Hindawi Publishing Corporation
Mobile Information Systems
Volume 2016, Article ID 9621925, 17 pages
http://dx.doi.org/10.1155/2016/9621925

Research Article

Perceived Image Quality on Mobile Phones with Different Screen Resolution

Wenjie Zou, Jiarun Song, and Fuzheng Yang

Collaborative Innovation Center of Information Sensing and Understanding and ISN, Xidian University, Xi'an 710071, China

Correspondence should be addressed to Jiarun Song; jrsong@xidian.edu.cn

Received 10 October 2015; Revised 6 January 2016; Accepted 6 March 2016

Academic Editor: Weisi Lin

The diverse display screen imposes significant challenges for assessing the perceptual media quality across different mobile devices. In this paper, the perceived image quality on different mobile phones is investigated. Firstly, subjective experiments for image quality evaluation are implemented on 9 popular mobile phones and a broadcast-quality monitor to evaluate the impact on perceived image quality regarding the screen resolution, screen size, image resolution, and image coding quality. Furthermore, the effect of mutual interaction between the image resolution and screen resolution is analyzed and an integrated assessment parameter is proposed to establish a device-dependent image quality assessment model. This model can be used to predict the user's perceptual quality of the images displayed on different mobile devices. Experimental results using twofold cross-validation indicate that the proposed model can accurately estimate users' perceived image quality on mobile devices.

1. Introduction

Mobile phones have now become a crucial part in one's daily life, leading to rapid growth of mobile multimedia applications and fast development of related hardware technologies. Lots of manufacturers have been focused on improving the screen size and resolution to provide a preferable viewing experience for consumers. For this purpose, the screen resolution of a mobile phone has been significantly increased, from the incipient QCIF (176 × 144) or even smaller to the current Quad HD (2560 × 1440) or even Ultra HD (3840 × 2160). Meanwhile, the screen size has also been enlarged from smaller than 1 inch to 5.7 inches or even larger as the flagship of industrial standards. It is reported that nearly one-third of smartphones sold in 2012 had a screen size larger than 4.5 inches [1], and the smartphones with 4.5-inch screens or more have represented almost 80% of all new models till May 2014 [2]. The screen resolution and size are now regarded as two key choice factors for smartphones. However, whether the increase in the screen size and resolution is helpful for improving the perceived quality visually still remains unclear, and if yes, how much is the related gain? Therefore, an objective and accurate assessment of user's perceived visual quality is needed regarding videos and images displayed on the mobile phone.

In the past decades, a number of objective image quality assessment (IQA) and video quality assessment (VQA) algorithms have been proposed to evaluate the image/video quality, as summarized in [3–6]. However, these traditional works only focused on evaluating the quality of image and video sources, losing sight of the effect from the display or the specific mobile device. Although some literatures have been involved with subjective experiments on mobile devices for media quality evaluation [7–11], they still addressed the characteristics of image or video sources or only focused on the impact of the assessment technologies. Moreover, the influence of usage location of mobile devices has been also investigated in [12–16].

In practice, the mobile device significantly influences the user's perceived quality in terms of the screen resolution, screen size, and devices type. In [17], four screen resolutions were studied in video quality evaluation, that is, QCIF (176 × 144), CIF (352 × 288), VGA (640 × 480), and HD (1920 × 1200), where the mobile display was simulated on a high quality monitor. The quality of scalable video was evaluated on actual mobile devices in [18], using a mobile phone with

Table 1: Parameters of display devices.

Display device	Screen size (inch)	Resolution	Screen type	PPI	Contrast ratio	Brightness cd/m^2
P1	4″	1136 × 640	IPS LCD	326	1320 : 1	640
P2	4.3″	1280 × 720	IPS LCD	342	—	—
P3	4.9″	1920 × 1080	IPS LCD	445	967 : 1	526
P4	5.1″	1280 × 720	IPS LCD	294	810 : 1	378
P5	5.1″	1920 × 1080	AMOLED	432	—	442
P6	5.1″	2560 × 1440	AMOLED	576	—	563
P7	5.5″	1920 × 1080	TFT-LCD	401	867 : 1	562
P8	5.5″	2560 × 1440	SLCD	565	1657 : 1	487
P9	5.7″	1920 × 1080	AMOLED	386	—	402
M1	30″	4096 × 2160	OLED	—	—	—

a 4.3-inch screen (800 × 480) and a tablet with a 1280 × 800 10.1-inch screen. The authors in [19] studied the impact of screen size through subjective experiments using a mobile phone (28 × 35 mm screen), a Personal Digital Assistant (PDA, 3.5-inch screen), and a laptop (15-inch screen). The results presented in [18, 19] indicated that the perceived quality does change with the screen size. A similar test regarding multimedia quality evaluation was carried out using a mobile phone, PDA, and laptop in [20], from which it is found that the user's acceptable multimedia quality is different when the multimedia is displayed on different mobile devices. In [21], subjective tests were conducted using QCIF videos displayed on a personal computer (1280 × 1024, 20-inch screen) and a mobile handset (320 × 240, 2.6-inch screen). The test results have shown that viewers rated the video with high spatiotemporal activity much lower on the mobile handset than on the computer. Therefore, recent work started to take characteristics of the mobile device into account when performing IQA or VQA. For example, authors in [22] used the pixels per inch (PPI) on the screen to predict the users' acceptability and pleasantness in various scenarios of mobile video usage. Very recently, a full-reference objective VQA algorithm, named as SSIMplus, was proposed in [23], where properties of the display device and viewing conditions were considered. However, in these works reported in literature, the tested mobile devices were rather limited, and they were of low screen resolutions and small sizes, which cannot catch the state of the art of the rapidly growing resolution of smartphones. And it is noteworthy that nowadays the commonly used image resolutions on smartphones have greatly exceeded the image resolutions provided by most of public IQA database [24–28].

In this paper, the effect of screen size and resolution of mobile phones on perceived image quality is explored, building on which an objective IQA model for perceived image quality is proposed. In our work, 9 mobile phones in vogue and a broadcast-quality monitor have been employed to conduct the subjective experiments. The screen resolution and size of the selected mobile phones cover a wide range, from 1136 × 640 to 2560 × 1440 in resolution and 4 to 5.7 inches in size. Moreover, the highest image resolution in the test database is 4K (3840 × 2160), which is now becoming common in mobile phones. Different image qualities are considered, obtained using different coding parameters.

Firstly, a variety of impact factors on perceived image quality, that is, the screen resolution, screen size, image resolution, and image coding quality, are investigated based on subjective experiments. The influence of screen size on the perceived image quality is then checked by a statistical analysis among 4.9- to 5.7-inch screens in vogue. The statistical analysis is also conducted to check the perceived image quality gain among the 720P (1280 × 720), 1080P (1920 × 1080), and Quad HD screens. The results of statistical analysis is helpful in indicating whether the user's perceived image quality can be improved by enhancing the screen resolution from 1080P to Quad HD. Furthermore, the impacts of image resolution, screen resolution were evaluated and integrated into one assessment parameter, named as effectively displayed pixels per inch (ED-PPI). Combining the ED-PPI and the image coding quality (ICQ), a device-dependent image quality assessment model is then proposed to advise the perceived image quality on different mobile phones.

2. Subjective Experiment

Two subjective experiments have been conducted to investigate whether the user's viewing experience will be significantly improved by enhancing the screen resolution and how the image resolution will affect the user's perceived quality on different mobile devices, respectively.

2.1. Test Devices. As listed in Table 1, a total of 9 popular mobile devices were chosen as the test devices in the experiments, that is, P1 to P9, and a broadcast-quality monitor (M1) was used as the benchmark. The screen size and resolution of the mobile devices are widely used in practice. Particularly, M1 is a professional device which can deliver much more dynamic, natural, and subtly rendered colors and provide a superior color restoration.

A uniform image browser, named Tidy, was used to display the images on each mobile phone, and the Google Picasa 3.9 for windows was employed to show the images on device

Figure 1: Original images. (a) FLOWER. (b) DOG. (c) SCULPTURE. (d) STREET. (e) PHOTO_WALL. (f) CAT. (g) TREES. (h) VILLAGE. (i) PARIS. (j) OPERA_HOUSE.

M1, respectively. Considering the potential influence of brand recognition, we covered each device with a corresponding full-body protection cover to conceal the brand type and only expose the screen area during the tests.

2.2. Test Material. A total of ten 4K (3840 × 2160) resolution color images were selected from dozens of pictures that we shot as the original test images, as shown in Figure 1. The image contents included the nature and artificial scenes.

Then, two image databases were generated using these original test images. In Database I, these original images were downsampled into images with four smaller resolutions using FFmpeg 0.4.9 [29], including 854 × 480, 1280 × 720, 1920 × 1080, and 2560 × 1440. The aspect ratio of images was 16 : 9. Moreover, to include a broad range of image quality, both the derived and original images were compressed into the "low," "medium," and "high" quality versions (using MATLAB *imwrite* function). The quality factor Q of the *imwrite* function was set as 5, 15, and 75, respectively. Consequently, a total of 150 images (10 contents × 5 resolutions × 3 quality levels) were obtained in Database I. In Database II, the 1280 × 720, 1920 × 1080, 2560 × 1440, and 4K images with five contents (i.e., OPERA_HOUSE, PHOTO_WALL, STREET, CAT, and VILLAGE) in Database I were selected to be the original images. They were further downsampled into different resolutions, respectively. Specifically, the 4K images were downsampled into 2560 × 1440, 1920 × 1080, 1280 × 720, and 1137 × 640. The 2560 × 1440 images were downsampled into 1920 × 1080, 1280 × 720, and 1137 × 640. The 1920 × 1080 images were downsampled into 1280 × 720 and 1137 × 640. The 1280 × 720 images were downsampled into 1137 × 640. These 150 downsampled images (5 contents × 10 resolutions × 3 quality levels) constituted Database II.

2.3. Experimental Procedures. Two subjective tests were designed for evaluation of the perceived image quality. For the first experiment, the influences of the screen resolution and image resolution on the image quality were checked, respectively, using all the display devices, that is, 9 popular mobile phones and 1 broadcast-quality monitor. For the second experiment, the necessity of Quad HD screen was investigated by comparing the user's perceived image quality on 1080P and Quad HD screens. The detail procedures of the two experiments are described as follows.

