



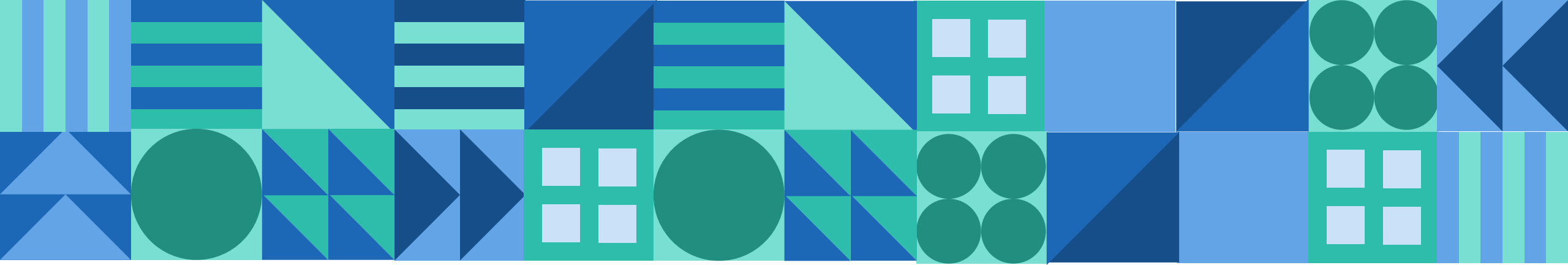
眼动追踪结题报告

《眼动追踪技术新方法及其在人机交互中的应用》

oalad

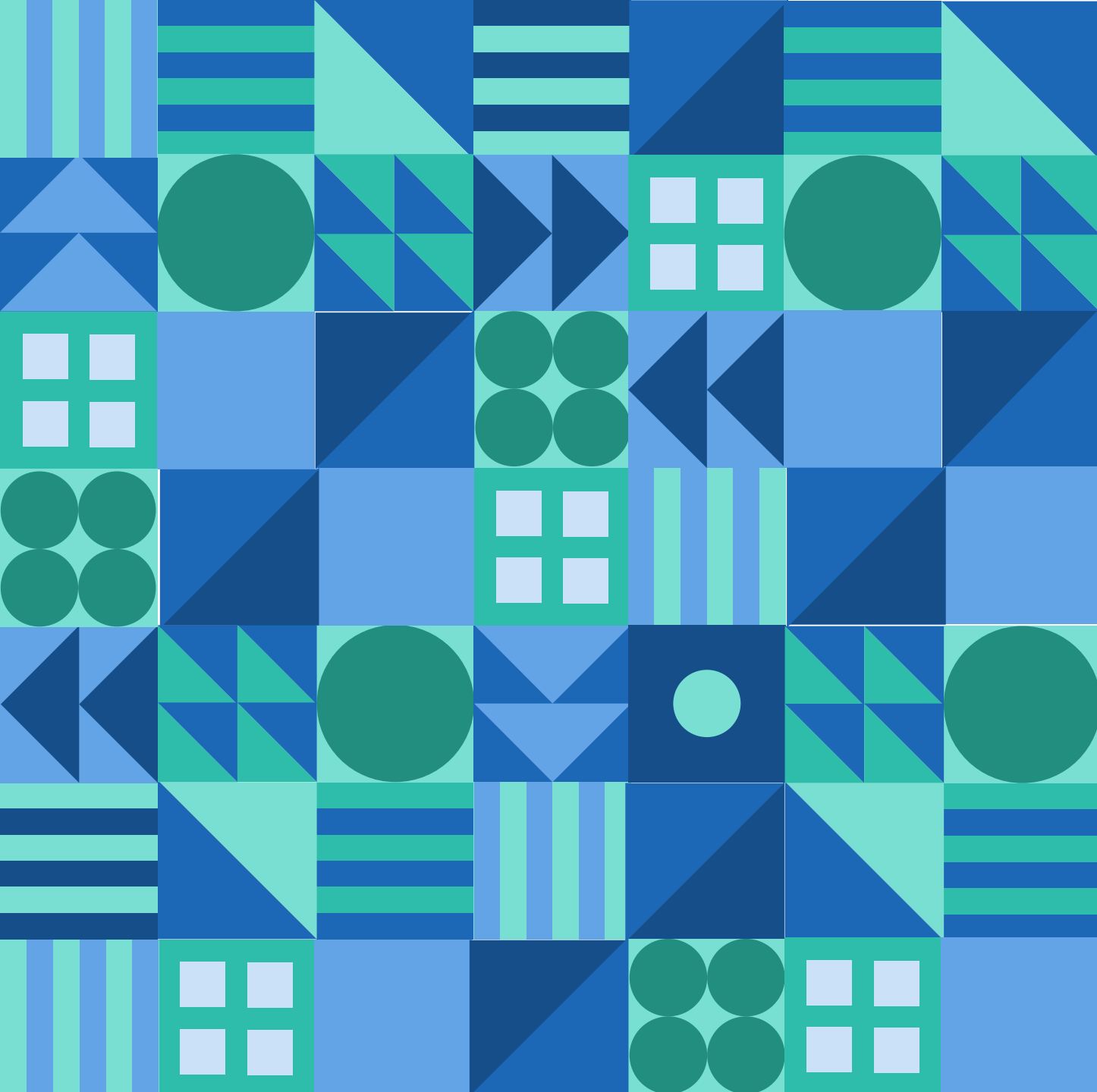
小组成员：邓龙 戴路 吴紫薇 徐煜森 于颖奇

2018年6月30日



CONTENTS

- | | |
|---|------|
| 1 | 项目回顾 |
| 2 | 总体结构 |
| 3 | 研究方法 |
| 4 | 成果展示 |
| 5 | 改进设想 |
| 6 | 参考文献 |



1

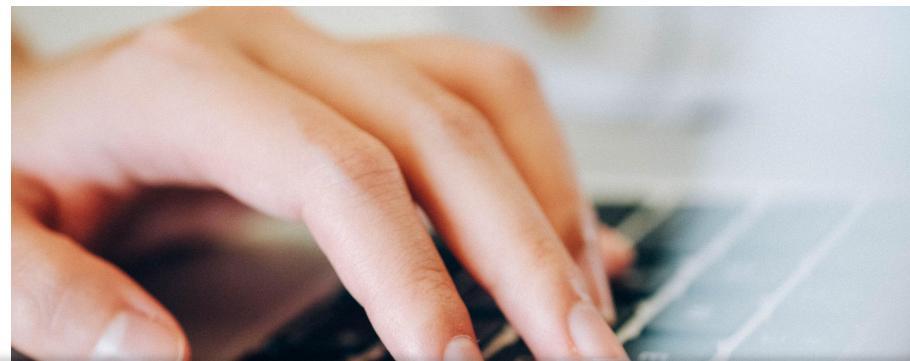
项目回顾

1 项目回顾



眼球追踪

眼球跟踪(Eye tracking)是测量注视点（人眼看的地方）或眼睛相对于头部的运动的过程。眼睛检测和跟踪研究侧重于两个领域：图像中的眼睛定位(eye localization)和注视跟踪(gaze estimation)。



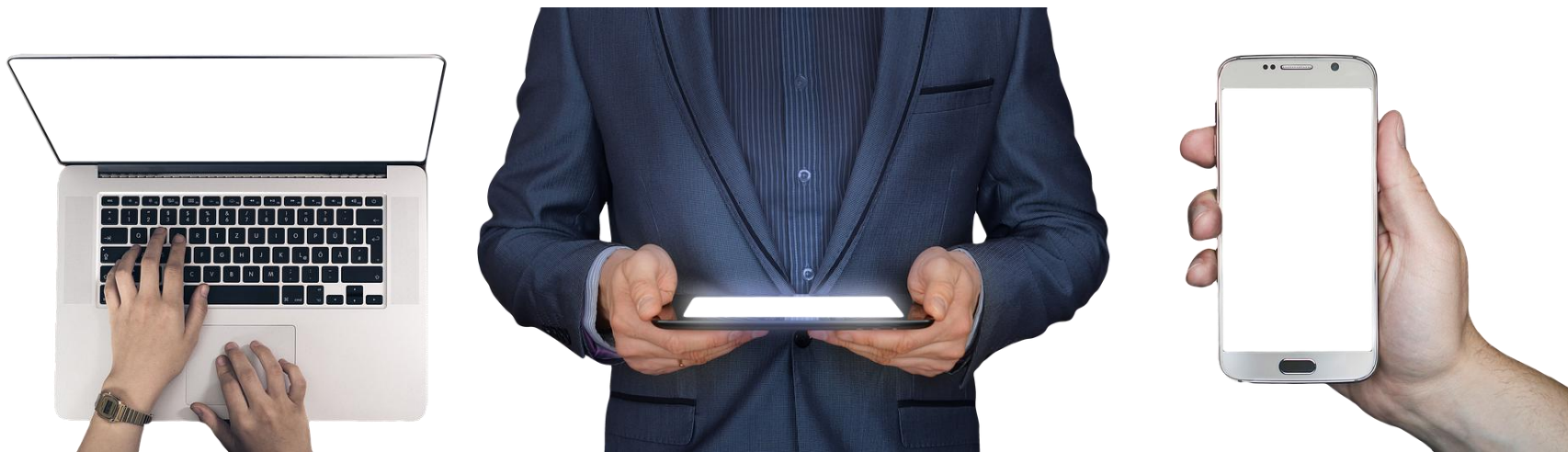
人机交互

人机互动（英语：human-computer interaction，缩写：HCI，或 human-machine interaction，缩写：HMI），是一门研究系统与用户之间的交互关系的学问。

