Contrast-sensitivity-based evaluation method of a surveillance camera’s visual resolution: improvement from the conventional slanted-edge spatial frequency response method

Zong Qin, Po-Jung Wong, Wei-Chung Chao, Fang-Cheng Lin, Yi-Pai Huang, and Han-Ping D. Shieh

Department of Photonics and Display Institute, National Chiao Tung University, Hsinchu 300, Taiwan
*Corresponding author: hpshieh@mail.nctu.edu.tw

Received 14 October 2016; revised 30 December 2016; accepted 16 January 2017; posted 17 January 2017 (Doc. ID 278518); published 9 February 2017

Visual resolution is an important specification of a surveillance camera, and it is usually quantified by linewidths per picture height (LW/PH). The conventional evaluation method adopts the slanted-edge spatial frequency response (e-SFR) and uses a fixed decision contrast ratio to determine LW/PH. However, this method brings about a considerable error with respect to subjectively judged results because the perceptibility of the human vision system (HVS) varies with spatial frequency. Therefore, in this paper, a systematic calculation method, which combines the contrast sensitivity function characterizing the HVS and e-SFR, is proposed to solve LW/PH.

Eight 720P camera modules in day mode, four 720P modules in night mode, and two 1080P modules in day mode are actually adopted. Corresponding to the three modes, mean absolute error between objective and subjective LW/PH are suppressed to as low as 26 (3.6% of 720P), 27 (3.8% of 720P), and 49 (4.5% of 1080P), while those of the conventional method are 68 (9.4% of 720P), 95 (13.2% of 720P), and 118 (10.9% of 1080P).

© 2017 Optical Society of America

OCIS codes: (040.1490) Cameras; (110.4100) Modulation transfer function; (330.5000) Vision - patterns and recognition; (330.1070) Vision - acuity; (110.3000) Image quality assessment; (330.6130) Spatial resolution.

https://doi.org/10.1364/AO.56.001464

1. INTRODUCTION

A camera module is an essential component in surveillance systems widely used in security, traffic, industry, etc. [1–4]. Unlike cameras used for photography, surveillance cameras directly output images on a monitor for users to watch, hence they usually do not have high resolutions (720P and 1080P are common specifications). To guarantee high-quality images provided to the users, the concept of visual resolution is introduced, which is defined in the ISO-12233 standard [5,6] as “spatial frequency at which all of the individual black and white lines of a test pattern frequency can no longer be distinguished by a human observer.” For surveillance cameras that produce images for visual observation, visual resolution is of paramount importance and is commonly quantified by linewidths per picture height (LW/PH), which is also known as television lines (TVL). Because imaging quality of a camera usually substantially differs in sagittal and tangential dimensions, specifications of horizontal (sagittal) and vertical (tangential) LW/PH are given separately. Figure 1 schematically shows the framework in which LW/PH of a surveillance camera is subjectively judged by a human inspector. In such a framework, a monitor shows a captured test chart, which is usually an ISO-12233 test chart containing horizontal and vertical line-pair patterns simultaneously [5,6], and then the maximum number of alternating bright and dark lines per picture height that can be humanly resolved on the monitor is determined to be the LW/PH index for either direction. Visual resolution in terms of LW/PH is the final specification requirement from customers; on the other hand, imaging quality of a surveillance camera module may be affected by lens quality, module assembly, built-in image processing, etc. Therefore, quality inspection is necessary to obtain the specification of visual resolution, with the aid of which, manufacturers can check product quality before delivery, optimize imaging quality, and so on.

Visual resolution is a subjective metric that needs human inspectors; however, corresponding costs of human resource
and inspection time are undesired in mass production. To judge visual resolution objectively, a number of studies have been dedicated to automated evaluation methods. Currently, the most popular method is the slanted-edge spatial frequency response (e-SFR) method [4–16], which has been introduced in ISO standards [5,6] and widely adopted in industry. This method adopts slanted edges in the ISO-12233 chart, as shown in Fig. 2, to calculate SFR. In this method, a spatial frequency (in cycle/pixel), at which SFR in terms of contrast ratio descends to a fixed decision value, is found, then this spatial frequency is converted to LW/PH value according to the vertical pixel number of the camera’s sensor, as LW/PH = spatial frequency × pixel number/0.0136. Conventionally, fixed decision SFR of 10% has been recommended in some literature, including the ISO standards [5–7,13,14], and is also widely used in industry.

