

A Subjective Study of Multi-Dimensional Aesthetic Assessment for Mobile Game Image

Suiyi Ling, Junle Wang, Patrick Le Callet



UNIVERSITÉ DE NANTES





Join work with Tencent


- Tencent: More than WeChat and QQ
 - Currently the largest video game company in the world
 - Own Riot Games, big part of Supercell and Epic games, small part of Bluehole, Activision Blizzard, Ubisoft and so on.
 - Focus on mobile game



Background

- Image Aesthetic Assessment
 - Application:
 - Selection of photos/cover images
 - Image enhancement
 - Challenge
 - Higher level and more subjective factors
 - Usually no reference
 - Progress
 - Either by specific rules, or by neural network
 - Natural content / photographic (AADB, AVA, ...)
 - Little study is about CG, particular video game image
- Our motivation
 - Game evaluation
 - Game design
 - Abnormal detection
 - Game image quality assessment and cloud gaming

What we do to evaluate a game?

- Visual analysis (subjective scoring, 5 scales, similar ACR)
 - **Overall aesthetic score**
 - **Color(harmony, colorfulness)** 
 - **Fineness**
 - Distant view
 - Interaction effect
 - Lighting/shadow effect
 - Model size consistency
 - Action/movement
 - Image style
 - Visual Effect
 - ...
 - Audio (subjective scoring, 5 scales)
 - Immersive(questionnaire)
 - Control (objective + subjective scoring)
 - Mobile performance (objective, i.e. CPU, FPS, Memory, Temperature)
- less subjective
less prior knowledge
less top-down related factors

Four aesthetic dimensions:

- The fineness (details)
- The colorfulness
- The color harmony
- The overall aesthetic quality:

Fineness

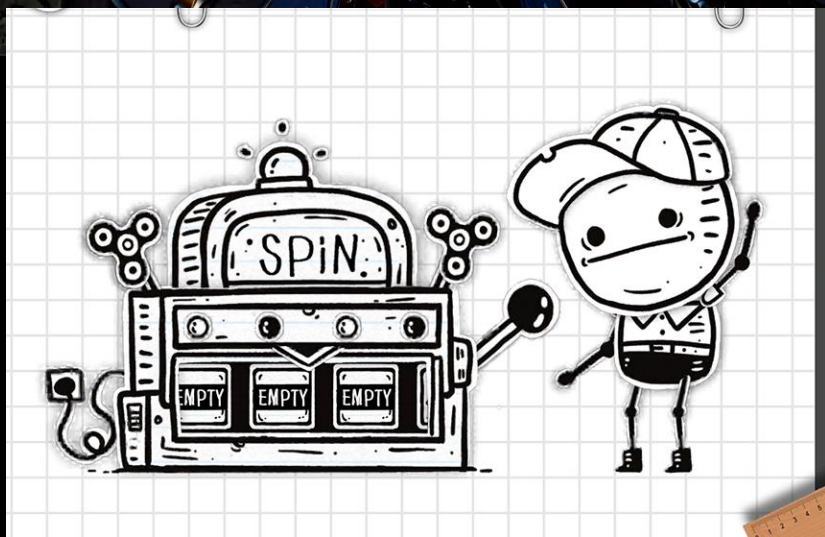
Do you think the details are enough and clearly presented?
Score 1

Score 5



Colorfulness

Score 1



Score 5



• Color harmony •

Do you feel comfortable/pleasant about the collocation of color?