1 项目回顾

人机交互变迁：命令行-视窗、键鼠-触摸-？

- 视觉作为人感知外界的最重要感觉，在人机交互中不可或缺。在传统的人机交互设备中，视觉仅仅作为机器信息的一种输出方式，其在信息输入方面的作用一方面不被重视，一方面也由于技术原因难以利用。
- 随着硬件技术（如GPU）和软件技术（如机器学习）的快速进步，以及包括VR、物联网等新的计算设备和概念的提出，眼球追踪作为一种新的人机交互方式，可以预见其潜力将被逐渐发掘，成为未来人机交互模式中不可缺少的一部分。



1 项目回顾

项目目标

- 本项目旨在建立一个低成本的新方式，利用简单的头戴设备或者桌面设备，对用户的视线信息做出分析，提取出有效信息作为机器输入，从而作为一种新的机器控制方法。
- 我们希望克服目前眼动追踪技术中设备成本高、便携性差的问题。另外，我们希望方法尽可能的简单，而不是消耗巨大的计算资源。
- 我们不会将重点放在精确性和高帧率上，原因是目前眼动追踪技术已经足够精确，采样频率也足够高。



低成本、设备简单



低计算资源需求



低精度、帧率要求

1 项目回顾

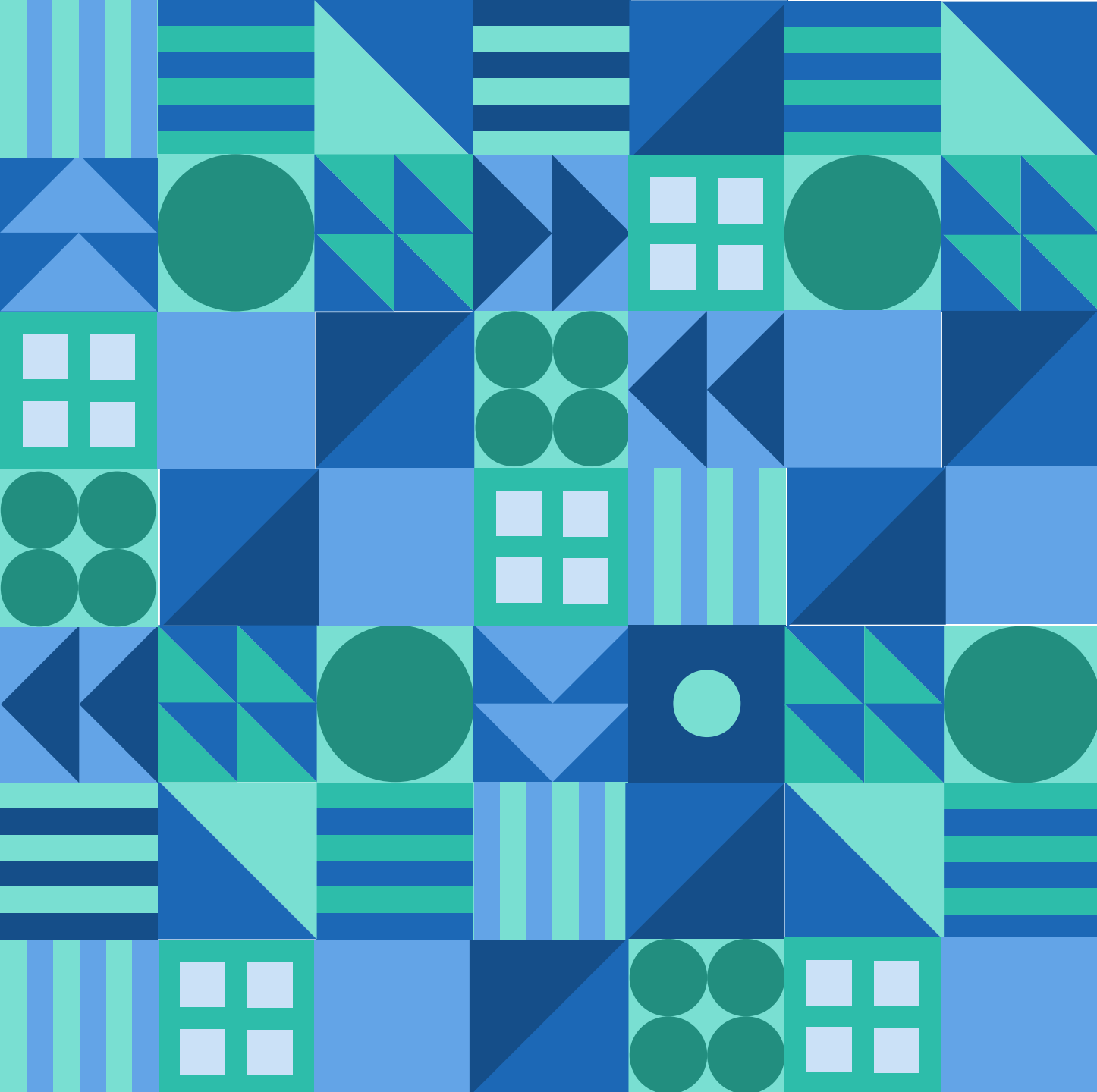


应用前景

开发人机交互，细致的用户界面

市场研究与消费者调研方面的应用

运动研究、心理学与神经科学研究、婴幼儿研究、体育运动研究、教育研究、临床研究



2

总体结构



2 总体结构

数据收集

快速数据采集模块

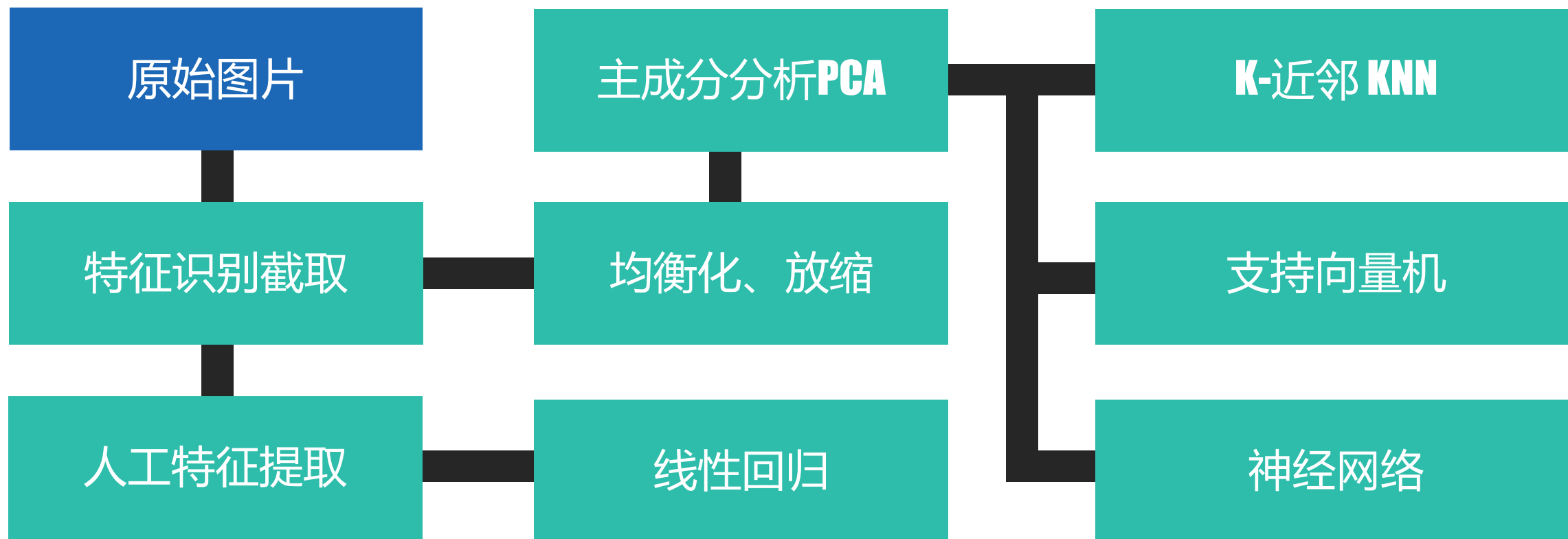
数据处理和预测

数据处理、学习训练、结果计算模块