LW/PH evaluated by this e-SFR method along with a fixed decision SFR is a useful indicator of a surveillance camera’s visual resolution [5–7,14]. However, as reported by some manufacturers, a considerable error still exists between objective LW/PH calculated by an automated program (LW/PH_program) and that subjectively judged by inspectors (LW/PH_human). For a kind of 720P (vertical pixel number = 720) camera product, mean absolute error (MAE) was reported to be nearly 100, which will be demonstrated in Section 3.A. As LW/PH is the final specification required by customers, the conventional method that can only predict visual resolution of the order of numbers is insufficient, and a new objective method that can directly predict visual resolution matched well with subjective results is desired.

In fact, the perceptibility of the human vision system (HVS) varies with spatial frequency [17–19], whereas the conventional e-SFR method adopts a fixed decision SFR for camera modules with different visual resolutions. Figures 3(a) and 3(b) schematically show the difference between the conventional method and the physical truth. It can be clearly seen that finding the intersection between the SFR curve and a straight line representing the fixed decision SFR inevitably brings about some errors because a curved line characterizes HVS in reality. Therefore, it is the fixed decision SFR where the error essentially comes from.

In this paper, with the purpose of suppressing the error of the conventional evaluation method of a surveillance camera’s visual resolution, we consider both HVS characterized by the contrast sensitivity function (CSF) and the e-SFR method. A systematic calculation method is introduced in detail to determine LW/PH value based on a slanted-edge pattern and inspection conditions actually used. Eight 720P camera modules working in day mode, four 720P modules working in night mode, and two 1080P modules working in day mode, each of which has seven different focusing statuses, are adopted to implement visual tests with 10 testees. Resultantly, MAEs between objective LW/PH_program and subjective LW/PH_human corresponding to 720P’s day mode, 720P’s night mode, and 1080P’s day mode are 26 (3.6% of 720P), 27 (3.8% of 720P), and 49 (4.5% of 1080P), while those of the conventional e-SFR method are 68 (9.4% of 720P), 95 (13.2% of 720P), and 118 (10.9% of 1080P). The proposed method suppresses the error to about one-third and a half for 720P and 1080P camera modules, respectively. In addition, the essential reason why the proposed method can achieve such better performance is discussed by comparing with the conventional method with different decision SFR.

2. EVALUATION METHOD

A. Calculation of SFR Based on a Slanted Edge

The calculation method of SFR based on a slanted-edge pattern has been developed for many years [4–16], hence here we follow the well-developed method, whose flow chart is shown in Fig. 4. At the beginning of the flow chart, it should be noticed that the optoelectronic conversion function is not applied when the image is input because produced electronic images are directly used for visual observation. Then, by taking a ROI of 50 pixels by 50 pixels stored as a JPEG image file as an example, the flow chart is explicated with the aid of Fig. 5. (1) Detect the slanted
edge in ROI and linearly fit the points on the edge to find out its slant angle, as shown in Fig. 5(a). (2) Convert gray levels in the image file to luminance levels with the display gamma value \(\gamma\), as luminance level = (gray level/255)^\gamma, where \(\gamma\) is 2.2 in our study. (3) Based on the slant angle, several raw edge spread functions (ESFs) are interpolated to obtain oversampled ESFs, as shown in Figs. 5(a) and 5(b), where 4 times oversampling is used for this example. Detailed approach of interpolation for oversampling can be found in [20]. (4) Use Fermi function to fit oversampled ESF to acquire a smooth ESF [21], as shown in Fig. 5(b). (5) Compute discrete derivative of a fitted ESF to acquire line spread function (LSF); here 2-point derivative is implemented in MATLAB. (6) Apply Hamming window function on LSF to compensate for the truncation error caused by the finite length of LSF, as shown in Fig. 5(c). (7) Compute discrete Fourier transform of compensated LSF to acquire uncorrected SFR. (8) Use cardinal sine function given in Eq. (1) to divide uncorrected SFR to correct for the error caused by discrete derivative [22–24]. (9) Output SFR within Nyquist frequency (0.5 cycle/pixel), as shown in Fig. 5(d).