2.3.1. Experiment I: Rating the Perceived Image Quality and Image Coding Quality. A total number of 28 nonexpert subjects participated in this experiment, including 16 males and 12 females aged between 20 and 30 years. All of them had normal or correct-to-normal sight. Each subject viewed the images in Database I with a random order on each mobile device and viewed both Databases I and II on M1. He/she rated his/her perceived image quality in the Absolute Category Rating (ACR) 5-point scale (corresponding to the perceived quality of "excellent," "good," "fair," "poor," and "bad"). The environment of the experiments was set following the suggestion of ITU-R recommendation BT.500-13 [30].

It is noted that there were two differences during the test on the mobile phones and the broadcast-quality monitor, that is, the display scale and viewing distance. Firstly, the images were displayed in full screen on the mobile phones, where the aspect ratio of the image was in accordance with that of the original image. Comparatively, on M1, the images were displayed in their original resolutions. Secondly, the viewing distance to the mobile phones' screen was determined by the viewers themselves. They were informed to choose a comfortable distance according to their preference [31], within a distance limitation that the upper bound was four times of the displayed image height. Differently, the distance to the device M1 was set to be 3.5 times of the image height as specified in [32]. Here, the rating scores of image quality on M1 were employed as the image coding quality, which was used for comparison purposes.

During the test, the subjects were encouraged to take breaks after rating all image quality on each two devices to avoid visual fatigue. Before the formal test, the subjects were asked to rate a few example images to get familiar with the scoring scale and the image browsers.

2.3.2. Experiment II: Comparison Test on Devices with the Same Screen Size. In the second experiment, only 10 high quality 4K images in the database were tested. The general principle of double stimulus comparison scale (DSCS) method [31] was followed to compare the user's perceived image quality of the same image on two different devices. A five-level scale was used as the comparison scale, shown as follows.

Five-Level Rating Scale

−2: worse.

−1: slightly worse.

0: the same.

1: slightly better.

2: better.

The same subjects in Experiment I were required to view one image on two mobile phones successively, and they would rate the impairment scores after careful comparison. There were no specific viewing orders between two devices, so the subjects could determine which device to view first. The experiment consisted of two comparisons, that is, P5 versus P6 and P7 versus P8. For instance, the resolutions of P5 and P6's screens were 1080P and Quad HD, respectively, while the screen size of both devices was 5.1 inches.

2.4. Raw Data Processing. After the subjective tests, the credibility of assessment results was checked using the linear Pearson correlation coefficient (LPCC) suggested by ITU-T Recommendation P.913 [31]. The LPCC is calculated as follows:

$$\text{LPCC} = \frac{\sum_{i=1}^{n} x_i \cdot y_i - \sum_{i=1}^{n} x_i \cdot \sum_{i=1}^{n} y_i / n}{\sqrt{\left(\sum_{i-1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2 / n\right)\left(\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2 / n\right)}}, \quad (1)$$

where x_i was the average value of all subjects' rating scores for the ith image and y_i was the individual rating score of one subject for the ith image. Here n represented the number of images.

The values of LPCC for each subject in Experiment I were calculated and finally results from two subjects were discarded since LPCC values of their rating scores were lower than the discard threshold, that is, 0.75 [31]. As a result, the number of the valid subjects (i.e., 26) meets the requirement

Table 2: Minimum LPCC on each device.

Devices	P1	P2	P3	P4	P5	P6	P7	P8	P9	M1
LPCC	0.86	0.87	0.83	0.85	0.89	0.86	0.86	0.85	0.89	0.88

Table 3: Results of internal consistency checking.

Devices	P1	P2	P3	P4	P5	P6	P7	P8	P9	M1
Alpha	0.978	0.976	0.980	0.982	0.982	0.980	0.976	0.983	0.978	0.976

of the Video Quality Experts Group (VQEG). Table 2 lists the minimum LPCC of viewer's rating scores on each device after the screening process. The perceived image quality of each image was measured in terms of the average score of all valid subjects, also known as the Mean Opinion Score (MOS) [31].

Moreover, the subjects in Experiment II were also screened according to the screening result in Experiment I. If the subject was discarded in Experiment I, all the values they rated in Experiment II were discarded as well. The perceived image quality difference for each image pair was measured in terms of the average score of all valid subjects, also known as the Differential Mean Opinion Score (DMOS) [31].

Then, Cronbach's alpha value was computed to measure the internal consistency of the valid scores on each device. As per the results illustrated in Table 3, the value of alpha for each device is considerably large, which indicates that there is a strong internal consistency among the valid subjects.

In summary, a total of 1360 data points were obtained on the mobile phones apart from 300 data points on the monitor in two experiments. The results on the monitor were utilized as the image coding quality in further analysis.

2.5. Characteristic of the Image Coding Quality. The relationship between the vertical resolution of the image and the image coding quality is illustrated in Figure 2, where different types of points represent different image contents, and each type of lines links the average value of image coding quality for different contents at the same quality level (i.e., the same Q). It can be seen that the image coding quality under the same quality factor Q is nearly constant for different resolutions. This trend is significantly different with that of the corresponding perceived image quality on mobile phones (discussed in Section 5), which means that it is inappropriate to directly employ the image coding quality as the perceived image quality on mobile phones. Hence, it is necessary to further estimate the perceived image quality based on the image coding quality.

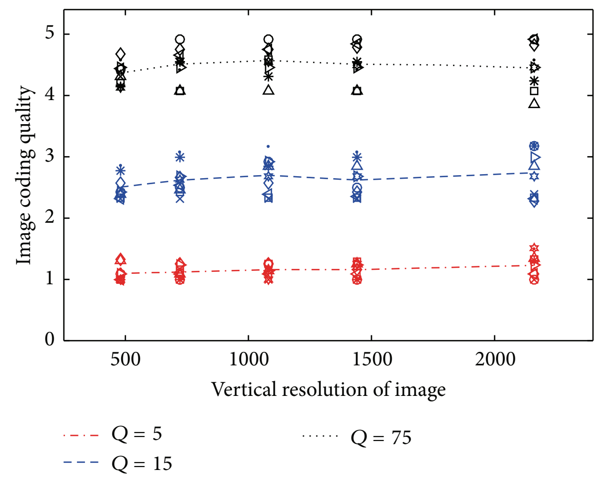


Figure 2: Relationship between the vertical resolution of image and the image coding quality under three quality levels.

3. Perceived Image Quality on Diverse Screen Sizes

To meet different preferences of consumers, the manufactures tend to produce mobile phones with various screen sizes. In recent years, the mobile phones are likely to have larger size of display screen. However, there is lack of an investigation on the user's perceived image quality on screens with different sizes, without considering the comfortable grip feeling and convenient one-handed operation.

In this section, the perceived image quality on diverse screen sizes is firstly investigated based on the rated scores, that is, MOS, for the images displayed on the P3, P5, P7, and P9, respectively. Considering the possible influence of the screen resolution, these screens have the same resolution (i.e., 1080P) but in different sizes (i.e., 4.9, 5.1, 5.5, and 5.7 inches). Take the high and low quality images with five randomly selected contents (FLOWER, OPERA_HOUSE, PARIS, PHOTO_WALL, and STREET) as an example; the relationship between the MOS and the screen size is shown in Figure 3. It can be seen that there is no significant increase or decrease in the perceived quality, when the screen is increased from 4.9 inches to 5.7 inches. The viewer's perceived quality is not significantly influenced by the change of screen size during the viewing process when the devices have the 1080P resolution, and the large screen does not show its superiority on providing better perceived image quality.

In a general sense, the MOS for the images displayed on all mobile phones are used to illustrate the difference of perceived image quality across 4-inch to 5.7-inch screens. Figure 4 takes the 4K images with PARIS and PHOTO_WALL as an example. It can be seen that there is still no significant increase or decrease in the perceived quality though the little fluctuation of perceived image quality presents on the devices with different screen resolutions.

Furthermore, a statistical analysis, that is, the one-way analysis of variance (ANOVA), is further performed to check the significance of influence of the screen size on