Score 1

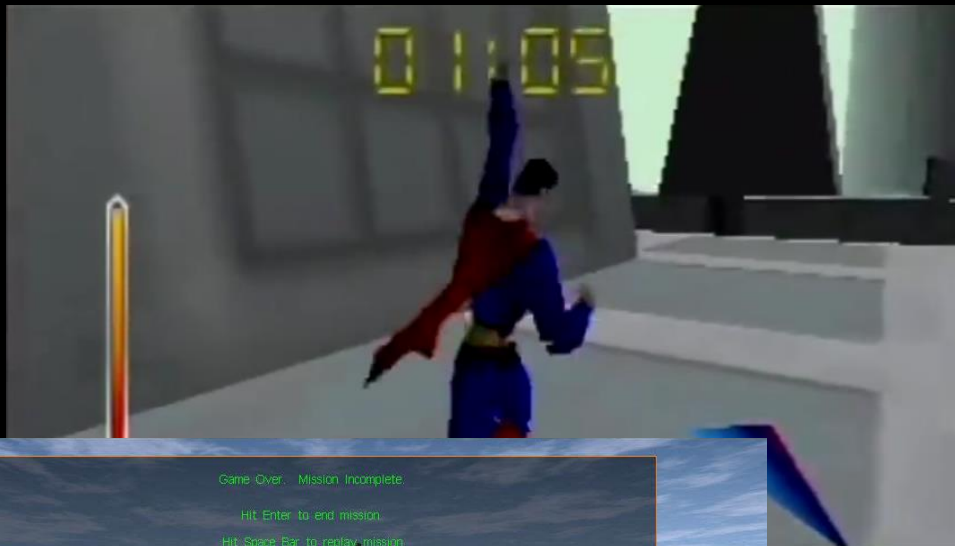


Score 5

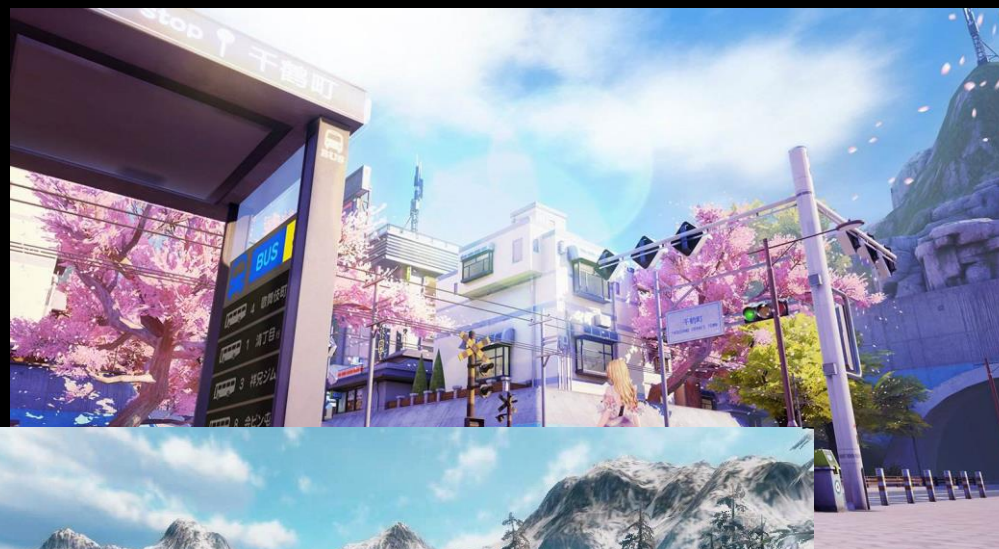


Overall score

Score 1

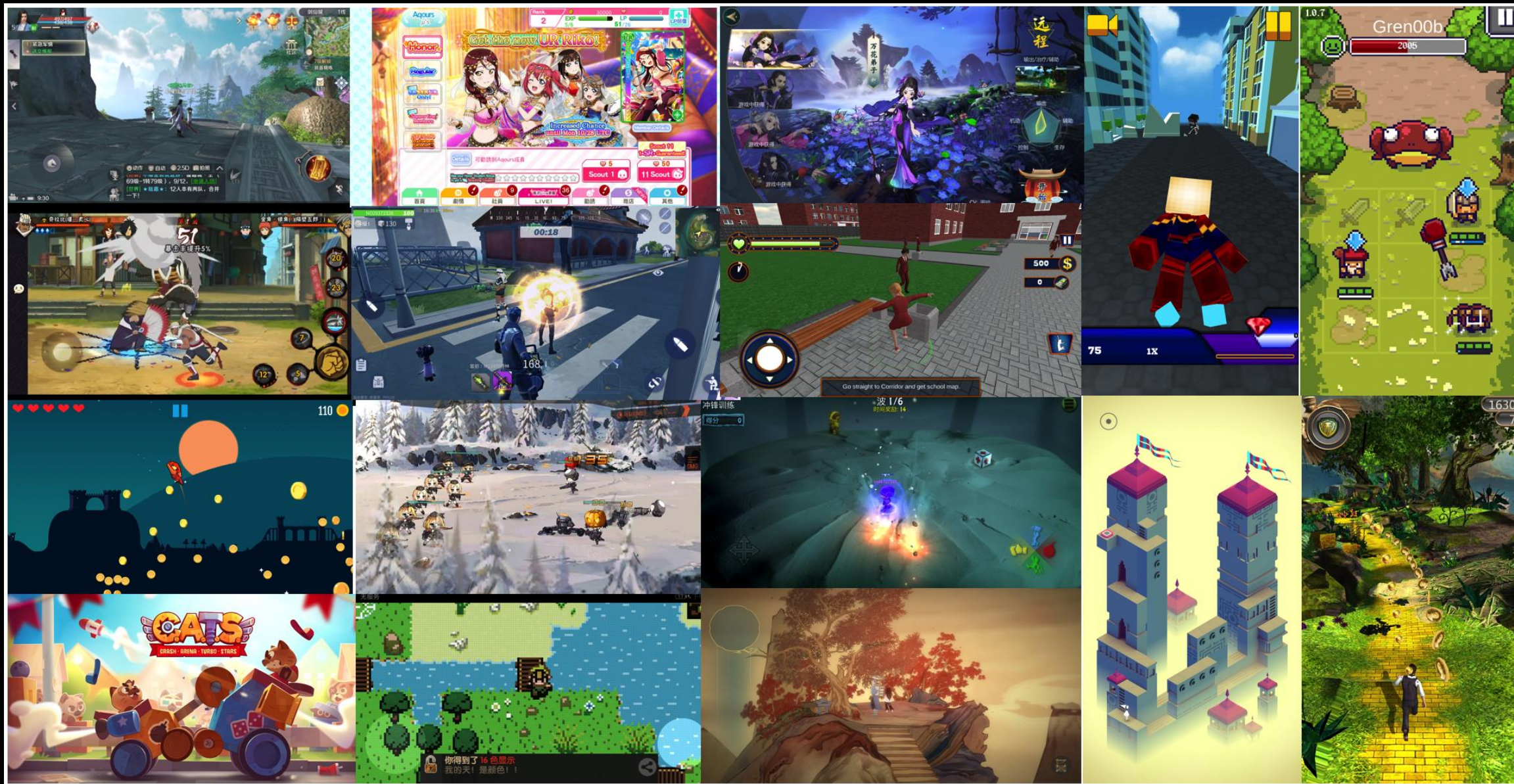


Score 5



Experiment

- Stimuli
 - 100 games collected from Play Store
 - Some game types are filtered out:
 - Casino
 - Teaching
 - AR/MR
 - Q&A challenge
 - 1091 screenshot images in total
 - Some images are filtered out:
 - UI for setup,
 - Advertising image included,
 - Natural content,
 - OS image related