The following discussions of the conventional and proposed methods are both strongly relegated to this flow chart of SFR calculation. To make our study reasonable and repeatable, a MATLAB P-file is provided in supplementary materials, which provides a function to calculate SFR from a slanted edge in an image file, as we show in Code 1, Ref. [25]:

\[
T(\omega) = \frac{\sin(\omega X)}{\omega X},
\]

(1)

where \(\omega\) is spatial frequency in cycle/pixel, \(N\) is oversampling factor in 1/pixel, and \(k\) is 1 (2-point derivative) or 2 (3-point derivative).

**B. HVS Characterized by CSF and Its Combination with e-SFR**

Perceptibility of HVS is usually characterized by the CSF, which provides a spatial-frequency-dependent contrast-sensitivity threshold (reciprocal of contrast ratio) at which HVS can just distinguish a sinusoidal luminance modulated grating from a uniform pattern [17–19]. The Barten model is the most commonly used CSF model for normal vision [19]. For vertical or horizontal grating patterns, its mathematical model is shown in Eq. (2). To reproduce the typical test scenario in which camera manufacturers evaluate a surveillance camera’s visual resolution, we adopt a 19 in. (22.86 cm) monitor with display luminance \(L\) of 200 cd/m² and surrounding luminance \(L_s\) of 5 cd/m². Viewing distance \(D\) is 80 cm, and then field of view (FOV) \(X \times X\) is 33.6° according to the monitor’s diagonal size and viewing distance. In this way, a specific CSF curve can be drawn over a range of spatial frequencies and contrast sensitivities, as shown in Fig. 6. Additionally, as verified in previous studies, the perception of a square grating is actually dominated by the perception of its fundamental harmonic, which is acquired by decomposing into Fourier series [26,27]. Equation (3) shows the Fourier series of a square grating where the contrast ratio of the fundamental harmonic is \(4/\pi\) times larger than that of the original grating. Correspondingly, CSF for a line-pair pattern is amplified by \(4/\pi\) times based on the CSF used for a sinusoidal grating, as also shown in Fig. 6. Equation (2) is

\[
S(\omega) = \frac{5200 \exp[-0.0016\omega^2(1 + 100/L)^{0.08}]}{\sqrt{1 + \frac{144}{X^2} + 0.64\omega^2}} \left[1 + \frac{1}{1 - \exp(-0.02\omega^2)}\right] f,
\]

(2)

\[
f = \exp\left\{-\frac{\ln^2\left[\frac{L}{X}(1 + \frac{144}{X^2})^{0.25}\right] - \ln^2\left[\frac{(1 + \frac{144}{X^2})^{0.25}}{2\ln 2^{32}}\right]}{2\ln 2^{32}}\right\},
\]

where \(S(\omega)\) is the spatial-frequency-dependent CSF, \(\omega\) is spatial frequency in cycle/degree, \(L\) is the pattern’s luminance in cd/m², \(X\) is FOV of the pattern’s diagonal in degrees, and \(f\) is a

---

**Fig. 4.** Flow chart of calculating SFR based on a slanted-edge pattern.

**Fig. 5.** Example of SFR calculation. (a) ROI with detected edge marked with a red line and four raw ESFs for interpolation marked with cyan lines. (b) Four times oversampled and fitted ESF. (c) LSF with window function applied. (d) Outputted SFR within Nyquist frequency.