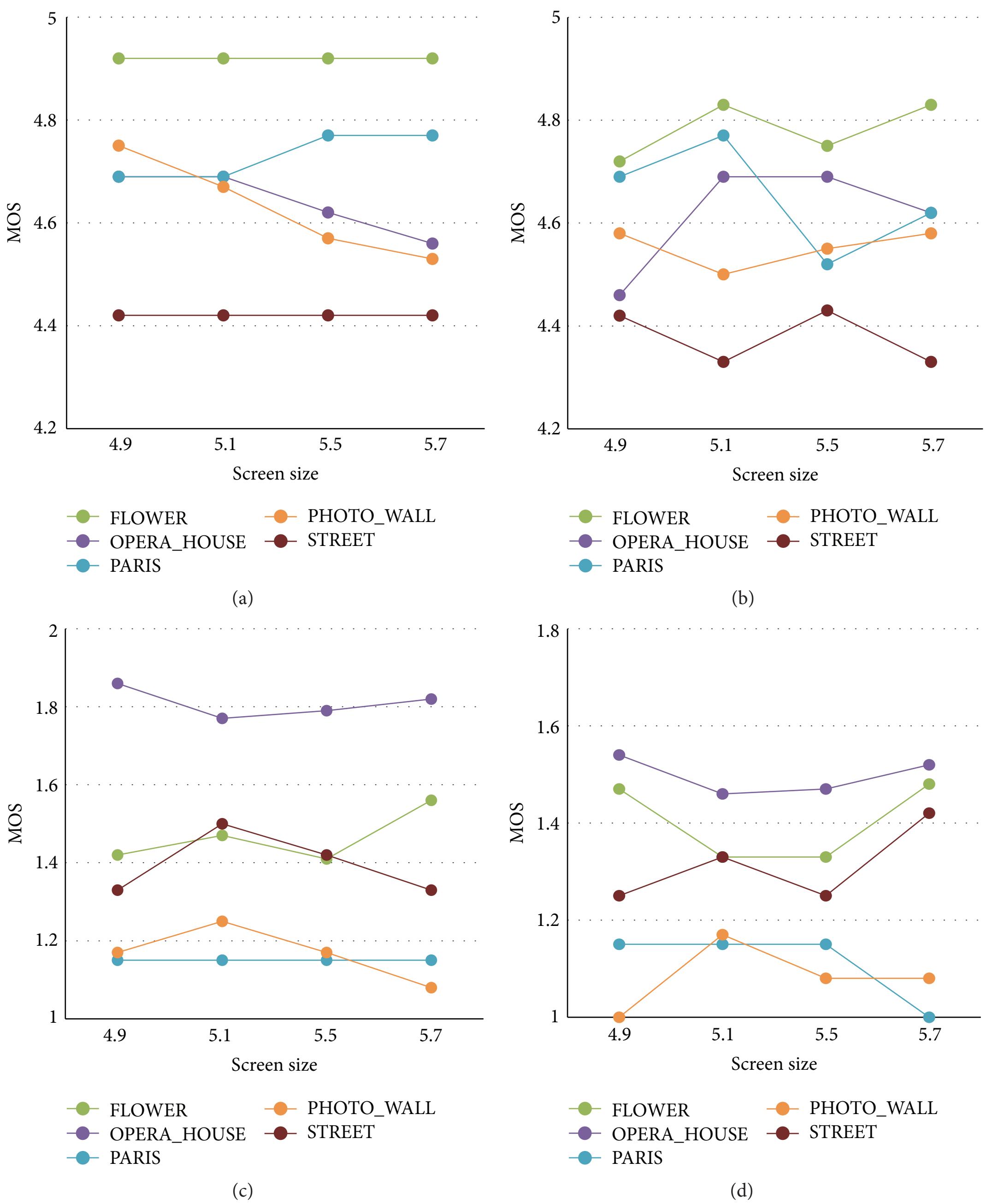


Figure 3: MOS versus screen size. (a) 4K high quality images. (b) 1080P high quality images. (c) 4K low quality images. (d) 1080P low quality images.

the perceived image quality. The test is firstly implemented on P3, P5, P7, and P9. The analysis is conducted under different image resolutions and quality levels. The corresponding F and p values are listed in Table 4. It can be seen that all the p values are significantly larger than 0.05 at 95% level. It indicates that the screen size does not have a significant correlation with the perceived image quality on P3, P5, P7, and P9, that is, 4.9- to 5.7-inch screens. Similarly, the one-way ANOVA test is also performed on the rated scores on all mobile phones, that is, 4- to 5.7-inch, for a common conclusion. The F value is 0.048, and its corresponding p value is 1 which is larger than 0.05 at 95% level. Therefore, it is safe to conclude that impact of the screen size on the perceived image quality is not significant from 4- to 5.7-inch screen, and the fluctuation of perceived image quality is mainly caused by another difference of devices, that is, screen resolution.

In conclusion, there seems no direct relevance between the screen size and the perceived image quality when the screen size is from 4 to 5.7 inches. The reason for this phenomenon may be due to the flexible (i.e., unfixed) viewing

Table 4: The results of one-way ANOVA for P3, P5, P7, and P9 under different image resolution and quality levels. The F and p value are denoted as (F, p).

Quality level	2160P	1440P	1080P	720P	480P
High (Q = 75)	(0.200, 0.896)	(0.447, 0.721)	(0.901, 0.450)	(0.942, 0.430)	(0.116, 0.950)
Medium (Q = 15)	(0.310, 0.993)	(0.266, 0.850)	(0.190, 0.902)	(0.020, 0.996)	(0.172, 0.915)
Low (Q = 5)	(0.146, 0.931)	(0.158, 0.924)	(0.072, 0.975)	(0.132, 0.940)	(0.277, 0.841)

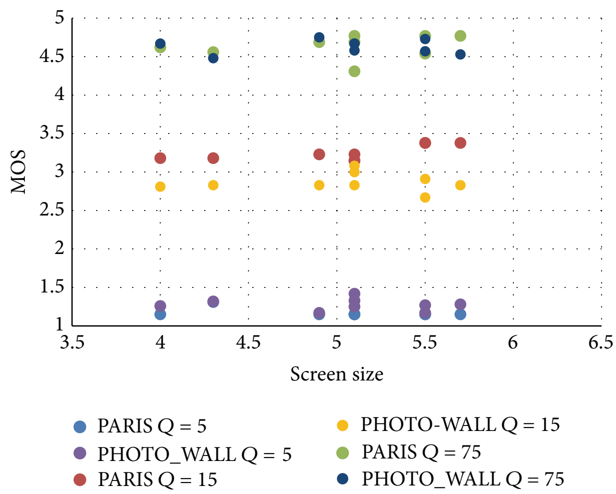


Figure 4: MOS versus screen size.

distance during the subjective experiment. The subjects can adopt a distance by themselves which can preview the image most legibly. For example, we noticed that the device with a larger screen was generally put at a longer distance from the subject than the device with a smaller screen during the subjective experiment. Considering viewing images only, this "self-adaptive" viewing distance tends to mitigate the influence of the screen size.

4. Benefit of Increasing the Screen Resolution

In this section, the impact of another crucial characteristic of screen on the perceived image quality, that is, screen resolution, is investigated. The benefit of increasing the screen resolution for improving the user's experience is then evaluated among the popular 720P, 1080P, and Quad HD screens.

4.1. Observation and Quantitative Analysis on the Impact of Different Screen Resolutions. The impact of different screen resolutions on the perceived image quality is evaluated by analyzing the MOS rated on P4 (720P, 5.1 inches), P5 (1080P, 5.1 inches), P6 (Quad HD, 5.1 inches), P7 (1080P, 5.5 inches), and P8 (Quad HD, 5.5 inches) in Experiment I. The analysis is performed individually for the same image resolution, since it is needed to avoid the impact of image resolution when investigating the impact of screen resolution. It is noted that the image with a high quality level has been paid more attention during the analysis, while the image with medium and low quality levels are studied on the way.

The 4K images, which can be captured by the camera on the mobile phone conveniently nowadays, are firstly selected to observe the difference of the perceived image quality on the screen with different resolutions. Figure 5(a) shows the perceived image quality of 4K high quality images displayed on the 5.1-inch screens, that is, P4–P6. The screen resolutions of the mobile phones are 720P, 1080P, and Quad HD, respectively. It can be seen that the perceived image quality for most images can obtain a slight improvement (average 0.15) when the screen resolution is increased from 720P to 1080P. However, this increasing trend seems to be not obvious when the screen resolution is increased from 1080P to Quad HD on the 5.1-inch screen. In another word, the Quad HD screen does not guarantee significantly better users' experience. This phenomenon can also be observed on the 5.5-inch screens, as shown in Figure 5(b). It can be found that the perceived image quality of some images (e.g., "PARIS" and "OPERA_HOUSE" on the 5.1-inch screen) even decreases when the screen resolution is increasing. Consequently, when the subjects view the high quality 4K images on the 5.1- and 5.5-inch screen, it seems that the subjects may not perceive a higher image quality with the Quad HD screen than with the 1080P screen.

For a more precise analysis, Table 5 lists the increase in the perceived quality (denoted as ΔPQ) for each image in Figure 5 when they are displayed on the screen with a higher screen resolution. It can be seen that there is no decrease of the perceived quality for any 4K high quality image when increasing the screen resolution from 720P to 1080P (P4 and P5). This phenomenon indicates that the 1080P screen can provide a meaningful gain on user's perceived image quality for the high quality 4K images. This improvement brought by the screen resolution can be distinguished by the viewers. However, not all 4K images can obtain a quality increment on the Quad HD screen compared to the 1080P screen. Although 80% of 4K images obtain a performance gain (average 0.06) on P7, the remaining 20% have a much more decrease (average −0.12) in perceived image quality in this condition. Likewise, similar phenomenon can be observed from the data of images with medium and low quality levels as shown in Table 5. The improvement of screen resolution from 1080P to Quad HD is not necessary, since the human visual system cannot identify the difference when humans watch the screen from a common distance.

Moreover, the results of Experiment II also indicated the previous results. Figure 6 shows the DMOS rated on two groups of the 1080P and Quad HD screens, where DMOS > 0 or DMOS < 0 represents that the subjects can perceive a better or worse image quality on the screen with a higher resolution than with a lower resolution, respectively. The average values of DMOS in Figures 6(a) and 6(b) are 0.01 and 0.09, respectively, which are very close to zero. This results further indicate that the subjects experienced similar perceived

Table 5: Comparison between perceived image quality for 4K images on P4, P5, P6, P7, and P8.

Quality level	Test devices	Percentage of ΔPQ ≥ 0	Mean ΔPQ	Percentage of ΔPQ < 0	Mean ΔPQ
High (Q = 75)	P4 versus P5	100%	0.12	0%	—
	P5 versus P6	80%	0.06	20%	−0.12
	P7 versus P8	60%	0.19	40%	−0.14
Medium (Q = 15)	P4 versus P5	90%	0.12	10%	−0.07
	P5 versus P6	60%	0.11	40%	−0.18
	P7 versus P8	40%	0.18	60%	−0.13
Low (Q = 5)	P4 versus P5	80%	0.13	20%	−0.09
	P5 versus P6	70%	0.12	30%	−0.18
	P7 versus P8	60%	0.13	40%	−0.16

*P4 (720P, 5.1 inches), P5 (1080P, 5.1 inches), P6 (Quad HD, 5.1 inches), P7 (1080P, 5.5 inches), and P8 (Quad HD, 5.5 inches).

