如何从资源文件中，分析提取交互元素信息、交互元素约束关系，生成操作序列。针对当前交互式教育资源标准不统一的问题，对交互元素信息进行抽象化定义，详细列出了如何从初步约束关系中提取资源元素约束关系的算法，并对操作事件进行抽象化定义，列出了生成操作序列的算法。然后介绍了模拟人与资源交互的方法，主要是借助自动化软件工具实现自动交互，并获取交互过程中的截图。紧接着介绍了从众多截图选取有代表性图片即关键图片的方法，包括特征提取和聚类分析两个步骤。在特征提取部分，介绍了一些图像处理技术，提取截图的颜色直方图特征与边缘特征并组合这两种特征。在聚类分析部分，介绍了使用 K-Means 方法进行聚类的步骤，针对 K-Means 聚类结果可能会改变原始图片排列顺序的问题，提出解决方案。最后，将关键图片组合为一张 GIF 动图，作为图片预览展现形式。

第四章：基于 IERVP 方法的网络画板资源可视化预览。本章具体介绍了将 IERVP 方法应用于动态数学平台网络画板：（1）阐释了如何从资源的 JSON 文件中手动分析 JSON 文件的结构，从中总结存储特征。（2）介绍了交互元素坐标值获取方式。介绍了全局坐标值与屏幕像素坐标值的关系，详细介绍了如何获取可交互元素的屏幕像素坐标。（3）解释了如何获取交互元素间的初始约束关系，主要是根据“arg”属性列表中出现数字和变量获取。（4）详细介绍了自动交互过程中截图的关键时间点给定参考值，针对不同类型的交互元素，采用了有针对性的截图方法，特别是对于按钮类元素，根据其播放信息，精准控制播放时间，在特定时间进行截图，以展现一次完整播放过程，更加准确全面地获取资源隐藏的动态内容。（5）介绍了 K-Means 聚类方法预设 K 值的给定以及 GIF 动图帧速率的给定。

第五章：实验结果及分析。本章选取了动态数学平台网络画板上三个类别中的资源，按 IERVP 方法生成图片预览。在这些资源中，选取 15 个有代表性的资源，取出其预览图片，邀请熟悉网络画板操作的师生参与问卷评估，参与者首先手动与这些资源进行交互，了解资源内容并记录手动交互时间，然后对预览图片的预览正确性、预览完整性、

能在多大程度上反映资源内容进行评分，并分析图片预览节约时间比。实验结果表明了本文提出的 IERVVP 方法的可操作性和有效性。

第六章：总结与展望。本章总结了全文的研究内容，剖析了当前研究中仍存有的不足，并阐明了后续工作展望。

图 1-3 展示了本文的整体研究框架。

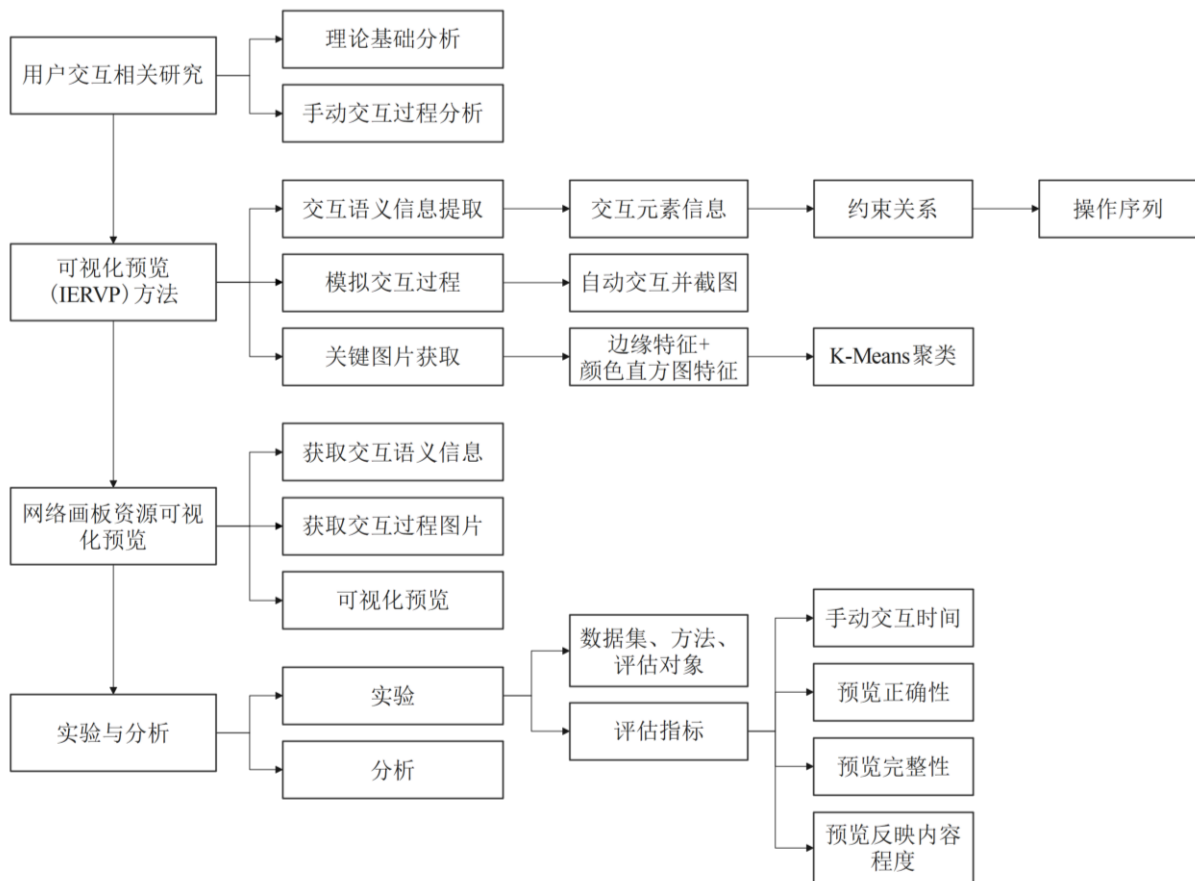


图 1-3 本文的研究框架图

第二章 用户交互相关研究与分析

2.1 用户交互过程理论基础分析

如今网络资源日益丰富、数量规模日益庞大，交互行为已成为用户获取各类信息资源最普遍的行为^[35]。交互指信息主体之间进行信息沟通、信息交换、信息传递的过程。

交互式教育资源即包含交互元素等信息的教育资源，它不仅包含了常见的文本、图形、音频和视频，而且包含了具有更复杂语义信息的交互元素。一般而言，交互元素包括按钮、滑竿和可拖拽对象等^[36]。通过点击按钮、拖拽元素等方式能与交互式教育资源进行交互，在这个过程中可以提高用户对于资源中知识的理解程度，理解的过程生动形象，便于用户接受资源所传达的知识。

但是，由于交互式教育资源的特殊性，它含有一些交互元素，而这些交互元素只能通过鼠标或键盘操作才能触发，因此不能使用传统资源预览的方式去获取交互式教育资源的内容。基于此种情况，本文将交互式教育资源预览问题视作对交互过程预览的问题。交互式教育资源预览是对交互过程的预览，最终的目标是为用户服务，帮助用户快速了解资源内容，帮助用户查找期望资源。因此，本文有必要首先分析用户在交互过程中的行为，以从中总结手动交互过程的特点。

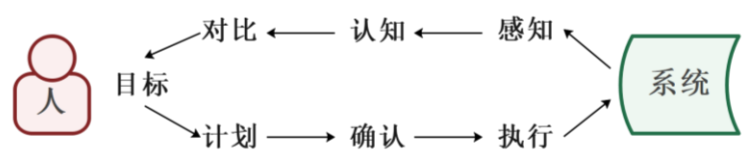


图 2-1 行动七阶段图

认知心理学家 Norman 提出行动七阶段理论^[37]，如图 2-1 所示，分别为目标、计划、确认、执行、感知、认知、对比。分析该理论可以帮助我们从抽象层面理解用户与资源交互过程的整体逻辑。目标阶段即明确要从交互过程中得到什么信息，在用户与交互式教育资源交互场景中，一般用户交互的目的是为了理解资源内容，以找到满足知识查找意愿的资源；计划阶段指在大脑中构思完成目标需要哪几步，比如要理解资源内容，首先要打开该资源，然后要观察资源内容，……；确认阶段指构思现在能够做什么，比如现在可以打开资源；执行阶段即执行大脑中所想动作，比如点击鼠标左键打开资源；感知阶段通过感官感知外部环境当前状况，例如通过眼睛观察资源状况；认知阶段处理感

觉信息，转换为知觉信息，然后大脑对信息做出判断；对比阶段将当前结果与目标进行对比，以确定结束行动或再展开下一次行动。

用户在操作界面中的交互行为通常有以下几个规律：

费茨定律^[38]：心理学家 Fitts 分析研究了人们在交互操作过程中的运动特征，在经一系列实验验证后提出该定律，这个定律现在被广泛应用于人机交互领域，作为交互研究不可或缺的基础定律。如图 2-2 所示，该定律定义用户操作点在当前界面上，由初始位置移动到目标位置所需时间由两个因素决定，分别为距目标区域的距离和目标区域宽度^[39]，其数学公式可抽象为

$$T = a + b \log_2 \left(1 + \frac{D}{W} \right) \quad (2-1)$$

其中， T 是用户操作点由初始位置移动到目标区域所需的时间； a 是操作开始时间，一般为常量； b 代表操作点移动速度； D 为起始操作点到目标位置的距离； W 是目标区域的宽度。费茨定律说明用户从当前操作点移动到目标操作点的时间与距目标的距离成正比，与目标区域宽度大小成反比，这意味着在一定范围内目标区域越近、目标区域宽度值越大，操作时间越短，因此用户操作难度越低。

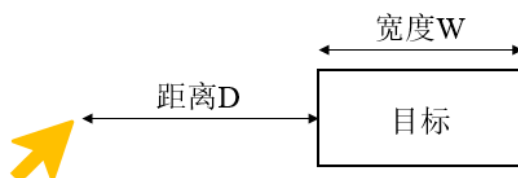


图 2-2 费茨定律示意图

席克定律^[40]：由 Hick 于 1952 年提出。如图 2-3 所示，该定律指出用户反应时间与面临的选择数量正相关，选择数量越多对应的反应时间越长，可被抽象为

$$T = a + b \log_2 N \quad (2-2)$$

其中， a 是与决策无关的总时间，例如观察时间，一般为常量； b 为认知处理时间，一般为常数； N 为用户面临的选择数。因此当用户面对复杂交互式教育资源时，如果资源内可交互元素的数量越多，用户交互耗费时间也随着增加。

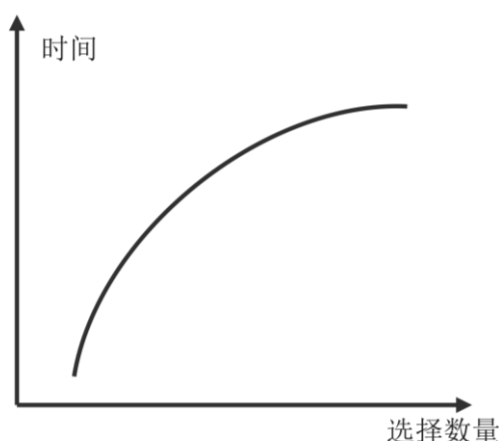


图 2-3 席克定律示意图

接近法则：描述的是在空间上邻近的物体会被用户视为一个整体。研究表明，当用户通过用户界面与资源交互时，用户与按钮交互的顺序与按钮元素的布局位置有一定的关系。当按钮元素被安排在相邻区域时，用户通常按照从左到右、从上到下的顺序与按钮进行交互^[41]。这种规律不止适用于按钮类元素，几乎所有同类型的元素若在空间上相邻，即坐标位置邻近，也存在这种交互习惯。

2.2 用户手动交互过程分析

2.2.1 手动交互过程分析

人与单个交互资源进行手动交互的过程整体可划分为三个阶段，分别为打开资源阶段、信息交互阶段、关闭资源阶段。首先用户需要打开交互资源，然后观察资源的特征，按照一定的顺序选择交互元素进行交互，与此同时，在交互的过程中观察资源的变化状况，学习动态过程中所传达的知识，最后关闭交互资源。

在上述三个阶段中，最重要同时也最为复杂的是信息交互阶段，用户需按照一定的顺序选择交互元素进行交互。这一过程可再被细化，首先用户需要观察资源的图形界面特征，识别界面中的可交互元素信息，包括元素的类别、坐标位置、交互方式。其次，进入操作阶段，选择某个交互元素，根据坐标位置将鼠标光标移动到可交互元素上方，根据元素类别决定交互方式，例如对于按钮，一般通过在按钮位置上点击鼠标左键与之交互，对于滑竿和可拖动对象，一般通过按住鼠标左键然后拖动鼠标的方式与这些交互元素交互。再接着，进入反馈阶段，观察资源页面随着交互操作的变化过程，理解资源的动态内容，判别元素之间的约束关系，重新调整交互顺序。然后再次进入操作阶段，

选择下一个交互元素进行交互操作，再进入反馈阶段，循环下去，直到所有的交互元素都被执行完、用户充分理解资源内容后结束这一阶段。

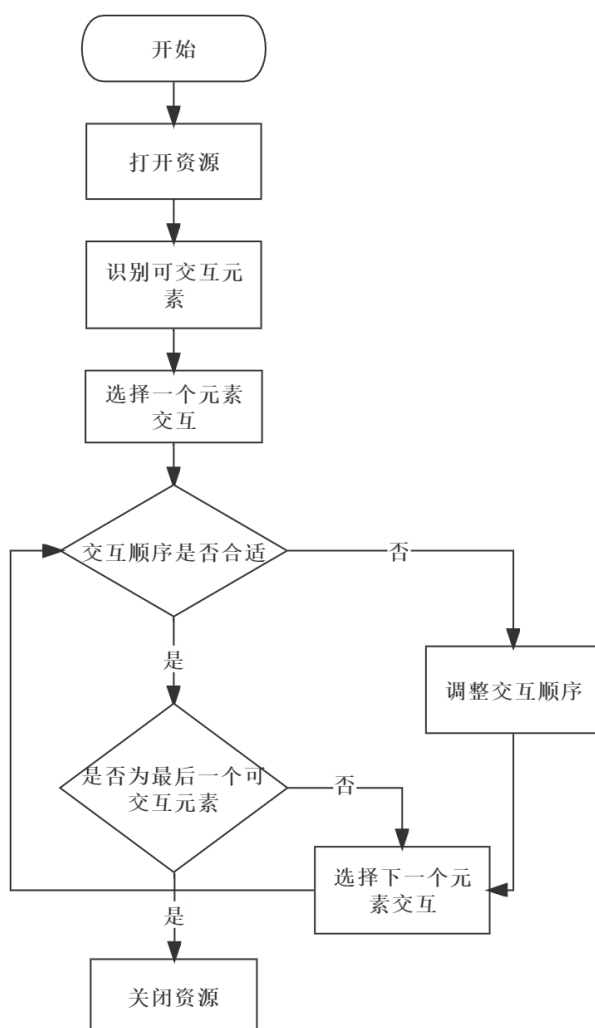


图 2-4 用户手动交互过程

由此可以观察到用户在交互阶段需完成两个任务：识别资源中可交互元素信息，然后按照一定的交互顺序选择交互元素进行交互操作。本文将用户选择交互顺序的原则归纳为两个原则：用户心理学原则和逻辑反馈原则。用户心理学原则即根据费茨定律和接近法则，用户会优先选择距离接近的元素交互，对于邻近元素一般会将它们视作一个整体依次进行交互，在邻近元素间按照从左到右、从上到下的顺序交互。逻辑反馈原则指用户会根据操作行为执行后资源界面变化情况，判别元素之间的约束关系，重新调整交互顺序，直至达到较好知识呈现效果。

2.2.2 手动交互问题分析

不同的交互顺序会导致不同的知识呈现效果，而用户在交互过程中获得的知识反馈效果直接影响用户体验感的好坏^[42]，从人与资源手动交互的过程中可以看到该过程会给用户带来两方面的问题：