**Fig. 6.** CSF curves corresponding to our test scenario, where the blue and orange lines denote CSF for sinusoidal and square gratings, respectively. The line-pair patterns schematize the variation of spatial frequency and contrast sensitivity.
factor between 0 and 1, which depicts the influence of the surrounding and contains another parameter of surrounding luminance \( L_s \) in cd/m\(^2\). Equation (3) is

\[
S(x) = X_0 + \sum_{n=1}^{\infty} \frac{4A}{(2n - 1)\pi} \cdot \sin \left( \frac{2\pi(2n - 1)x}{T} \right),
\]

where \( S(x) \) is a square wave pattern with period \( T \), mean value \( X_0 \), and peak-to-peak amplitude \( 2A \).

Finally, spatial frequency \( \omega \) in cycle/degree in CSF needs to be converted to \( \text{LW/PH} \). Here 1° FOV simply corresponds to a length of \( 2 \cdot D \cdot \tan 0.5° \) on the screen, where \( D \) is viewing distance, thus the length of each cycle is \( (2 \cdot D \cdot \tan 0.5°)/(\omega \cdot 1°) \) and \( \text{LW/PH} = (H \cdot \omega \cdot 1°)/(D \cdot \tan 0.5°) \), where \( H \) is monitor height. In addition, contrast sensitivity should be converted to its reciprocal, as contrast ratio. On the other hand, the corresponding spatial frequency in cycle/pixel in SFR also needs to be converted to \( \text{LW/PH} \) with the method introduced before, as \( \text{LW/PH} = \text{spatial frequency} \times \text{pixel number} \times 2 \). In this way, SFR and CSF curves having the same horizontal and vertical coordinates can be plotted in the same figure. Finally, the intersection of the two curves indicates the desired \( \text{LW/PH} \) value.

3. TESTS AND VERIFICATION

To verify the accuracy of the proposed method combining CSF and e-SFR, visual resolution tests need to be implemented to acquire subjective \( \text{LW/PH}_{\text{human}} \) with some testees (persons who are tested). On the other hand, the corresponding \( \text{LW/PH}_{\text{program}} \) needs to be objectively calculated.

A. Visual Resolution Tests

The visual resolution tests adopt an inspection machine consisting of an ISO-12233:2000 chart illuminated by a uniform white-light backlight, a mounting position for camera modules, a black housing, and a 19 in. (22.86 cm) monitor with an aspect ratio of 4:3 (monitor height \( H = 290 \text{ mm} \)). Figure 7(a) shows the monitor in the inspection machine while a captured test chart is provided to the testees. (The whole machine is not disclosed for commercial reasons.)

While implementing the visual resolution tests, the surveillance camera to be tested is mounted at the mounting position and then captures the test chart. The monitor shows the captured chart in full-screen mode and a mark is set 80 cm away from the monitor for the testees to place their eyes. Figure 7(b) shows an image captured by a 720P camera module where two hyperbolic wedges at central and peripheral fields are enlarged. At the central field, actual frequencies of the line-pair pattern certainly consist with the text labels annotated aside. Nevertheless, the peripheral line-pair pattern is severely distorted, and its actual frequencies are significantly changed, thus the text labels actually give wrong information. By considering that testees rely on these labels to tell \( \text{LW/PH} \) values, subjective judgment at peripheral fields makes little sense due to such an inconsistency. (Correction measures can be taken to make peripheral fields adapted to subjective judgment for surveillance cameras that usually have ultrawide FOV, though it is a different topic from our study.) Then only the central field is investigated in our study. By watching the hyperbolic wedges marked in Fig. 7(b), testees are asked to tell vertical and horizontal \( \text{LW/PH}_{\text{human}} \) they can just distinguish. For instance, vertical \( \text{LW/PH}_{\text{human}} \) is judged to be around 650 according to the enlarged wedge in Fig. 7(b). To suppress the influence of individual difference, 10 persons aged from 22 to 37, including six males and four females, are invited, and the median of 10 test results is determined to be \( \text{LW/PH}_{\text{human}} \).