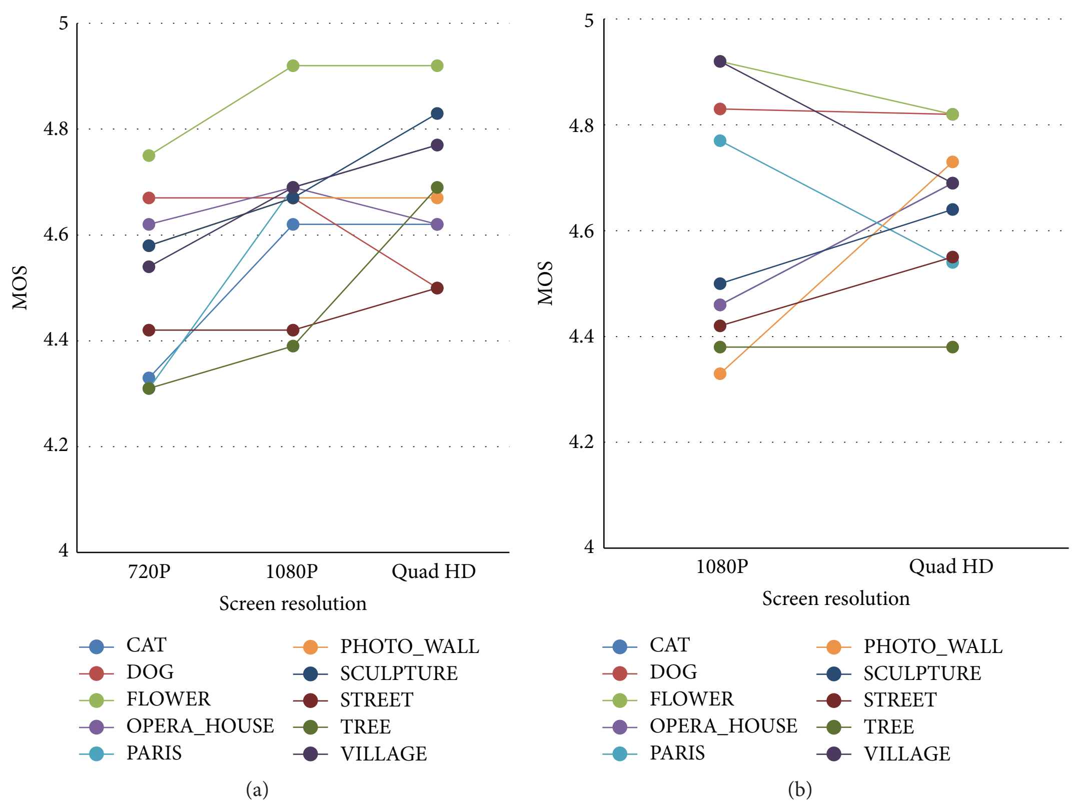


Figure 5: MOS for ten high quality 4K images displayed on 720P, 1080P, and 2K screen. (a) On P4, P5, and P6 with 5.1-inch screen. (b) On P7 and P8 with 5.5-inch screen.

image quality on the 1080P and Quad HD screen regarding the high quality 4K image. Hence, when the user views the high quality images, such as the images captured by the camera of mobile phone, the 1080P screen can provide almost the same user experience as the Quad HD screen.

For the 1440P, 1080P, and 720P high quality images, the variation of perceived image quality for these image is similar to that for the 4K images. Compared with the perceived image quality on the 720P screen, a slight improvement on perceived image quality is obtained when the images are displayed on the 1080P screen, as shown in Figures 7(a), 7(c), and 7(e). Table 6 lists the details of the variation of perceived quality for different image quality levels, from which it can be seen that the user's perceived image quality on both 5.1- and 5.5-inch Quad HD screen has no advantage compared to that on the 1080P screen, with different quality levels considered.

However, for the high quality 480P images, no significant increase or decrease in the perceived quality can be observed among the 720P, 1080P, and Quad HD screen, as plotted in Figure 8. The results in Table 7 also show this phenomenon. Furthermore, all three screen resolutions do not provide a "good" viewing experience, which corresponds to 4 points. In this case, the viewers cannot even distinguish the experience

Table 6: Comparison between perceived image quality for 1440P, 1080P, and 720P images on P4, P5, P6, P7, and P8.

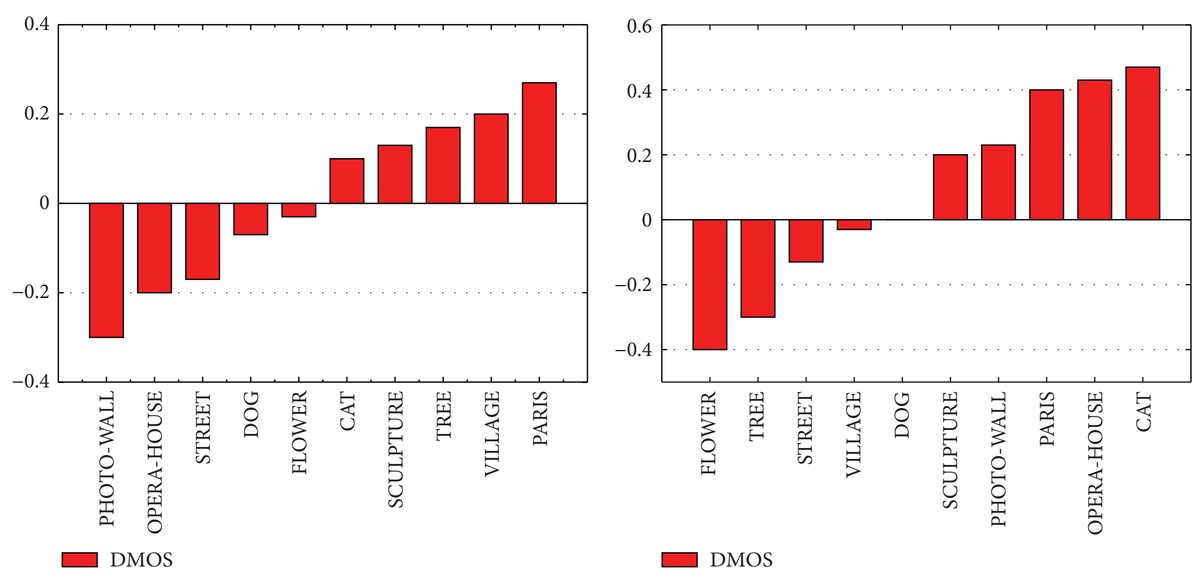
Quality level	Test devices	1440P				1080P				720P			
		ΔPQ ≥ 0 (%)	Mean ΔPQ	ΔPQ < 0 (%)	Mean ΔPQ	ΔPQ ≥ 0 (%)	Mean ΔPQ	ΔPQ < 0 (%)	Mean ΔPQ	ΔPQ ≥ 0 (%)	Mean ΔPQ	ΔPQ < 0 (%)	Mean ΔPQ
Q = 75	P4, P5	90%	0.16	10%	−0.04	80%	0.22	20%	−0.09	30%	0.03	70%	−0.13
	P5, P6	80%	0.05	20%	−0.11	50%	0.15	50%	−0.17	40%	0.11	60%	−0.19
	P7, P8	60%	0.11	40%	−0.07	80%	0.11	20%	−0.15	50%	0.09	50%	−0.07
Q = 15	P4, P5	80%	0.11	20%	−0.09	80%	0.21	20%	−0.13	30%	0.08	70%	−0.19
	P5, P6	60%	0.12	40%	−0.16	40%	0.12	60%	−0.11	70%	0.08	30%	−0.21
	P7, P8	40%	0.16	60%	−0.12	40%	0.13	60%	−0.22	50%	0.14	50%	−0.13
Q = 5	P4, P5	70%	0.13	30%	−0.10	80%	0.16	20%	−0.05	30%	0.06	70%	−0.13
	P5, P6	70%	0.15	30%	−0.2	60%	0.12	40%	−0.17	80%	0.06	20%	−0.19
	P7, P8	50%	0.13	50%	−0.15	50%	0.08	50%	−0.19	70%	0.04	30%	−0.20

*P4 (720P, 5.1 inches), P5 (1080P, 5.1 inches), P6 (Quad HD, 5.1 inches), P7 (1080P, 5.5 inches), and P8 (Quad HD, 5.5 inches).

Table 7: Comparison between perceived image quality for 480P images on P4, P5, P6, P7, and P8.

Quality level	Test devices	Percentage of ΔPQ ≥ 0	Mean ΔPQ	Percentage of ΔPQ < 0	Mean ΔPQ
High (Q = 75)	P4 versus P5	70%	0.18	30%	−0.16
	P5 versus P6	70%	0.15	30%	−0.21
	P7 versus P8	50%	0.24	50%	−0.16
Medium (Q = 15)	P4 versus P5	40%	0.03	60%	−0.19
	P5 versus P6	70%	0.08	30%	−0.09
	P7 versus P8	60%	0.17	40%	−0.08
Low (Q = 5)	P4 versus P5	50%	0.02	50%	−0.01
	P5 versus P6	80%	0.02	20%	−0.08
	P7 versus P8	70%	0.04	30%	−0.08

*P4 (720P, 5.1 inches), P5 (1080P, 5.1 inches), P6 (Quad HD, 5.1 inches), P7 (1080P, 5.5 inches), and P8 (Quad HD, 5.5 inches).



(a) (b)

Figure 6: Illustration of the DMOS for ten 4K high quality images. (a) On 5.1-inch screen, that is, P5 and P6. (b) On 5.5-inch screen, that is, P7 and P8.

