Journal of Electronic Imaging

JElectronicImaging.org

Preferred image quality metric for shifted superimposition-based resolution-enhanced images

Svein Arne Jervell Hansen
Muhammad Nadeem Akram
Jon Yngve Hardeberg
Marius Pedersen

Svein Arne Jervell Hansen, Muhammad Nadeem Akram, Jon Yngve Hardeberg, Marius Pedersen, "Preferred image quality metric for shifted superimposition-based resolution-enhanced images," *J. Electron. Imaging* **27**(3), 033017 (2018), doi: 10.1117/1.JEI.27.3.033017.

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Electronic-Imaging on 22 Mar 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Journal of Electronic Imaging 27(3), 033017 (May/Jun 2018)

Preferred image quality metric for shifted superimposition-based resolution-enhanced images

Svein Arne Jervell Hansen,[a,b,*] Muhammad Nadeem Akram,[a] Jon Yngve Hardeberg,[c] and Marius Pedersen[c]
[a]University of South-Eastern Norway, Borre, Norway
[b]Barco Fredrikstad AS, Gamle Fredrikstad, Norway
[c]Norwegian University of Science and Technology, Gjøvik, Norway

Abstract. Shifted superimposition is a resolution-enhancement method that has gained popularity in the projector industry the last couple of years. This method consists of shifting every other projected frame spatially with subpixel precision, and by doing so creating a new pixel grid on the projected surface with smaller effective pixel pitch. There is still an open question of how well this technique performs in comparison with the native resolution, and how high the effective resolution gain really is. To help investigate these questions, we have developed a framework for simulating different superimposition methods over different image contents, and evaluate the result using several image quality metrics (IQMs). We have also performed a subjective experiment with observers who rate the simulated image content, and calculated the correlation between the subjective results and the IQMs. We found that the visual information fidelity metric is the most suitable to evaluate natural superimposed images when subjective match is desired. However, this metric does not detect the distortion in synthetic images. The multiscale structural similarity metric which is based on the analysis of image structure is better at detecting this distortion. *© 2018 SPIE and IS&T* [DOI: 10.1117/1.JEI.27.3.033017]

Keywords: resolution; image quality; image processing; displays; projectors; projection systems.
Paper 170937P received Oct. 30, 2017; accepted for publication Apr. 26, 2018; published online May 21, 2018.

1 Introduction

Resolution is one of the key performance parameters of a projector, and the projector industry continuously aims to increase it. Superimposition of projected images is a cost-effective way of enhancing the resolution in a projector above the native resolution of the spatial light modulator (SLM). Superimposition may be implemented either with a multiprojector setup or with an opto-mechanical system within a single projector. As long as a superimposition consists of two or more images superimposed on one projected surface, the resulting image will be an additive function of the projected images.

Currently, resolution enhancement has gained some momentum because of the market drive for 4K images and video. Not all SLM technologies have cost-efficient 4K modulators available. For these kinds of modulator technologies, it is appealing to push the resolution above the native resolution of the SLM. Even though the actual pixel count on the canvas will increase, this method also introduces some artifacts in the image. As the optical overlap of superimposed images acts like a low-pass filter, some high-frequency content is lost in the image. The spatial artifacts manifest as blurring in the image, and these artifacts impact both the visual quality and the resolution measurements. The introduced artifacts raise the question of whether the resulting image on the canvas really has a higher resolution and a higher quality than downscaling the high-resolution image and displaying it at the native resolution of the SLM.

This paper investigates different methods of superimposition and explores how these methods compare to each other in quality. Then we seek out to find the most suitable image quality metric that correlates with how we subjectively rate the quality as observers. The goal of this paper is to find a way to evaluate the quality of superimposition algorithms through simulations, so that it is possible to achieve a level of confidence before building up a complete physical system.

The rest of this paper is organized as follows: the Related Work section provides an insight into the prior work done in the field of superimposition and how the quality is evaluated in these papers. The next section presents a set of relevant quality metrics, while the Methodology section describes the experimental setup that is used to tests different metrics and superimposition methods. The simulated results and the subjective experiment are presented in the following section. Thereafter, the findings are summarized and discussed in the Discussion section. Finally, the last section concludes on how to evaluate the resulting image quality for this particular application and the future work still to be done.