结果可视化

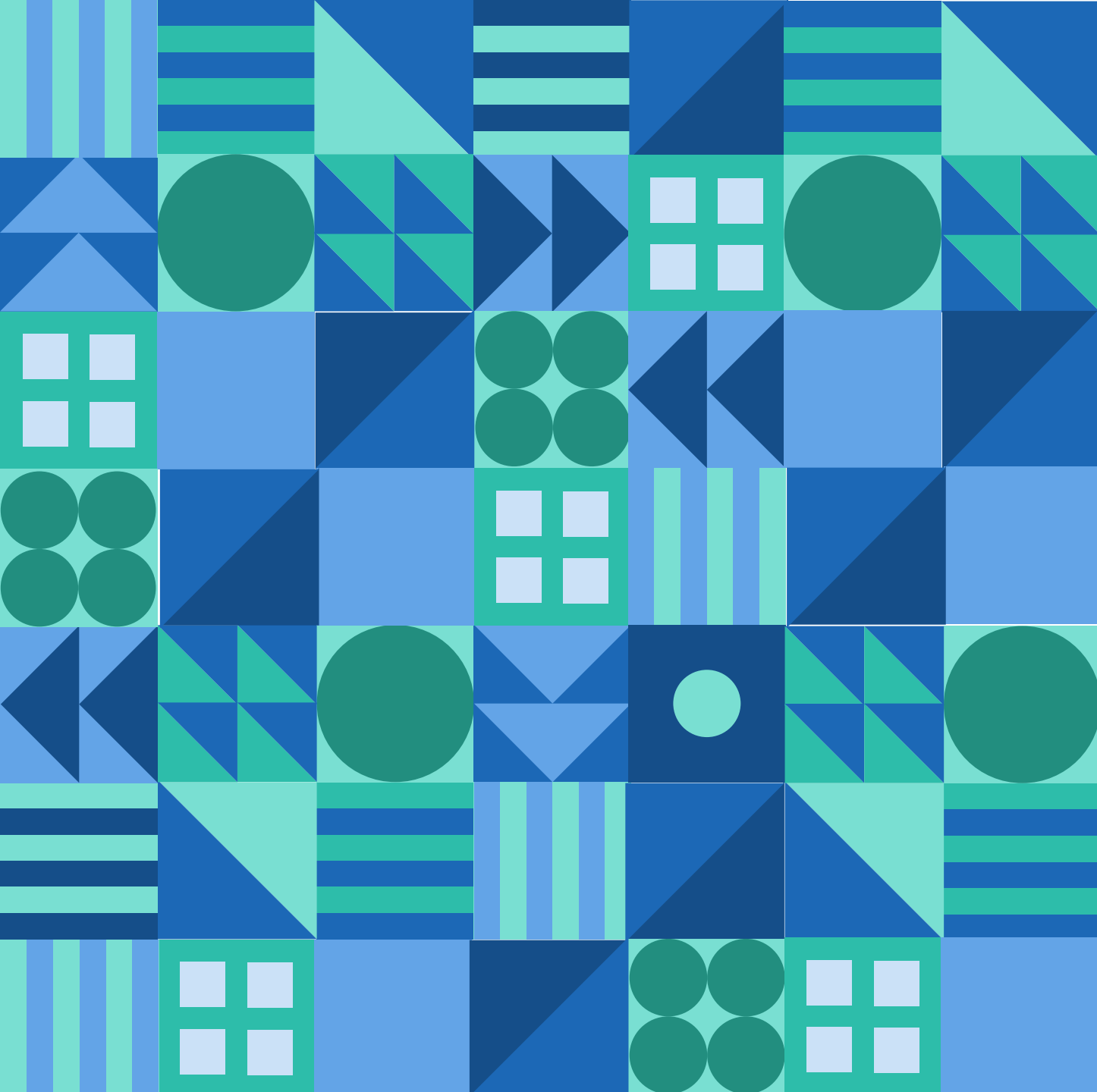
交互界面

2 总体结构



时间安排

[illegible]



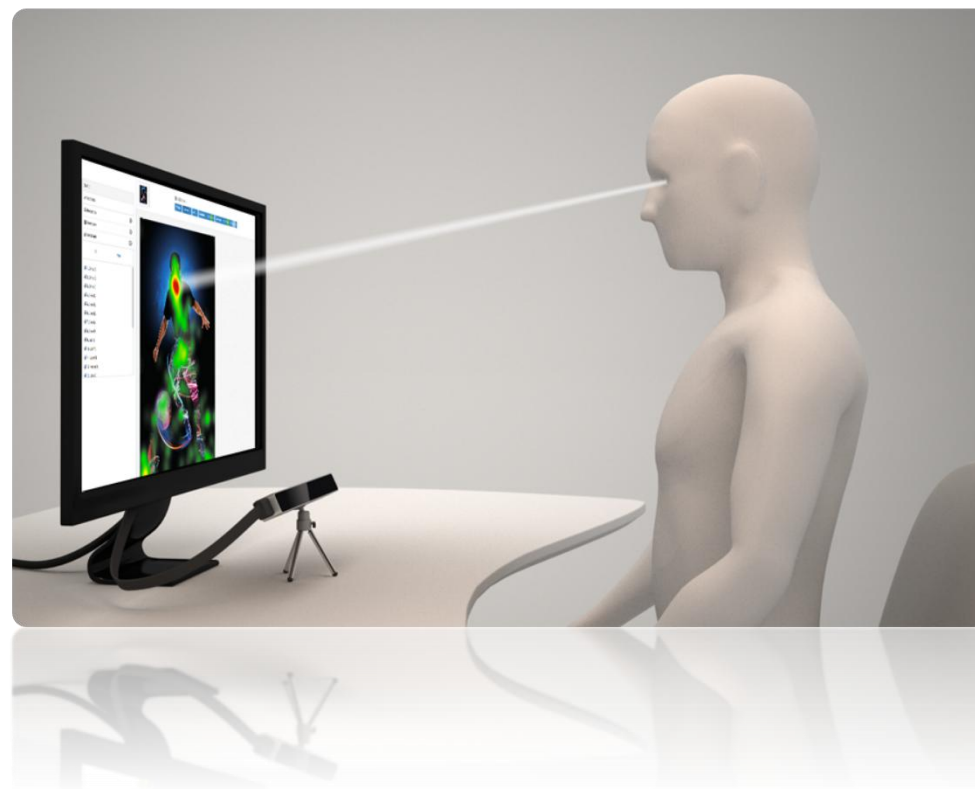
3

研究方法

3 研究方法--数据收集

如何进行数据收集？

- 通过定位标记固定头部和设备的相对位置
- 在屏幕上随机位置显示目标点
- 测试者确认看向目标点
- 记录数据，切换到下一个随机点



3 研究方法——确定瞳孔

人工特征提取 + 线性回归

- 依据
- 在采集数据中发现，发现x坐标与瞳孔相对位置高度相关，而y坐标和眼睛高度高度相关。故提取这两个特征。
- 主要步骤
- 利用OpenCV及Dlib库，将眼睛图片截出。
- Dlib库的特征点识别可以计算出眼睛大小。
- 使用梯度方法寻找瞳孔位置。[1]
- 在参考了一些论文[2]之后，
- 我们还尝试了RR和SVR等回归方法。

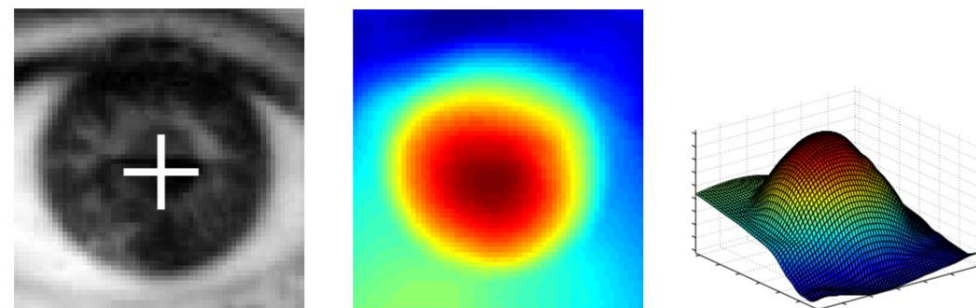


Figure 3: Evaluation of (1) for an exemplary pupil with the detected centre marked in white (left). The objective function achieves a strong maximum at the centre of the pupil; 2-dimensional plot (centre) and 3-dimensional plot (right).