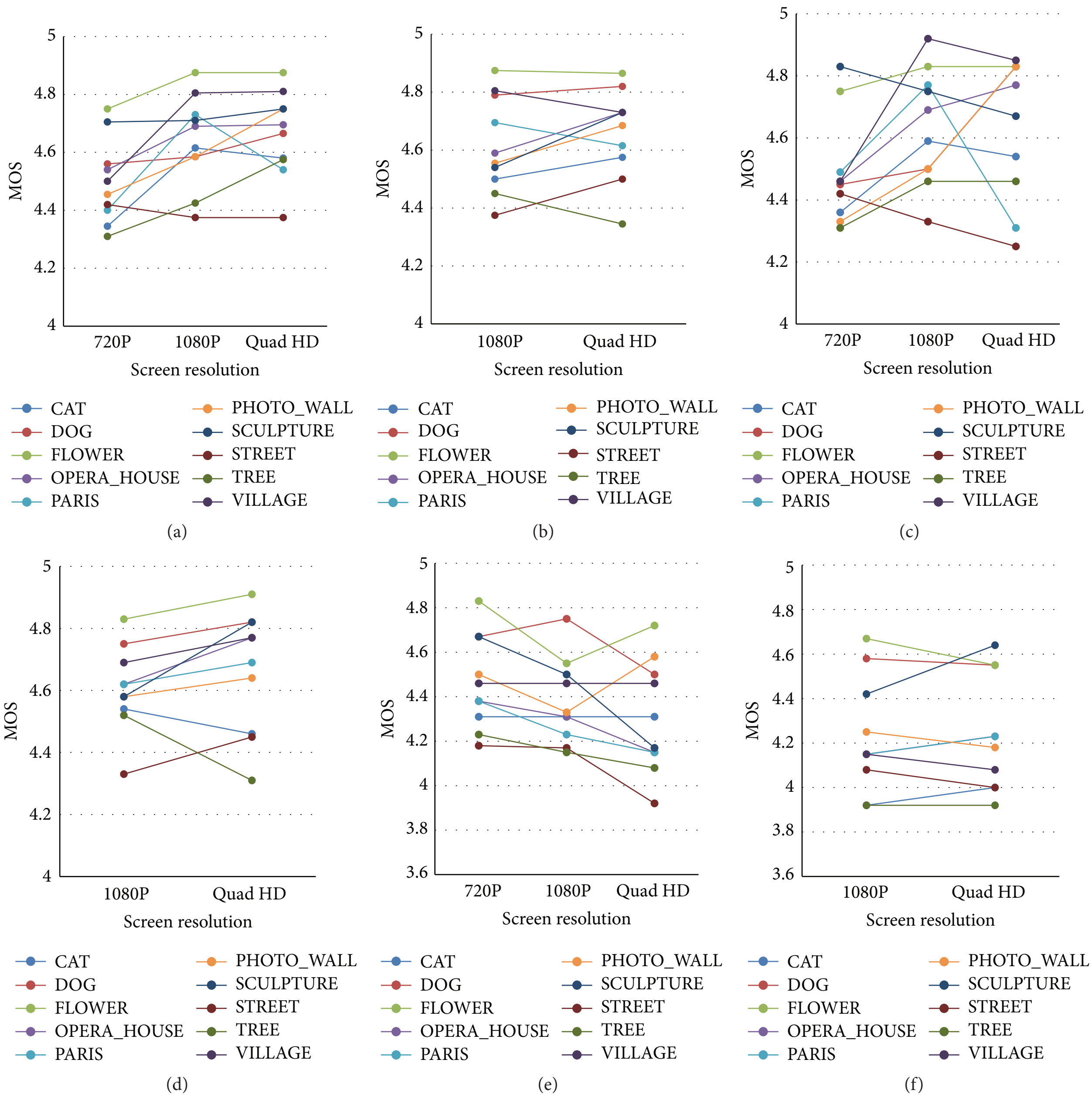


Figure 7: MOS for ten 1440P, 1080P, and 720P high quality images displayed on 720P, 1080P, and 2K screen, respectively. (a) 1440P images displayed on P4, P5, and P6. (b) 1440P images displayed on P7 and P8. (c) 1080P images displayed on P4, P5, and P6. (d) 1080P images displayed on P7 and P8. (e) 720P images displayed on P4, P5, and P6. (f) 720P images displayed on P7 and P8.

between watching the 720P and 1080P screen, let alone the Quad HD screen.

4.2. Statistical Significance Analysis. A hypothesis testing is further conducted to verify whether the improvement of perceived image quality is statistically significant on a higher resolution screen, where the MOS rated on two specific screens with different resolutions (e.g., P4 and P5, P5 and P6) for the same images in Experiment I are employed. The assumption of normality of the ΔPQs on two screens is checked firstly by the Kolmogorov-Smirnov (K-S) test [33]. In the analysis, we find that all the null hypothesis (i.e., the ΔPQs have a normal distribution) cannot be rejected at the 5% level and hence our assumption of normality is valid for each set of ΔPQs. Then, a paired samples t-test [34] is performed to assess whether the mean of ΔPQs is statistically different with zero. The test results are given in Table 8 for three pairs of devices with the same screen size. It illustrates that in most cases the difference of perceived image quality on the 720P and 1080P screen (i.e., P4 versus P5) is statistically

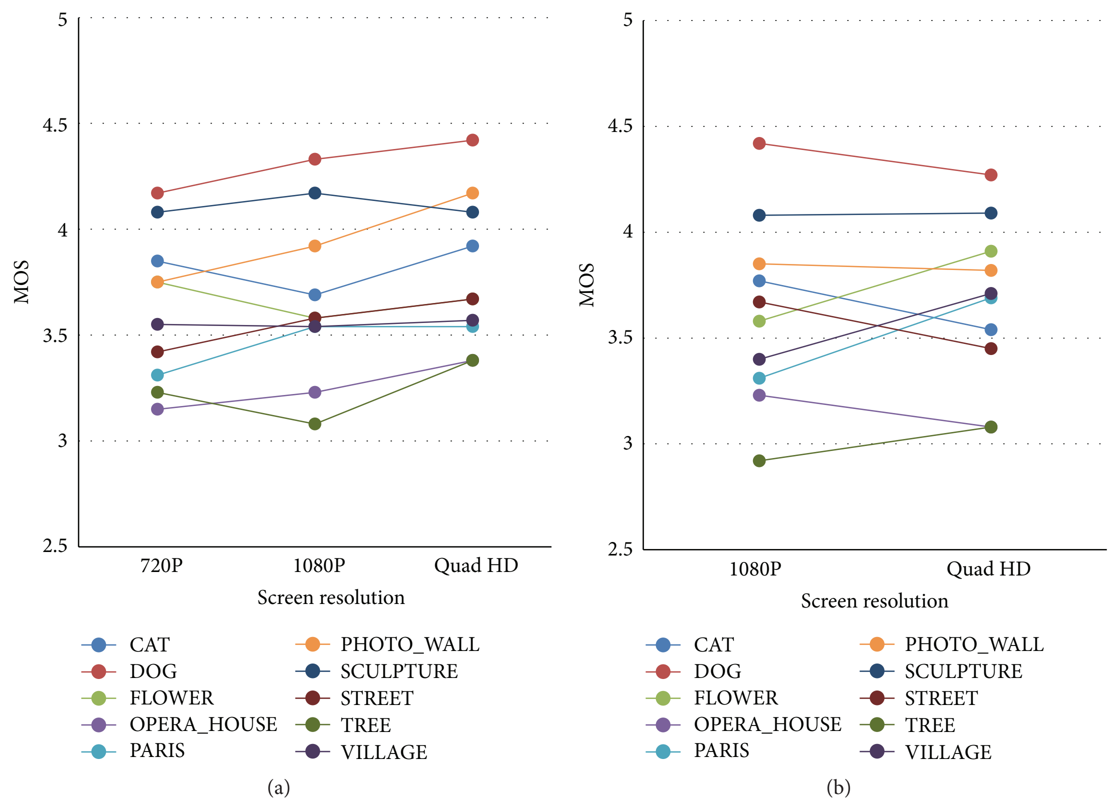


FIGURE 8: MOS for ten 480P high quality images displayed on 720P, 1080P, and 2K screen. (a) On P4, P5, and P6 with 5.1-inch screen. (b) On P7 and P8 with 5.5-inch screen.

TABLE 8: Results of paired samples t-test based on MOS rated on different screens under 5 image resolutions.

Image quality level	Image resolution	Significance		
		P4, P5	P5, P6	P7, P8
High (Q = 75)	4K	1	0	0
	1440P	1	0	0
	1080P	1	0	0
	720P	1	0	0
	480P	0	0	0
Medium (Q = 15)	4K	1	0	0
	1440P	1	0	0
	1080P	1	0	0
	720P	1	0	0
	480P	0	0	0
Low (Q = 5)	4K	0	0	0
	1440P	0	0	0
	1080P	0	0	0
	720P	0	0	0
	480P	0	0	0

*1 means that the difference in definition is statistically significant. 0 means that the difference in definition is not statistically significant.

different. However, there is no statistical difference between the perceived image quality on the 1080P and Quad HD screen (i.e., P5 versus P6 and P7 versus P8) in any cases.

In conclusion, the experimental results indicate that it will make a considerably meaningful improvement on the user's perceived image quality for high and medium quality image when increasing the resolution of 5.1-inch screen from 720P to 1080P. However, increasing the screen resolution from 1080P to Quad HD on 5.1- or 5.5-inch screen is not useful for improvement of the user's perceived image quality.

5. Modeling the Perceived Image Quality on Mobile Phones

In this section, four impact factors, that is, image resolution, screen resolution, screen size, and image coding quality, are investigated to establish an objective quality assessment model. The impact of the screen size and resolution on user's perceived image quality has been discussed in Sections 3 and 4. Here, the influence of image itself will be checked at first. Then, the mutual interaction of these four impact factor is evaluated.