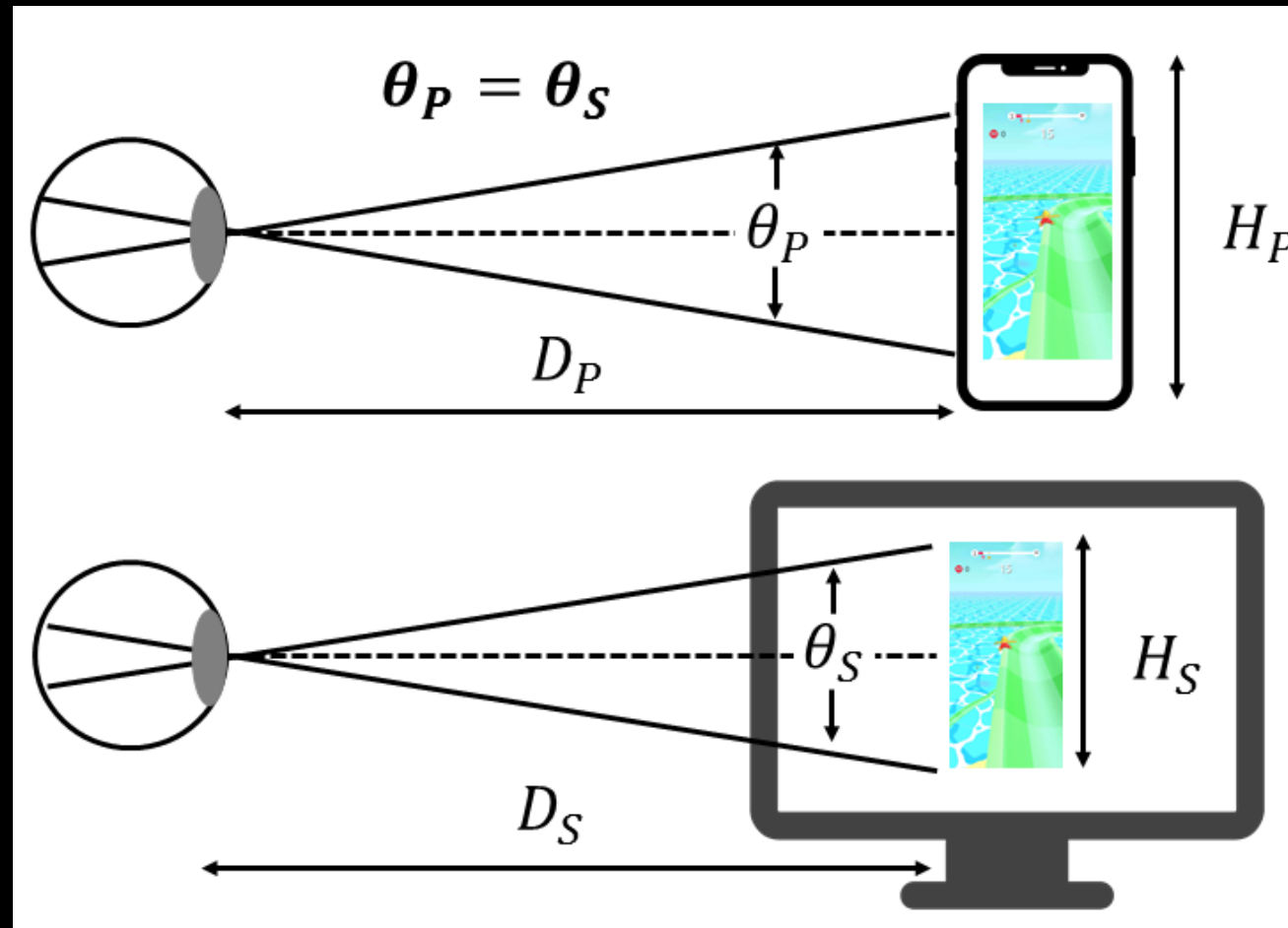
Experiment

- Task
 - Evaluate each image based on four dimensions
 - Fineness
 - Colorfulness
 - Color harmony
 - Overall score
 - Observers are asked to give a 1 to 5 score to each dimension