2 Related Work

Takahashi et al.[1] proposed a setup in 1995 with four liquid crystal display (LCD) projectors projecting on the same screen with an elaborate mirror-setup. By taking advantage of the small fill factor in the LCD pixels, the overlap between the pixels is very low in this case. By interleaving the pixels from all of the projectors, the idea here is to fill out the blocked area of the pixels with the other projector channels, and together double the resolution both horizontally and

*Address all correspondence to: Svein Arne Jervell Hansen, E-mail: svein.a.hansen@usn.no

1017-9909/2018/$25.00 © 2018 SPIE and IS&T

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Electronic-Imaging on 22 Mar 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

vertically. This setup is very cumbersome and requires careful adjustment in the installation phase. Over time, the fill factor of LCD panels has also increased, leaving one of the main prerequisites of this method obsolete. They used the modular transfer function (MTF) as a main parameter to evaluate the resolution enhancement. The MTF is obtained in this case through optical simulations of the projector prism and projection lens, and then calculating the resulting MTF based on the pixel overlap, amount of projectors and the projection lens, and prism performance.

Jaynes and Ramakrishnan[2] proposed a system where several projectors project at the same screen, and then they are calibrated to determine the relative subpixel shift for each projector. The goal of this calibration is to derive an accurate mapping of each projector's framebuffer coordinates to the high-resolution target frame. Such a calibration needs to be very accurate and represents a significant challenge in practice, and the system is quite fragile when fully calibrated. Jaynes et al. verified their work by printing close-up photographs of the superimposed resolution enhancement showing the quality improvement. The authors presented the gained image quality as visual results printed side-by-side for the reader to compare them, and they do not quantify the quality gain. The images presented are close-up photography of the two projected scenes from natural images and two projected images containing text.

Allen and Ulichney[3] made a breakthrough with their idea to keep the whole system within one projector unit, and instead include an optomechanical image shifter to shift every nth image frame spatially on the projected surface. This method, called wobulation, ensures uniform pixel shift and a controlled overlap of the pixels. Wobulation allows each pixel in the SLM to address multiple locations (pixels) in the final projected image. The cost of using the same SLM to show different image positions is that the temporal resolution decreases with a factor equal to the number of image positions used in the wobulation. In the paper by Allen and Ulichney, the same subframe is used in both positions resulting in a slightly blurred image. The authors present the gained image quality as visual results printed side by side for the reader to compare them, and they do not quantify the quality gain. Two natural images were used in this evaluation.

Said[4] presented in 2006 an extensive work on how to generate the subframes. The focus of his work was to establish a theoretical framework for understanding the potential and limitations of the superimposition method. The objective in Said's work is not to obtain the most optimal generation of the subframes, but to understand the mathematical properties that define the quality of the solution. Said used peak signal-to-noise ratio (PSNR) as a quality metric and also printed the native resolution and the superimposed resulting images side by side for the reader to compare them. Two natural images were used to showcase the enhanced quality of the superimpositioning methods.

Damera-Venkata and Chang[5] proposed the year after a method to produce superimposed images through multiprojector systems. This work proves that the superimposition method is valid for displaying frequencies above the Nyquist frequency of a single projector. Other than these theoretical results, the work lacks real quality measurements besides printing the results for the reader to visually inspect the superimposed results. Damera-Venkata and Chang used two computer-generated images as test-scenes in their evaluation.

Okatani et al.[6] explored the theory from Damera-Venkata and Chang[5] further, and showed how the quality of the superimposed images changes with the maximum brightness of the system. In this work, the quality decisions are also made by printing the resulting images for the reader to judge the enhanced quality, and no quality metric is used. Okatani et al. used a low-resolution image of computer-generated text and a natural image of a horse to evaluate their method.

Sajadi et al.[7] presented in 2012 a different image enhancement approach, where two cascaded SLMs are used for enhancing the edges of the image, and by that approach also enhancing the resolution. Between the SLMs an optical pixel sharing unit is introduced to create smaller pixels in the spatial domain. This approach seems to work quite well, and they use just noticeable difference (JND) in CIELAB ΔE to analyze the image for local variance and to identify the edges of interest in the image. But the quality evaluation of their algorithm is determined only by printing the resulting images, and encouraging the reader to zoom in on the images to observe the quality enhancement. Sajadi et al. used six different natural scenes, one computer-generated image of a building, and a technical drawing as test scenes in their work. Some of the resulting images were simulated results and other results were photographs taken from test setups.

The year after, Sajadi et al.[8] proposed a low-cost approach that shifts the whole image with subpixel precision and superimposes the shifted image on top of the original image. This may seem similar to the wobulation method proposed by Allen and Ulichney,[3] but the method proposed by Sajadi et al. does not time-multiplex the images, but rather superimposes the image on a shifted version of itself. When it comes to spatial quality this method may be suboptimal, but it is very computationally cost efficient. The quality gain of this method is quantified through the structural similarity (SSIM)[9] metric, and they used the CIELAB ΔE to check if the colors have drifted. Sajadi et al. also evaluated the content preservation in the image by calculating histogram of gradients for different combinations of pixel-shift and numbers of superimposed frames. Six natural images, mostly buildings, and one map were used as test scenes in this work.