3 研究方法——特征提取

主成分分析 **Principal Component Analysis**

- 主成分分析是对于原先提出的所有变量，将重复的变量（关系紧密的变量）删去多余，建立尽可能少的新变量，使得这些新变量是两两不相关的，而且这些新变量在反映问题的信息方面尽可能保持原有的信息。
- 图片是典型的高维信息，在从眼睛图片转换到坐标的过程中，大部分内容是无用的，我们希望用主成分分析的方法提取出图片中真正的信息。

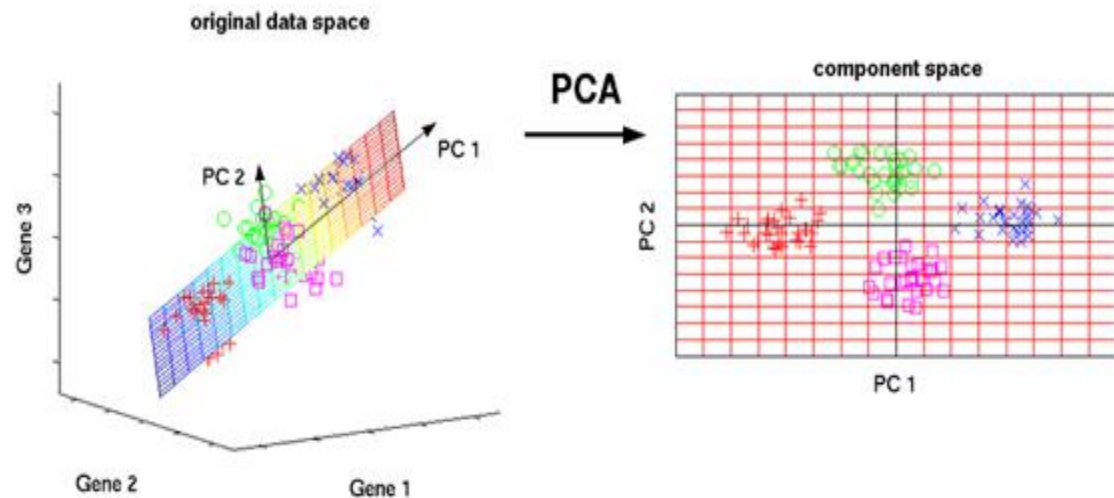
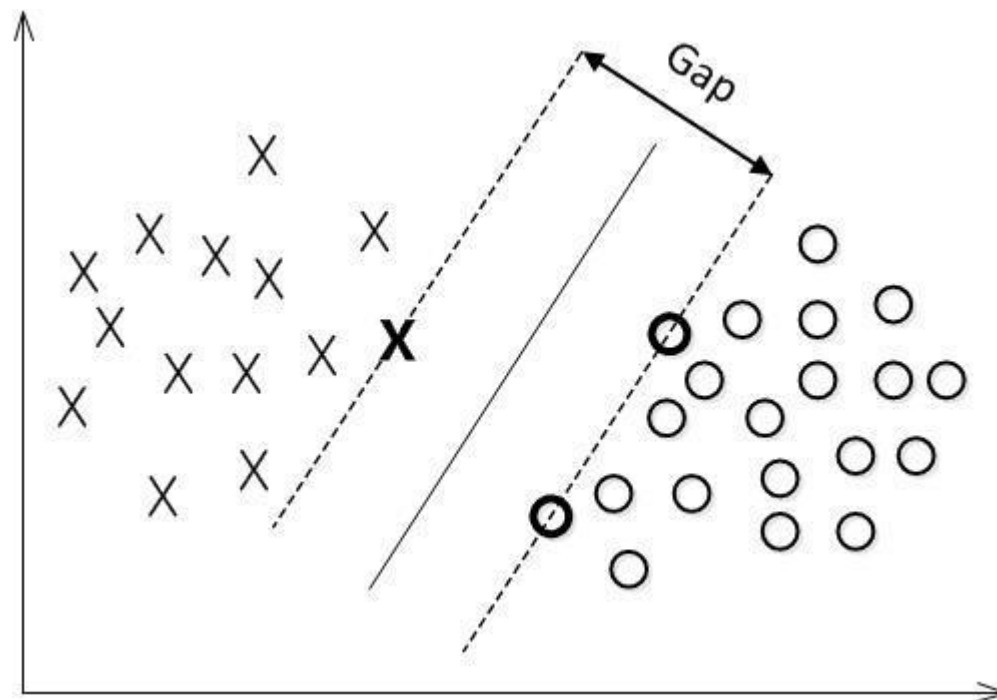


图 1

3 研究方法——视线预测

支持向量机 **Support Vector Machine**

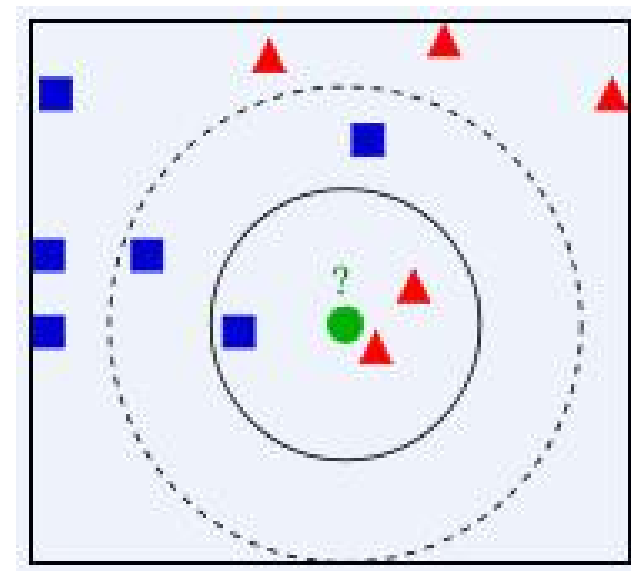
- 支持向量机是一种监督式学习的方法，可广泛地应用于统计分类以及回归分析。支持向量机属于一般化线性分类器，这族分类器的特点是他们能够同时最小化经验误差与最大化几何边缘区，因此支持向量机也被称为最大边缘区分分类器。
- 难以对大样本进行训练，且解决多分类问题存在困难，不具备可拓展性。



3 研究方法——视线预测

K-近邻算法 k-Nearest Neighbor

- 给定一个训练数据集，对新的输入实例，在训练数据集中找到与该实例最邻近的K个实例（K个邻居），这K个实例的多数属于某个类，就把该输入实例分类到这个类中。
- 屏幕是一个小范围的空间，在精度要求不高时，无需太多的样本，样本点的密度已经足够大。若样本分布均匀，选取适当的k已经能足够反映新输入实例的位置。
- 此处我们改进了k近邻的思想，我们并非用附近的实例进行分类，而是利用它们的坐标加权平均来计算新实例的坐标。

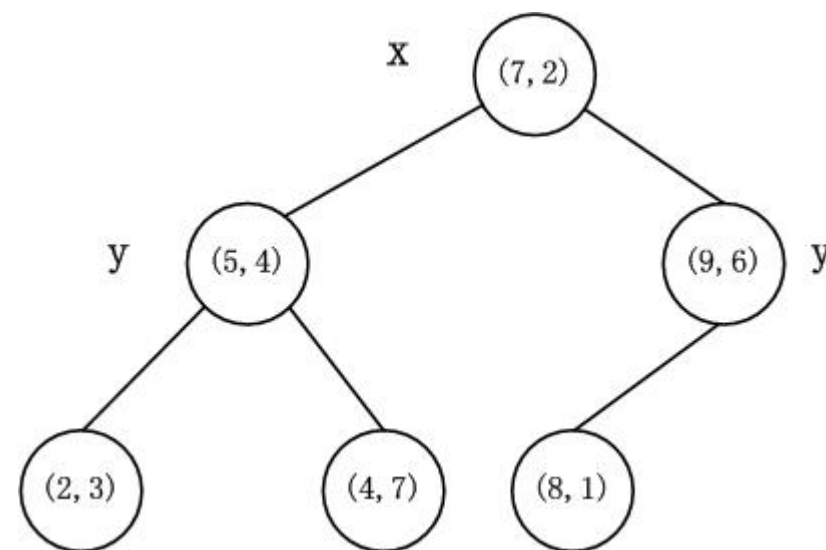
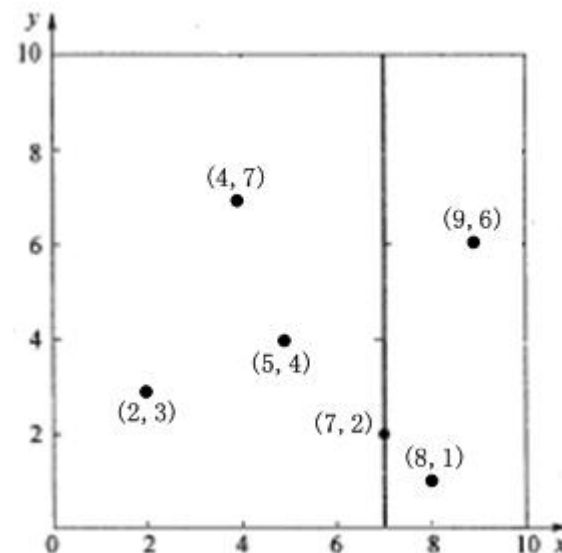