第一个问题是手动交互的过程比较耗费用户时间。用户需要识别在该资源中存在哪些可与之进行交互的元素，根据席克定律，决策时间会随着资源中可交互元素数量的增加而增加。然后用户需要手动依次与这些元素交互，等待每个交互元素展示完它的整个动态过程。例如某个交互式教育资源中包含 5 个动画按钮元素，每个动画按钮播放时间为 20 秒，那么手动交互时间至少需要 100 秒，这已是一个不短的时间。

第二个问题是手动交互的过程耗费用户的脑力。识别在该资源中存在哪些可与之进行交互的元素，这一步需要用户具有一定的知识储备，对于不了解该教育资源平台、不了解这种教育资源基础构成的用户来说，这一步更是难上加难，如果让他们从头学习这种教育资源基础构成知识，会极大增加学习成本。此外，用户在识别出当前资源中的交互元素后，还需考虑按何种顺序、何种交互方法与这些元素进行交互，而不同交互顺序会呈现不同的知识展示效果，只有按照正确的操作顺序才能展示出该资源本应传达的知识，这会给用户带来交互困惑。

正是基于上述问题，因此迫切需要提出针对交互式教育资源的预览方法。

2.2.3 手动交互过程分析意义

在手动交互过程中，用户需完成两个任务：识别资源中可交互元素信息，然后按照一定的交互顺序选择交互元素进行交互操作。在这个过程中，存在如下几点信息要素：可交互元素类别、坐标位置、交互方式、元素间约束关系、交互顺序。只有提取出资源中的这些信息要素，才能重现交互过程，这将会迈出交互式教育资源预览的一大步。

2.3 本章小结

本章从认知心理学的视角，对手动交互过程的理论储备基础进行了分析，然后具体分析用户与交互式教育资源进行的手动交互过程，对此过程对用户造成的时间负担和脑力认知负荷等问题展开说明，从而证明进行交互式教育资源预览的必要性。此外，从手动交互过程中分析出信息要素，为交互式教育资源预览提供指导方向。

第三章 交互式教育资源可视化预览（IERVP）方法

交互式教育资源即包含交互元素等信息的教育资源，用户可通过点击按钮、拖拽元素等方式与交互式教育资源进行交互，在这个过程中可以提高用户对于资源中知识的理解程度，理解的过程生动形象，便于用户接受资源所传达的知识。但是，由于交互式教育资源的特殊性，它含有一些交互元素，而这些交互元素只能通过鼠标或键盘操作触发，因此不能使用传统资源预览的方式获取交互式教育资源的内容。基于此种情况，本文将交互式教育资源预览问题视作对交互过程预览的问题，因此提出了一种生成交互式教育资源可视化预览（IERVP）的方法：首先，提取交互语义信息。从资源文件中获取交互元素信息，然后提取约束关系，生成操作序列。其次，模拟交互过程。使用自动化软件工具，自动执行操作序列，模拟用户与资源之间的交互过程，并保存交互过程中的截图。最后，获取关键图片。提取这些截图的特征并进行聚类，分析提取具有代表性的关键图片，生成交互式教育资源的图片预览。IERVP 方法总体流程图如图 3-1 所示。

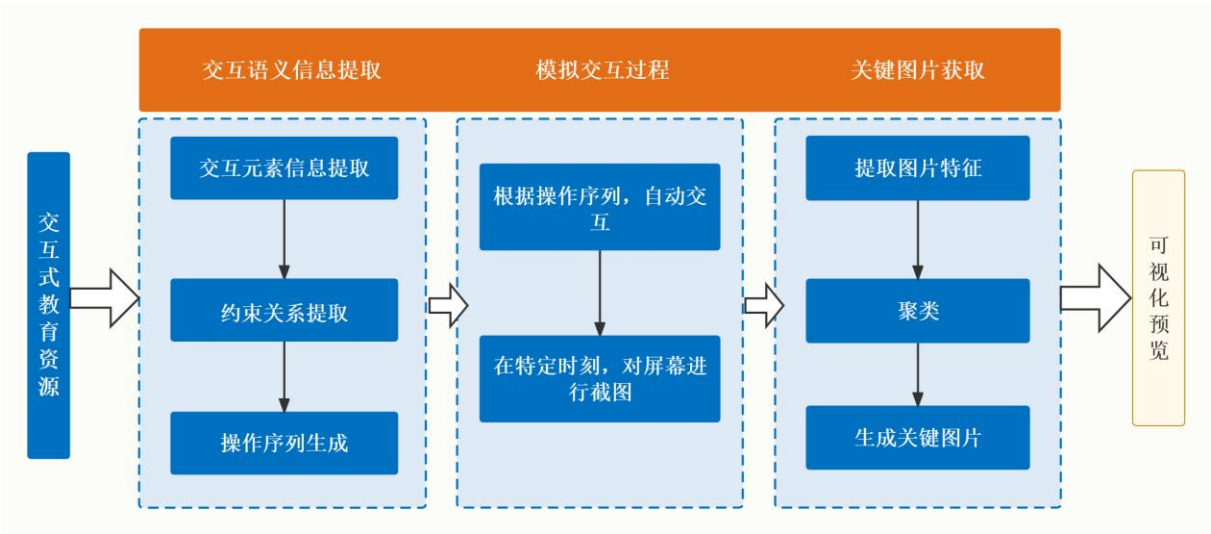


图 3-1 IERVP 方法总体流程图

3.1 交互语义信息提取

3.1.1 交互元素信息提取

国内外众多教育组织纷纷开发特定的交互式教育资源库，建设在线教育平台，但是每个交互资源库都有独特的建设标准。由于交互资源的类型较多，标准不统一，在提取交互元素信息之前，需要对交互式教育资源进行统一定义。一个交互式教育资源可以被抽象为

$$IER = (id, E, CR) \quad (3-1)$$

其中, id 为该交互式教育资源在资源库中的编号, 该编号唯一。 E 为所有在该资源中出现的元素 e 的集合:

$$E = \{e_1, e_2, \dots, e_n\} \quad (3-2)$$

CR 为元素之间的约束关系集合, 定义如下:

$$CR = \{\langle e_i, e_j \rangle, e_i, e_j \in E\} \quad (3-3)$$

$\langle e_i, e_j \rangle$ 是有序对, 意味着前面的元素 e_i 受后面的元素 e_j 的约束。

单个元素 e 可被抽象为:

$$e = (eid, name, class, type, position, actInfo, playInfo) \quad (3-4)$$

其中, eid 表示当前元素在资源中的编号, 一般由元素构建顺序决定; $name$ 表示元素的名称; $class$ 表示元素的类别, 在不同的资源库中元素类别定义不同; $type$ 表示元素的视觉及交互特征: $type$ 的值可以是 *invisible*、*visible&static* 和 *visible&interactive*, 分别表示不可见、可见但不可交互、可见且可交互。可见且可交互元素又可被分为 *visible&interactive&independent* (可见、可交互且独立元素) 和 *visible&interactive&dependent* (可见、可交互且受约束元素) 两类, 前者不受其它元素约束, 而后者可能受其它元素约束; $position$ 表示元素的坐标信息; $actInfo$ 是元素的交互信息, 可被抽象为:

$$actInfo = (boolType[, actType]) \quad (3-5)$$

其中, $boolType$ 判定该元素是否为交互元素, 取值为 1 或 0, 值为 1 意味着可与之交互, 否则值为 0; $actType$ 是元素的交互方式, 仅当元素为交互元素时, $actType$ 的值域才有效, 取值为“click”或“drag”, 前者表示用户可以通过点击鼠标左键的方式与元素进行交互, 后者表示用户可以通过按下鼠标左键然后拖动鼠标的方式与元素进行交互。

$playInfo$ 是按钮类元素播放时的相关信息, 可被抽象为:

$$playInfo = (boolType[, playValue]) \quad (3-6)$$

其中, $boolType$ 值为 1, 表示该元素是按钮类元素, 是可播放的, 否则值为 0; $playValue$ 是按钮播放时的相关值, 包含播放步数、播放时间步长、总播放时间等信息。

在完成对交互式教育资源的抽象定义后，对于一种具体给定的交互式教育资源库，由于资源中隐藏的动态内容仅在交互元素被触发时才显示，为了获取隐藏的动态内容，需要获取上述这些抽象信息，因此如何获取上述这些抽象信息是一个重要的研究任务。

为了获取上述这些抽象信息，首先需要了解在交互式教育资源中有哪些元素是可交互的。为了适应一般情况，本文假设整个教育资源库中的交互元素信息没有被整合到资源数据库中，即不知道在整个教育资源库中有哪些交互元素，不知道交互元素的名称、类别、交互方式、坐标提取方法等信息。但是通常可以在给定的教育资源库中获取对应存储资源信息的文件。一切从头开始，根据这些存储资源信息的文件，需要整合整个教育资源库中的交互元素的信息。

为了整合整个教育资源库中的交互元素的信息，首先需要手动分析一些存储资源信息的文件，总结出交互信息的存储结构特征，然后编写代码，通过遍历资源库中一定数量的资源文件，提取出交互元素在资源文件中的名称。接着，根据交互元素的名称，打开对应编号的资源，分析该元素的坐标提取方法、交互方式和播放信息，得到一个交互元素信息表，该表列出了交互元素的类别、名称、坐标提取方法、交互方式和播放信息。有了交互元素信息表，就整合好整个教育资源库中的交互元素的信息。以此为参照，可以获取特定资源中交互元素的具体信息。如何针对一种特定的交互式教育资源库提取交互元素信息表，这一步骤将在第四章网络画板资源可视化预览中详细描述。

3.1.2 约束关系提取

在提取了交互元素信息后，还需获取元素间的约束关系。本文将约束关系（CR）分为数据约束关系（Data Constraint，简称 DC）和控制约束关系（Control Constraint，简称 CC）。

数据约束关系表示元素在数据上的约束关系，如果元素 e_i 的数据受元素 e_j 的约束，那么对元素 e_i 的操作受限于元素 e_j 的输出数据，要先执行 e_j ，再执行 e_i 。例如，一个交互式教育资源中有两个变量按钮 A 和 B，当前值分别为 a 和 b，最大值分别为 max_a 和 max_b，其中变量按钮 A 的最大值 max_a 受 b 值的约束，即需要先获取另一个变量按钮的当前值 b，才能得到当前变量按钮 A 的最大值 max_a。这两个按钮间有两种操作顺序，第一种先执行变量按钮 A，此时 max_a 等于 b，由于变量变化的步长一般是固定的，所以变量按钮 A 的播放时间较短，造成不太理想的知识呈现效果。第二种先执行变量按钮

B, 一轮播放完成后, 当前值 b 等于 $\max_b$, 再执行变量按钮 A, 此时 $\max_a$ 等于 $\max_b$, 变量按钮 A 的播放时长达到最长, 播放效果达到最佳。因此, 若元素 e_i 的数据受元素 e_j 的约束, 要先执行 e_j , 才能得到比较好的交互体验及知识呈现效果。

控制约束关系表示一个元素控制改变另一个元素坐标位置的约束关系, 如果元素 e_i 对元素 e_j 存在控制约束关系, 那么 e_i 必须在 e_j 之后执行。例如, 一个交互式教育资源中有一个自由点 (即通过鼠标可以任意移动的点) 和一个按钮, 点击该按钮会改变自由点的位置, 造成自由点之前的坐标位置失效, 那么此时如果再按以前的坐标位置寻找这个自由点, 会发现该点已不在原先位置, 导致不能与它进行交互操作。所以, 应该先与这个自由点交互, 然后再点击按钮, 这样才能保证与每个交互元素都有充分的互动, 才能充分展现资源的动态内容。

由于用户只对人眼能看到的元素进行操作, 所以本文根据元素的视觉及交互特征来考虑元素之间的约束关系。前文已提及, 根据元素的可见性和交互性, 将元素分为 *invisible* (不可见)、*visible&static* (可见但不可交互)、*visible&interactive* (可见且可交互) 三种类型, 其中根据元素的约束关系, *visible&interactive* 元素又可被分为 *visible&interactive&independent* (可见、可交互且独立元素) 和 *visible&interactive&dependent* (可见、可交互且受约束元素) 两类。

由于 *invisible* 元素是不可见的, 不能看到该元素在交互过程中的变化, 用户不可能与之交互, 因此本文不考虑这类元素与其它元素的约束关系; 除此之外, *visible&static* 元素可能受相同类型的元素或 *visible&interactive* 元素的约束, 例如一个交互式教育资源中有一个自由点 A, 点 B 由点 A 旋转得到, 点 B 属于 *visible&static* 元素, 点 B 受点 A 的约束。而 *visible&interactive* 元素只能受相同类型的元素的约束, 其中, *visible&interactive&independent* 元素不受其它元素的约束, 这是由于这类元素是一种基类元素, 大部分元素都是在此类元素的基础上构建的, 而 *visible&interactive&dependent* 元素可能受其他 *visible&interactive* 元素的约束。例如在一个资源中有一条线段, 这条线段是在两个端点自由点 A 与 B 的基础上构造的, 这条线段上还有一个点 C, 那么点 A 与点 B 属于 *visible&interactive&independent* 元素, 点 C 属于 *visible&interactive&dependent* 元素, 点 C 是在点 A 和点 B 的基础上构建的, 点 C 受点 A 和点 B 的约束。

由于存储资源信息的文件不会明确记录交互元素之间的约束关系, 且不同教育资源

库中约束关系类型的分类也存在差异。因此,在实际应用中,为了抽取相互关联的元素,对特定资源库,必须分析存储资源信息文件的结构,总结元素间约束关系的存储特征,对元素关系类型进行分类,获得初步的约束关系集 CR 。如何针对一种特定的交互式教育资源库提取初始约束关系,这一步骤将在第四章网络画板资源可视化预览中详细描述。

假设已获得了初始约束关系集 CR ,由于用户最终直接对可见的交互元素进行操作,所以需要进一步过滤约束关系,去除冗余,得到最终对可见但不可交互元素和可见且可交互元素实施约束的可见且可交互元素,得到最简洁的约束关系。例如,元素 A 受约束于元素 B ,但元素 B 受约束于元素 C ,……,元素 $N-1$ 受约束于可见的交互元素 N 。在这个循环中,需要找到最终对元素 A 实施约束的交互元素 N 。

表 3-1 获取简洁约束关系算法

算法 3-1 generateCR (E, CR)
输入: 元素集 E ; 初始约束关系集 CR
输出: CR
1: $interactElementToDepElements = \{\}$
2: $staticElementToDepElements = \{\}$
3: for $e_i \in E$ and $e_i.type == visible \& interactive \& dependent$ do
4: $dfsTimes = 0$
5: $depElements = []$
6: $temp = findRelationForInteract(e_i, CR, depElements, dfsTimes)$
7: if $depElements \neq Null$
8: $interactElementToDepElements[e_i] = depElements$
9: end if
10: end for
11: for $e_i \in E$ and $e_i.type == visible \& static$ do
12: $dfsTimes = 0$
13: $depElements = []$
14: $temp = findRelationForStatic(e_i, CR, depElements, dfsTimes)$
15: if $depElements \neq Null$
16: $staticElementToDepElements[e_i] = depElements$
17: end if
18: end for
19: clear CR
20: $CR = interactElementToDepElements \cup staticElementToDepElements$
21: return CR

表 3-1 列出了获取简洁约束关系的算法。该算法的输入为当前资源中元素的集合 E 、初始元素间约束关系集 CR ,算法输出经过简化后的约束关系集 CR 。首先,第 1-2 行初始化两个字典 $interactElementToDepElements$ 和 $staticElementToDepElements$,分别用来存

放对可见且可交互元素实施约束的元素、对可见但不可交互元素实施约束的元素。第 3-10 行得到可见且可交互元素之间的约束关系。遍历元素集 E 中的每个可见、可交互且受约束元素，得到最终对该元素实施约束的可见且可交互元素。首先初始化部分变量，其中 $dfsTimes$ 是迭代次数， $depElements$ 用来存放对该元素实施约束的可见且可交互元素，然后将约束关系放入字典 $interactElementToDepElements$ 中，其中第 6 行是此块的核心部分，是一个递归函数，将在表 3-2 中详细介绍。第 11-18 行获取对可见但不可交互元素实施约束的可见且可交互元素，并将获得的约束关系放入另一个字典 $staticElementToDepElements$ 中，第 14 行是此块的核心部分，是一个递归函数，将在表 3-3 中详细介绍。第 19 行清空 CR ，然后将存储在两个字典中的约束关系放到 CR 中。