5.1. Perceived Image Quality Assessment Model for High Quality Images. As the basic quality of the perceived image quality, the image coding quality and its characteristic have been discussed in Section 2. Another significant impact factor of image quality is the image resolution. The influence of image resolution on the perceived image quality is investigated for the high quality images (coded under $Q = 75$) at first. The MOS on three mobile phones, that is, P4 (5.1-inch 720P

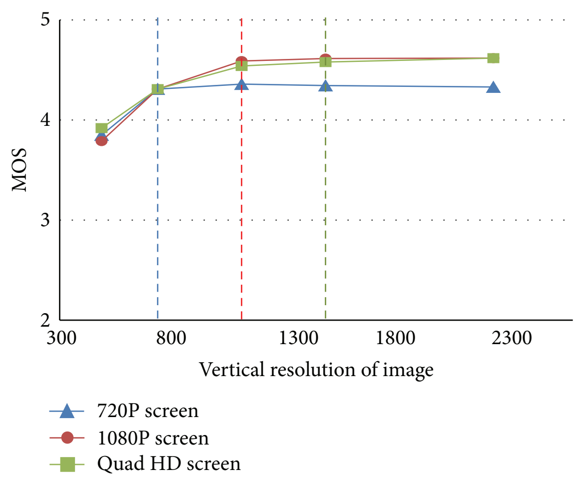


FIGURE 9: Relationship between the vertical resolution of image and MOS for "CAT" high quality images displayed on P4–P6.

screen), P5 (5.1-inch 1080P screen), and P6 (5.1-inch Quad HD screen), are selected to check the relationship between the image resolution and the MOS. The interactive impact of image and screen resolutions can be investigated on the way. Figure 9 takes this relationship for the "CAT" images as an example. For the high quality images displayed on the 720P screen, the values of MOS will keep increasing with the increment of the image resolution when the image resolution is smaller than 720P. However, when the image resolution is larger than that of screen (the blue vertical dashed line), the value of MOS will remain nearly constant. Similarly, when displaying these images on the 1080P screen, the value of MOS will not increase when the image resolution is larger than 1080P. Likewise, for the images displayed on the Quad HD screen, the values of MOS also keep nearly constant after the image resolution is larger than the screen resolution, that is, 1440P. Consequently, a preferable perceived image quality can be provided by improving the image resolution, but this improvement of perceived image quality will be limited by the screen resolution. It is noteworthy that this trend of user's perceived image quality conforms to the famous Weber-Fechner Law [35] and the corresponding logarithmic/negative-exponential behavior in QoE discussed in [36, 37]. Specifically, QoE ~ $\log(R)$ is the solution of the differential equation $\mathrm{d}\text{QoE}/\mathrm{d}R \sim 1/R$ where R is the image resolution. It means that the more the resources (i.e., higher image resolutions in this paper) that are present, the less the increment of the QoE becomes. Moreover, the trend of curves of the image on 1080P and Quad HD screen almost overlaps in Figure 9, which conforms to the conclusion discussed in Section 4.

To reflect the limitation caused by the screen resolution on the perceived image quality, an integrated assessment parameter, that is, the density of the effective image pixels per inch displayed on the screen (ED-PPI), is proposed. The effective pixels do not include the pixels that interpolated by the upsample or lost by the downsample. When both the horizontal and vertical resolutions of the image are less than or equal to that of the screen, which means that the effective image pixels on the screen is nonsaturated, the ED-PPI will increase with the image resolution until the image resolution is larger than the screen resolution. However, when one of the horizontal or vertical resolution of the image is larger than that of the screen, which means that the effective image pixels on the screen is saturated, the ED-PPI will be equal to the physical pixel per inch (PPI) on the screen and not increase with the image resolution. Hence, the ED-PPI conforms to the trend of the perceived image quality in Figure 9.

Figure 10 illustrates the condition that the effective image pixels is nonsaturated. H_I, V_I, H_S, and V_S are the horizontal and vertical resolutions of image and screen, respectively. L_{SW} and L_{SH} are the length along the width and height of the screen in inch, respectively. When the aspect ratio of the screen resolution is larger than that of the image, two lateral mattes (the dark grey areas in Figure 10(a)) will be added to the top and bottom of the image to fill up the screen. Conversely, the mattes (the dark grey areas in Figure 10(b)) will be added to the left and the right of the image. In this case, that is, $H_I \le H_S$ and $V_I \le V_S$, the ED-PPI is expressed as follows:

$$\text{ED-PPI} = \begin{cases} \dfrac{H_I}{L_{SW}}, & \dfrac{H_S}{V_S} \le \dfrac{H_I}{V_I}, \\ \dfrac{V_I}{L_{SH}}, & \dfrac{H_S}{V_S} > \dfrac{H_I}{V_I}. \end{cases} \quad (2)$$

However, when the effective image pixels are saturated, that is, $H_I > H_S$ or $V_I > V_S$, the image needs to be downsampled to fit the screen resolution. Each physical pixel in the actually used area is corresponding to an effective pixel on the image. Hence, the ED-PPI will be equal to the physical PPI on the screen and be expressed as follows:

$$\text{ED-PPI} = \frac{\sqrt{H_S^2 + V_S^2}}{L_{\text{diag}S}}, \quad (3)$$

where $L_{\text{diag}S}$ represents the diagonal length of screen in inch, which can be calculated as follows:

$$V_{\text{diag}S} = \sqrt{L_{SW}^2 + L_{SH}^2}. \quad (4)$$

By using the proposed ED-PPI, the trends of perceived image quality on different screens in Figure 11, especially for the 720P screen, can be unified by the relationship between the ED-PPI and MOS. The increasing trend of MOS becomes slow with the increase of ED-PPI, which reveals the restriction from the ED-PPI on the perceived image quality. Simultaneously, the trend in Figure 11 also conforms to the Weber-Fechner Law mentioned above. This law can also be observed on other high quality images.

Then, the impact of ED-PPI is normalized to represent the variation trend of MOS. The perceived image quality of high quality images on mobile phones can then be estimated by

$$\begin{aligned} \text{PQ}_{\text{img}} &= f(\text{ICQ}, \text{ED-PPI}) \\ &= \text{ICQ} \cdot \left(v_1 + v_2 \cdot \ln(\text{ED-PPI})\right), \end{aligned} \quad (5)$$

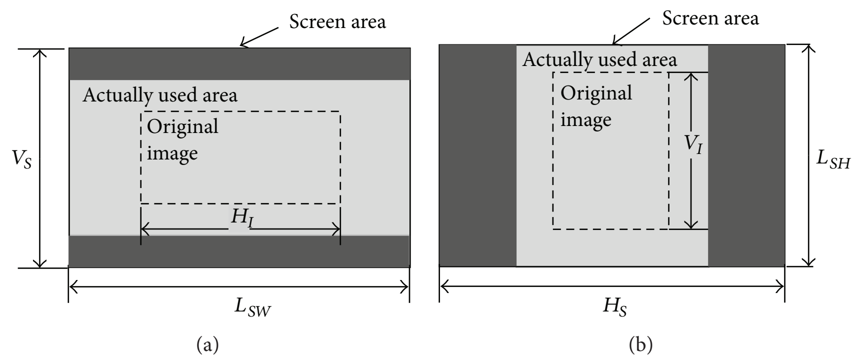


FIGURE 10: Illustration of parameters utilized to calculate ED-PPI, when the effective image pixels are nonsaturated. (a) The aspect ratio of screen resolution is larger than the image resolution. (b) The aspect ratio of screen resolution is smaller than the image resolution.

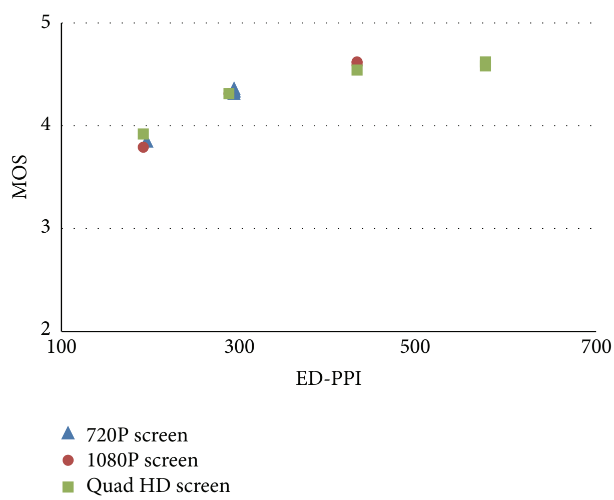


FIGURE 11: Relationship between ED-PPI and MOS for "CAT" in high quality displayed on P4–P6.

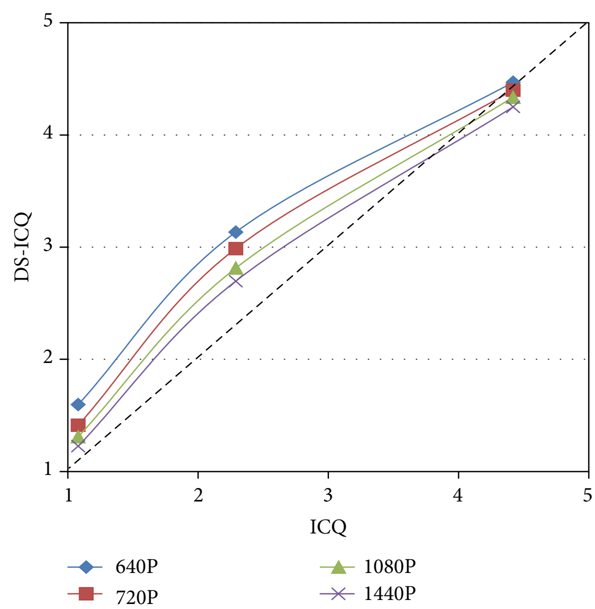


FIGURE 12: Relationship between DS-ICQ and ICQ for "CAT."

where PQ_{img} is the perceived image quality on the specific mobile phone. ICQ is the image coding quality. v_1, v_2 are two model coefficients that can be obtained by regression.

To avoid the case that PQ_{img} exceeds the minimum and maximum value, that is, 1 and 5 points, PQ_{img} is mended as

$$PQ_{img} = \min\left(\max\left(PQ_{img}, 1\right), 5\right). \quad (6)$$

5.2. Modification and Complement of the Assessment Model for Low and Medium Quality Images. For the low and medium quality images, the relationship between MOS and the image resolution is the same as that of the high quality images when the effective image pixels are nonsaturated. However, when the effective image pixels are saturated, the perceived image quality still rises while not keeping constant like the perceived image quality for high quality images. In such cases, the images need to be downsampled by the mobile phone to fit the screen. This phenomenon is caused by the downsampling process, which decreases the distortion region in the medium and low quality images and increases the image coding quality, correspondingly.

The rated scores of the downsampled images in Database II and their corresponding original images in Database I on M1 were utilized to investigate the mapping relation between the image coding quality of the downsampled images (denoted as DS-ICQ for the convenience) and the ICQ of original images. Figure 12 illustrates the relationship between DS-ICQ and ICQ for "CAT," where the different types of lines indicates four different resolution images downsampled from 4K images, respectively. It can be seen that the DS-ICQ is higher than the ICQ, especially for the low and medium quality images. Hence, it is necessary to map the ICQ into DS-ICQ, when the image resolution is higher than the screen resolution.