Experiment

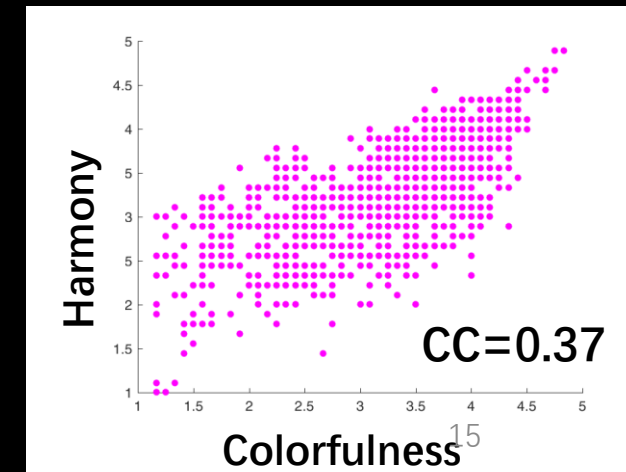
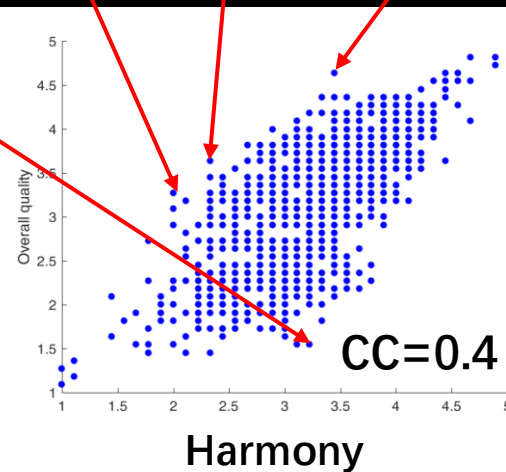
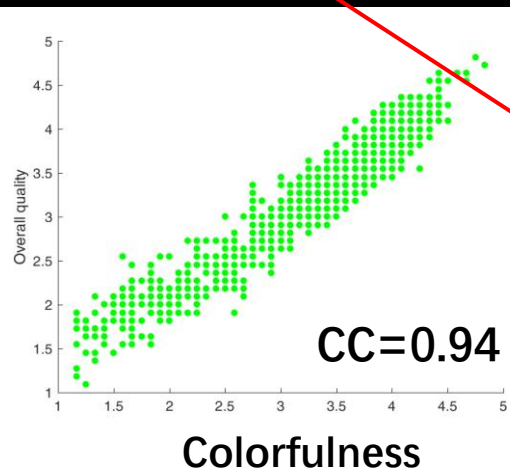
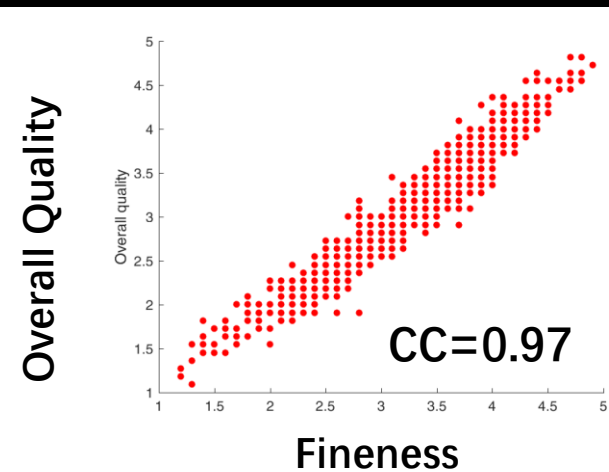
- Participants
 - 20 observers, 14 male and 6 female
 - an ongoing task, more observers will join
 - Game testers
 - Not designers, not artist, not working in game evaluation
- Environment
 - Indoor, office environment
 - Display: HP P223A
- Procedure [Rosenfield,et al, 2017]
 - Adjust image size to simulate the visual field of mobile phone
 - 5.9 inch at a 33.95-cm viewing distance
 - Evaluate four dimensions at the same time