表 3-2 获取可见且可交互元素之间约束关系算法

算法 3-2 findRelationForInteract ($e, CR, opElement, dfsTimes$)
输入：元素 e ；初始约束关系集 CR ；对当前元素 e 实施约束的交互元素 $opElement$ ；迭代次数 $dfsTimes$
输出：对当前元素 e 实施约束的交互元素 $opElement$
1: if $dfsTimes \geq 20$
2: return
3: end if
4: $dfsTimes = dfsTimes + 1$
5: if $e.type == invisible$
6: return
7: end if
8: if $e.type == visible \& interactive \& independent$
9: return e
10: end if
11: if $e.type == visible \& interactive \& dependent$
12: $e_i = e$
13: for $e_k \in \{e \mid e \in E \wedge \langle e_i, e \rangle \in CR\}$ do
14: $answer = findRelationForInteract(e_k, CR, opElement, dfsTimes)$
15: if $answer \neq Null$
16: $opElement.append(answer)$
17: end if
18: end for
19: if $opElement == Null$
20: $opElement.append(e_i)$
21: end if
22: else
23: return
24: end if

表 3-2 列出了获取可见且可交互元素之间约束关系的算法,该算法是一个递归算法。第 1-4 行限制了递归的次数,以防止无限递归造成死循环。第 5-7 行表示,如果当前元素是不可见元素,则返回到上层递归。第 8-10 行表示,如果当前元素是一个可见、可交互且独立元素,这种类型的元素不可能再受其它元素的约束,因此返回当前元素至上层递归。第 11-21 行是递归算法的主要部分:如果当前元素是一个可见、可交互且受约束元素,那么在 CR 中找到对当前元素实施约束的元素,并将该元素作为新元素,进入下一层递归。第 22-24 行表示,如果不满足上述列出的所有条件,则直接返回到上层递归。

表 3-3 获取可见但不可交互元素约束关系算法

算法 3-3 findRelationForStatic ($e, CR, opElement, dfsTimes$)

输入: 元素 e ; 初始约束关系集 CR ; 对当前元素 e 实施约束的交互元素 $opElement$; 迭代次数 $dfsTimes$

输出: 对当前元素 e 实施约束的交互元素 $opElement$

```

1: if  $dfsTimes \geq 20$ 
2:   return
3: end if
4:  $dfstimes = dfstimes + 1$ 
5: if  $e.type == invisible$ 
6:   return
7: end if
8: if  $e.type == visible \& interactive \& independent$ 
9:   return  $e$ 
10: end if
11: if  $e.type == visible \& interactive \& dependent$ 
12:   if  $e$  in  $interactElementToDepElements$ 
13:     return  $interactElementToDepElements[e]$ 
14:   else
15:     return  $e$ 
16:   end if
17: end if
18: if  $e.type == visible \& static$ 
19:    $e_i = e$ 
20:   for  $e_k \in \{e \mid e \in E \wedge \langle e_i, e \rangle \in CR\}$  do
21:      $answer = findRelationForStatic(e_k, CR, opElement, dfsTimes)$ 
22:     if  $answer \neq Null$ 
23:        $opElement.append(answer)$ 
24:     end if
25:   end for
26: else
27:   return
28: end if

```

表 3-3 中的算法列出了如何获得最终对可见但不可交互元素实施约束的可见且可交互元素，该算法也是一个递归算法，其基本思想与算法 3-2 相似。第 1-4 行限制了递归的次数，以防止无限递归造成死循环。第 5-7 行表示，如果当前元素是不可见元素，则返回到上层递归。第 8-10 行表示，如果当前元素是一个可见、可交互且独立元素，这种类型的元素不可能再受其它元素的约束，因此返回当前元素至上层递归。第 11-17 行表明，如果当前元素是一个可见、可交互且受约束元素，将直接返回到上层递归，返回值是由算法 3-2 获得的对当前元素实施约束的元素。这一步可以减少递归次数，加快算法运行速度。第 18-25 行是递归的主要部分：如果当前元素是一个可见但不可交互元素，那么在 CR 中查找对当前元素实施约束的元素，然后使用这个元素作为新元素来执行递归。第 26-28 行表示，如果不满足上述条件，则直接返回到上层递归。

3.1.3 操作序列生成

本文定义操作序列（Operate Sequence，简称 OS）由用户在用户界面中触发的事件组成。在这里，将事件定义为：

$$event = (eid, actType, position, playInfo) \quad (3-7)$$

其中， eid 是该事件作用的交互元素在当前资源中的编号， $actType$ 表示该交互元素的交互方法， $position$ 是该交互元素的坐标， $playInfo$ 是按钮类元素播放时的相关信息。具体来说，操作序列由一系列在用户界面中触发的事件组成，即一系列对可见且可交互元素的交互操作：

$$OS = \langle event_1, event_2, \dots, event_n \rangle \quad (3-8)$$

不同的操作顺序会产生不同的动态内容和知识呈现效果，操作序列直接决定了在自动交互过程中，按照何种顺序与元素进行交互，因此获取合适的操作序列的重要性不言而喻。

为了更加全面地表达资源的动态内容，本文设定一个规则：与所有可见且可交互元素相关的事件均需被执行一次，即在操作序列出现一次。接下来，需要考虑这些事件在操作序列中的次序。

已得到了最终对可见但不可交互元素和可见且可交互元素实施约束的可见且可交互元素，将考虑从三个方面来调整事件次序、生成操作序列，分别为元素的优先级、元素在用户界面中的布局位置以及元素之间的约束关系（CR）。

优先级高的元素将被优先执行。在前面的章节中提到交互元素之间可能存在控制约束关系, 如果元素 e_i 对元素 e_j 存在控制约束关系, 那么 e_j 必须先执行, 然后再执行 e_i , 所以 e_j 这种类型的元素的优先级要设置得比较高。给定可见且可交互元素的优先级很重要, 因此需要根据不同资源库, 进行具体分析, 对不同类别的可见且可交互元素, 分别给出优先级值。如何针对一种特定的交互式教育资源库给出优先级值, 这一步骤将在第四章网络画板资源可视化预览中详细描述。给定可见且可交互元素的优先级后, 将所有可见且可交互元素按照类别优先级由高到低的顺序排列, 得到操作序列中交互元素的初始序列。

为了更好地模拟用户与资源的交互过程, 使得自动交互过程更贴近现实, 本文将第二章中提及的费茨定律和接近法则这两个交互习惯应用到操作序列生成过程中, 在邻近元素间按照从左到右、从上到下的顺序交互。

元素之间的约束关系通常是一对多的关系, 也就是说, 对于一个可见且可交互元素, 它可能受约束于多个元素。本文结合约束关系来调整操作序列, 具体来说, 如果一个元素 e_i 的数据受约束于元素 e_j , 两者之间的约束关系为数据约束关系, 而且这两个元素的优先级是相同的, 但是在操作序列中, e_j 对应的事件在 e_i 对应的事件后面, 那么, 将 e_j 对应的事件放在 e_i 对应的事件前面, 这意味着如果 e_i 对应的事件在操作序列中的下标是 i , 那么将 e_i 对应的事件移动到第 $i+1$ 位, 并将 e_j 对应的事件移动到第 i 位。如果元素 e_i 对元素 e_j 存在控制约束关系, 而且这两个元素的优先级是相同的, 但是在操作序列中, e_j 对应的事件在 e_i 对应的事件后面, 那么, 将 e_j 对应的事件放在 e_i 对应的事件前面, 这意味着如果 e_i 对应的事件在操作序列中的下标是 i , 那么将 e_i 对应的事件移动到第 $i+1$ 位, 并将 e_j 对应的事件移动到第 i 位。

表 3-4 中的算法 3-4 给出了生成适当的操作序列的步骤。首先, 第 1-3 行将所有可见且可交互元素都加入操作序列中, 然后再对元素进行排序。具体来说, 第 4 行按优先级从高到低的顺序对交互元素进行初步排序。第 5-11 行, 对于同类别元素, 若元素在空间上较为接近, 通过从上到下和从左到右的坐标位置调整序列。第 12-28 行, 根据元素间的约束关系调整序列。

表 3-4 生成操作序列算法

算法 3-4 $\text{getOperateList}(E, CR, \text{elementToPri}, \text{threshold})$

输入：元素集 E ；约束关系集 CR ；<元素，优先级>映射 elementToPri ；距离阈值 threshold

输出：操作序列 operateList

```

1: for  $e_i \in E$  and  $e_i.\text{type} == \text{visible} \& \text{interactive}$  do
2:    $\text{operateList.append}(e_i)$ 
3: end for
4: array  $\text{operateList}$  by  $\text{elementToPri}$  from high to low
5: for  $e_i$  in  $\text{operateList}$  do
6:   find all  $e_j$  when  $e_j.\text{class} = e_i.\text{class}$ 
7:   if  $\text{distance}(e_i, e_j) < \text{threshold}$ 
8:     array  $\text{operateList}$  by  $e_j.\text{position}(\text{left})$  from low to high
9:     array  $\text{operateList}$  by  $e_j.\text{position}(\text{top})$  from low to high
10:   end if
11: end for
12: for  $\langle e_i, e_j \rangle \in CR$  do
13:   if  $e_i, e_j$  in  $\text{operateList}$  and  $\text{elementToPri}(e_j) == \text{elementToPri}(e_i)$ 
14:     if  $\langle e_i, e_j \rangle \in DC$  and  $e_j$  is behind  $e_i$  in  $\text{operateList}$ 
15:       for  $e_k$  in front of  $e_j$  and  $e_k$  behind  $e_i$  do
16:         put  $e_k$  in  $\text{operateList}$  with  $\text{index}(e_k) + 1$ 
17:       end for
18:       put  $e_j$  in  $\text{operateList}$  with  $\text{index}(i)$ 
19:       put  $e_i$  in  $\text{operateList}$  with  $\text{index}(i + 1)$ 
20:     else if  $\langle e_i, e_j \rangle \in CC$  and  $e_j$  is in front of  $e_i$  in  $\text{operateList}$ 
21:       for  $e_k$  behind  $e_j$  and  $e_k$  in front of  $e_i$  do
22:         put  $e_k$  in  $\text{operateList}$  with  $\text{index}(e_k) + 1$ 
23:       end for
24:       put  $e_i$  in  $\text{operateList}$  with  $\text{index}(j)$ 
25:       put  $e_j$  in  $\text{operateList}$  with  $\text{index}(j + 1)$ 
26:     end if
27:   end if
28: end for
29: return  $\text{operateList}$ 

```

3.2 模拟交互过程

3.2.1 自动交互

手动交互过程耗时耗力，而自动交互，即利用计算机技术自动与资源进行交互，自动执行操作序列，不再需要用户直接参与交互过程，是手动交互问题的解决方案。从手动交互过程中，了解到用户和资源之间的交互是通过用户操作鼠标或键盘来进行的^[43]，那么接下来需考虑如何与资源进行自动交互，并要考虑以何种形式保留这个交互过程。

自动化软件工具可以执行鼠标和键盘的操作，也就是说，通过创建一个脚本程序，自动化软件工具可以自动执行鼠标及键盘动作。在第三章第一节中已获取了交互元素的坐标信息、交互方式并生成了操作序列，因此可使用自动化软件工具，按照这些信息，编写程序与每个交互元素交互。常用的自动化软件工具包括 AutoHotKey、AutoIt、脚本精灵等。

AutoHotKey 由 Chris Mallett 及其小组研发，是一款开源且免费的、简洁便捷却功能强大的、以脚本为基础的热键，运行在 Windows 平台上。AutoHotKey 具有多样功能，例如可实现热键、热字符串、窗口置顶、快速输入、处理系统管理相关工作等功能。用户利用这款工具，能够通过编写自定义脚本实现几乎所有 Windows 平台操作的自动化，而且仅占用极小内存，是一个十分灵活，可被高度自定义、可用性极高的工具。Chris 设计 AutoHotKey 的初衷是简化用户的日常操作，使用户从重复机械的活动中解放出来。

AutoIT 是由 Jonathan Bennett 带领他的小组编写的 Windows GUI 自动化操作软件。这个软件一开始的设计目的是为了帮助用户进行 IT 管理和维护，完成个人电脑在此方面的批处理任务，例如自动进行软件的安装及更新。后来，随着日益更新的计算机环境的发展，用户已不再满足于仅仅利用它进行软件安装，产生了更多对于自动操作计算机其他任务的需求。由于自动化操作在设计思想、执行过程等方面与软件自动安装类似，都需要用户编写脚本，然后代替用户进行自动操作。因此，AutoIT 发展为一款自动化测试软件工具。

按键精灵是一款国产自动化软件工具，由福建创意嘉和软件有限公司开发并运营。这款软件的界面简洁，操作便捷，富含操作指引，支持插件共享，对用户而言十分友好。

这些自动化软件工具的共同思想为控制鼠标和键盘，执行相应操作，从而执行一系列自动化任务。由于隐藏的动态内容和交互过程既是交互式教育资源的基本特征，又是资源的核心内容，所以需要展现这些动态内容和交互过程。因此在获取交互语义信息后，可使用自动化软件工具，与各个交互元素进行交互，展现这些动态内容和交互过程。

3.2.2 获取截图

在与交互元素进行交互时，交互式教育资源的内容将发生显著而有意义的变化。只有通过捕获并记录这些变化，才能对交互式教育资源所传递的知识有更深入的了解。因此，在自动交互过程中，根据资源的动态特性，需要利用自动化软件工具在特定时刻对用户界面进行截图，并保存截图。如何确定具体时刻将在第四章网络画板资源可视化预览中详细介绍。

3.3 关键图片获取

针对一个交互式教育资源，经过自动交互并保存截图，已可获得数量可观的截图。但是如果选择获取到的所有截图作为最终的预览输出，会造成内容过多，导致用户难以抓住资源重点内容。因此，本文从这些截图中选择具有代表性的关键图片作为预览输出，选取关键图片包括提取图片特征和聚类两个步骤。

3.3.1 提取图片特征

3.3.1.1 图像灰度化

彩色图像中每个像素点的值由 R（红）、G（绿）、B（蓝）三种颜色分量值组合而成^[44]。对于一幅大小为 $M \times N$ 的彩色图像，计算机通常用三维数组存储该图像的像素值，三维数组的大小为 $M \times N \times 3$ ，假设该数组名为 P ， $P[i, j, k]$ 索引中的前两位标志像素点的像素位置，第三位值为 1 时，数组中对应的值表示该像素点 R 分量的值，为 2 时表示 G 分量的值，为 3 时表示 B 分量的值，数据类型可以为整型或浮点型，若为 8 位无符号整型变量，则这三种颜色分量的值都为定义在 $[0, 255]$ 区间内的整数。

灰度图像是 R（红）、G（绿）、B（蓝）三种颜色分量值相同的一种特殊图像^[45]。对于一幅大小 $M \times N$ 的灰度图像，计算机使用大小为 $M \times N$ 的数组存储该图像中像素点对应的灰度值，若为 8 位无符号整型变量，则像素点的灰度值为 $[0, 255]$ 区间中的整数。

大部分交互式教育资源都是彩色的，因此经过自动交互获得的截图也基本都是彩色图像，彩色图像所含的信息量更丰富，但在存储时占用的空间也更大。对交互式教育资源来说，其动态内容变化主要体现在图形的变化。在提取交互式教育资源关键图片时，最重要的技术是获取图形的变化、获取图片的纹理特征，因此很多彩色信息一般不具有太大作用，而且由于灰度图像存储体积更小，因此在提取交互式教育资源关键图片时，

将彩色图像转为灰度图像可在保留主要特征的前提下，减少信息量，降低存储负荷，从而加快图像处理速度。

图像灰度化处理过程即将彩色图像转换为灰度图像的过程^[46]，对每个像素点，将其对应的三个颜色分量值转换为一个灰度值。一般有三种方法进行灰度化处理^[47]。

(1) 取最大值法

取最大值法选取彩色图像中 R（红）、G（绿）、B（蓝）三种颜色分量中的最大值作为灰度图像在该点处的灰度值。

$$Gray(i, j) = \max\{R(i, j), G(i, j), B(i, j)\} \quad (3-9)$$

(2) 平均值法

平均值法^[48]获取彩色图像中 R（红）、G（绿）、B（蓝）三种颜色分量的平均值，将该值作为灰度图像在该点处的灰度值。

$$Gray(i, j) = \frac{1}{3} \times [R(i, j) + G(i, j) + B(i, j)] \quad (3-10)$$

(3) 加权平均值法

加权平均值法^[49]根据彩色图像中 R（红）、G（绿）、B（蓝）三种颜色分量的重要性，对三种分量赋予不同的权值，进行加权平均运算，以该值作为灰度图像在该点处的灰度值。

$$Gray(i, j) = w_r \times R(i, j) + w_g \times G(i, j) + w_b \times B(i, j) \quad (3-11)$$