Heide et al.[10] made an interesting twist in 2014 to project the image on a new SLM instead of superimposing the images on the projected surface. By shifting the second SLM with subpixel accuracy, the second SLM is subtracting light instead of adding it. This method is named multiplicative superimpositioning as opposed to the regular additive superimpositioning, where the light from the subimages is added on top of each other. This method apparently provides good results, which is verified by PSNR, SSIM, and MTF analysis. Heide et al. used seven natural images, mostly motorsport scenes with commercial decals in them, and two computer-generated images, as test scenes in their work.

Barshan et al.[11] proposed their own superimposition scheme in 2015 named shifted superposition. This method is quite similar to the wobulation method proposed by Allen and Ulichney,[3] but the generation of the subimages is done independently instead of using the same subimage for both positions. The quality improvement in this work is verified by visual inspection and using the SSIM[9] metric

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Electronic-Imaging on 22 Mar 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

as well. Barshan et al. used two computer-generated test images and one natural image as test scenes in their work.

Hansen et al.[12] presented in 2017 a study of how a selected set of IQMs detect distortion and loss of detail in shifted superimposition. The IQMs included in that work was PSNR, PSNR- human visual system (HVS),[13] PSNR-HVSM,[14] ESSIM,[15] feature-SIM[16] (FSIM), DCTex,[17] visual information fidelity (VIF),[18] SSIM,[9] SRSIM,[19] and multi-scale structural similarity metric (MSSSIM).[20] As that work is based solely on simulations and not on subjective experiments, most of the quality criteria are objective. Hansen et al. used three computer-generated test images and two natural test images as test scenes in their work.

3 Image Quality Metrics

As seen in the previous section, there are some variations of how the quality is evaluated by different authors in the field of superimpositioning. The most common method is to present different resulting images representing the improvement in visual quality of the superimpositioning, but this is a poor method for comparing different algorithms objectively. This section will look briefly into different quality metrics mentioned in the previous section, and present some other quality metrics that may also be suitable.

As we have the reference image available, we will focus on full-reference metrics for evaluating the superimposed images. We categorize these metrics mainly into two main categories: raw error-based calculations and HVS-inspired metrics.

The error-based calculations are mathematical metrics based on error quantification between two images. They are popular as they are simple to understand, easy to use, and have a low computational cost. Typical examples of these metrics are mean square error and different versions of SNR. SNR and PSNR are based on the principle that the distorted image consists of the original image and a noise component in addition to an independent signal. SNR is defined as the ratio of average signal power to noise signal power, whereas PSNR is defined as the ratio of peak signal power to noise signal power.

The weighted SNR (WSNR) was developed to take the HVS contrast-sensitivity function into account.[21] WSNR is defined as the ratio of the averaged weighted signal power to the average weighted noise power. The WSNR is a hybrid between the raw error-based calculations and the HVS-inspired metrics as it is an error-based metric (SNR) modified slightly using some of the HVS attributes. Other metrics such as PSNR-HVS[13] and PSNR-HVSM[14] use the principles from PSNR and modify this metric based on the frequency-based contrast sensitivity of the HVS. PSNR-HVS is calculated utilizing the mean shift and contrast stretching to highlight the areas of the image that the HVS is most sensitive to. The PSNR-HVSM on the other hand uses DCT to calculate contrast masking, and by taking the contrast-sensitivity function of the HVS into account the metric ignores the same contrast steps that the HVS will also ignore.