On the other hand, captured images are stored, and ROI of slanted edges shown in Fig. 7(b) are inputted into a MATLAB program where SFR is first calculated and then the \( \text{LW/PH}_{\text{program}} \) is solved by finding the intersection between SFR and CSF curves. Here CSF corresponding to our tests has already been specialized from the Barten model, as the orange line in Fig. 6.

To verify the proposed method comprehensively, a series of commercial surveillance camera modules are provided by our manufacturer. Eight 720P camera modules working in day mode using visible light (noted as 720P-D) and four 720P modules working in night mode using infrared ray (noted as 720P-N) are adopted. Moreover, to investigate the effectiveness of the proposed method under different resolutions, two 1080P modules working in day mode (noted as 1080P-D) are also adopted. To generate camera modules with different imaging qualities, seven different focusing statuses are set for each module by using a stepper motor to slightly tune the distance between lens module and camera sensor. (This distance is in fact where the assembly error primarily comes from.) The seven focusing statuses include the sharp focusing one and six defocusing ones, as called -3, -2, -1, 0, +1, +2, and +3 statuses where “0” means the sharp one, as shown in Fig. 8. In addition,
vertical and horizontal LW/PH are both investigated for each module. By considering that the evaluation method does not differ in the vertical or horizontal direction, results of the two directions are blended for investigation. In this way, there are 112, 56, and 28 test samples corresponding to 720P-D, 720P-N, and 1080P-D, respectively.

B. Evaluation with the Conventional Method

First, the conventional e-SFR method with a decision SFR of 10% is adopted to calculate $LW/PH_{program}$. For the first 720P-D module, Figs. 9(a) and 9(b), respectively, show the solving process of the horizontal and vertical $LW/PH_{program}$. Calculation results and corresponding subjective $LW/PH_{human}$ are shown in Table 1, where significant errors are indicated. Positive and negative errors tend to occur at high and low spatial frequency, respectively. With the same method, the 112, 56, and 28 $LW/PH_{program}$ corresponding to 720P-D, 720P-N, and 1080P-D are calculated. Next, for these three modes, the relationships between calculated $LW/PH_{program}$ and corresponding subjective $LW/PH_{human}$ are plotted in Fig. 10. It can be seen that data points are quite scattered with respect to the line of "$y = x\)" which reveals that subjective and objective results do not match well. By referring to $LW/PH_{human}$, MAEs of 720P-D, 720P-N, 1080P-D are 68 (9.4% of 720P), 95 (13.2% of 720P), and 118 (10.9% of 1080P). As mentioned before, such an error of around 10% cannot satisfy camera manufacturers.

C. Evaluation with Proposed Method

The proposed method combining e-SFR and CSF is adopted with the expectation of improved performance of error. Similarly, $LW/PH_{program}$ is first solved for the first 720P-D module, as shown in Fig. 11. Moreover, $LW/PH_{program}$ and corresponding $LW/PH_{human}$ are shown in Table 2, where errors are much suppressed in comparison with the results in Table 1, as only two errors larger than 30 can be found.

![Table 1. LW/PH_{program} Calculated with the Conventional Method and Corresponding LW/PH_{human}^a](image)

<table>
<thead>
<tr>
<th>Direction</th>
<th>Focusing Status</th>
<th>$LW/PH_{human}$</th>
<th>$LW/PH_{program}$</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horizontal</td>
<td>−3</td>
<td>305</td>
<td>255</td>
<td>−50</td>
</tr>
<tr>
<td></td>
<td>−2</td>
<td>478</td>
<td>381</td>
<td>−97</td>
</tr>
<tr>
<td></td>
<td>−1</td>
<td>500</td>
<td>524</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>605</td>
<td>659</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>520</td>
<td>535</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>403</td>
<td>369</td>
<td>−34</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>318</td>
<td>271</td>
<td>−47</td>
</tr>
<tr>
<td></td>
<td>−3</td>
<td>310</td>
<td>261</td>
<td>−49</td>
</tr>
<tr>
<td></td>
<td>−2</td>
<td>347</td>
<td>295</td>
<td>−52</td>
</tr>
<tr>
<td></td>
<td>−1</td>
<td>525</td>
<td>504</td>
<td>−21</td>
</tr>
<tr>
<td>Vertical</td>
<td>0</td>
<td>575</td>
<td>609</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>473</td>
<td>430</td>
<td>−43</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>438</td>
<td>362</td>
<td>−76</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>348</td>
<td>301</td>
<td>−47</td>
</tr>
</tbody>
</table>

*aBold and italic values denote positive and negative errors larger than 30, respectively.