其中 $w_r + w_g + w_b = 1$ ，本文采用的权值是 $w_r = 0.3$ ， $w_g = 0.59$ ， $w_b = 0.11$ 。

3.3.1.2 颜色直方图

对大小为 $M \times N$ 的灰度图像来说，计算机中存储该图像的数组大小为 $M \times N$ ，对应的特征维度即 $M \times N$ ，常见的计算机分辨率为 1366×768 或 1920×1080 ，因此对这些计算机屏幕截图得到的图像经过图像灰度化后，对应的灰度图像的维度非常大，如果直接将这种维度的特征向量直接输入计算机，会极大降低信息处理的速度，造成计算资源的浪费，实际上，提取交互式教育资源的关键图像并不需要如此高维的特征。因此需要进一步提取灰度图像的特征。

颜色直方图能够反映某个灰度值在图像中出现的次数^[50]。美国学者 Swain 和 Ballard 在上世纪九十年代提出了提取图像的颜色直方图特征^[51]，首先分离图像 RGB 颜色空间

三个分量，得到三个分量上的颜色直方图，然后，他们通过一系列实验，比如对图像进行平移、旋转、缩放等几何变化、对图像进行滤波处理，一系列实验结果证实颜色直方图特征对这些改变不敏感，证明了颜色直方图特征是一种稳定的特征，因此常用颜色直方图特征评估两幅图像间的全局颜色差异^[52]。

对彩色图像而言，分离三种颜色分量，每种颜色分量的值都定义在 $[0,255]$ 区间内，对每种颜色分量，将每个颜色值对应的像素点数累加在一起，即可得到彩色图像在某一颜色分量上的颜色直方图特征，对应的特征维度为 1×256 。对灰度图像而言，整幅图像每个像素点的灰度值在 $[0,255]$ 区间内，将每个颜色值对应的像素点数累加在一起，就是灰色图像的颜色直方图特征，特征维度为 1×256 。

与灰度图像维度相比，颜色直方图特征显著降低了特征维度。

3.3.1.3 图像边缘特征

图像边缘指相邻区域灰度值产生突变的像素点集合^[53]，目前主要用灰度值差异来提取图像边缘。

梯度原指函数值变化速度最快的方向，在图像中，指图像在某点处灰度变化最快的描述，图像在 (x, y) 处的梯度向量可表示为：

$$\nabla g = \text{grad}(g) = \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} f(x+1, y) - f(x, y) \\ f(x, y+1) - f(x, y) \end{bmatrix} \quad (3-12)$$

获取了某点处的梯度向量，可以获得图像在该点处最大的灰度变化率以及对应的方向。其幅值为：

$$|g(x, y)| = \sqrt{g_x^2 + g_y^2} \quad (3-13)$$

梯度向量的方向为：

$$\alpha(x, y) = \arctan\left(\frac{g_y}{g_x}\right) \quad (3-14)$$

因此，只要能够找出像素点的梯度，就能够获取图像的边缘信息。常见的边缘检测算子有：Sobel 算子、Roberts 算子和 Canny 算子^[54]。

(1) Sobel 算子与 Roberts 算子

Sobel 算子中， g_x 、 g_y 的模板为：

$$g_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad (3-15)$$

$$g_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad (3-16)$$

g_x 的方向沿水平方向、 g_y 的方向沿垂直方向，因此 Sobel 算子检测直线的效果较好。

Roberts 算子中， g_x 、 g_y 的模板为：

$$g_x = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (3-17)$$

$$g_y = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad (3-18)$$

因此，该算子对呈对角线样式的边缘检测效果较好。

(2) Canny 算子

Canny 边缘检测算法的步骤如下：

- 1、对输入图像，使用高斯滤波器滤波，减少图像中的噪声。
- 2、计算梯度的幅值和方向。
- 3、在梯度方向上，找梯度中的局部最大值点，并把其它点像素值置为 0。
- 4、使用两个阈值检测并连接边缘。这两个阈值能够检测出两种边缘结果，当这两种边缘比较靠近时，连接边缘时就会将它们连到一起。

对比众多边缘检测算子，由于交互式教育资源中资源内容各不相同，对应的边缘特点不能简单抽象为直线或对角线，可知 Canny 算子检测交互式教育资源内容边缘的效果最好。

对大小为 $M \times N$ 的灰度图像，利用 Canny 算子检测边缘后得到的图像为一幅二值图像，每个像素点的像素值仅为 0 或 255，边缘处像素值为 255，特征维度为 $M \times N$ ，此时特征维度没有改变。对此，在二值图像中，初始化一个 m 个 bin 的直方图，从上到下扫描图像，将像素值为 255 的点投射到水平方向，对应的 $\text{bin} + 1$ ，最后得到 $1 \times m$ 维特征。

为了更好地提取交互式资源图像特征，将颜色直方图特征与 Canny 边缘检测特征组合在一起，将这 $1 \times m$ 维特征放置到颜色直方图 1×256 维特征后，得到 $1 \times (256 + m)$ 维特征。

3.3.2 聚类方法

聚类是一种无监督学习方法^[54]，基本思想为：给定一组数据，按照某个度量标准，将数据自动划分为不同类别，使得同类中数据的度量值较为接近，而不同类数据之间的度量值尽可能相差较大。从空间上看，聚类后同一类的数据相互之间距离较近，比较紧凑，而不同类数据在空间上距离较远。

面对获取到的众多截图，要想减少图片数，就要从中挑选相互之间差异较大的图片，而聚类方法能够很好解决这个诉求，将图片集划分为不同的类，各类内部的图片相似度高，类间图片差异度高。本文使用 K-Means 聚类方法划分图片集。

K-Means 聚类方法流程如下：

- 1、随机创建 k 个点作为初始聚类中心。
- 2、遍历数据集，计算每个数据点到这 k 个聚类中心的距离，将当前数据点加入离它最近的聚类中心所在类中。
- 3、对每个类，更新聚类中心，一般取该类中所有数据点的均值作为新的聚类中心。
- 4、重复步骤 2-3，直到达到收敛条件。

K-Means 聚类达到收敛条件后，每类中包含的图片之间相似度较高，选取最接近聚类中心的图片作为代表该类的关键图片，可以减少图片总数。但这有可能导致关键图片之间的次序改变。比如某图片集经聚类后，第 1、2、5、6 张图片为一类，其中第 6 张图离聚类中心距离最近，选择这张图作为该类的代表图，第 3、4 张图片为一类，第 3 张图离聚类中心距离最近，所以选取该图作为该类的代表图。最终生成的关键图片为第 6 张图、第 3 张图，打乱了之前的图片顺序，而交互式教育资源的交互过程是有顺序的，关键图片乱序会导致打乱交互过程次序，不利于用户了解交互式教育资源的内容，甚至可能造成理解歧义。

针对关键图片乱序问题，本文将聚类结果中每类中非连续图片分割开来，划分为几个区间，使得每区间中图片下标连续。例如对于上个例子，将第一类划分为[1,2]和[5,6]两个区间。具体做法为引入标签 *label*，结合图片顺序和聚类编号将图片划分为一个个小区间，即按顺序遍历图片集，如果当前图片所属类别不同于上张图片，切分新区间，每

个区间内图片的 *label* 等于该区间第一张图片的编号。从每个区间内，选取与聚类中心最近的图片作为该区间的图片，这样生成的关键图片不会打乱次序，保留了原始动态过程的时间顺序特征。

3.4 关键图片组合

自动交互过程所获截图在经过聚类后，可得到几张关键图片。需考虑如何将这几张图片组合在一起，作为最终的可视化预览展现形式。由于大多数交互式教育资源库都以网站为载体，实现交互式教育资源的设计与共享等服务，目前基于 JavaScript 技术，网页能很便捷地展示动态图片，将鼠标放置在图片上方即可展示其动态过程，鼠标离开恢复静态状态，通过动态图片用户能够快速了解一组静态图片的变化过程。因此本文选择动态图片作为最终的可视化预览展现形式。

动态图片即一组静态图片按照特定的频率依次展现的图片，它能够呈现静态图片的动态变化过程。最常用的动态图片格式为 GIF 即图形交换格式，GIF 动画针对多张图片，按时间顺序依次播放以生成动画效果。GIF 图片支持多平台跨平台使用，应用场景较为广泛。

GIF 图片实质上是一种压缩图片，原始每张静态图片都被转换为位图，图片中每个像素点都由指定颜色构成，一般生成 GIF 图片采用 LZW（Lempel-Zev-Welch）压缩算法，每个像素点由 RGB 颜色空间中的三种分量值构成，每种分量值最多占有 8bit，即每种颜色分量值最大表示区间为[0,255]，因此 GIF 图片能够表达的颜色数较少，画质不够精细。GIF 可以调整颜色分量占据的比特位数，从而减少图片使用的颜色数量，这样能够减少动态图片存储体积。

GIF 具有文件体积小、兼容性强等优点。将获取到的关键图片按照交互顺序，采取合适的播放频率组合为一张 GIF 动图，是一种合适的图片预览形式。

3.5 本章小结

本章首先针对当前交互式教育资源标准不统一的问题，对交互元素信息进行抽象化定义。介绍了在提取资源交互语义信息过程中用到的一些方法，包括如何从资源文件中，分析提取交互元素信息、交互元素间约束关系，生成操作序列。详细列出了如何从初步约束关系中提取简洁资源元素约束关系的算法，并对操作事件进行抽象化定义，列出了生成操作序列的算法。

接下来介绍了模拟人与资源交互的方法，主要是借助自动化软件工具实现自动交互，并获取交互过程中的截图。

然后介绍了从众多截图选取有代表性图片即关键图片的方法，包括特征提取和聚类分析两个步骤。在特征提取部分，介绍了一些图像处理技术，提取了图像颜色直方图和边缘特征，并组合了这两种特征。在聚类分析部分，介绍了使用 K-Means 方法进行聚类的步骤，针对 K-Means 聚类结果可能会改变原始图片排列顺序的问题，提出解决方案。

最后，将关键图片组合为一张 GIF 动图，作为图片预览展现形式。

第四章 基于 IERVP 方法的网络画板资源可视化预览

网络画板是一个以动态数学为重点的在线教育资源平台^{[56][57][58]}，它不仅包含丰富的作图功能，还提供了一个开放的平台来分享作品和探索数学问题^{[59][60][61]}，且支持资源共享。网络画板因其丰富的教学资源 and 便捷的动态几何作图环境受到广大师生的喜爱，因此该平台上的资源数量日益增多。

然而，该平台上资源数量的快速增加给用户查找资源带来很多不便，目前该平台仅支持通过关键词搜索目标资源，然而用户在输入关键词后，界面展示出的资源数量可能较多，这时候需要用户手动打开资源，手动与资源交互，从而了解资源的内容，才能确定最终使用哪个资源。这个过程耗时耗力，对用户不友好。因此，目前迫切需要实现对该平台资源的预览。

与其它资源相比，网络画板上的动态几何资源特点鲜明，含有较多几何元素和隐藏的交互信息。如果这些几何元素与动态动作相关联，就是可交互的，这时只有与几何元素交互，才能使隐藏的交互信息显露出来，才能理解资源的内容。用户可通过点击按钮或直接拖拽可交互几何元素与该资源进行交互，同时加深对于数学知识的理解程度。图 4-1 是一个“多边形的平移”的案例。用户可通过点击动画按钮或拖拽下方线段上的绿点实现对多边形的平移。

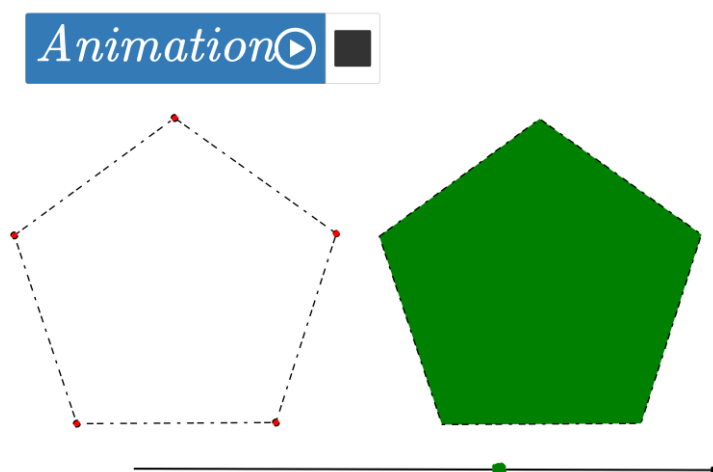


图 4-1 “多边形的平移”案例

4.1 获取交互语义信息

4.1.1 资源信息文件分析

在第三章中，已抽象地定义交互式教育资源，接下来，需要获取实际资源信息(包括元素的 *eid*、元素的类别、元素在画布中的位置，元素的交互方式以及元素之间的约束关系等)。由于交互式教育资源的特殊性，需要借助资源库中存储资源信息的文件获得上述资源信息。在网络画板中，资源对应的 JSON 文件存储了该资源的所有基本信息，通过 HTTP GET 请求即可获得该文件。JSON 文件中的“content”字段，存储了资源的重要内容信息。在解析资源内容时，应重点关注这部分字段，如表 4-1 所示。

表 4-1 JSON 文件中 content 字段重点内容

字段名	包含信息	数据结构
defaultsDisplay	该资源中所有元素的默认显示样式	数组
p	该资源中所有变量信息	数组
r	该资源中所有元素信息	数组

在 r 字段中，每个元素信息又存有如下字段与属性，如表 4-2 所示：

表 4-2 JSON 文件中 r 字段重点内容

字段名	属性名	信息
fun	name	元素名称
	arg	受约束元素信息
display	show	该元素是否可见
	position	元素坐标信息

```
684  "fun": [{
685    "name": "AnimationVars",
686    "arg": [
687      [
688        {
689          "varName": "u000",
690          "s": "0",
691          "num": "100",
692          "e": "1",
693          "timer": "50"
694        }, "2", "0", "动画"
695      ]
696    ],
697    "display": {
698      "fontProp": { ... },
699      "fillProp": {},
700      "penProp": {},
701      "private": {
702        "zIndex": 1012,
703        "position": {
704          "top": 14.730364906152602,
705          "left": 25.252054124832966
706        },
707        "selectable": true,
708        "show": true
709      }
710    }
711  ]
712 }
```

图 4-2 “多边形的平移”案例中动画按钮存储信息

图 4-2 给出了图 4-1 所示“多边形的平移”资源中左上角动画按钮，在对应的 JSON 文件中 r 字段的存储信息。在 fun 字段中，从 name 属性可获取该按钮在资源中的存储名“AnimationVars”，在 arg 属性列表中，varName 给出了动画按钮控制的变量名称，s 给出了变量的初始值，e 给出了变量的终止值，num 指出了变量值变化的步数，timer 给出了变量值变化一次所需的时间，单位为毫秒。在 display 字段中，position 中给出了动画按钮的屏幕坐标位置，show 的值为 true 说明该元素可见。

在分析完资源的 JSON 文件结构特征后，需要获取资源的实际信息。但是整个资源库中的所有交互元素以什么样的名称存储在 JSON 文件中还未知（即“name”属性值未知），而该名称对提取交互元素的其它信息十分重要，只有知道交互元素存储名称，才能通过英文值了解该元素大致是什么类型的元素，才能去分析对应的 JSON 文件存储信息，才能从中总结存储规律，获取元素的约束关系信息、坐标信息、是否可见等信息，因此要尽可能多地获取资源库中所有类别的交互元素名。

通过爬虫技术读取网络画板中前 10000 个（id 编号 1-10000）资源的 JSON 文件，保存了所有“name”属性值，在去除重复值后，分析得到了交互元素的名称。然后根据交互元素的名称，分析对应的 JSON 文件中的实例，得到表 4-3，该表详细列出了可交互元素的类别，可交互元素的“name”属性值，坐标提取方法，以及可交互元素的交互方式、播放信息。

