Supplementary Material for Paper "An Automated Estimator of Image Visual Realism Based on Human Cognition"

Shaojing Fan[1,2], Tian-Tsong Ng[2], Jonathan S. Herberg[4], Bryan L. Koenig[3], Cheston Y. -C. Tan[2], and Rangding Wang[*1]

[1]Ningbo University
[2]Institute for Infocomm Research

[3]Lindenwood University
[4]Institute of High Performance Computing

A. Pilot study for psychophysics study I

We performed a pilot study to determine how many participants we needed to reliably estimate human judgments of visual realism. We had 160 workers from Mechanical Turk ($> 95\%$ approval rating in Amazon's system) to judge 60 images as a "photo" or "CG". We calculated a *realism score* (ranges from 0 to 1) for each image as the proportion equal to the number of judgments indicating that the image is a photo over the total number of judgments for that image.

We next used bootstrapping to evaluate how reliable the judgments were for various numbers of participants. For multiple group sizes, we randomly split the participants into two equal-sized groups and calculated the Spearman's rank-order correlation (ρ) between the two groups' realism scores. We did so 25 times per group size. We also calculated the root mean square errors (RMSE) of each image's realism scores in the similar way, using the data of all 160 participants as ground truth. When the number of participants was over 30, ρ was close to 0.8 and RMSE was around 0.075 (Fig. 1), suggesting 30 judgments per image is sufficient to reliably estimate visual realism.

B. Questionnaire for psychophysics study II

Following are the questions we used for image attributes annotation on Amazon Mechanical Turk.

Please look at the image on the left and answer the following questions:

1. Please rank the degree that the image appears to be a photograph versus a computer-generated image:

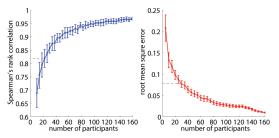   (1) Definitely a photo

    (2) Probably a photo

Figure 1. Human performance analysis on pilot study. (a) Spearman's rank correlation between two random splits of participants as a function of participants size. (b) Root mean square error of image scores as a function of participants size. All results are averaged over 25 random splits.

    (3) Not clearly a photo or a computer-generated image

    (4) Probably a computer generated-image

    (5) Definitely a computer generated-image

2. Is the scene in the image familiar to you?

    (1) Yes, I think I have seen similar things/scenes before.

    (2) No, I have never seen similar things/scenes before.

3. Is this image attractive to you?

    (1) Very attractive

    (2) Moderately attractive

    (3) Neither attractive nor unattractive

    (4) Moderately unattractive

    (5) Very unattractive

4. How does the lighting effect appear to you?

    (1) Natural

    (2) Not clearly natural or unnatural.

---
[*]Corresponding author. E-mail:wangrangding@nbu.edu.cn

(3) Unnatural

5. Is the image colorful?

    (1) Very colorful

    (2) Moderately colorful

    (3) Not colorful

6. Does the color in the image appear natural?

    (1) Yes, the color appears natural.

    (2) No, the color appears strange or unusual.

7. Do the colors appear to go well together?

    (1) Yes, the colors go well together.

    (2) No, the colors seem strange/unusual together.

8. How sharp is the image?

    (1) Very sharp

    (2) Moderately sharp

    (3) Neither sharp nor blurry

    (4) Moderately blurry

    (5) Very blurry

9. What's the quality of the image?

    (1) Very high quality

    (2) Moderately high quality

    (3) Medium quality

    (4) Moderately low quality

    (5) Very low quality

10. Do you see shadows in the image?

    (1) Yes, there are obvious shadows in the image.

    (2) Yes, but the shadows are not obvious. They are only in a small region, are very light, or something similar.

    (3) No, there are no shadows in the image.

11. If you see shadows in the image, would you characterize them as sharp or soft?

    (1) Most shadows are sharp.

    (2) Some shadows are sharp, and some are soft.

    (3) Most shadows are soft.

12. Are there living things in the image?

    (1) Yes

    (2) No

13. Does the image show a naturally-occurring combination/arrangement of objects?

    (1) Yes, both combination and arrangement are natural.

    (2) Natural combination, but unnatural arrangement.

    (3) Natural arrangement, but unnatural combination.

    (4) No, both combination and arrangement are unnatural.

14. Are the objects in the scene look familiar to you?

    (1) Yes, all objects are familiar.

    (2) Some are familiar, some are unfamiliar.

    (3) No, all objects are unfamiliar.

15. Is the appearance of the objects natural?

    (1) Mostly natural

    (2) Some are natural, some are unnatural.

    (3) Mostly unnatural

16. Does the image contain fine details?

    (1) A lot of fine details

    (2) Some fine details

    (3) No fine details

17. Do the scene and objects look clean?

    (1) Yes

    (2) No

18. Is this image unusual or strange?

    (1) Highly unusual/strange

    (2) Moderately unusual/strange

    (3) Not at all unusual/strange

19. Does this image look like it is a photograph taken by a professional photographer?

    (1) Definitely yes

    (2) Probably yes

    (3) Not clearly yes or no.