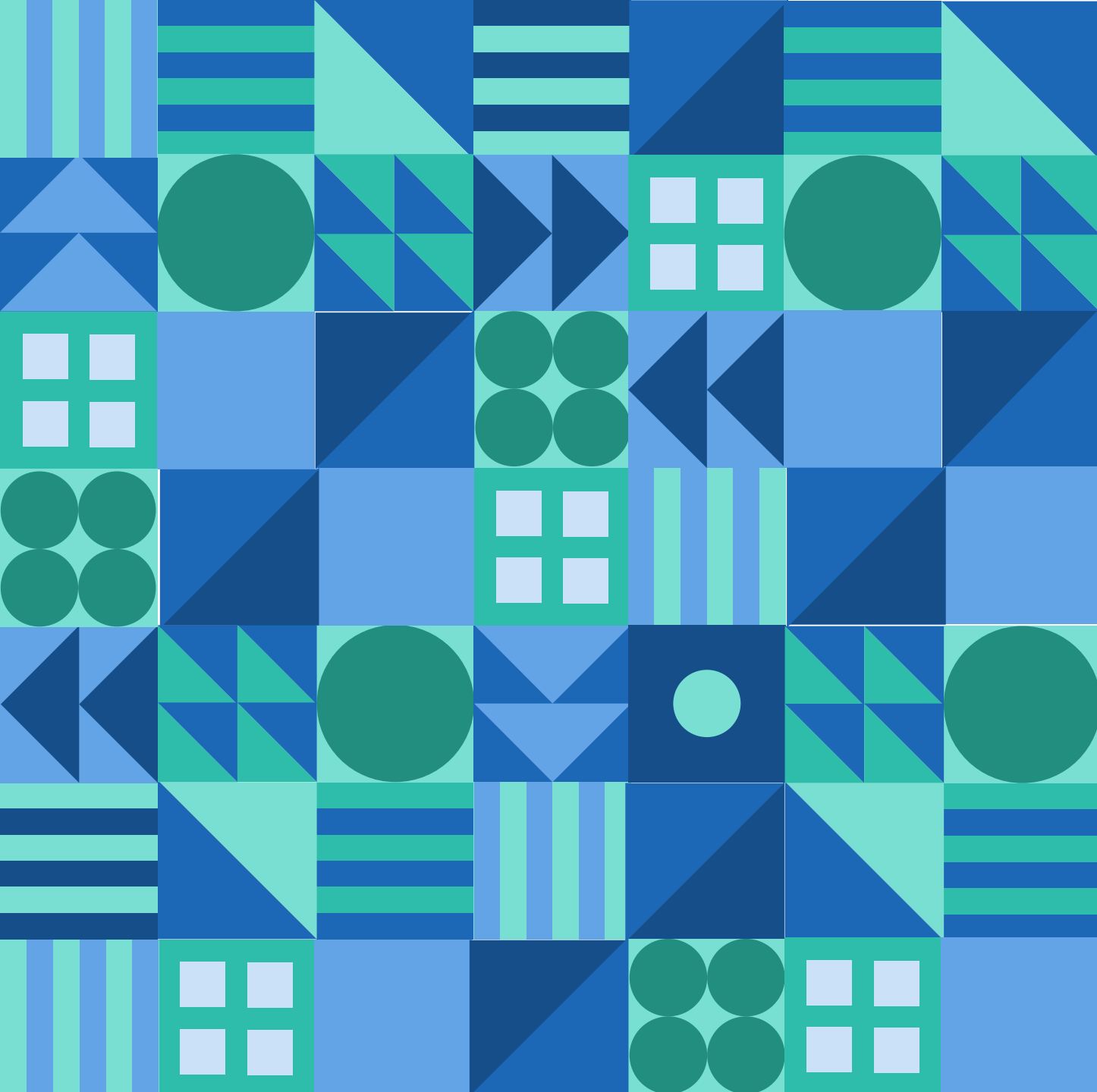
3 研究方法——视线预测

KD Tree

- 利用KD Tree, K近邻算法能够有极高的效率。
- Kd树是一种对k维空间中的实例点进行存储以便对其进行快速检索的树形数据结构。Kd树是二叉树, 表示对k维空间的一个划分。
- 构造kd树相当于不断地用垂直于坐标轴的超平面将k维空间切分, 构成一系列的k维超矩形区域。Kd树的每个结点对应于一个k维超矩形区域

- **主要步骤**
- 构造平衡kd树
- 用kd树完成最近邻搜索
- 最终将二维数据映射到二维坐标





4

成果展示

4 成果展示

数据收集

- 在所有组员的不懈努力下，我们收集了数百张的原始数据。



0.bmp



1.bmp



2.bmp



3.bmp



4.bmp



5.bmp



6.bmp



9.bmp



10.bmp



11.bmp



12.bmp



13.bmp



14.bmp



15.bmp



18.bmp



19.bmp



20.bmp



21.bmp



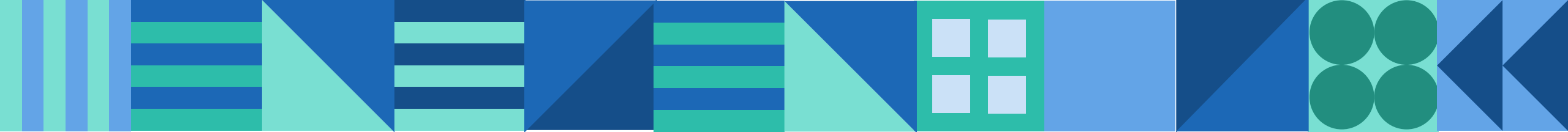
22.bmp



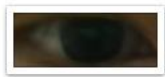
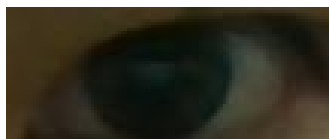
23.bmp