表 4-3 交互元素信息表

可交互类别		可交互元素名称	坐标提取	交互方式及播放信息
自由点		Point	“arg”属性值:(x, y) 需转换为(left, top)	沿 x 轴,y 轴拖动该点 1 次
元素上的点	线段上的点	PointOn PointPar PropElement	通过约束关系分析, 找到线段的两个端点标签, 然后根据变量值, 计算出(x, y), 需转换为(left, top)	
	圆上的点		通过约束关系分析, 找到圆的标签, 然后根据变量值, 计算出(x, y), 需转换为(left, top)	
	垂线上的点	PointOnVLine PointPar	通过约束关系分析, 找到线段的两个端点标签, 以及垂线的另一点标签, 然后根据变量值, 计算出(x, y), 需转换为(left, top)	
变量按钮		Variable	“position”属性值: (left, top)	获取变量的最小值、最大值、当前值、增量。均分变量播放时间, 在特定时间点点击按钮
动画按钮		AnimationVar		获取按钮的步数, 一步所需时间, 总时长。均分总时长, 在特定时间点点击按钮
动作按钮		ButtonFun		点击按钮 2 次

4.1.2 获取交互元素坐标值

表 4-3 中, 有两个坐标值: (x, y) 与 $(left, top)$, 前者指的是: 在资源本身坐标系上的坐标值, 即全局坐标, 后者指的是: 在以显示屏屏幕左上角为原点的坐标系上的坐标值, 即屏幕像素坐标。由于自动化软件工具根据屏幕像素坐标值进行操作, 因此最终需要获得每个交互元素在屏幕像素坐标系上的坐标值。

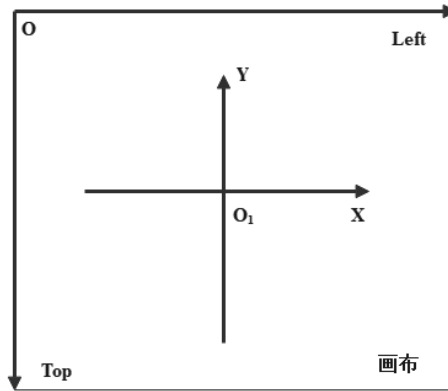


图 4-3 网络画板中的两个坐标系

通过一系列数学计算可将全局坐标值转换为屏幕像素坐标值:

$$\begin{aligned} left &= x \times scaleX + originPosX \\ top &= y \times scaleY + originPosY \end{aligned} \quad (4-1)$$

其中, $scaleX$ 与 $scaleY$ 分别为全局坐标系上 X、Y 坐标轴的像素比例尺, $(originPosX, originPosY)$ 是全局坐标系坐标原点的屏幕像素坐标。

在表 4-3 中, 有两大类交互元素需进行坐标转换: 自由点和元素上的点。根据上述公式 4-1, 可以得到自由点的屏幕像素坐标。此外, 元素上的点(包括线段上的点、圆上的点、垂线上的点)的屏幕像素坐标值还需通过约束关系分析得到。

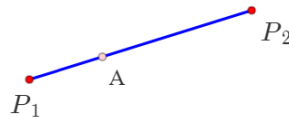


图 4-4 线段上的点 A

要求线段上的点的屏幕像素坐标, 首先要找到线段的两个端点的全局坐标, 这两个端点坐标可由约束关系分析得到。图 4-4 中, 线段上的点 A 的全局坐标 (x, y) 可由公式 4-2 获取:

$$\begin{aligned} x &= x_1 + v \times (x_2 - x_1) \\ y &= y_1 + v \times (y_2 - y_1) \end{aligned} \quad (4-2)$$

其中, v 为线段上的点对应的变量值, 通过查找资源的 JSON 文件的 p 字段获取, (x_1, y_1) 、 (x_2, y_2) 分别为线段两个端点的全局坐标。再由公式 4-1 完成坐标系的转换, 即可得到线段上的点 A 的屏幕像素坐标。

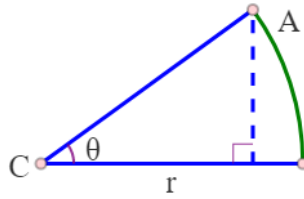


图 4-5 圆上的点 A

要求圆上的点的屏幕像素坐标, 首先要找到圆心的全局坐标, 这个坐标由约束关系分析可以得到。图 4-5 中, 圆上的点 A 的全局坐标 (x, y) 可由公式 4-3 至 4-4 得到:

$$\begin{aligned} \Delta x &= r \times \cos \theta \\ \Delta y &= r \times \sin \theta \end{aligned} \quad (4-3)$$

$$\begin{aligned} x &= centerX + \Delta x \\ y &= centerY + \Delta y \end{aligned} \quad (4-4)$$

其中, θ 为该点与圆心连线和 X 轴正方向的夹角, $\theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, $(centerX, centerY)$

是圆心 C 在全局坐标系上的坐标值。再由公式 4-1 完成坐标系的转换, 即可得到圆上的点 A 的屏幕像素坐标。

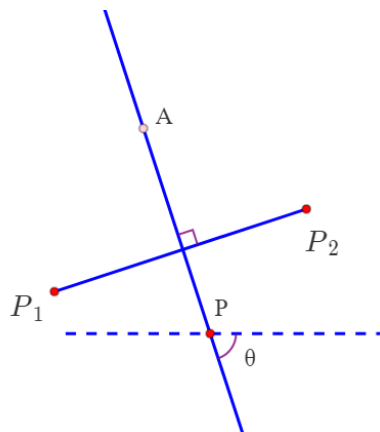


图 4-6 垂线上的点 A

要求垂线上的点的屏幕像素坐标,首先要找到垂线上另一点以及垂线垂直于的线段的两个端点的全局坐标,这三个坐标由约束关系分析可以得到。图 4-6 中,垂线上的点 A 与垂线上另一点 P 的距离 $|AP|$,以及垂线和 x 轴正方向的夹角 θ 可由公式 4-5 得到:

$$\begin{aligned} |AP| &= v \times |P_1P_2| \\ k_2 &= \frac{-1}{k_1} (k_1 \neq 0) \\ \theta &= \arctan k_2 \end{aligned} \quad (4-5)$$

其中, v 为垂线上的点对应的变量值,通过查找资源的 JSON 文件的 p 字段获取, $|P_1P_2|$ 为垂直于该垂线的线段的长度, k_1 、 k_2 分别为该线段、垂线的斜率。有了距离 $|AP|$ 及夹角 θ 后,且已知垂线上另一点 P 的全局坐标 (x_p, y_p) ,那么由公式 4-6 至 4-7,即可获取垂线上的点 A 的全局坐标 (x, y) :

$$\begin{aligned} \Delta x &= \begin{cases} |AP| \times \cos \theta, x_p < x_A \\ -|AP| \times \cos \theta, x_p \geq x_A \end{cases} \\ \Delta y &= \begin{cases} |AP| \times \sin \theta, x_p < x_A \\ -|AP| \times \sin \theta, x_p \geq x_A \end{cases} \end{aligned} \quad (4-6)$$

$$\begin{aligned} x &= x_p + \Delta x \\ y &= y_p + \Delta y \end{aligned} \quad (4-7)$$

再由公式 4-1 完成坐标系的转换,即可得到垂线上的点的屏幕像素坐标。

4.1.3 获取约束关系

在第三章第一节约束关系提取中,默认已存在初步约束关系,介绍了如何从初步的约束关系中进一步过滤约束关系。本小节将阐述如何获取网络画板平台资源中交互元素的初步约束关系。

在资源的 JSON 文件中,元素的“arg”属性中存储了当前元素的约束关系,该属性列表中出现数字可能是对当前元素实施约束的元素的索引 eid ,出现的字母字符串可能是对当前元素实施约束的变量的名称,因此根据在该属性中出现的数字与变量提取初步约束关系,如表 4-4 中的算法 4-1 所示。首先,第 5-7 行得到变量名到元素索引的映射,当且仅当当前元素为变量按钮时,其它元素才可能受约束于它对应的变量,遍历元素集中的每个变量按钮,建立变量名到元素索引的映射。第 8-20 行遍历元素集,遍历每个元

素对应的“arg”属性列表，获取约束当前元素的元素，其中第 10-12 行表明若当前遍历对象为数字，且该数字不是该元素对应的索引，那么该数字对应的索引元素即为对当前元素实施约束的元素，它们之间的约束关系为控制约束关系；第 13-18 行表明若当前遍历对象为纯字母字符串，若能在变量名到元素索引的映射找到该字符串，那么该映射对应的索引元素即为对当前元素实施约束的元素，它们之间的约束关系为数据约束关系。

表 4-4 提取初步约束关系算法

算法 4-1 generateOriCR ($E, argList$)
输入：元素集 E ; arg 属性行列表 $argList$
输出：初步约束关系集 CR
1: $CR = \{\}$
2: $DC = \{\}$
3: $CC = \{\}$
4: $variableToIndex = \{\}$
5: for $e_i \in E$ and $e_i.name == Variable$ do
6: $variableToIndex[argList[i][0]] = i$
7: end for
8: for $e_i \in E$ do
9: for j in $argList[i]$ do
10: if j is a number and $j \neq i$
11: add $\langle e_i, e_j \rangle$ into CC
12: end if
13: if j is an alphabetic string
14: find j in $variableToIndex, k = variableToIndex[j]$
15: if $k \neq i$
16: add $\langle e_i, e_k \rangle$ into DC
17: end if
18: end if
19: end for
20: end for
21: $CR = DC \cup CC$
22: return CR

4.1.4 生成操作序列

网络画板中元素上的点包括三类点：线段上的点、圆上的点和垂线上的点，这些点都是其他元素的基础上构建的，例如线段上的点是在线段的基础上构建的，而线段在两端的自由点的基础上构建的。分析约束关系可知，这类元素受其它自由点的约束，它们之间存在控制约束关系，若在自动交互过程中，先改变了自由点的坐标，则会导致元素上的点的屏幕交互坐标失效，因此将元素上的点类别的元素的优先级设为 5，将自由点

元素优先级设为 4，将变量元素、动画元素和动作元素的优先级分别设置为 3、2 和 1。按照算法 4-1 的思想，按优先级由高到低排列元素，并根据约束关系和布局位置得到元素操作序列。

以图 4-1 中的案例“多边形的平移”为例，在提取交互元素信息、获取约束关系、获取屏幕坐标信息并生成操作序列后，操作序列中元素编号排序为(15, 5, 6, 13, 19)。如表 4-5 所示。

表 4-5 “多边形的平移”资源中的操作序列信息

元素编号(eid)	元素名	坐标信息(left, top)	播放信息
15	PointOnLine	(691, 703)	无
5	Point	(11, 352)	无
6	Point	(251, 175)	无
13	Point	(1056, 704)	无
19	AnimationVar	(46, 26)	(100, 50, 5000.0)

4.2 获取交互过程图片

在得到交互元素的位置信息、交互信息、播放信息以及操作序列后，下一步需要考虑选用哪款自动化软件工具与资源进行自动交互。

AutoHotKey 由 Chris Mallett 及其小组研发，是一款开源且免费的、简洁便捷却功能强大的、以脚本为基础的热键，运行在 Windows 平台上。AutoHotKey 具有多样功能，例如可实现热键、热字符串、窗口置顶、快速输入、处理系统管理相关工作等功能。用户利用这款工具，能够通过编写自定义脚本实现几乎所有 Windows 平台操作的自动化，而且仅占用极小内存，是一个十分灵活，可被高度自定义、可用性极高的工具。Chris 设计 AutoHotKey 的初衷是简化用户的日常操作，使用户从重复机械的活动中解放出来。

因为 AutoHotKey 能够向计算机发送按键和模拟鼠标操作，与其他自动化软件工具相比，AutoHotKey 具有轻量级的特点，本文选用 AutoHotKey 来模拟用户与交互式教育资源进行交互。

由于图像与视频相比，占用的内存更小，且表达信息的方式更加直接，因此选用图片的形式来记录交互过程。在交互的过程中利用 AutoHotKey 的第三方库文件 Gdip.ahk 来进行截图并保存图片。在第三章中提到，在自动交互过程中，会利用自动化软件工具在特定时刻对用户界面进行截图。截图的时间点选取如下：

对自由点，拖拽该点一次，等待 1500ms 后截取一张图；

对元素上的点，拖拽该点一次，等待 1500ms 后截取一张图；

对动画按钮，获取动画的步数 $stepNums$ 、一步的步长间隔时间 $timer$ 及总播放时间 $totalTime$ ，将总播放时间等分为 n 份，从而获取 $n+1$ 个关键时间点。在这些时间点处点击动画按钮两次，点击一次暂停，截取一张图，然后再点击一次播放，共截取 $n+1$ 张图， n 的值由以下公式确定：

$$n = \begin{cases} \text{floor}(stepNums \div 10), & stepNums \geq 100 \\ 20 + \text{floor}(stepNums \div 5), & stepNums \in (50, 100) \\ stepNums, & stepNums \leq 50 \end{cases} \quad (4-8)$$

当变量元素为变量按钮时，获取变量的最小值 $minValue$ 、当前值 $curValue$ 和最大值 $maxValue$ ，以及每步变量的增加值 $addValue$ 、一步的步长间隔时间 $timer$ 。从而可获得当前位置播放到最大值位置的步数 $stepNums$ 及所需时间，将播放时间均分为 k 份，在这些时间点处点击变量按钮两次，点击一次暂停，截取一张图，然后再点击一次播放，共截取 $k+1$ 张图， k 的值由以下公式确定：

$$k = \begin{cases} \text{floor}(stepNums \div 10), & stepNums \geq 200 \\ 20 + \text{floor}(stepNums \div 5), & stepNums \in (50, 200) \\ stepNums, & stepNums \leq 50 \end{cases} \quad (4-9)$$

对动作按钮，点击按钮两次，截取两张图，两次点击间隔时间为 4500ms。

使用 AutoHotKey 与资源进行自动交互过程如表 4-6 所示。

表 4-6 AutoHotKey 与资源交互过程

使用 AutoHotKey 自动交互过程
a、读取操作序列中的交互元素信息（名称、坐标信息、交互方式、播放信息）。
b、打开浏览器，进入资源浏览页面。
c、进入画布最大化模式。
d、按照操作序列，依次与交互元素进行交互，并在特定时刻使用 Gdip 截图。
e、保存截图，关闭浏览器。

4.3 可视化预览

经过自动交互过程，可获得交互过程中的截图。由于此时截图数量仍较多，因此需提取其中有代表性的关键截图。在将图片灰度化后，提取图片的颜色直方图特征和边缘特征，将边缘特征投影到水平方向后，与颜色直方图特征组合在一起作为特征向量，然

后使用 K-Means 方法对这些特征向量进行聚类。由于 K-Means 聚类方法的结果受 K 值影响较大, 经过实验, 本文根据网络画板资源中的可交互元素数量, 给出预估 K 值。

$$K = pn + pon + 5 \times vn + 2 \times fn + 1 \quad (4-10)$$

其中, pn 为自由点的数量, pon 为元素上的点的数量, vn 为变量按钮及动画按钮的数量, fn 为动作按钮数量。

将聚类结果中每类中非连续图片分割开来, 划分为几个区间, 使得每个区间内图片所属类别相同且连续, 从每个区间内选取一张图片, 这些图片构成了关键图片集合。

以“多边形的平移”资源为例, 在经过可交互元素信息提取、操作序列生成后, 利用 AutoHotKey 与该资源进行交互, 分析提取关键图片后, 部分关键图片如图 4-7 所示。