    (4) Probably no

    (5) Definitely not

20. This is an image of:

    (1) Very neat space

    (2) Moderately neat space

(3) Not clearly neat or cluttered space

(4) Moderately cluttered space

(5) Very cluttered space

21. This is an image of:

(1) Very empty space

(2) Moderately empty space

(3) Not clearly empty or full space

(4) Moderately full space

(5) Very full space

22. Does the image appear to have objects of focus?

(1) Definitely yes

(2) Probably yes

(3) Not clearly yes or no

(4) Probably not

(5) Definitely not

23. Does the perspective of the image appear natural?

(1) Definitely natural.

(2) Moderately natural.

(3) Not clearly natural or unnatural

(4) Moderately unnatural.

(5) Definitely unnatural.

24. How exciting is this image?

(1) Very exciting

(2) Moderately exciting

(3) Neither exciting nor boring

(4) Moderately boring

(5) Very boring

25. Is this image mysterious?

(1) Very mysterious

(2) Moderately mysterious

(3) Not at all mysterious

26. How happy does this image make you?

(1) Very happy

(2) Moderately happy

(3) Neither happy nor unhappy

(4) Moderately unhappy

(5) Very unhappy

27. Does this image make you sad?

(1) Very sad

(2) Moderately sad

(3) Not at all sad

28. Is there a storyline in the picture?

(1) Definitely yes

(2) Probably yes

(3) Probably not

(4) Definitely not

29. Is the scene dynamic/energetic/in motion?

(1) Very dynamic/energetic

(2) Moderately dynamic/energetic

(3) Not at all dynamic/energetic

30. Does the image appear to be a close-range shot or distant-view shot?

(1) Very close range

(2) Moderately close range

(3) Between close and distant

(4) Moderately distant view

(5) Very distant view

If the image contains people, please answer the following questions, otherwise leave them BLANK.

31. Are the faces of any of the people visible?

(1) Yes

(2) No

32. Does any person appear to make eye-contact with a viewer of the image?

(1) Yes

(2) No

33. How many people are in the image?

(1) One

(2) Two

(3) More than two, but not a crowd.

(4) A large assembling of people

34. Do any persons who seem to be the focus of the image appear to be posing for the image?

(1) Definitely yes

(2) Probably yes

(3) Definitely not

35. What are the persons who seem to be the main focus of the image doing? Check the box by each appropriate activity.

    (1) Standing, but not walking.

    (2) Sitting or lying down

    (3) Walking

    (4) Running

    (5) Riding/driving

    (6) Talking/Shouting

    (7) Other activities

If the faces of the persons in the main focus are visible, please answer the following questions, otherwise leave them BLANK.

36. Are the expressions of the persons who are the main focus of the image genuine?

    (1) Definitely genuine

    (2) Probably genuine

    (3) Perhaps genuine, perhaps fake.

    (4) Probably fake

    (5) Definitely fake

37. What are the expressions of the persons who are the main focus of the image? Check the box by each appropriate expression.

    (1) Happiness

    (2) Anger

    (3) Surprise

    (4) Sadness

    (5) Fear

    (6) Disgust

    (7) Contempt

    (8) Neutral (no expression)

    (9) Can't tell

38. How attractive are the persons who are the main focus of the image?

    (1) Very attractive

    (2) Moderately attractive

    (3) Not at all attractive

C. Feature dimensions

The dimensions of all features used in the paper are listed in Table 1 and 2.

Table 1. Dimensions of the features in our method.

Attribute	Feature	Dimension
Naturalness	Statistics of intrinsic components	5
	Natural image statistics	98
	CG characteristics	108
Familiarity	Content-based similarity measure	50
	Color compatibility	77
	Color name	5376
Aesthetics	Ke's metrics on aesthetics	11
	SSIM	5376
Semantics	GIST	512
	Object Bank	5576

Table 2. Dimensions of other computer features used in comparison.

Category	Feature type	Dimension
Signal feature	Wavelet [5]	216
	Geometry feature [6]	196
	Camera noise [2]	4
	Color compatibility [4]	77
Object & scene feature	SIFT [3]	1280
	GIST [3]	512
	HOG2x2 [3]	2100
	LBP [3]	1239
Feature learning	K-means encoding [1]	2560

References

[1] A. Coates, A. Y. Ng, and H. Lee. An analysis of single-layer networks in unsupervised feature learning. In *Conference on Artificial Intelligence and Statistics*, 2011. 4

[2] E. Dirik, S. Bayram, H. Sencar, and N. Memon. New features to identify computer generated images. In *ICIP, 2007*. 4

[3] A. Khosla, J. Xiao, A. Torralba, and A. Oliva. Memorability of image regions. In *NIPS*, 2012. 4

[4] J. Lalonde and A. Efros. Using color compatibility for assessing image realism. In *ICCV*, 2007. 4

[5] S. Lyu and H. Farid. How realistic is photorealistic? *IEEE Transactions on Signal Processing*, 2005. 4

[6] T.-T. Ng and S.-F. Chang. Discrimination of computer synthesized or recaptured images from real images. In *Digital Image Forensics*. 2013. 4