Pure HVS-inspired metrics take the attributes of the HVS into account and aim to measure specific image attributes that the HVS is particularly sensitive to. SSIM[9] is such a metric, which compares the luminance, contrast, and structure in both images to measure the similarity between them. The approach of taking the HVS fully or partially into account has fostered several quality metrics, such as multiscale SSIM[20] (MSSSIM), ESSIM,[15] SR-SIM,[19] FeatureSIM[16] (FSIM), DCTex,[17] VIF,[18] and VSNR.[22] MSSSIM is a multi-scale structural similarity method, which supplies more flexibility than single-scale methods in incorporating the variations of viewing conditions. ESSIM aims to model the perceptual fidelity of semantic information between two images by assuming that the semantic information of images is fully represented by edge-strength of each pixel. SR-SIM is based on a specific visual saliency model, spectral residual visual saliency. This metric follows the theory that an image's visual saliency map is closely related to its perceived quality. FSIM is based on the fact that the HVS understands an image mainly according to its low-level features. By considering the phase congruency and the gradient magnitude of the image, the image quality is calculated. DCTex is based on a key assumption that the signal error in each subband and local region contributes to the entire distortion independently. This assumption is reasonable as most typical distortions have few (linear) correlations both between the subbands and between the neighborhoods at large spatial scales. The HVS contrast-sensitivity function and texture mapping property is used to weight the contribution from the different subbands into a global metric for the distortion over the whole image. VIF quantifies the information that is present in the reference image, and also quantifies how much of this reference information can be extracted from the distorted image. Combining these two quantities, the visual information fidelity measurement is calculated. VSNR quantifies the visual fidelity of natural images based on near-threshold and suprathreshold properties of the HVS. In addition the metric operates on physical luminances and visual angle (rather than on digital pixel values and pixel-based dimensions) to accommodate different viewing conditions.

The metrics included in this setup are the following: PSNR is one of the most widely used error calculation metrics. For metrics taking the HVS into account, we have included the metrics PSNR-HVS, PSNR-HVSM, ESSIM, FSIM, VSNR, DCTex, and VIF. In addition to these categories, we also have used metrics that are purely looking at the structure in the image, that is SSIM, SR-SIM, and MSSSIM.

4 Simulation Framework

In the first part of this research, we concentrate on verification of the objective quality differences through simulation, thus the entire part of this setup is carried out within a simulation environment written in MATLAB.

We have implemented four different ways of generating the superimposed image in this simulation framework, and are also comparing the superimposed methods with images presented in the native SLM resolution. We are not aiming to develop the best method for superimposing images in this work, so we have picked some methods that are distinguishable from each other, with different properties.

The superimposition methods described in this paper are either mentioned in previous papers[3,4,8] or methods that build further upon them. The Naïve and the filtered Naïve methods are actual methods that are quite close to what is actually in use in several products on the market. Several vendors use the idea of upscaling the input image to double the resolution

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Electronic-Imaging on 22 Mar 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

of the SLM, and then do some image processing in that doubled resolution domain.[23,24] The main idea here is that they are not downscaling the input image to the SLM resolution with the loss of detail that downscaling gives. Instead they are upscaling the input image for preserving more details, before the subframe pixels are chosen among the resulting upscaled pixels.

In theory, iterative algorithms like the one presented by Sajadi et al.[8] may achieve higher quality of the superimposed image. But since processing latency is crucial in a lot of high end projection applications, such algorithms are not used as they typically introduce frame(s) of latency and the hardware implementation is very expensive. This is the reason why this class of algorithms is not taken into account in this paper.

4.1 *Downscaled*

This method is included for reference. The goal of the superimpositioning is to enhance the resolution above the native resolution of the SLM, so the downscaled image represents the SLM resolution. The resulting output image will then be given by

$$\text{OutImage} = \text{Resize}(\text{RefImage}, \text{SLMresolution}), \tag{1}$$

where Resize in this example is the corresponding MATLAB function and SLMresolution is the resolution of the SLM in use. RefImage is the original high-resolution image.

4.2 *Downscaled Superimposed*

This method generates the subimages as the downscaled method, but then these subimages are spatially shifted and superimposed on themselves. It is not an ideal method, but as seen in Sec. 6 it is a step up in perceived quality from the regular downscaled version in some instances. Allen and Ulichney[3] used this version to verify the superimpositioning in their wobulation paper. In this method, both subframes will be equal and given by

$$\begin{aligned}\text{Subframe}A &= \text{Subframe}B \\ &= \text{Resize}(\text{RefImage}, \text{SLMresolution}).\end{aligned} \tag{2}$$

4.3 *Naïve*

In the Naïve method, we upscale the input image to the double horizontal and vertical resolution of the SLM, then we pick the pixels for the different subframes directly from the up-scaled frame. The Naïve method produces quite sharp images, but some details will be lost as we just select every other pixel

$$\begin{aligned}\text{IntermediateFrameNaive} &= \text{Resize}(\text{RefImage}, 2 \\ &\quad * \text{SLMresolution}) \\ \text{Subframe}A(i,j) &= \text{IntermediateFrame}(2*i-1,2 \\ &\quad * j-1) \\ \text{Subframe}B(i,j) &= \text{IntermediateFrame}(2*i, 2*j).\end{aligned} \tag{3}$$

4.4 *Gaussian*

This method starts out with the same intermediate frame as the Naïve, but in addition we have filtered the up-scaled image with a Gaussian filter. By doing this, we produce an image that is slightly more blurred, but we will not lose as much details as in the Naïve method. This operation is performed in MATLAB as

$$\begin{aligned}&\text{IntermediateFrameGauss} \\ &= \text{imfilter}[\text{IntermediateFrameNaive}, \text{fspecial}(\text{'gaussian'})],\end{aligned} \tag{4}$$

and then selecting the pixels in the same way as in the Naïve method. The imfilter function in MATLAB filters the multidimensional array IntermediateFrameNaive with a multidimensional filter, in this case a Gaussian filter.