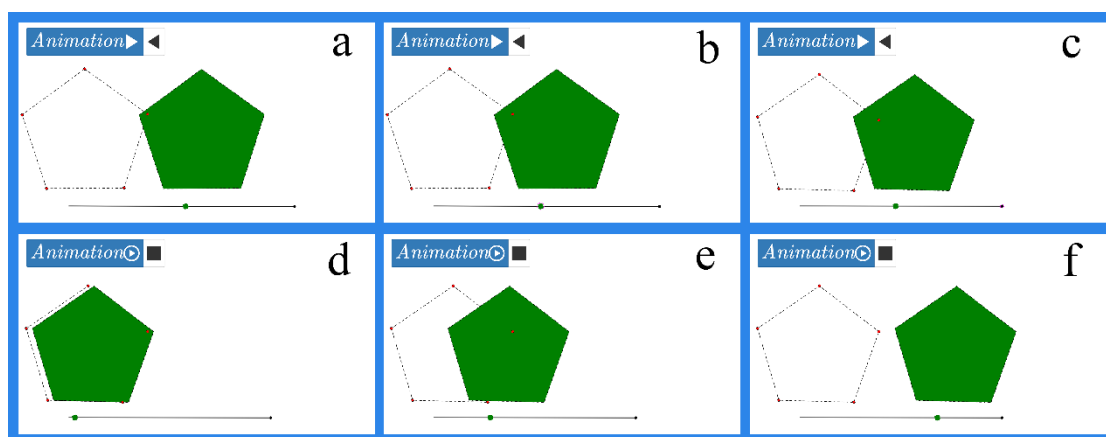


图 4-7 “多边形的平移”资源的部分关键图片

在获取到关键图片后, 按关键图片的顺序, 将帧速率调整为 3 帧每秒, 将关键图片组合为 GIF 动图, 作为最终的图片预览形式。

4.4 本章小结

本章具体介绍了将 IERVP 方法应用于动态数学平台网络画板, 相比较第三章重点阐释了几点: 第一点, 阐释了如何从资源的 JSON 文件中手动分析 JSON 文件的结构, 从中总结存储特征。第二点, 介绍了交互元素坐标值获取方式。介绍了全局坐标值与屏幕像素坐标值的关系, 详细介绍了如何获取可交互元素的屏幕像素坐标。第三点, 解释了如何获取交互元素间的初始约束关系, 主要是根据“arg”属性列表中出现数字和变量获取。第四点, 详细介绍了自动交互过程中截图的关键时间点给定参考值。针对不同类型的交互元素, 采用了有针对性的截图方法, 特别是对于按钮类元素, 根据其播放信息, 精准控制播放时间, 在特定时间进行截图, 以展现一次完整播放过程, 更加准确全面地

获取资源隐藏的动态内容。第五点，介绍了 **K-Means** 聚类方法预设 **K** 值的给定以及 GIF 动图帧速率的给定。

第五章 实验与分析

5.1 实验

5.1.1 实验数据集

本文获取网络画板平台下，“小学数学”、“初中数学”、“高中数学”这三个类别下 7263 个动态几何数学教育资源。其中，小学数学资源有 1561 个，主要包括常见的量、图形的认识、图形的变换等类别；初中数学资源有 3955 个，主要包括图形的性质、图形的变化、函数等类别；高中数学资源有 1747 个，主要包括平面向量、平面解析几何、三角函数等类别。样本资源的数据分布如图 5-1 所示。

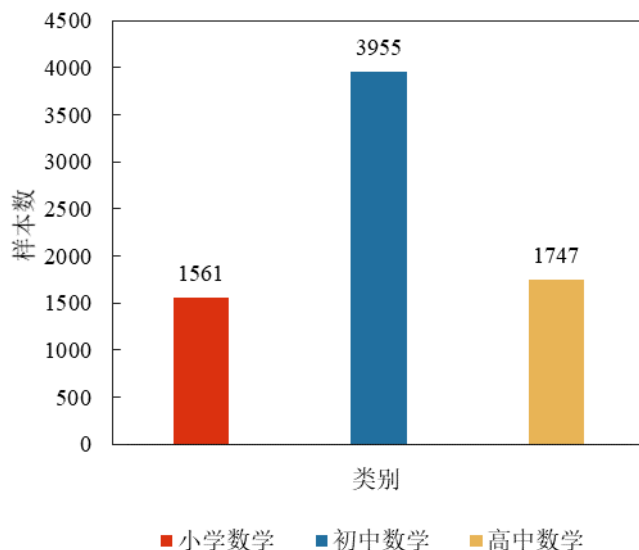


图 5-1 资源样本数据分布

首先，以这 7263 个资源为样本，验证 IERVP 方法是否能够获取预览图片。然后，考虑到网络画板上资源的质量参差不齐，例如有些资源由作者随意创作，导致内容意义不明确，而资源标题与资源内容高度相关的资源由于作者创作意图明确，因此传达的知识较为清晰。为了确保实验资源及实验的质量，本文从上述 7263 个资源中选取标题与内容相关度高的资源，然后从中随机选取 15 个资源，以这 15 个资源为样本进行第二轮实验。

5.1.2 实验方式

第一轮实验使用 IERV 方法，从资源对应的 JSON 文件中提取交互语义信息后，使用 AutoHotKey 工具执行操作序列，自动与这 7263 个资源进行交互，并保存过程中的截图。然后针对一个资源对应的截图，获取两个特征：图片的颜色直方图特征和边缘特征，将这两个特征组合在一起，使用 K-Means 方法进行聚类，提取有代表性的关键图片，生成图片预览。生成的所有预览图片都存储在服务器端的数据库中。

然后进行第二轮实验，以随机挑选出的 15 个资源作为样本，以发放问卷的方式进行调研。问卷中列出这 15 个资源对应的网络画板网址链接及对应的预览图片。让每个参与者点开链接，手动与样本资源进行交互，理解资源内容，记录手动与资源交互所需的时间，然后观察该资源对应的预览图片，对预览图片的正确性、完整性、能够在多大程度反映资源内容分别进行评估，再分别给出评分。具体问卷内容见附录。

5.1.3 实验评估对象

由于与网络画板上的资源进行交互需要一定的动态几何知识储备，本文邀请了来自广州大学的 21 名师生参与实验，他们熟悉网络画板的交互操作，能客观评价图片预览效果。

5.1.4 评估指标

1、手动交互时间。实验参与者分别打开资源的链接，手动与资源交互，了解资源的内容，并记录手动交互所需的时间。

2、预览正确性。评估使用 AutoHotKey 进行自动交互过程中，与元素交互的顺序、交互方式是否正确。评分区间为[1, 5]。分数越高，代表着图像预览反映的自动交互顺序、交互方式越合理。

3、预览完整性。评估自动交互过程是否与所有可交互元素都进行了交互。评分区间为[1, 5]。分数越高，代表着图像预览反映的自动交互过程中交互操作越完整，与该进行交互的元素都进行了交互操作。

4、预览反映内容程度。参与者观察对应的图片预览，评估资源的图片预览能在多大程度上反映资源的内容。评分区间为[1, 5]。分数越高，代表着图像预览反映资源内容的效果越好。

5.2 实验结果及分析

5.2.1 图片预览案例

选取“正多边形的迭代”、“圆面积”、“万花筒”和“缩放变换”四个资源，展示图片预览效果。

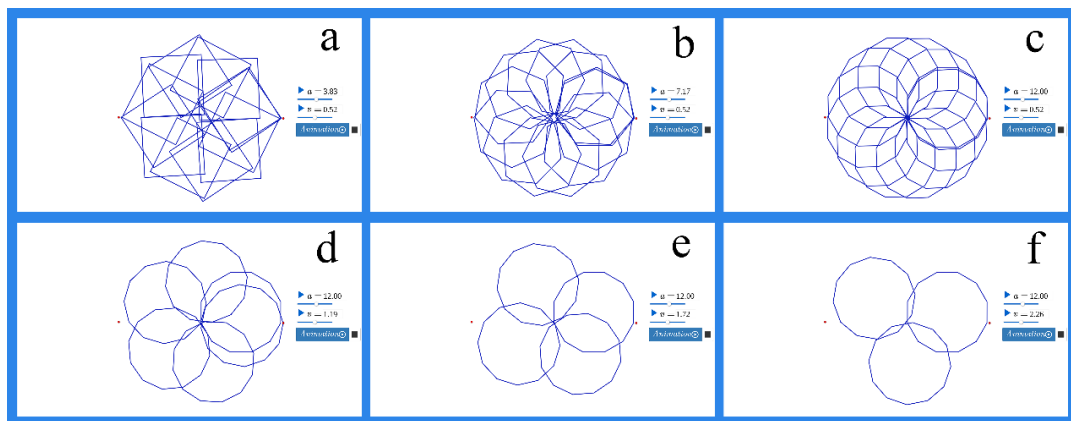


图 5-2 “正多边形的迭代”资源的部分关键图片

“正多边形的迭代”资源是一个迭代在正多边形上的应用，通过分析 JSON 文件，可获取到交互元素的坐标信息和操作信息，在“正多边形的迭代”资源中，有自由点、两个变量按钮、一个动画按钮。根据生成的操作序列与这些交互元素交互，可获取该资源对应的图片预览，在图 5-2 中，可以清晰地观察到迭代过程中的多边形变化情况，可以有效地帮助用户选择所需的案例，节省与资源进行手工交互的时间。

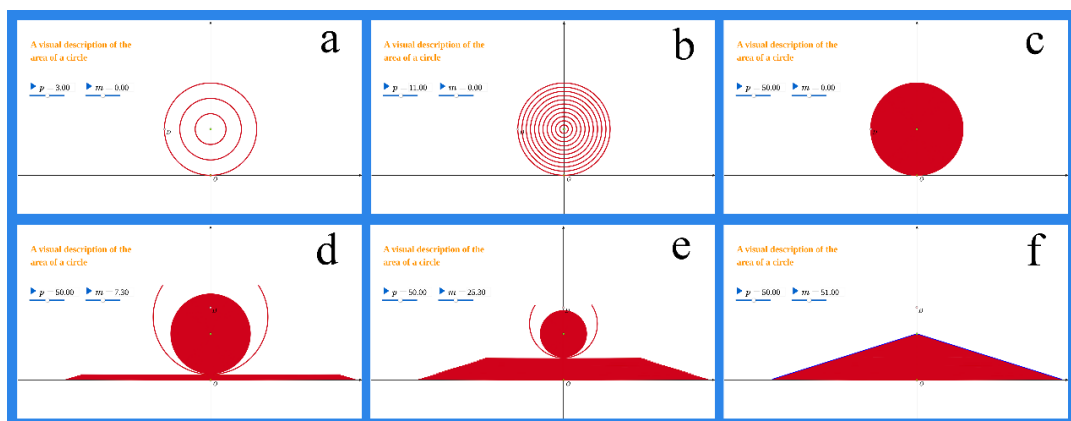


图 5-3 “圆面积”资源的部分关键图片

“圆面积”资源是为了帮助学生理解“圆的面积”的概念而创建，该资源里有两个变量按钮，变量尺 p 控制划分圆的圆环个数，变量尺 m 控制打开和关闭圆环，如图 5-3 所示，该资源的图片预览可以清晰地反映实心圆的构建过程——由无数个半径不等的圆

环构成，然后将每个圆环拆开，可以得到一条线，将这些圆环拆成的线叠加在一起，构成一个三角形，那么圆的面积即为三角形的面积。从图片预览中可以反映这个资源的有趣性。

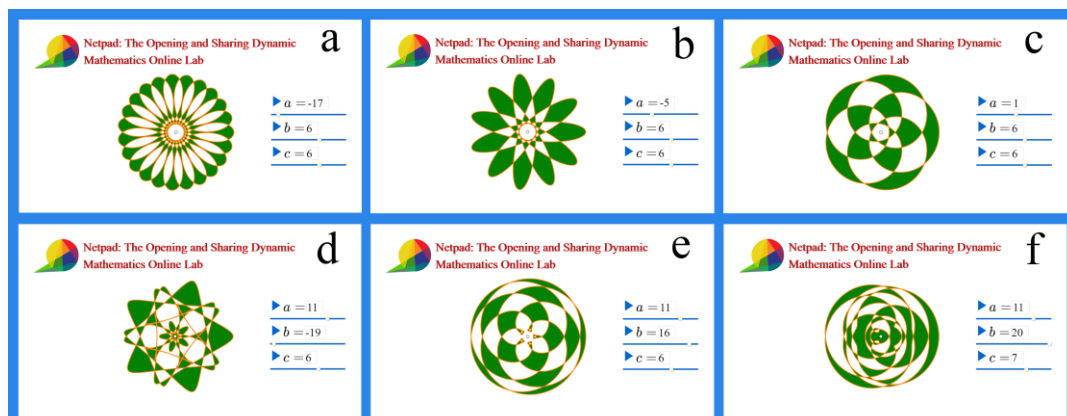


图 5-4 “万花筒”资源的部分关键图片

“万花筒”资源通过描绘某个点的运动轨迹展现几何变换之美，因运动轨迹如花般绚丽多姿而命名。在“万花筒”资源中，有三个变量按钮。根据生成的操作序列依次与这些变量按钮交互，对交互界面进行截图，获取截图中的关键图片，可获取该资源对应的图片预览。在图 5-4 中，可以清晰地观察到随着变量值的改变，运动轨迹发生改变，描绘生成的“花”的形状也随之变化，通过图片预览能有效地帮助用户快速了解资源的动态内容。

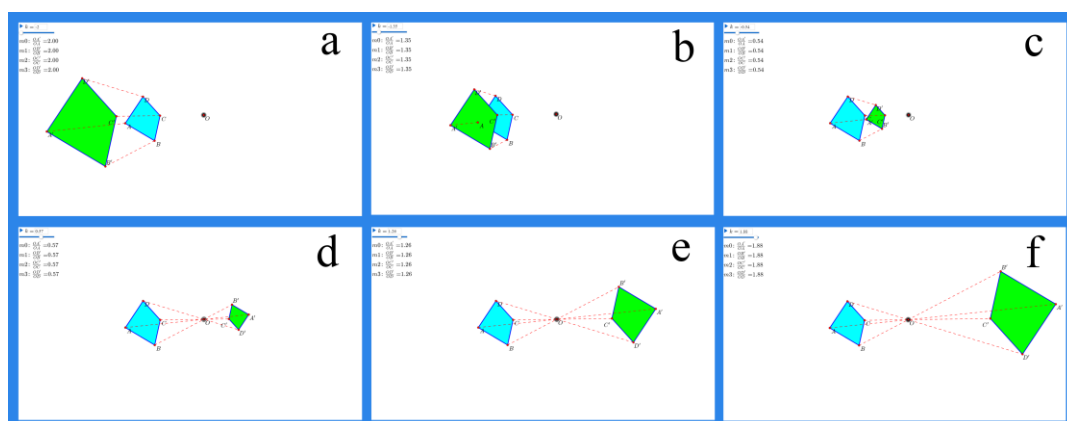


图 5-5 “缩放变换”资源的部分关键图片

“缩放变换”资源展现了四边形经缩放变换后得到的图形形状，在该资源中，有一个变量按钮，点击这个按钮，可以改变缩放值的大小以及缩放的方向，如图 5-5 所示，初始状态经缩放变换后的图形在基准四边形的左边，慢慢移动到了基准四边形的右边，在穿过图中某一点后，变换后的图形上下左右颠倒，与基准四边形呈现中心对称缩放状态。通过图片预览，能够快速看到缩放变换过程。

从上述四个资源案例中,可以看到,图片预览能够很好地帮助教师在节约时间的前提下,选择适合作为教案的素材资源,并且由于动态数学交互式教育资源具有动态性、交互性的特点,可以使学生在互动中接受数学的概念知识,生动且形象。