Figure 12 shows the relationships between $LW/PH_{program}$ and corresponding $LW/PH_{human}$ for 720P-D, 720P-N, and 1080P-D. It can be seen that all the data points are quite close
to the line of \( y = x \), which reveals that \( \text{LW/PH}_{\text{program}} \) and \( \text{LW/PH}_{\text{human}} \) match quite well. MAEs of 720P-D, 720P-N, and 1080P-D are 26 (3.6\% of 720P), 27 (3.8\% of 720P), and 49 (4.5\% of 1080P). By comparing with the error of the conventional method (9.4\% of 720P-D, 13.2\% of 720P-N, and 10.9\% of 1080P-D), the proposed method reduces the error to about one-third for 720P modules and more than a half for 1080P modules.

Although CSF is believed to accurately depict the characteristic of HVS, residual errors, as smaller than 5\%, can be still found even if the proposed method is used. In fact, in the ISO-12233:2000 test chart, a hyperbolic wedge for subjective judgment and its counterpart slanted edge for objective evaluation separate a little spatially, as shown in Fig. 7. By considering that imaging quality of a camera varies from spatial positions, visual resolutions at the wedge and slanted edge substantially differ a little. More important, in the image region hyperbolic wedges occupy, lens distortion occurs more or less. As mentioned before, lens distortion causes inconsistency between actual frequencies and text labels, hence subjective results essentially have a little bit of error. The error caused by such two factors is inevitable if the ISO-12233:2000 test chart is used, although it can still meet the requirement of most of surveillance camera manufacturers. If additional improved performance is desired, a redesigned test chart that considers spatial imaging quality variation and lens distortion is recommended.

### Table 2. LW/PH<sub>program</sub> Calculated with Proposed Method and Corresponding LW/PH<sub>human</sub>

<table>
<thead>
<tr>
<th>Direction</th>
<th>Focusing Status</th>
<th>LW/PH&lt;sub&gt;human&lt;/sub&gt;</th>
<th>LW/PH&lt;sub&gt;program&lt;/sub&gt;</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horizontal</td>
<td>-3</td>
<td>305</td>
<td>315</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>-2</td>
<td>478</td>
<td>422</td>
<td>-56</td>
</tr>
<tr>
<td></td>
<td>-1</td>
<td>500</td>
<td>517</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>605</td>
<td>588</td>
<td>-17</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>520</td>
<td>523</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>403</td>
<td>413</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>318</td>
<td>330</td>
<td>12</td>
</tr>
</tbody>
</table>

| Vertical  | -3              | 310                    | 321                      | 11    |
|           | -2              | 347                    | 351                      | 4     |
|           | -1              | 525                    | 505                      | -20   |
|           | 0               | 575                    | 564                      | -11   |
|           | 1               | 473                    | 457                      | -16   |
|           | 2               | 438                    | 406                      | -32   |
|           | 3               | 348                    | 357                      | 9     |

<sup>a</sup>Bold and italic values denote positive and negative errors larger than 30, respectively.

---

**Fig. 11.** Calculation of LW/PH<sub>program</sub> with the proposed method for the first 720P-D camera module: (a) horizontal direction, and (b) vertical direction.