4.5 *Gaussian Sharpened*

The Gaussian sharpened method is the same as Gaussian, but in addition, we apply a sharpening filter after applying the Gaussian filter. This will remove some of the blur added, but with the possibility of amplifying noise in the image. This operation is performed in MATLAB as

$$\begin{aligned}S\text{filter} &= b * [0, a, 0; a, (-4) * a, a; 0, a, 0] \\ \text{InputImageMask} &= \text{imfilter}(\text{IntermediateFrameGauss}, S\text{filter}) \\ \text{IntermediateFrameSharpened} &= \text{IntermediateFrameGauss} + k * \text{InputImageMask},\end{aligned} \tag{5}$$

and then selecting the pixels in the same way as in the Naïve method. The constants in the sharpening filter are set to $a = -1$, $b = 0.25$, and $k = 0.5$.

4.6 *Superimpositioning*

The superimpositioning is done by shifting every other image half a pixel in the up-left/down-right diagonal of the image. This results in a two-position additive superimpositioning scheme, which is the use case for our experiments. We have not investigated more than two positions or other techniques, but only additive superimpositioning in our experiments.

4.7 *Test Images*

We have used 13 different test-images to test different image properties. Ten natural images that are included in the subjective experiment, and three synthetic images that are generated to provoke different types of errors in the algorithms mentioned above. The natural images are images containing different types of high-frequency contents, ranging from buildings and architecture to random edge structure in waves. The natural images are shown in Fig. 1, whereas the synthetic images are shown in Fig. 2.

Natural images are in itself a very diverse group of images, and it is not given that one IQM will be optimal

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Electronic-Imaging on 22 Mar 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