5.2.2 实验结果及分析

收集问卷并将评分情况汇总,每个资源平均预览正确性得分如图 5-6 所示。

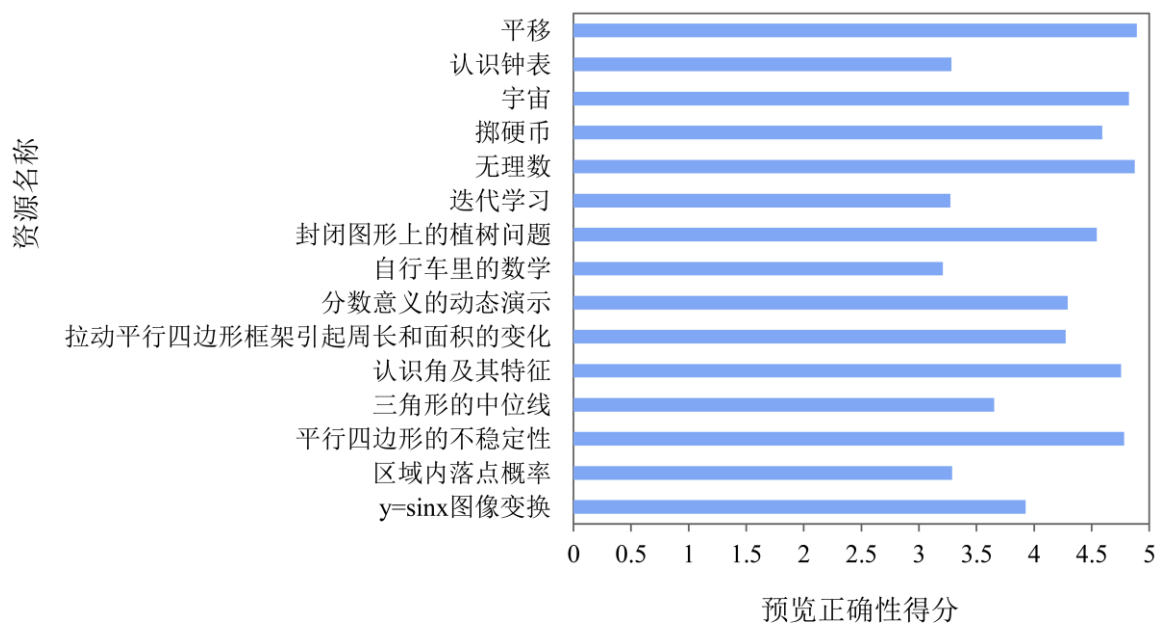


图 5-6 预览正确性得分图

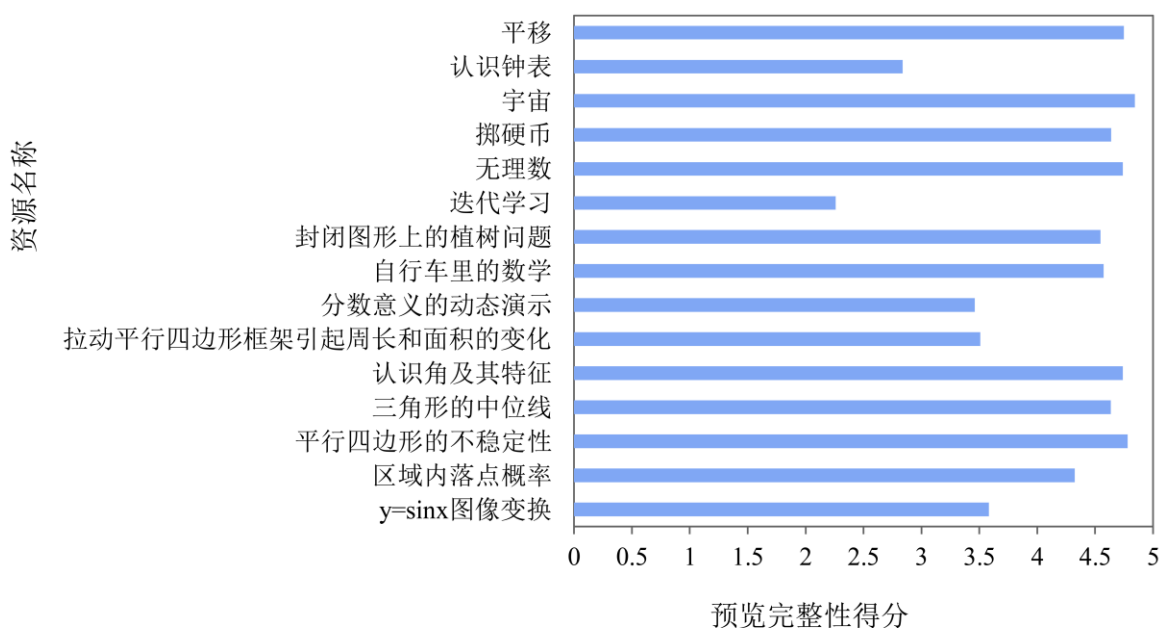


图 5-7 预览完整性得分图

分析图 5-6 中的数据, 可以看到所有资源的正确性评分都在 3.2 分以上, 大多数资源的得分在 3.5 分以上, 所有资源平均得分为 4.164 分。这证明了使用本文提出的 IERVP 方法, 在自动交互过程中与元素交互的顺序、交互方式较为合理。

图 5-7 列出了每个资源平均预览完整性得分, 分析图 5-7 中的数据, 可以看到所有资源的评分都在 2.2 分以上, 大多数资源的得分在 3.5 分以上, 所有资源平均分为 4.147 分。这证明了使用本文提出的 IERVP 方法, 在自动交互过程中能与大部分交互元素进行充分的交互操作。

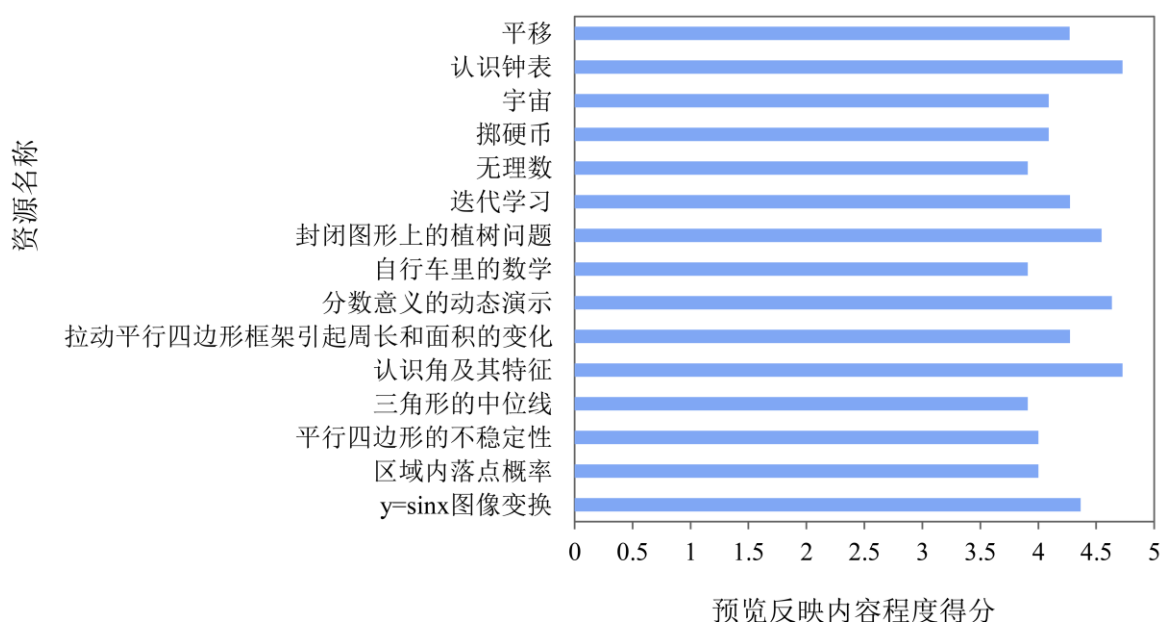


图 5-8 预览反映内容程度得分图

图 5-8 列出了每个资源平均预览反映内容程度得分, 分析图 5-8 中的数据, 可以看到所有资源的评分都在 3.7 分以上, 大多数资源的得分在 4 分以上, 所有资源平均分为 4.248 分。这证明了使用本文提出的 IERVP 方法, 生成的图片预览能够代表动态几何数学教育资源的大部分内容。

接下来将解释图片预览能为用户节省多长时间。根据参与者的意见, 将阅读一张图片所需的时间设置为 1 秒, 阅读一个资源对应的资源预览所需的时间为 GIF 图片中静态图片数乘以阅读单张图片所需时间。在实验过程中, 记录每个参与者手动与资源交互所需的时间。图片预览节省的时间百分比如图 5-9 所示。

如图 5-9 所示, 利用本文提出的图片预览, 几乎理解所有资源内容的时间都能节约在 85% 以上, 平均节约 87.3% 的时间。平均节约时间和平均预览正确性、完整性、反映内容程度得分证实了本文提出的 IERVP 方法的有效性。

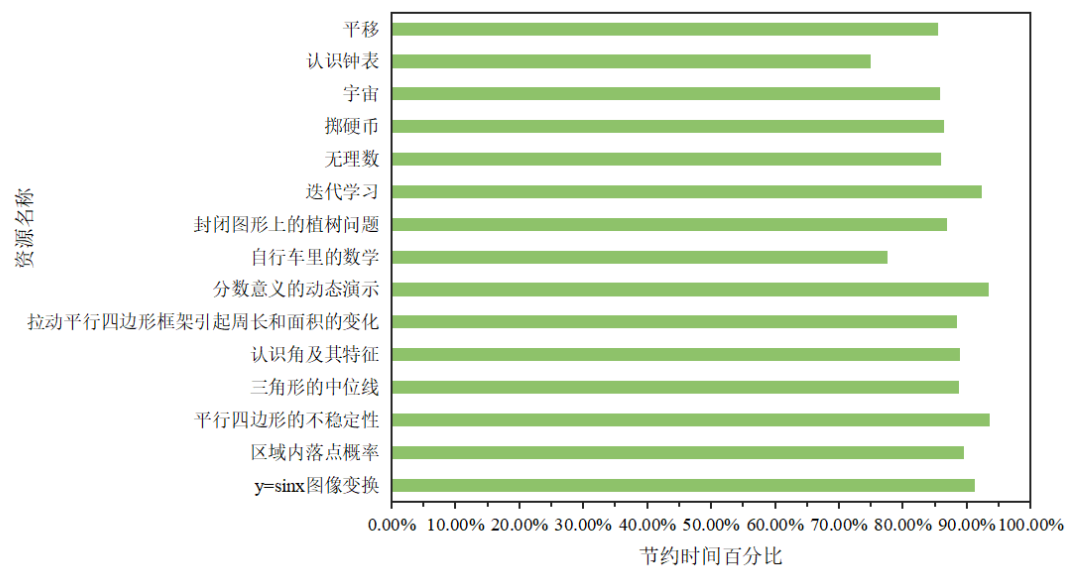


图 5-9 节约时间百分比图

5.3 本章小结

本章进行了两轮实验，首先选取了动态数学平台网络画板上三个类别中的资源，按 IERV P 方法生成图片预览。然后在这些资源中，选取 15 个有代表性的资源，取出其预览图片，邀请熟悉网络画板操作的师生参与问卷评估，参与者首先手动与这些资源进行交互，了解资源内容并记录手动交互时间，然后对预览图片的预览正确性、预览完整性、能在多大程度上反映资源内容进行评分，并分析图片预览节约时间比。网络画板的实验证明了本文提出的 IERV P 方法的可操作性和有效性。

第六章 总结与展望

6.1 研究总结

由于交互式教育资源的互动性,理解资源的内容需要付出一定的成本。因此,本文在对用户交互理论与手动交互过程进行分析的基础上,提出了一种生成交互式教育资源可视化预览(IERV)的方法。该方法首先提取了交互语义信息,包括从资源文件中提取交互元素信息,然后分析约束关系,最后生成操作序列;然后利用自动化软件工具自动执行操作序列,从而模拟交互过程,并在特定时间点对交互界面进行截图;最后组合了图像的颜色直方图特征和边缘特征,使用 K-Means 方法聚类,提取截图中的关键图片。将关键图片按顺序排列生成一张 GIF 动图,作为该资源最终的图片预览形式。

然后将 IERV 方法应用到动态数学资源平台——网络画板上,对具体技术实现过程展开详细阐释。最后,选取网络画板上资源进行了两轮实验,邀请熟悉网络画板操作的人员进行问卷测评,从节约时间比、预览正确性、预览完整性、预览反映内容程度四个维度评估图片预览效果。结果显示,本文提出的 IERV 方法能够有效提取资源中的交互语义信息、模拟交互过程、提取关键图片,生成的图片预览平均节省用户 87.3%的时间,帮助用户理解 88.6%的资源内容,对用户理解交互式教育资源的内容有较大帮助。

6.2 存在问题和不足

本文研究过程存在以下不足之处:

(1) 用户与资源进行手动交互过程相关研究不足。对于用户在打开资源后,按何种顺序与交互元素进行交互,仅从现有的研究文献及少量用户意见进行调研,缺少包含足够样本数的用户主观交互顺序调研。

(2) 对按钮类交互元素的交互操作考虑不全面。对这类元素,获取当前值到最大值变化过程中的截图,但没有考虑到当前值为最大值这一特殊情况,导致在这一特殊情况下,没有与按钮类元素进行交互,使得交互操作不全面。

(3) 图片特征选取的方法不完善。仅提取了图片的颜色直方图特征和边缘特征,再将边缘特征投影到水平方向后,组合这两种特征作为截图的特征。

(4) 对网络画板上资源的图片预览效果调查不够精细, 仅从节约时间比、预览正确性、预览完整性、预览反映内容程度四个维度评估图片预览效果, 应拓宽并细化评估维度, 使问卷调查结果更为直观。

6.3 展望

由于研究时间、调研用户数量等条件限制, 本文关于交互式教育资源可视化预览研究仍存有一些不足之处, 未来将进行进一步工作:

(1) 完善操作序列生成规则。邀请更多的用户参与调研, 收集用户主观交互顺序数据, 从中总结更加全面合理的交互顺序规则, 从而完善操作序列生成规则。

(2) 完善模拟交互过程中的交互操作。权衡交互时长, 使得每个交互元素都有充足的与之交互的时间, 从而充分展示资源的动态内容。

(3) 完善截图的特征提取方式。尝试从更多方面提取截图的特征, 例如尝试使用神经网络模型提取特征, 或者将更多的传统特征相结合得到更全面的特征。

参考文献

- [1] 张宝全, 黄祖源, 周枫, 等. 电力行业信息安全攻防演练研究及应用[J]. 软件, 2017, 38(9): 40-45.
- [2] 冯影. 智慧教育背景下师范生创新能力评价指标体系构建研究[D]. 甘肃: 西北师范大学, 2021.
- [3] 石晓红. 任务驱动的数学教育资源可视计算模型研究[D]. 广东: 广州大学, 2021.
- [4] 张景中, 葛强, 彭翥成. 教育技术研究要深入学科[J]. 电化教育研究, 2010, (2): 8-13.
- [5] Kopf J, Shamir A, Peers P. Content-adaptive image downscaling[J]. ACM Transactions on Graphics (TOG), 2013, 32(6): 173.1-173.8.
- [6] Oeztireli A C, Gross M. Perceptually based downscaling of images[J]. ACM Transactions on Graphics (TOG), 2015, 34(4): 77.1-77.10.
- [7] Weber N, Waechter M, Amend S C, et al. Rapid, detail-preserving image downscaling[J]. ACM Transactions on Graphics (TOG), 2016, 35(6): 205.1-205.6.
- [8] Gastal E S L, Oliveira M M. Spectral remapping for image downscaling[J]. ACM Transactions on Graphics (TOG), 2017, 36(4): 145.1-145.16.
- [9] Strobbe C, Frees B, Engelen J. An accessibility checker for libreoffice and openoffice.org writer[C]. International Conference on Computers for Handicapped Persons. 2012: 484-491.
- [10] Meena Y K, Gopalani D. Evolutionary algorithms for extractive automatic text summarization[J]. Procedia Computer Science, 2015, 48: 244-249.
- [11] Anh B T M, My N T, Trang N T T. Enhanced genetic algorithm for single document extractive summarization[C]. Proceedings of the Tenth International Symposium on Information and Communication Technology. 2019: 370-376.
- [12] Banerjee S, Mitra P, Sugiyama K. Multi-document abstractive summarization using ilp based multi-sentence compression[C]. Twenty-Fourth International Joint Conference on Artificial Intelligence. 2015: 1208-1214.
- [13] Nallapati R, Zhai F, Zhou B. Summarunner: A recurrent neural network based sequence model for extractive summarization of documents[C]. Thirty-first AAAI