**Fig. 12.** Relationship between LW/PH<sub>program</sub> calculated with the proposed method and corresponding LW/PH<sub>human</sub>.
the conventional method can be also validated in Fig. 14, where the data of 720P-D are investigated. In Fig. 14(a), the proposed method makes subjective and objective results match well, thus all the data points gather close to the line of $y = x$. In Fig. 14(c) where fixed decision SFR of 10% is used, the data points approximately distribute in an “S” shape that intersects with the line of $y = x$ just at the “accurate point,” which accords with our explanations above.

Actually, the reason why decision SFR of 10% is the most recommended in previous literature [5–7,13,14] can be also quantitatively explained. Besides 10%, SFR of 5%, 15%, and 20% are also illustrated in Fig. 13, and corresponding relationships between objective LW/PHprogram and subjective LW/PHhuman are shown in Figs. 14(b)–14(e) along with their MAEs. In Fig. 13, SFR of 5%, 10%, 15%, and 20% intersect with CSF’s reciprocal at LW/PH around 350, 500, 600, and 650, respectively, therefore the data points in Figs. 14(b)–14(d) approximately distributing in “S” shapes intersect with the line of $y = x$ at such “accurate points.” By considering that the CSF used in our study is specified with very typical viewing conditions [19 in. (22.86 cm) display at 200 cd/m² and viewing distance of 80 cm], our test scenario is quite close to most of the camera manufacturers’ Decision SFR of 10%, with SFR larger or smaller than which worse MAEs are obtained is reasonable to be the optimum. In addition, some manufacturers like to use SFR of 20% rather than 10%, which can be also explained by our results. According to Figs. 13 and 14(e), SFR of 20% leads to an “accurate point” of about 650. By considering that the sharp focusing status is important for cameras, it is a practical approach to guarantee calculation accuracy only for cameras with that high of LW/PH, though those low LW/PH ones are sacrificed. If higher resolution modules, like 1080P, are adopted, a larger “accurate point” is preferred, hence the decision SFR of 20% or even larger becomes more appropriate.

As explained above, the conventional method can only obtain an accurate objective LW/PHprogram around the “accurate point,” and it is just a makeshift to select a different decision SFR. The essential solution is to use the proposed method that can take varying perceptibility of HVS into consideration, i.e., the proposed method actually generates a dynamic “accurate point” for a different spatial frequency.

4. CONCLUSIONS
The conventional e-SFR method with a fixed decision SFR is a useable indicator for a surveillance camera’s visual resolution. However, significant absolute errors with respect to the subjective results make it only able to predict the order of numbers. To obtain an objective evaluation method that can directly predict the subjective visual resolution in LH/PW, e-SFR and HVS characterized by CSF was combined in this paper. Because perceptibility variation of HVS was considered via CSF, the error level has been much suppressed. By adopting eight 720P-D modules, four 720P-N modules, and two 1080P-D modules to implement a series of visual tests with 10 testees, MAEs of the three types of modules were as low as 26 (3.6% of 720P-D), 27 (3.8% of 720P-N), and 49 (4.5% of 1080P-D), while those of the conventional method using fixed decision SFR of 10% are 68 (9.4% of 720P-D), 95 (13.2% of 720P-D), and 118 (10.9% of 1080P-D). We also demonstrated where the error of the conventional method comes from and how the proposed method takes effect.

Because the error between calculated and humanly judged visual resolution has been significantly suppressed, the proposed method is an important improvement of the conventional e-SFR method in evaluating visual resolution of surveillance cameras. This study is on the basis of the ISO-12233:2000 chart; however, the test chart is not limited to this.
Any other chart based on which SFR can be correctly calculated can be adopted, e.g., the new ISO-12233:2014 chart that provides a more convenient way of calculating e-SFR [6] or even the newly introduced sinusoidal Siemens star that is believed to lead to more accurate sinusoidal SFR [28–30].

Funding. Ministry of Science and Technology of the People’s Republic of China (MOST) (MOST104-2628-E-009-012-MY3, NSC102-2221-E-009-168-MY3); Industry-university cooperative research project between Sercomm Corporation and National Chiao Tung University (104C041).

Acknowledgment. We thank Sercomm Corporation, Taiwan, for their financial and technical support.

REFERENCES