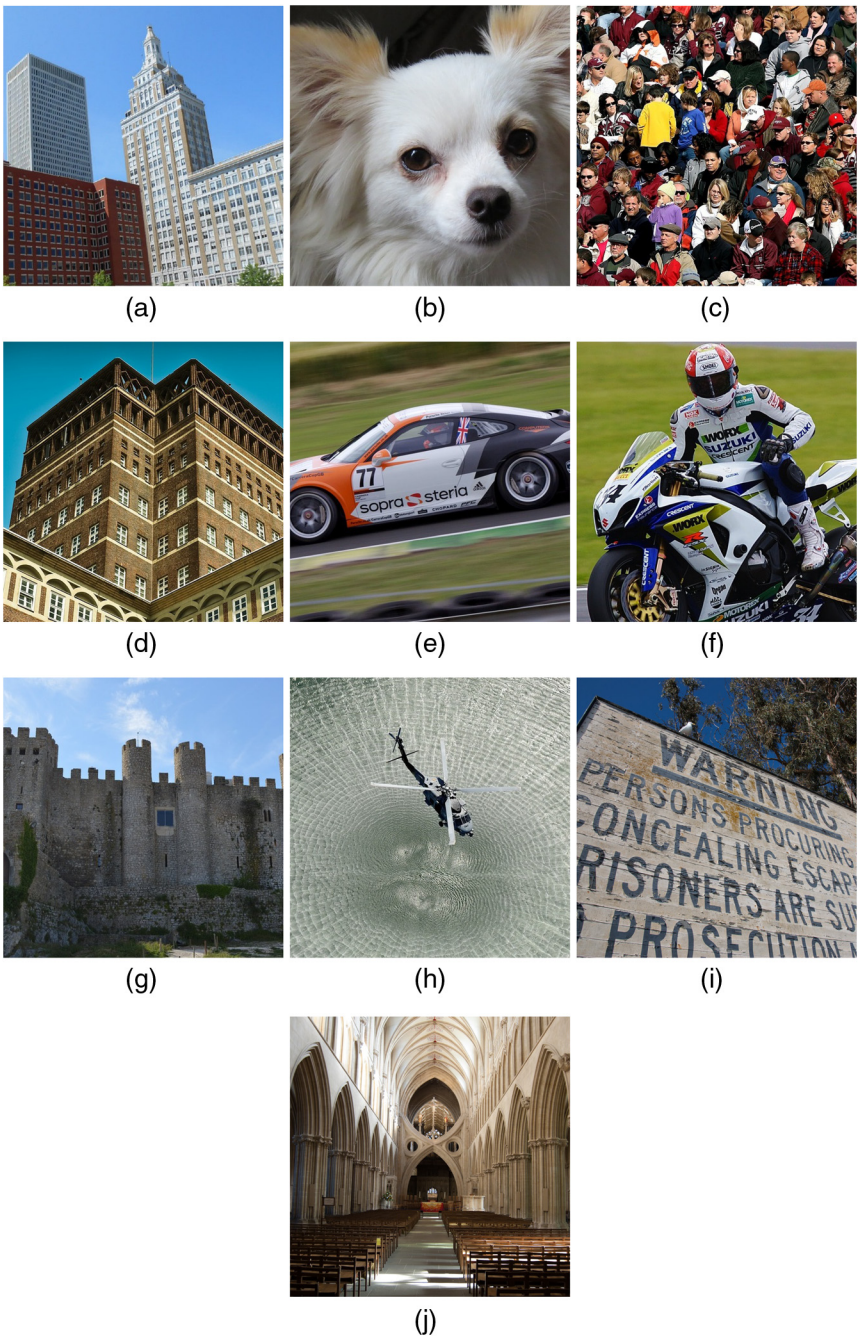

Fig. 1 Natural images used for the subjective experiment. (a) Downtown, (b) dog, (c) crowd, (d) architecture, (e) Porsche, (f) Michael Rutter, (g) medieval castle, (h) helicopter, (i) sign, and (j) church. The images (a, b, c, d, e, g) and (h) is from pixabay, (f) is from Wikipedia under the creative commons license, and (i) and (j) are from the CID:IQ database.[25]

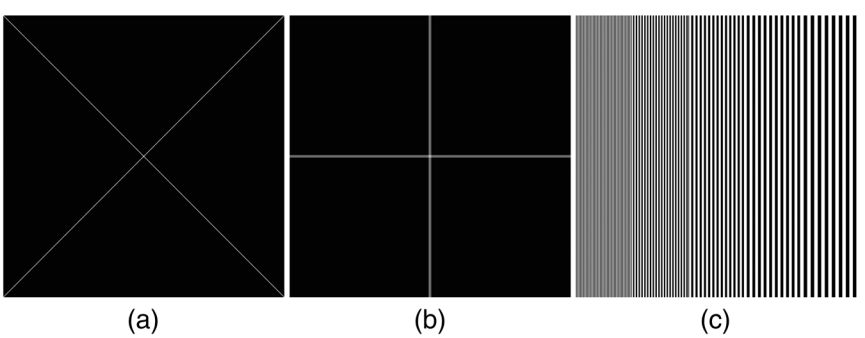

Fig. 2 Synthetic test images used (a) cross, (b) line pairs, and (c) H-frequency.

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Electronic-Imaging on 22 Mar 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

for the whole group of images. Different scenes have very different characteristics in textures, contrast, and frequency content. How the different IQMs react on these differences depends on what attributes in the image the IQM analyze. To accommodate these facts, we have included a broad selection of natural images that include different features, but also that include details that make the resolution enhancement worthwhile. Flat images with no high-frequency content will not benefit from a higher resolution, so we have not included any images of this kind. The natural images presented in Fig. 1 show that most of the images in this study have a lot of high-frequency details. The optical overlap in the superimposition acts as a low-pass filter, so the possible artifacts introduced by this method is that we may lose details in high frequency areas of the image. The selected images all have some kind of objects with a high degree of detail that the different superimposition algorithms may affect in different ways.

The three synthetic images included in this study are:

Cross. A white cross on a black background with single pixel diagonals. This is included to see how the metrics detect distortion of single pixel details.

Line pairs. A synthetic image consisting of three line pairs in a horizontal direction and three line pairs in a vertical direction. This is included to see how the metrics perform in detecting missing line pairs.

H-frequency. Synthetic image that includes bands of five different frequencies starting at the highest possible spatial frequency at the image native resolution. The natural images are presented in the subjective experiment section.

The synthetic images are rendered at the different resolutions given in the paper, so that the description of the images fits the given resolution. This means that single pixel details remain single pixel at all resolutions. In the experiments, we keep the SLM resolution fixed, so the ratio between the reference image pixel size and the SLM pixel size is changed in that sense.