Setup



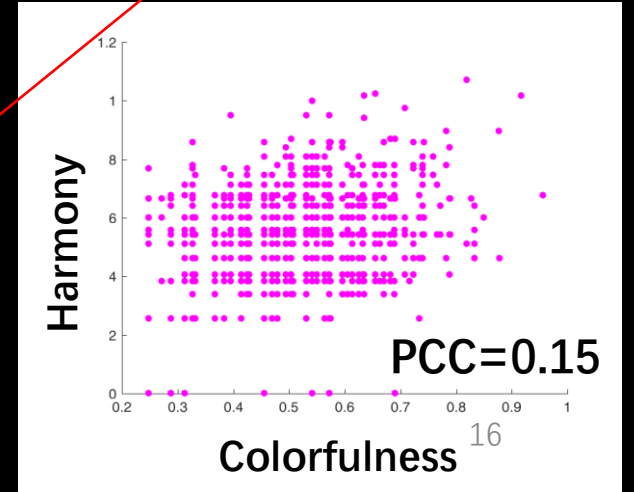
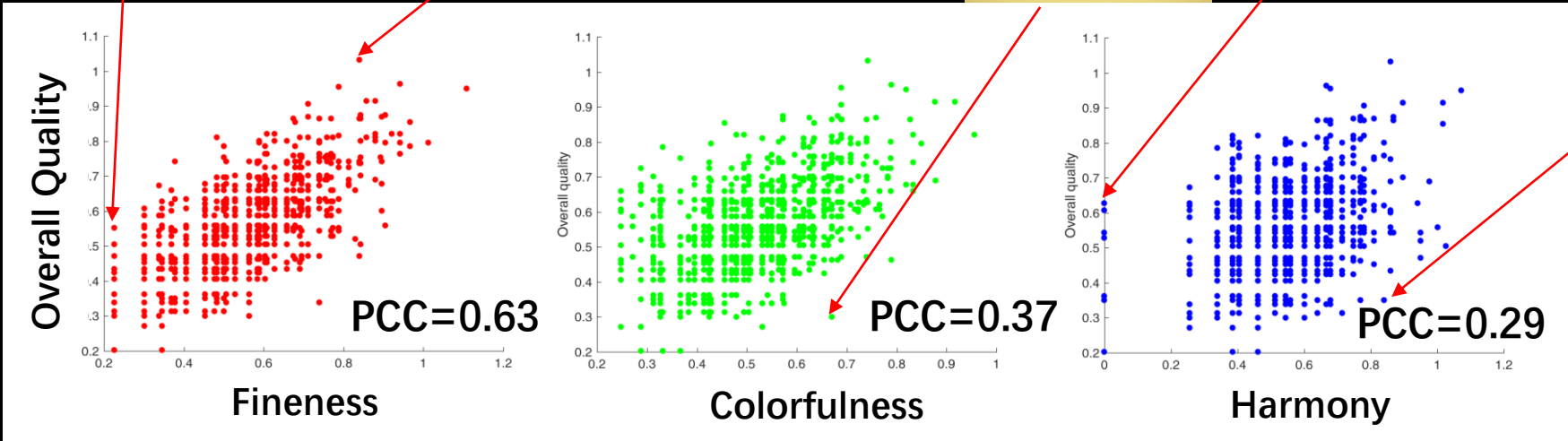
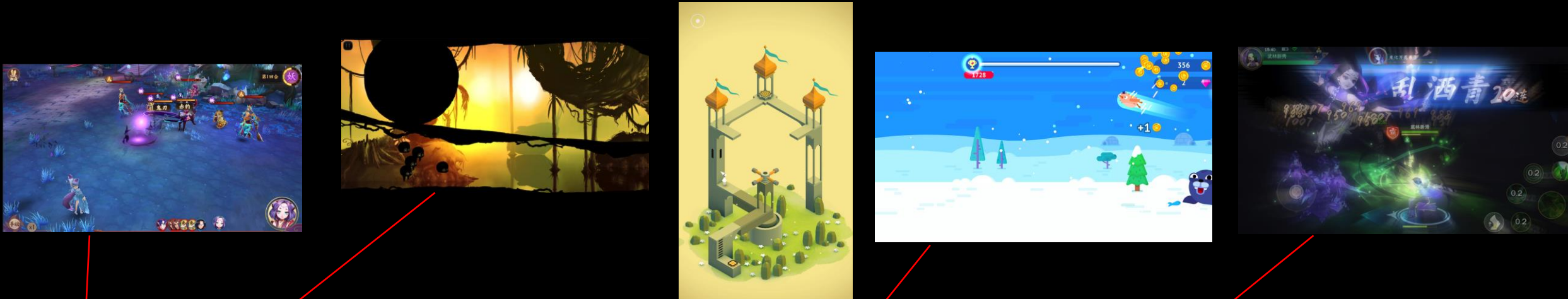
Analysis