The trend of these curves in Figure 12 can be uniformly expressed as follows:

$$\text{DS-ICQ} = v_3 \cdot \text{ICQ}^{v_4}, \quad (7)$$

where v_3 and v_4 are experimental parameters.

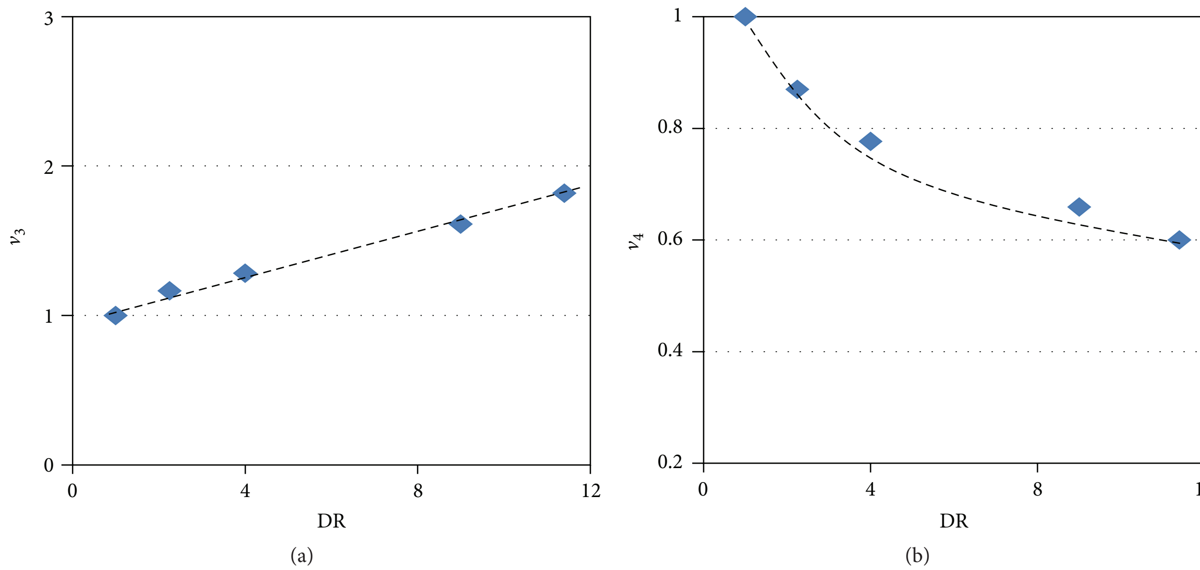


Figure 13: Relationship between DR and v_3, v_4. (a) DR and v_3. (b) DR and v_4.

Table 9: Parameters v_3, v_4, and DR.

Downsampled image resolution	Downsampling ratio (DR)	v_3	v_4
640P	11.4	1.82	2.77
720P	9	1.61	2.36
1080P	3	1.28	1.51
1440P	2.25	1.17	1.21

The parameters v_3 and v_4 adapted to different curves in Figure 12 are presented in Table 9. For convenience, the downsampling ratio (DR) is defined as the ratio of the resolution of the downsampled image and original image. It is clear that the values of v_3 and v_4 vary for different DRs. Figure 13 gives the relationship between DR and v_3 and v_4, respectively. It can be found that the values of v_3 is linearly related to DR, and their relationship can be expressed as follows:

$$v_3 = v_5 \cdot (\mathrm{DR} - 1) + 1, \tag{8}$$

where v_5 is model parameter which can be obtained by regression.

Simultaneously, the values of v_4 decrease with DR, and the relationship can be expressed as follows:

$$v_4 = \exp\left(-v_6 \cdot (\mathrm{DR} - 1)\right), \tag{9}$$

where v_6 is model parameter which can be obtained by regression. Submitting (8) and (9) into (7) and considering the range of image quality, DS-ICQ can be achieved by

$$\begin{aligned} &\text{DS-ICQ} \\ &\quad = \min\left(\left(v_5 \cdot (\mathrm{DR} - 1) + 1\right) \cdot \mathrm{ICQ}^{\exp(-v_6(\mathrm{DR}-1))}, 5\right). \end{aligned} \tag{10}$$

It should be noted that the mapping relationship also fits other images. All of the model parameters in (10), that is, v_5 and v_6, will be regressed by all the images in Database II. In summary, when one of the horizontal or vertical resolutions of the image is larger than that of the screen, the ICQ is firstly mapped into DS-ICQ as the coding quality of the high resolution images, and the DS-ICQ will be used as the input of (5) instead of the ICQ, as

$$\begin{aligned} \mathrm{PQ}_{\mathrm{img}} &= f\,(\text{DS-ICQ}, \text{ED-PPI}) \\ &= \text{DS-ICQ} \cdot \left(v_1 + v_2 \cdot \ln\left(\text{ED-PPI}\right)\right). \end{aligned} \tag{11}$$

Particularly, when the value of DR equals 1, which means that there is no downsample process, the DS-ICQ equals ICQ.

5.3. Performance Evaluation. The performance of proposed model was validated by applying a twofold cross-validation based on the results of Experiment I. Half of the images were randomly selected to form the dataset D1 (i.e., "CAT," "DOG," "FLOWER," "OPERA_HOUSE," and "PARIS" image), and the rest of the images were assigned to dataset D2. We then trained our model parameters based on the rating results of D1 and validated the performance based on the rating results for D2 and then exchanged the training set and validating set to repeat the process above. Specifically, the results of nonsaturated part of effective image pixels and the results for the high quality images of saturated part in the subset were used to conduct a three-dimensional curve fitting on function (5) by using the "sftool" in MATLAB 2014b. The trained parameters in (5) were the same as those in function (11).

To the best of our knowledge, there is no literature aiming at further estimating the perceived image quality for specific mobile device where the image coding quality has been evaluated. Hence, we can only benchmark the performance of the proposed method with the image coding quality.

Three commonly used performance criteria were employed to measure the performance of the proposed

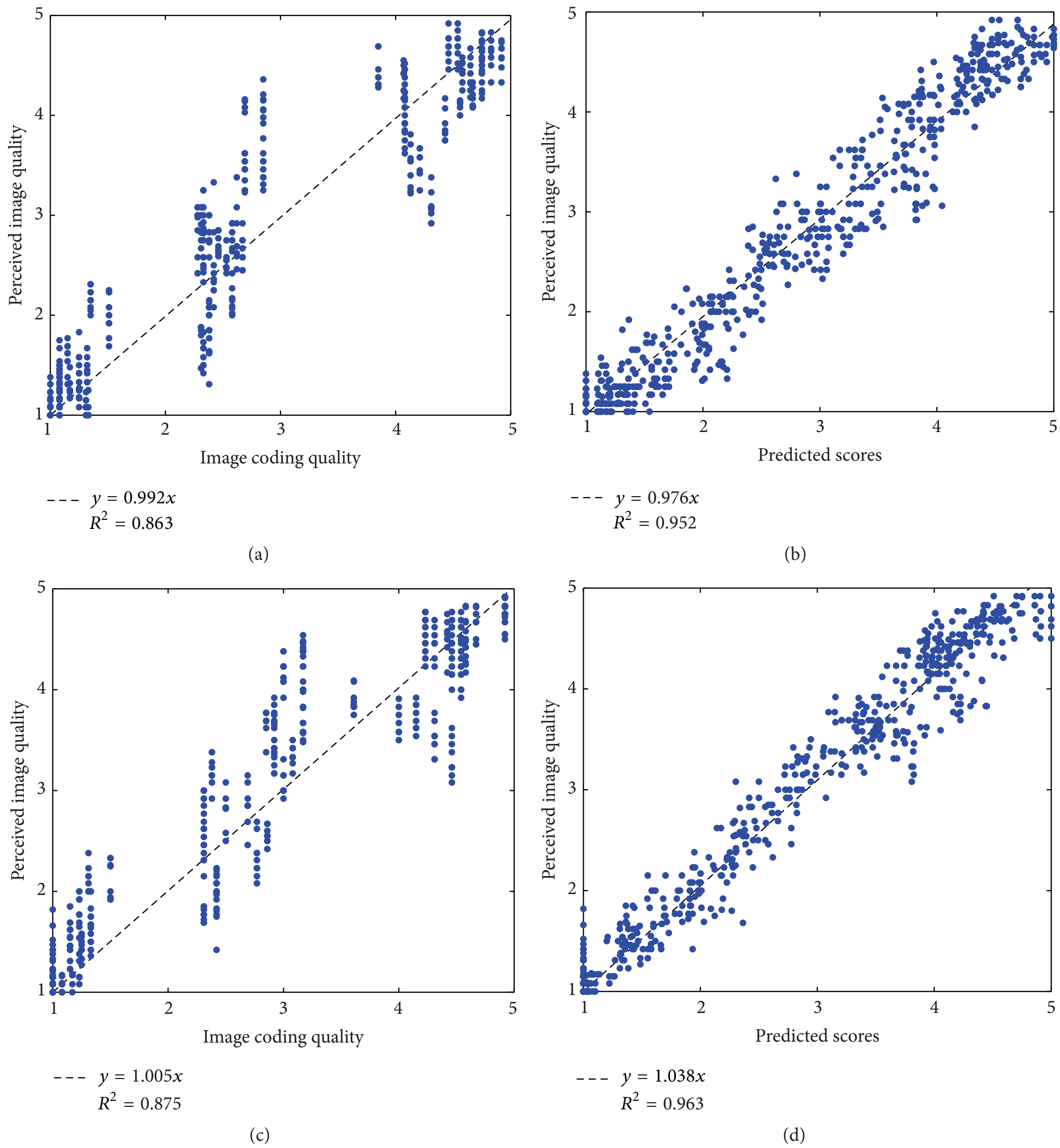


Figure 14: Scatter plots of the perceived image quality (PIQ) versus the ICQ and PQ_{img}. (a) ICQ versus PIQ on D2. (b) PQ_{img} versus PIQ on D2. The model is trained on D1. (c) ICQ versus PIQ on D1. (d) PQ_{img} versus PIQ on D1. The model is trained on D2.

model, namely, the Pearson correlation coefficient (PCC), Root-Mean-Squared Error (RMSE), and the Spearman Rank Order Correlation Coefficient (SROCC). A summary of performance evaluation in terms of these metrics is listed in Table 10. It can be found that an outstanding and reliable prediction performance is obtained when using the proposed quality assessment model in the cross-validation.