4.8 *Test Scenario*

In our work to identify the objective imperfections in the synthetic images, we have defined an SLM with the resolution of 250×250 pixels. We have chosen to set the resolution low for keeping the computational time down. We have then iterated the input resolution in 25 pixel steps from 225×225 to 600×600 to generate different input-resolution/output-resolution ratios, and use this as a parameter to provoke different behavior from both the subframe generation methods and the quality metrics. With this input resolution range, we are simulating input resolutions from below the native resolution to above double of the native resolution.

For the natural images, we have set up a subjective experiment as described in Sec. 6. In this experiment, the input images are 512×512 pixels, whereas the SLM resolution is kept at 256×256 pixels.

As we are simulating subpixel behavior, each pixel of the reference image is quadrupled into four pixels in the subimages. The superimpositioning system shifts every other frame half a pixel diagonally, and this is simulated by shifting one pixel diagonally after quadrupling each pixel. This doubles the resolution of the simulated result both horizontally and vertically, bringing the simulated resulting image into the same resolution as the reference image.

5 Simulation Results

For the synthetic images, we have concrete symptoms to look for. The line pair image has three distinguishable line pairs that will eventually fuse together when the input/output ratio gets too high. The goal of the superimpositioning is to

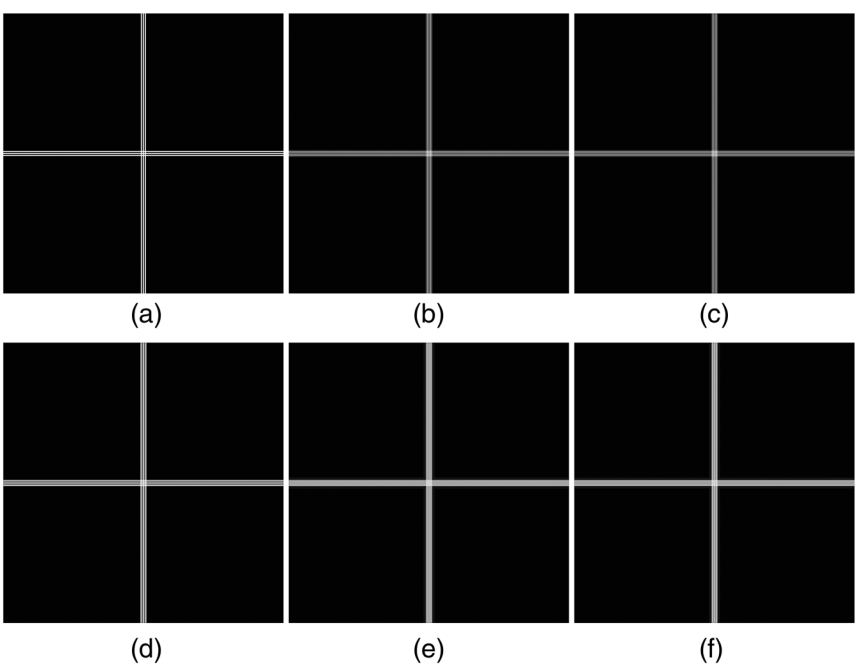

Fig. 3 Zoomed in on the resulting line pairs at 275 pixels input resolution. (a) Reference image, (b) downscaled, (c) downscaled superimposed, (d) Naïve, (e) Gaussian, and (f) Gaussian sharpened. Note how (b) and (c) have lost one line pair.

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Electronic-Imaging on 22 Mar 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

preserve the details in the image at frequencies above the spatial frequency of the SLM, so the superimpositioning methods should preserve the line pairs better than the downscaled method. In addition, we are looking for metrics that detect when we lose line pairs in the different superimpositioning methods. The different methods perform as the following; downscaled and downscaled superimposed both lose one line pair when the input resolution goes above the SLM resolution at 250 pixels. The Naïve method preserves the three line pairs up to 300 pixels, and the Gaussian and Gaussian sharpened preserves the line pairs up to around 350 pixels. Figure 3 shows how the different algorithms perform at 275 pixels input resolution. The line pairs have lost much of the local contrast when pushing the limits, but it is still distinguishable as three line pairs. None of the metrics detect these details, and some of the metrics even rate the two visually worst methods as the two best ones. We note the preference for the downscaled and the downscaled superimposed methods may be because they add less blur and preserve more local contrast in the image, even though they lose details in the image.

The test image cross is made to test single-pixel details. When given the cross image as an input, the Naïve method deteriorates the diagonal in the nonshifted direction. This diagonal gets worse at higher resolutions and is completely lost at 500 pixels and above. The loss of details is shown in Fig. 4, showing how the Naïve superimpositioning looks with and at input resolution of 300×300, 400×400, 450×450, and 500×500 pixels.