- Correlation of MOS among dimensions

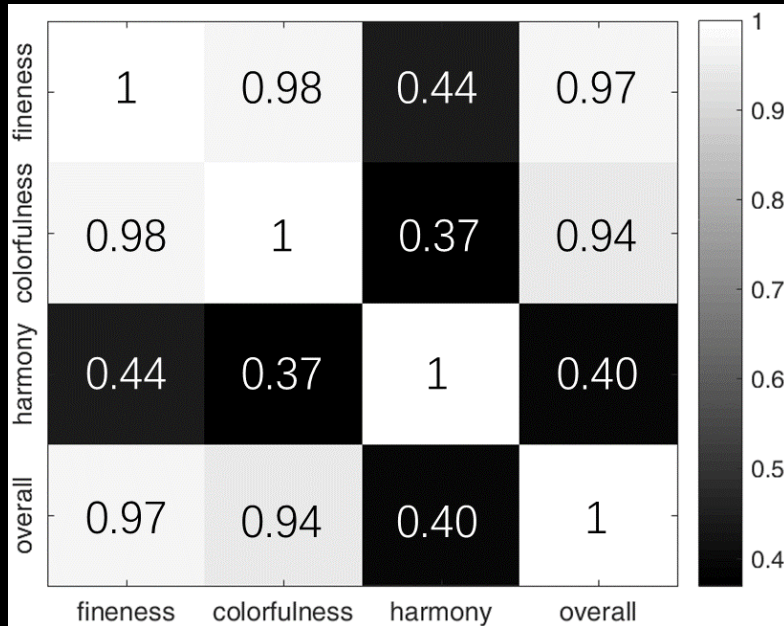


Analysis

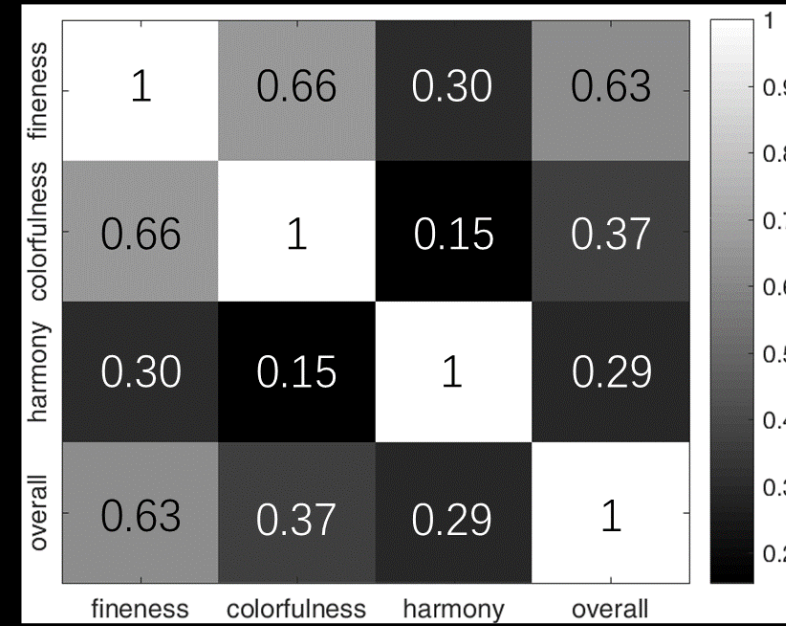
- Correlation of confidence interval between dimensions
 - Agreement among the observers



Analysis



PCC of MOS



PCC of CI

Take away

- Fineness and colorfulness are two important factors that affect the overall aesthetic assessment.
- Two color-related dimensions are not associated with each other.
- Harmony seems to be a totally different dimension that could provide other special information.
- the difficulty/uncertainty-level of the our dimensions are not associated. Observers consider they are different tasks when scoring the four dimensions.

Benchmarking objective metrics

PCC	Fineness	Colorful	Harmony	Overall
Color [ISOP, 2003]	0.3353	0.3624	0.6563	0.3679
ColorNet [ICIP, 2019]	0.0761	0.0843	0.1373	0.0527
CPBD [TIP, 2011]	0.5545	0.6007	0.3171	0.4868
Blur [ISOP, 2007]	0.1412	0.1293	0.1783	0.1408
NIMA [TIP, 2018]	0.1509	0.1459	0.3864	0.1401
MLSP [CVPR, 2019]	0.4258	0.4649	0.1971	0.3471

Take away: CPBD achieves the best performance in terms of predicting the fineness, colorfulness, and overall aesthetic quality score, when metric Color performs the best in predicting harmony.

Thank you!