- p>conference on artificial intelligence. 2017: 3075-3081.
- [14] Nguyen M T, Cuong T V, Hoai N X. Exploiting user comments for document summarization with matrix factorization[C]. Proceedings of the Tenth International Symposium on Information and Communication Technology. 2019: 118-124.
- [15] Schoeffmann K, Boeszoermenyi L. Video browsing using interactive navigation summaries[C]. 2009 Seventh International Workshop on Content-Based Multimedia Indexing. IEEE, 2009: 243-248.
- [16] Yuan Y, Ma L, Zhu W. Sentence specified dynamic video thumbnail generation[C]. Proceedings of the 27th ACM International Conference on Multimedia. 2019: 2332-2340.
- [17] Zhao B, Xu S, Lin S, et al. A new visual navigation system for exploring biomedical Open Educational Resource (OER) videos[J]. Journal of the American Medical Informatics Association, 2016, 23(e1): 34-41.
- [18] Zhao B, Lin S, Luo X, et al. A novel system for visual navigation of educational videos using multimodal cues[C]. Proceedings of the 25th ACM international conference on Multimedia. 2017: 1680-1688.
- [19] Devanne M, Wannous H, Berretti S, et al. 3-d human action recognition by shape analysis of motion trajectories on riemannian manifold[J]. IEEE transactions on cybernetics, 2014, 45(7): 1340-1352.
- [20] Wolf W. Key frame selection by motion analysis[C]. 1996 IEEE international conference on acoustics, speech, and signal processing conference proceedings. IEEE, 1996, (2): 1228-1231.
- [21] 马楠, 石祥滨, 代钦, 等. 一种音乐舞蹈视频关键帧提取方法[J]. 系统仿真学报, 2018, 30(7): 2801-2807.
- [22] Tang H, Liu H, Xiao W, et al. Fast and robust dynamic hand gesture recognition via key frames extraction and feature fusion[J]. Neurocomputing, 2019, 331: 424-433.
- [23] Zhang H J, Wu J, Zhong D, et al. An integrated system for content-based video retrieval and browsing[J]. Pattern recognition, 1997, 30(4): 643-658.
- [24] Chen L, Wang Y. Automatic key frame extraction in continuous videos from construction monitoring by using color, texture, and gradient features[J]. Automation

- in Construction, 2017, 81: 355-368.
- [25] Nasreen A, Roy K, Roy K, et al. Key frame extraction and foreground modelling using K-means clustering[C]. 2015 7th International Conference on Computational Intelligence, Communication Systems and Networks. IEEE, 2015: 141-145.
- [26] Gharbi H, Bahroun S, Zagrouba E. Key frame extraction for video summarization using local description and repeatability graph clustering[J]. Signal, Image and Video Processing, 2019, 13(3): 507-515.
- [27] 周舟, 韩芳, 王直杰. 面向手语识别的视频关键帧提取和优化算法[J]. 华东理工大学学报(自然科学版), 2021, 47(1): 81-88.
- [28] Wang Z, Kumar M, Luo J, et al. Extracting key frames from consumer videos using bi-layer group sparsity[C]. Proceedings of the 19th ACM international conference on Multimedia. 2011: 1505-1508.
- [29] Chen X, Zhang Y, Ai Q, et al. Personalized key frame recommendation[C]. Proceedings of the 40th international ACM SIGIR conference on research and development in information retrieval. 2017: 315-324.
- [30] Zhao A, Lai Y, Liu Y, et al. Key frame extraction algorithm for surveillance video based on golden section[C]. Proceedings of the 2019 International Symposium on Signal Processing Systems. 2019: 78-81.
- [31] Ma C, Chen X, Rao Y, et al. Generating Summarized Preview for Education Resource based on Exploring and Comparing GUIs[C]. In Proceedings of the KSI Research International Conference on Software Engineering and Knowledge Engineering. 2016: 141-146.
- [32] Song Y, Redi M, Vallmitjana J, et al. To click or not to click: Automatic selection of beautiful thumbnails from videos[C]. Proceedings of the 25th ACM international on conference on information and knowledge management. 2016: 659-668.
- [33] Chen J, Chen X, Ma L, et al. Temporally grounding natural sentence in video[C]. Proceedings of the 2018 conference on empirical methods in natural language processing. 2018: 162-171.
- [34] Liu M, Wang X, Nie L, et al. Attentive moment retrieval in videos[C]. The 41st international ACM SIGIR conference on research & development in information

retrieval. 2018: 15-24.

- [35] Bernard J J, 吴丹, 李凯, 等. 技术视野 人文情怀——国内首届“交互与信息行为研究学术研讨会”纪要[J]. 图书情报知识, 2019, (4): 122-129.
- [36] 杨琴. 基于交互元素的扁平化设计可用性问题研究[D]. 浙江: 浙江工业大学, 2017.
- [37] 唐纳德·A·诺曼. 设计心理学.1, 日常的设计. 增订版[M]. 北京: 中信出版社, 2015.
- [38] Fitts P M. The information capacity of the human motor system in controlling the amplitude of movement[J]. Journal of Experimental Psychology General, 1992, 121(3): 262.
- [39] 王佳珂. 基于轨迹球操控的数字界面用户情感体验研究[D]. 江苏: 东南大学, 2018.
- [40] Hick W E. On the rate of gain of information[J]. Quarterly Journal of experimental psychology, 1952, 4(1): 11-26.
- [41] Liu Z, Mao J, Wang C, et al. Enhancing click models with mouse movement information[J]. Information Retrieval Journal, 2017, 20(1): 53-80.
- [42] 刘未松. 基于交互行为的雷达控制界面设计[D]. 陕西: 西安工业大学, 2021.
- [43] 彭存银. 基于层次 MEA 图规划的 Web 用户界面自动化测试的研究[D]. 广东: 中山大学, 2009.
- [44] 宋晓莉. 嵌入式模具保护图像处理技术的研究与实现[D]. 四川: 电子科技大学, 2017.
- [45] 杨淑莹. 图像识别与项目实践:VC++、MATLAB 技术实现[M]. 北京: 电子工业出版社, 2014.
- [46] 韩晓军. 数字图像处理技术与应用[M]. 北京: 电子工业出版社, 2009.
- [47] 丁芝娟. 基于图像处理的车牌识别系统设计与应用[D]. 陕西: 长安大学, 2011.
- [48] 刘英波. 洗车店内车牌图像定位研究[D]. 黑龙江: 哈尔滨工程大学, 2014.
- [49] 白福忠. 视觉测量技术基础[M]. 北京: 电子工业出版社, 2013.
- [50] 孙新. 基于联合模型的目标跟踪[D]. 北京: 北京邮电大学, 2018.
- [51] Swain M J, Ballard D H. Indexing via color histograms[M]. Active perception and robot vision. Berlin: Springer, 1992: 261-273.

- [52] 柳杨. 数字图像物体识别理论详解与实战[M]. 北京: 北京邮电大学出版社, 2018.
- [53] 马伟. 基于机器视觉的泊车位检测关键技术研究[D]. 陕西: 西安理工大学, 2015.
- [54] 潘达, 史萍, 郑义. 基于分块形状特征匹配的台标识别系统[J]. 广播与电视技术, 2011, 38(6): 82-84.
- [55] 周培诚, 程臻, 姚西文, 等. 高分辨率遥感影像解译中的机器学习范式[J]. 遥感学报, 2021, 25(1): 182-197.
- [56] 管皓, 秦小林, 饶永生. 动态数学数字资源开放平台的研究与设计[J]. 哈尔滨工业大学学报, 2019, 51(5): 14-22.
- [57] Guan H, Rao Y, Zhang J, et al. Method for Processing Graph Degeneracy in Dynamic Geometry Based on Domain Design[J]. Journal of Computer Science and Technology, 2021, 36(4): 910-921.
- [58] 巩江源. 网络画板在逻辑推理素养培养中的应用策略研究[D]. 甘肃: 西北师范大学, 2021.
- [59] Wang Y, Rao Y, Guan H, et al. NetPad: An online DGS for mathematics education[C]. 2017 12th International Conference on Computer Science and Education (ICCSE). IEEE, 2017: 305-309.
- [60] Chen R, Rao Y, Cai R, et al. Design and Implementation of Human-Computer Interaction Based on User Experience for Dynamic Mathematics Software[C]. 2019 14th International Conference on Computer Science & Education (ICCSE). IEEE, 2019: 428-433.
- [61] 管皓, 秦小林, 饶永生, 等. 基于 Web 的动态几何软件领域模型及其应用[J]. 计算机应用, 2020, 40(4): 1127-1132.

附录 问卷调查

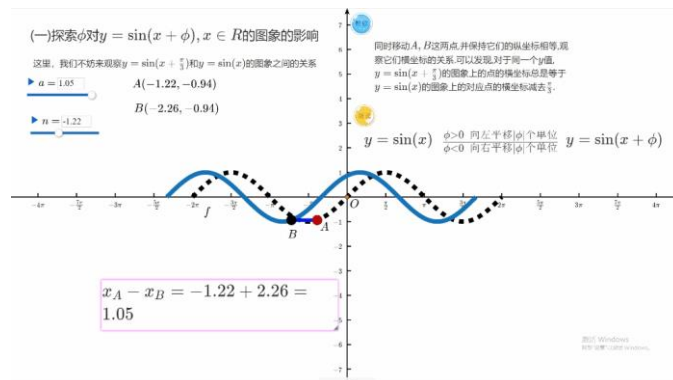
网络画板上资源的图片预览效果调查

您好，感谢您在百忙之中抽空填写该问卷，该问卷是关于网络画板上资源的图片预览效果调查，仅用于广州大学计算机科学与网络工程学院相关课题研究使用，我们承诺将对个人信息保密，请您放心填写。

本问卷选取了网络画板上的 15 个资源，请先手动点开资源链接，与资源进行手动交互，理解资源的内容并记录手动交互所需时间。

然后观察资源对应的预览图片，从三个维度：预览图片中交互顺序是否正确、交互操作是否完整、图片能反映多少资源的内容进行评分，打分区间为[1, 5]，分数越高，效果越好。

1. 资源 “ $y=\sin x$ 图像变换 ” 对应的网址为 https://www.netpad.net.cn/resource_web/course/#!/164274，手动与该资源交互所需时间为____秒。



您认为该资源对应的图片预览，预览正确性得分为：

☐非常差 ☐比较差 ☐一般 ☐比较好 ☐非常好

预览完整性得分为：

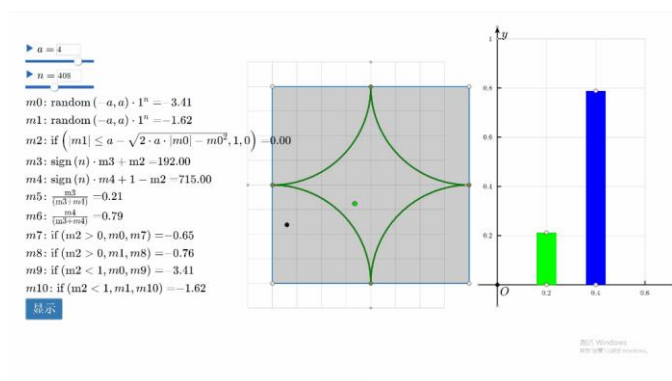
☐非常差 ☐比较差 ☐一般 ☐比较好 ☐非常好

预览能在多大程度上反映资源内容：

☐非常差 ☐比较差 ☐一般 ☐比较好 ☐非常好

2. 资源 “ 区域内落点概率 ” 对应的网址为 https://www.netpad.net.cn/resource_web/course/#!/205946，手动与该资源交互所需时间

为 秒。



您认为该资源对应的图片预览，预览正确性得分为：

☐非常差 ☐比较差 ☐一般 ☐比较好 ☐非常好

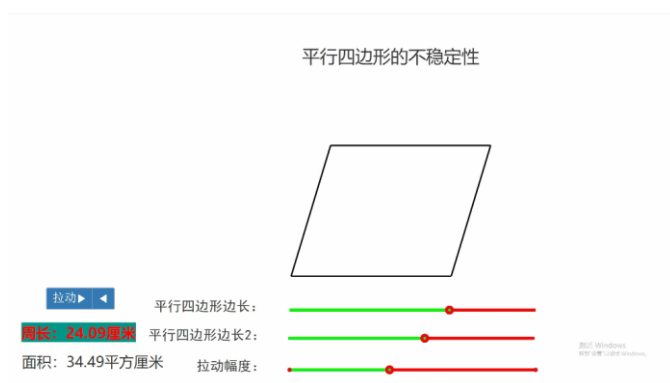
预览完整性得分为：

☐非常差 ☐比较差 ☐一般 ☐比较好 ☐非常好

预览能在多大程度上反映资源内容：

☐非常差 ☐比较差 ☐一般 ☐比较好 ☐非常好

3. 资源 “ 平行四边形的不稳定性 ” 对应的网址为 https://www.netpad.net.cn/resource_web/course/#!/179227，手动与该资源交互所需时间为 秒。



您认为该资源对应的图片预览，预览正确性得分为：

☐非常差 ☐比较差 ☐一般 ☐比较好 ☐非常好

预览完整性得分为：

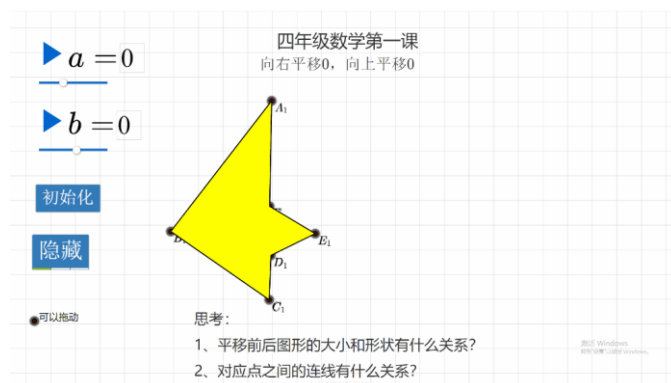
☐非常差 ☐比较差 ☐一般 ☐比较好 ☐非常好

预览能在多大程度上反映资源内容：

☐非常差 ☐比较差 ☐一般 ☐比较好 ☐非常好

...

15. 资源“平移”对应的网址为 https://www.netpad.net.cn/resource_web/course/#/181163，
手动与该资源交互所需时间为____秒。



您认为该资源对应的图片预览，预览正确性得分为：

☐非常差 ☐比较差 ☐一般 ☐比较好 ☐非常好

预览完整性得分为：

☐非常差 ☐比较差 ☐一般 ☐比较好 ☐非常好

预览能在多大程度上反映资源内容：

☐非常差 ☐比较差 ☐一般 ☐比较好 ☐非常好

备注：在文档中，只能展示 GIF 图片序列中的第一张静态图片，但在实际调查问卷中，
15 个资源对应的预览图片都为 GIF 格式。

攻读硕士学位期间取得的科研成果

科研论文：

- [1] **Jianxiong Wang**, Lanxing Xie, Xiaohong Shi, Xiangping Chen, Xiaoxia Li, Yongsheng Rao*, Preview Generation for Mathematical Interactive Educational Resources in Netpad. The 8th International Conference on Digital Home. IEEE, 2020: 221-226. (EI)
- [2] Lanxing Xie, **Jianxiong Wang**, Jie Wang, Ying Wang, Zhongwei Cui, Yongsheng Rao*, Design and Application of the Ruled Surface Function in NetPad. The 4th International Conference on Education and Multimedia Technology. ACM DIGITAL LIBRARY, 2020: 100-106. (EI)

申请专利：

- [1] 饶永生, **王剑雄**, 石晓红, 蔡瑞琦, 邹宇, 陈联. 一种动态几何资源预览的方法、系统、装置及存储介质, 中国发明专利, 申请号: 202011284002.X

致 谢

时光匆匆，转眼间我的研究生生活即将谢下帷幕。回首这三年的研究生生活，心中有太多不舍，在此，我要感谢所有关心和帮助我的人！

首先我要感谢我的导师张景中老师，老师渊博精深的专业知识、严谨细致的治学态度、活跃的学术思想以及谦逊的态度都给我留下深刻印象，受益匪浅，谨此向张老师致以我最崇高的敬意和最衷心的感谢！特别感谢饶永生老师，亦师亦友，一直帮助我，指引我前进的方向，从选题到答辩一直悉心指导我。感谢实验室邹宇老师、俞健老师、陈联老师、强小利老师以及石晓龙老师，感谢你们一直关心和帮助我们。感谢李龙飞老师、吴晓波老师和孙梦恬老师对我们的照顾，一直默默地照顾我们。

感谢团队师兄师姐和师弟师妹们，感谢石晓红师姐、管皓师兄、李静嫂子、陈如仙师姐、王婕师姐、蔡瑞琦师兄、雷斯然师姐、赵洁彤师姐给予的帮助，感谢你们带领我们学习、运动和玩耍；感谢我的室友陈婉薇、谢兰星、谭琳同学，和你们一起度过的三年时间很快乐。

衷心感谢答辩委员会的各位老师，感谢他们百忙之中审阅我的论文，为我指点迷津！