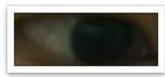
24.bmp



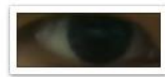
眼部提取



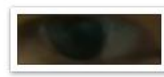
left0.bmp
↑ 14 KB



left1.bmp
↑ 14 KB



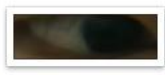
left2.bmp
↑ 15 KB



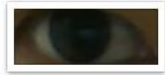
left3.bmp
↑ 14 KB



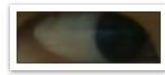
left4.bmp
↑ 12 KB



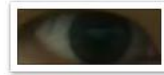
left5.bmp
↑ 12 KB



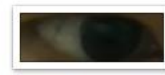
left6.bmp
↑ 14 KB



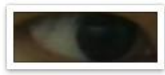
left7.bmp
↑ 15 KB



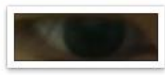
left8.bmp
↑ 15 KB



left9.bmp
↑ 13 KB



left10.bmp
↑ 14 KB



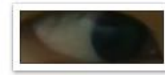
left11.bmp
↑ 14 KB



left12.bmp
↑ 12 KB

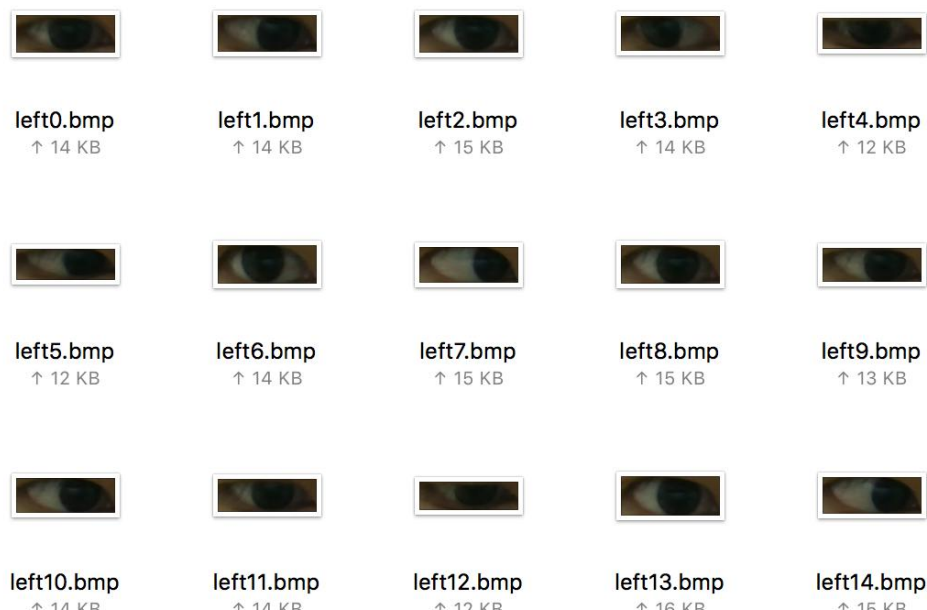


left13.bmp
↑ 16 KB



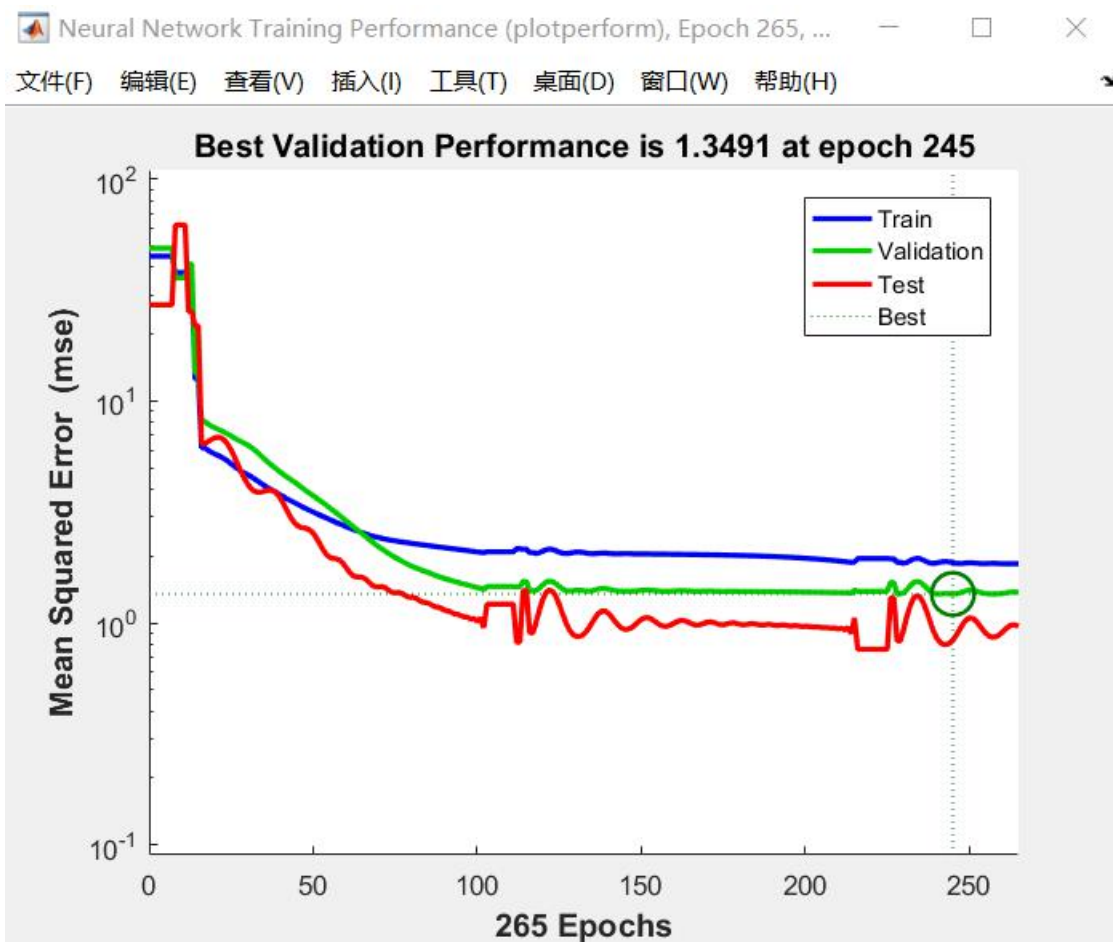
left14.bmp
↑ 15 KB

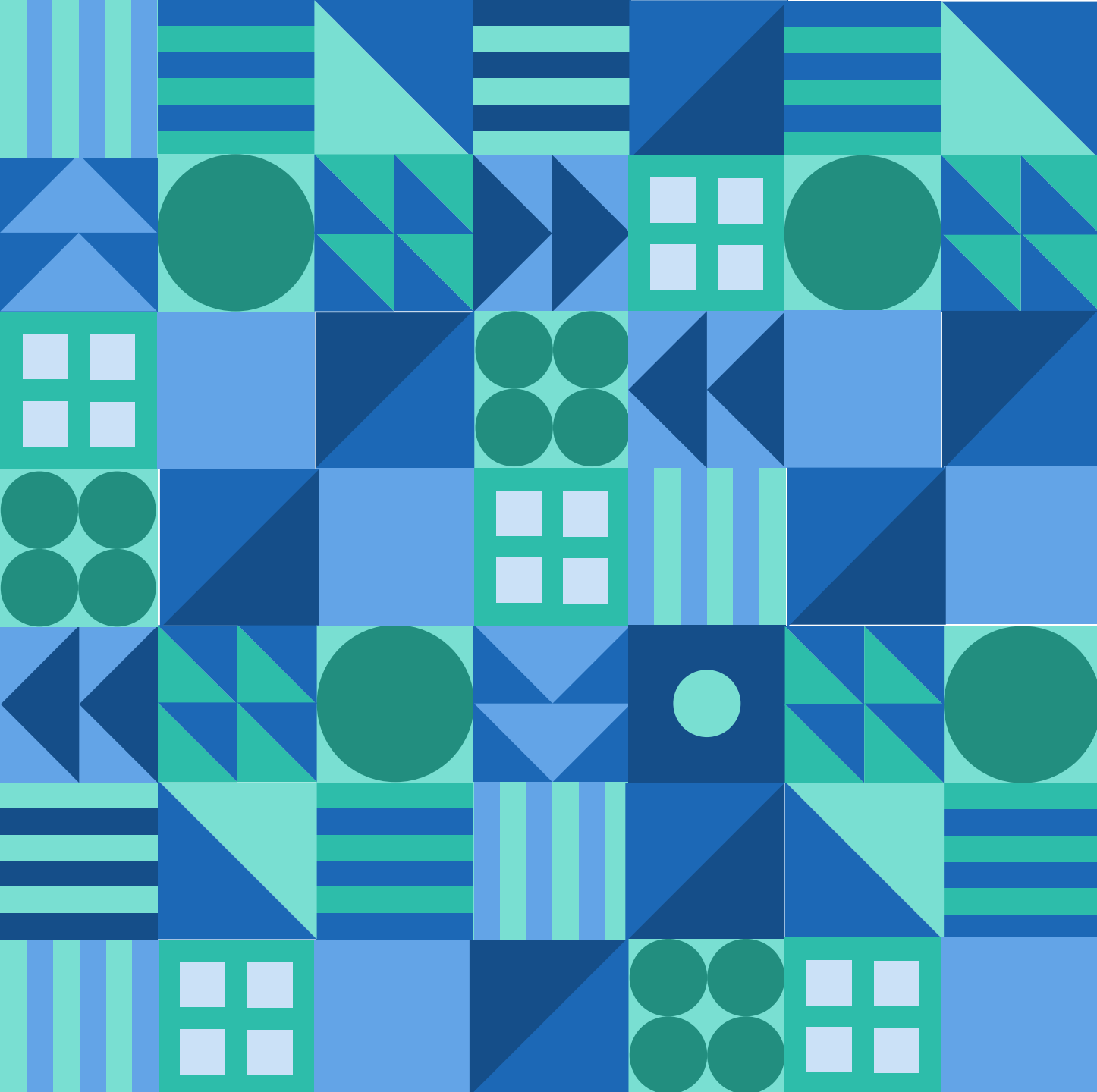
瞳孔位置的获取



eyeposition	--	For
right9_y.txt	1 byte	Pla
right9_x.txt	2 bytes	Pla
right8_y.txt	2 bytes	Pla
right8_x.txt	2 bytes	Pla
right7_y.txt	1 byte	Pla
right7_x.txt	2 bytes	Pla
right77_y.txt	1 byte	Pla
right77_x.txt	2 bytes	Pla
right76_y.txt	1 byte	Pla
right76_x.txt	2 bytes	Pla
right75_y.txt	1 byte	Pla
right75_x.txt	2 bytes	Pla
right74_y.txt	1 byte	Pla
right74_x.txt	2 bytes	Pla
right73_y.txt	1 byte	Pla
riaht73 x.txt	2 bvtes	Pla