To clarify the performance, the scatter plots of the image coding quality, the perceived image quality, and the predicted scores are shown in Figure 14, respectively. They are fitted by the same regression formula, and the corresponding R^2 are also listed in Figure 14. It is observed that the predicted results are much closer to the actual perceived image quality on mobile phones than the conventionally used image coding quality.

6. Conclusion

In this paper, a wide range of popular mobile phones are selected in the subjective experiments to investigate the impact of image resolution, screen size, and screen resolution on user's perceived image quality. The quantitative and

Table 10: Performance comparison of proposed image quality assessment model to the conventionally used original image coded quality.

Training	Validation	Comparison	PCC	RMSE	SROCC
D1	D2	ICQ	0.936	0.480	0.923
		PQ_{img}	**0.977**	**0.296**	**0.970**
D2	D1	ICQ	0.941	0.490	0.939
		PQ_{img}	**0.981**	**0.284**	**0.971**

statistical analyses are conducted to check whether the increase of the screen size and resolution will lead to improvement in user's perceived image quality. The finding is useful for the mobile phone industry to have a better understanding of the concrete benefit of enhancing the screen resolution. Furthermore, a device-dependent image quality assessment model is proposed to evaluate the perceived image quality on different mobile phones. The proposed quality assessment model is useful for image quality assessment on specific mobile phones.

Competing Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (no. 61371089, no. 61571337) and the 111 Project B08038.

References

[1] K. J. Kim and S. S. Sundar, "Does screen size matter for smartphones? Utilitarian and hedonic effects of screen size on smartphone adoption," *Cyberpsychology, Behavior, and Social Networking*, vol. 17, no. 7, pp. 466–473, 2014.

[2] "A comprehensive look at smartphone screen size statistics and trends," May 2014, https://medium.com/@somospostpc/a-comprehensive-look-at-smartphone-screen-size-statistics-and-trends-e61d77001ebe#.9rehfvbut.

[3] F. Porikli, A. Bovik, C. Plack et al., "Multimedia quality assessment," *IEEE Signal Processing Magazine*, vol. 28, no. 6, pp. 164–177, 2011.

[4] M. Lin, D. Chenwei, K. N. Ngan, and L. Weisi, "Recent advances and challenges of visual signal quality assessment," *China Communications*, vol. 10, no. 5, Article ID 6520939, pp. 62–78, 2013.

[5] F. Yang and S. Wan, "Bitstream-based quality assessment for networked video: a review," *IEEE Communications Magazine*, vol. 50, no. 11, pp. 203–209, 2012.

[6] T.-J. Liu, Y.-C. Lin, W. Lin, and C.-C. J. Kuo, "Visual quality assessment: recent developments, coding applications and future trends," *APSIPA Transactions on Signal and Information Processing*, vol. 2, article e4, 2013.

[7] A. K. Moorthy, L. K. Choi, A. C. Bovik, and G. de Veciana, "Video quality assessment on mobile devices: subjective, behavioral and objective studies," *IEEE Journal on Selected Topics in Signal Processing*, vol. 6, no. 6, pp. 652–671, 2012.

[8] A. Chan, A. Pande, E. Baik, and P. Mohapatra, "Temporal quality assessment for mobile videos," in *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking (MobiCom '12)*, pp. 221–232, ACM, Istanbul, Turkey, August 2012.

[9] A. K. Moorthy, L. K. Choi, G. De Veciana, and A. C. Bovik, "Subjective analysis of video quality on mobile devices," in *Proceedings of the 6th International Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM '12)*, Scottsdale, Ariz, USA, January 2012.

[10] C. Chen, L. Song, X. Wang, and M. Guo, "No-reference video quality assessment on mobile devices," in *Proceedings of the 8th IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB '13)*, pp. 1–6, London, UK, June 2013.

[11] R. Gong and H. Xu, "Impacts of appearance parameters on perceived image quality for mobile-phone displays," *Optik*, vol. 125, no. 11, pp. 2554–2559, 2014.

[12] T. Kallio and A. Kaikkonen, "Usability testing of mobile applications: a comparison between laboratory and field testing," *Journal of Usability Studies*, vol. 1, pp. 4–16, 2005.

[13] S. Jumisko-Pyykkö and M. M. Hannuksela, "Does context matter in quality evaluation of mobile television?" in *Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services*, pp. 63–72, ACM, September 2008.

[14] T. De Pessemier, K. De Moor, A. J. Verdejo et al., "Exploring the acceptability of the audiovisual quality for a mobile video session based on objectively measured parameters," in *Proceedings of the 3rd International Workshop on Quality of Multimedia Experience (QoMEX '11)*, pp. 125–130, Mechelen, Belgium, September 2011.

[15] A. Catellier, M. Pinson, W. Ingram, and A. Webster, "Impact of mobile devices and usage location on perceived multimedia quality," in *Proceedings of the 4th International Workshop on Quality of Multimedia Experience (QoMEX '12)*, pp. 39–44, IEEE, Melbourne, Australia, July 2012.

[16] T. De Pessemier, K. De Moor, W. Joseph, L. De Marez, and L. Martens, "Quantifying subjective quality evaluations for mobile video watching in a semi-living lab context," *IEEE Transactions on Broadcasting*, vol. 58, no. 4, pp. 580–589, 2012.

[17] G. Cermak, M. Pinson, and S. Wolf, "The relationship among video quality, screen resolution, and bit rate," *IEEE Transactions on Broadcasting*, vol. 57, no. 2, pp. 258–262, 2011.

[18] F. Agboma and A. Liotta, "Addressing user expectations in mobile content delivery," *Mobile Information Systems*, vol. 3, no. 3-4, pp. 153–164, 2007.

[19] J. P. López, M. Slanina, L. Arnaiz, and J. M. Menéndez, "Subjective quality assessment in scalable video for measuring impact over device adaptation," in *Proceedings of the IEEE EUROCON*, pp. 162–169, IEEE, Zagreb, Croatia, July 2013.

[20] F. Agboma and A. Liotta, "User centric assessment of mobile contents delivery," in *Proceedings of the 4th International Conference on Advances in Mobile Computing and Multimedia (MoMM '06)*, pp. 121–130, December 2006.

[21] A. Khan, L. Sun, J.-O. Fajardo, I. Taboada, F. Liberal, and E. Ifeachor, "Impact of end devices on subjective video quality assessment for QCIF video sequences," in *Proceedings of the 3rd International Workshop on Quality of Multimedia Experience (QoMEX '11)*, pp. 177–182, Mechelen, Belgium, September 2011.

[22] W. Song and D. W. Tjondronegoro, "Acceptability-based QoE models for mobile video," *IEEE Transactions on Multimedia*, vol. 16, no. 3, pp. 738–750, 2014.

[23] A. Rehman, K. Zeng, and Z. Wang, "Display device-adapted video quality-of-experience assessment," in *Human Vision and Electronic Imaging XX*, vol. 939406 of *Proceedings of SPIE*, International Society for Optics and Photonics, San Francisco, Calif, USA, March 2015.

[24] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, "LIVE Image Quality Assessment Database Release 2," 2005, http://live.ece.utexas.edu/research/quality/.

[25] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti, "TID2008—a database for evaluation of full-reference visual quality assessment metrics," *Advances of Modern Radioelectronics*, vol. 10, no. 10, pp. 30–45, 2009.

[26] E. C. Larson and D. M. Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy," *Journal of Electronic Imaging*, vol. 19, no. 1, Article ID 011006, 2010.

[27] J. Y. Lin, R. Song, C.-H. Wu, T. Liu, H. Wang, and C.-C. J. Kuo, "MCL-V: a streaming video quality assessment database," *Journal of Visual Communication and Image Representation*, vol. 30, pp. 1–9, 2015.

[28] N. Ponomarenko, O. Ieremeiev, V. Lukin et al., "Color image database TID2013: peculiarities and preliminary results," in *Proceedings of the 4th European Workshop on Visual Information Processing (EUVIP '13)*, pp. 106–111, IEEE, Paris, France, June 2013.

[29] FFmpeg, 2013, http://sourceforge.net/projects/ffmpeg.

[30] ITU, "Methodology for the subjective assessment of the quality of television pictures," ITU-R Recommendation BT 500-13, ITU Radiocommunication Sector, 2012.

[31] ITU-T, "Methods for the subjective assessment of video quality, audio quality and audiovisual quality of Internet video and distribution quality television in any environment," Recommendation P.913, ITU-T, 2014.

[32] H. R. Sheikh, Z. Wang, A. C. Bovik, and L. K. Cormack, "Image and video quality assessment research at live," 2003, http://live.ece.utexas.edu/research/quality/.

[33] Kolmogorov-Smirnov Test, http://www.physics.csbsju.edu/stats/KS-test.html.

[34] D. C. Montgomery and G. C. Runger, *Applied Statistics and Probability for Engineers*, John Wiley & Sons, 2010.

[35] P. Reichl, S. Egger, R. Schatz, and A. D'Alconzo, "The logarithmic nature of QoE and the role of the Weber-Fechner law in QoE assessment," in *Proceedings of the IEEE International Conference on Communications (ICC '10)*, pp. 1–5, Cape Town, South Africa, May 2010.

[36] M. Fiedler and T. Hossfeld, "Quality of experience-related differential equations and provisioning-delivery hysteresis," in *Proceedings of the 21st ITC Specialists Seminar on Multimedia Applications—Traffic, Performance and QoE*, Miyazaki, Japan, March 2010.

[37] T. Zinner, T. Hoßfeld, T. N. Minhas, and M. Fiedler, "Controlled vs. uncontrolled degradations of QoE—the provisioning-delivery hysteresis in case of video," in *Proceedings of the EuroITV Workshop: Quality of Experience for Multimedia Content Sharing*, Tampere, Finland, June 2010.