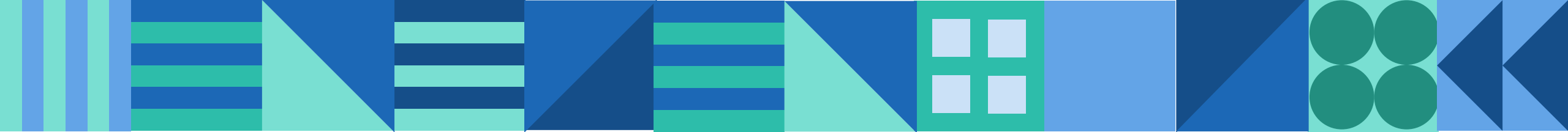
前馈神经网络





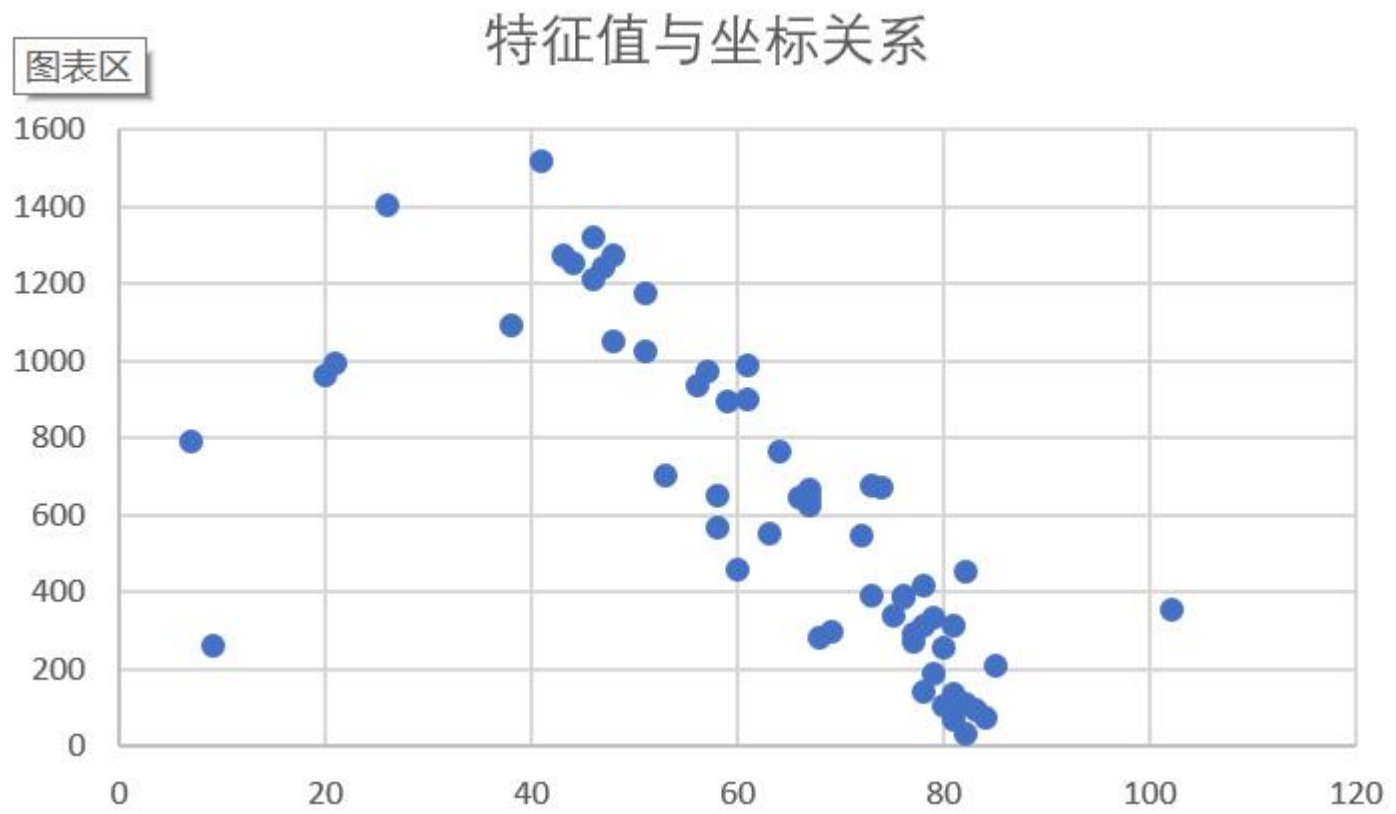
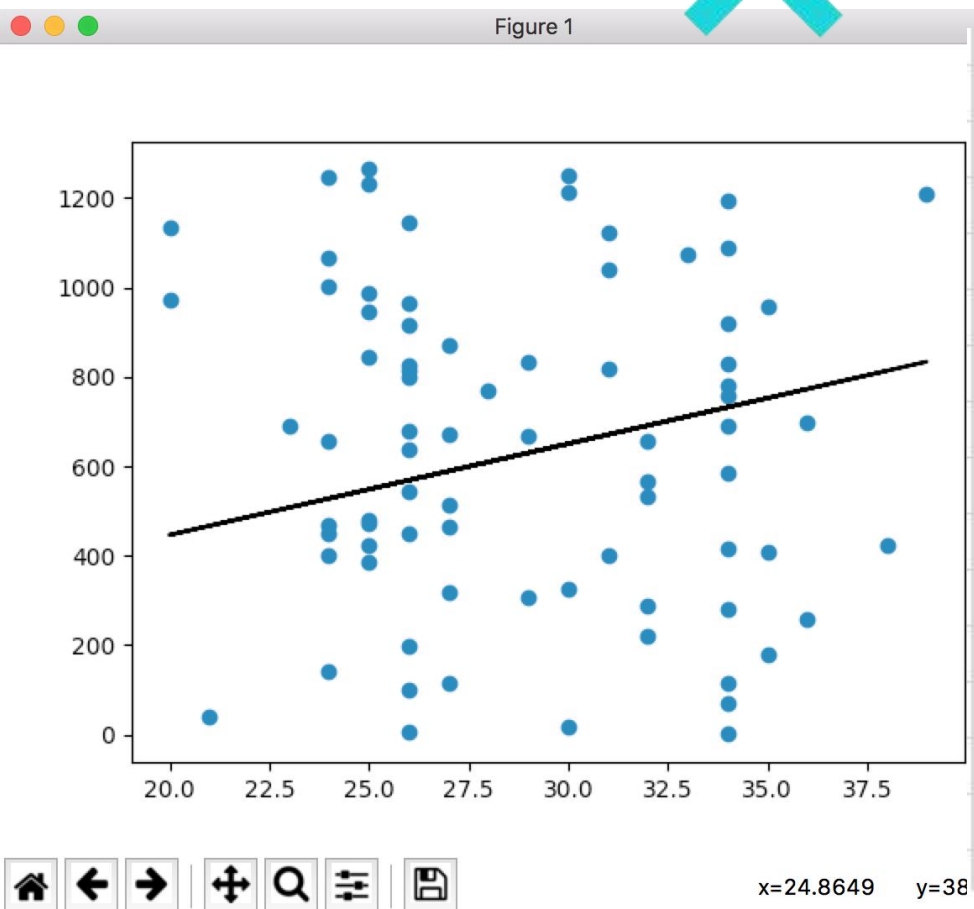
5

改进设想

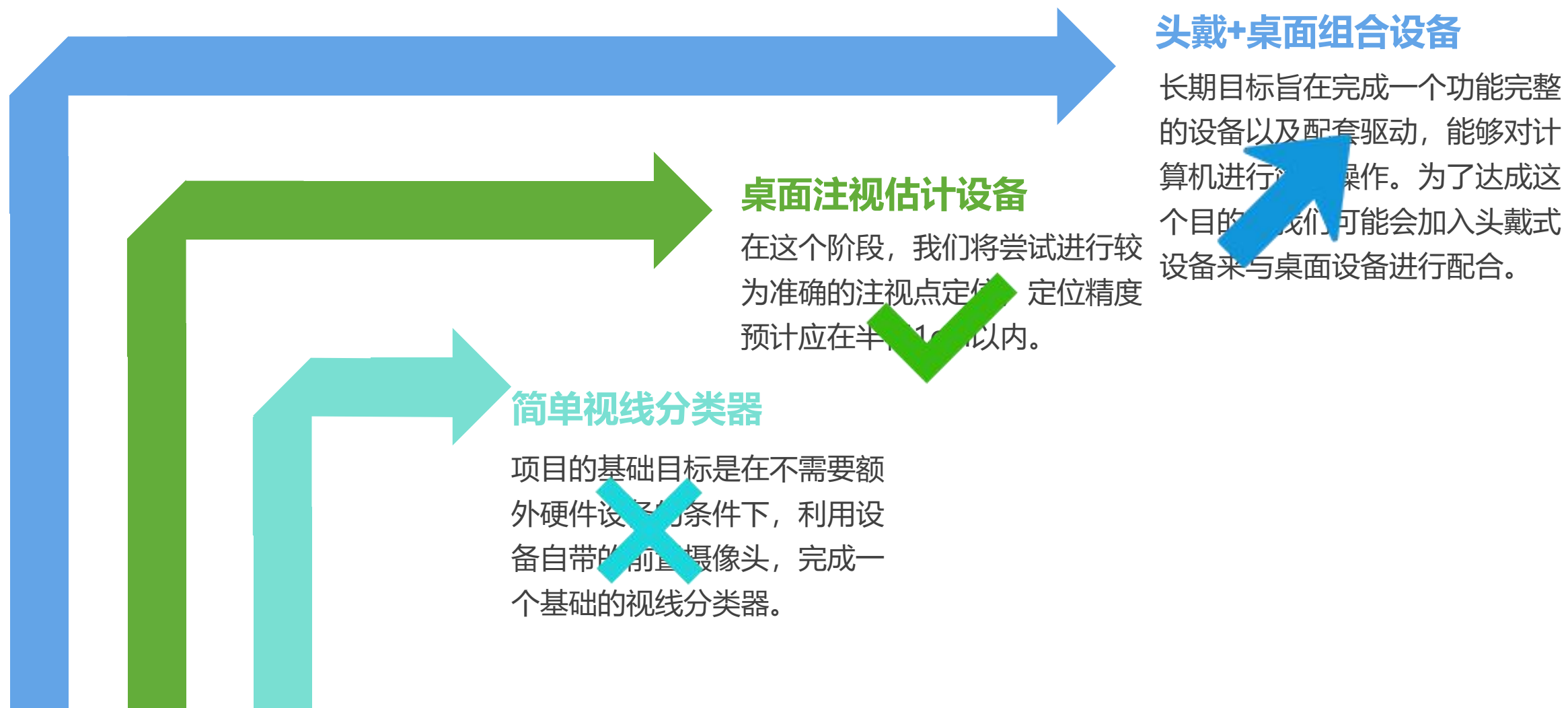


瓶颈：数据集不够可靠！

数据集可靠！



5 改进设想





5 改进设想

同步性仍然可以改进

我们在同步性方面已经做得很好了，但是若要用于游戏等方面，其延迟仍然可以减少。

计算头部运动的影响

我们目前只能处理头部固定的情况，其实单个摄像头也可以处理很多头部不固定的运动，如距离、偏角等。

更方便各种程序的整合

我们使用了matlab, c++, c#, python四种语言进行不同功能的实现，在整合时遇到了一些困难。

人眼的标准化

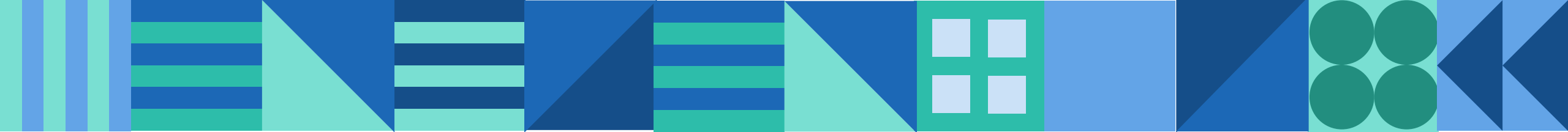
提高识别瞳孔的精度，有更广泛的适用面

速度的提高

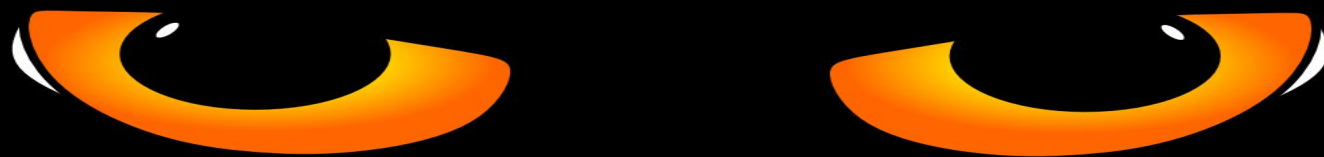
精度的提高

统一语言的使用

人眼的统一形式化

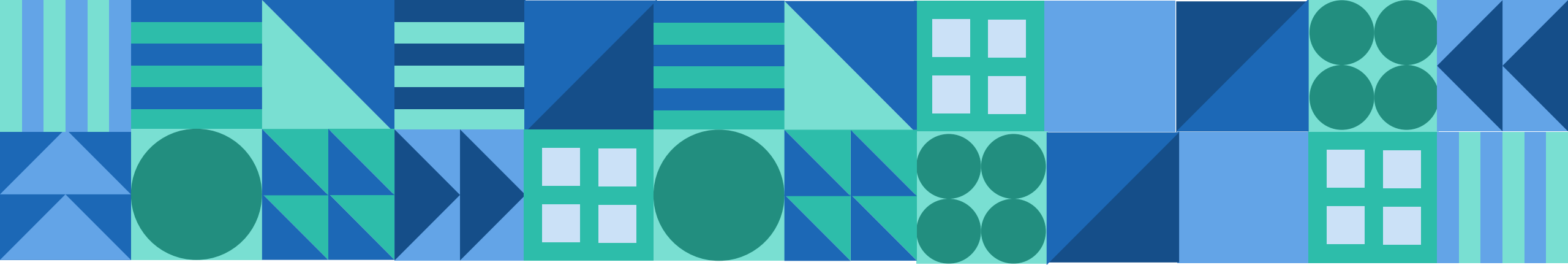


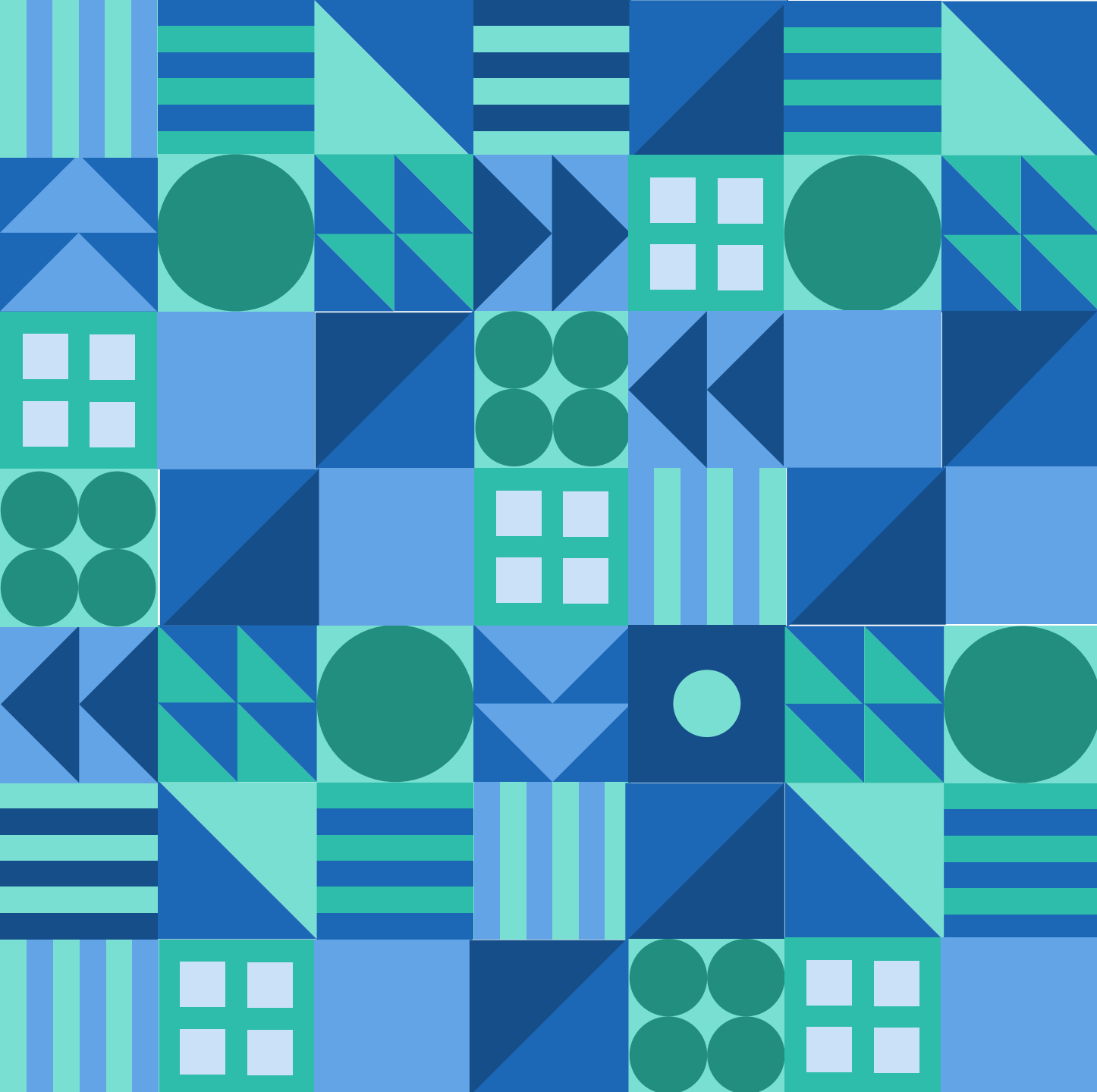
Q & A



感谢聆听!

小组成员：邓龙 戴路 吴紫薇 徐煜森 于颖奇





06

参考文献



06 参考文献

- [1] Timm F, Barth E. Accurate Eye Centre Localisation by Means of Gradients[J]. Visapp, 2011, 11: 125-130.
- [2] Lu F, Sugano Y, Okabe T, et al. Inferring human gaze from appearance via adaptive linear regression[C]//Computer Vision (ICCV), 2011 IEEE International Conference on. IEEE, 2011: 153-160.
- [3] 作者名称. 参考文献名称. 该文献出版社. 出版日期.
- [4] 作者名称. 参考文献名称. 该文献出版社. 出版日期.
- [5] 作者名称. 参考文献名称. 该文献出版社. 出版日期.
- [6] 作者名称. 参考文献名称. 该文献出版社. 出版日期.
- [7] 作者名称. 参考文献名称. 该文献出版社. 出版日期.
- [8] 作者名称. 参考文献名称. 该文献出版社. 出版日期.
- [9] 作者名称. 参考文献名称. 该文献出版社. 出版日期.
- [11] 作者名称. 参考文献名称. 该文献出版社. 出版日期.
- [12] 作者名称. 参考文献名称. 该文献出版社. 出版日期.
- [13] 作者名称. 参考文献名称. 该文献出版社. 出版日期.
- [14] 作者名称. 参考文献名称. 该文献出版社. 出版日期.
- [15] 作者名称. 参考文献名称. 该文献出版社. 出版日期.
- [16] 作者名称. 参考文献名称. 该文献出版社. 出版日期.
- [17] 作者名称. 参考文献名称. 该文献出版社. 出版日期.
- [18] 作者名称. 参考文献名称. 该文献出版社. 出版日期.
- [19] 作者名称. 参考文献名称. 该文献出版社. 出版日期.
- [20] 作者名称. 参考文献名称. 该文献出版社. 出版日期.

标注

字体使用

英文 Impact

中文 微软雅黑

行距

正文 1.3

背景图片出处

cn.bing.com

声明

本网站所提供的任何信息内容（包括但不限于 PPT 模板、Word 文档、Excel 图表、图片素材等）均受《中华人民共和国著作权法》、《信息网络传播权保护条例》及其他适用的法律法规的保护，未经权利人书面明确授权，信息内容的任何部分(包括图片或图表)不得被全部或部分的复制、传播、销售，否则将承担法律责任。