Several of the metrics do not detect this severe loss of details, but ESSIM, SR-SIM, FeatureSIM, and MSSSIM pick this up. Figure 5 shows how the different metrics evaluate the degradation of quality with the synthetic cross as an input image. Note how the Naïve method drops in performance after 400 pixels input resolution at the metrics mentioned above.

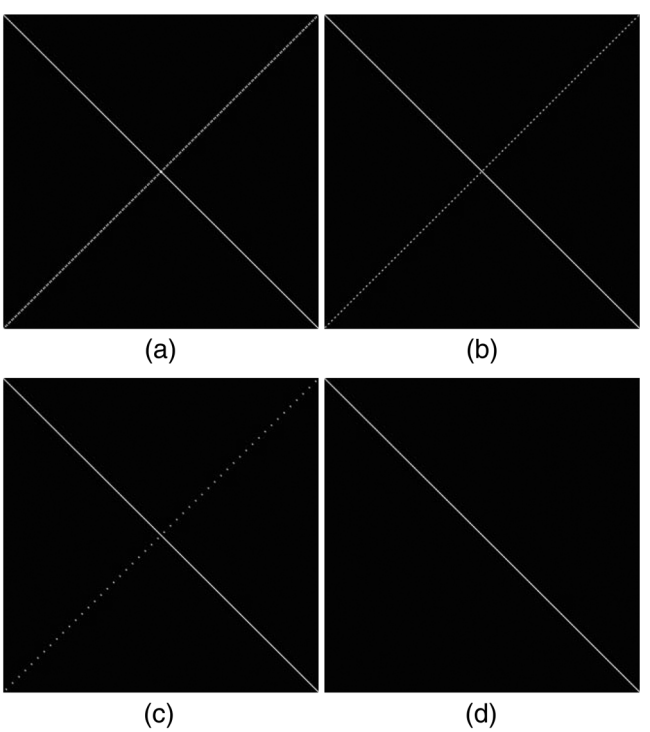

Fig. 4 Results from Naïve superimpositioning with (a) 300, (b) 400, (c) 450, and (d) 500 pixels input resolution.

The synthetic H-frequency image is by nature problematic for the superimpositioning method to represent correctly. This is because when the input resolution increases, the frequency of the patterns goes above the frequency the SLM is naturally able to reproduce, and in some cases this introduce aliasing. We see in Fig. 6 that the Gaussian and the Gaussian sharpened methods are less prone to the aliasing effect than the other methods. We do not find any metrics picking up this feature. The metrics seem to favor the methods that introduce less blur instead, even though these methods introduce quite severe aliasing in some instances. Figure 6 shows how the different superimpositioning methods produce varying amounts of aliasing.

6 Subjective Experiment

The subjective experiment is designed to see which of the IQMs that correlate best with our subjective opinion of how the different methods of superimpositioning perform over natural images. For this experiment, we have selected natural images with a variety of content and structure, and we have included images of persons, texture, nature, buildings, vehicles, and text. Please see Fig. 1 for the complete set of scenes used in this experiment.

In this experiment, we have simulated the superimposed image for all five algorithms over the 10 images. We have kept the input resolution over SLM resolution ratio at a factor of 2×, meaning that an input image of 512×512 is simulated with an SLM resolution at 256×256. We then made a paired comparison test, testing each superimpositioning algorithm against each other once, resulting in 10 pair-tests for each image. Twenty six participants took part in this test, which was performed using the online evaluation platform QuickEval.[26] The participants were told to select the visually preferred image in each image pair. As this was an online experiment conducted on the subjects' own computer, the viewing conditions were different for the different subjects. QuickEval makes sure that no images are scaled, and that all of the images are presented in fixed resolutions. As this experiment tests the perceived resolution and spatial quality enhancement of the image, this condition was deemed to be good enough. The z-scores[27] from this test are shown in Fig. 7. The general trend in Fig. 7 shows that the methods, where each superimposition subframe is calculated individually give a much better perceived image than the downscaled and downscaled superimposed methods.

The correlation of each IQM toward these z-scores was then calculated for both Pearson and Spearman coefficients, and the results from these calculations are shown in Tables 1 and 2, and these results are shown in Figs. 8 and 9. The Pearson coefficient is a measurement of the linear correlation between the different metric results and the subjective ratings, and the Spearman coefficient is a measurement of how well the different metric results and the subjective ratings may be described using a monotonic function. VIF is the IQM that performs best according to these correlation coefficients when looking at the mean values in Figs. 8 and 9.

Several of the IQMs perform well in the subjective experiment. Based on the Pearson and Spearman correlation

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Electronic-Imaging on 22 Mar 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

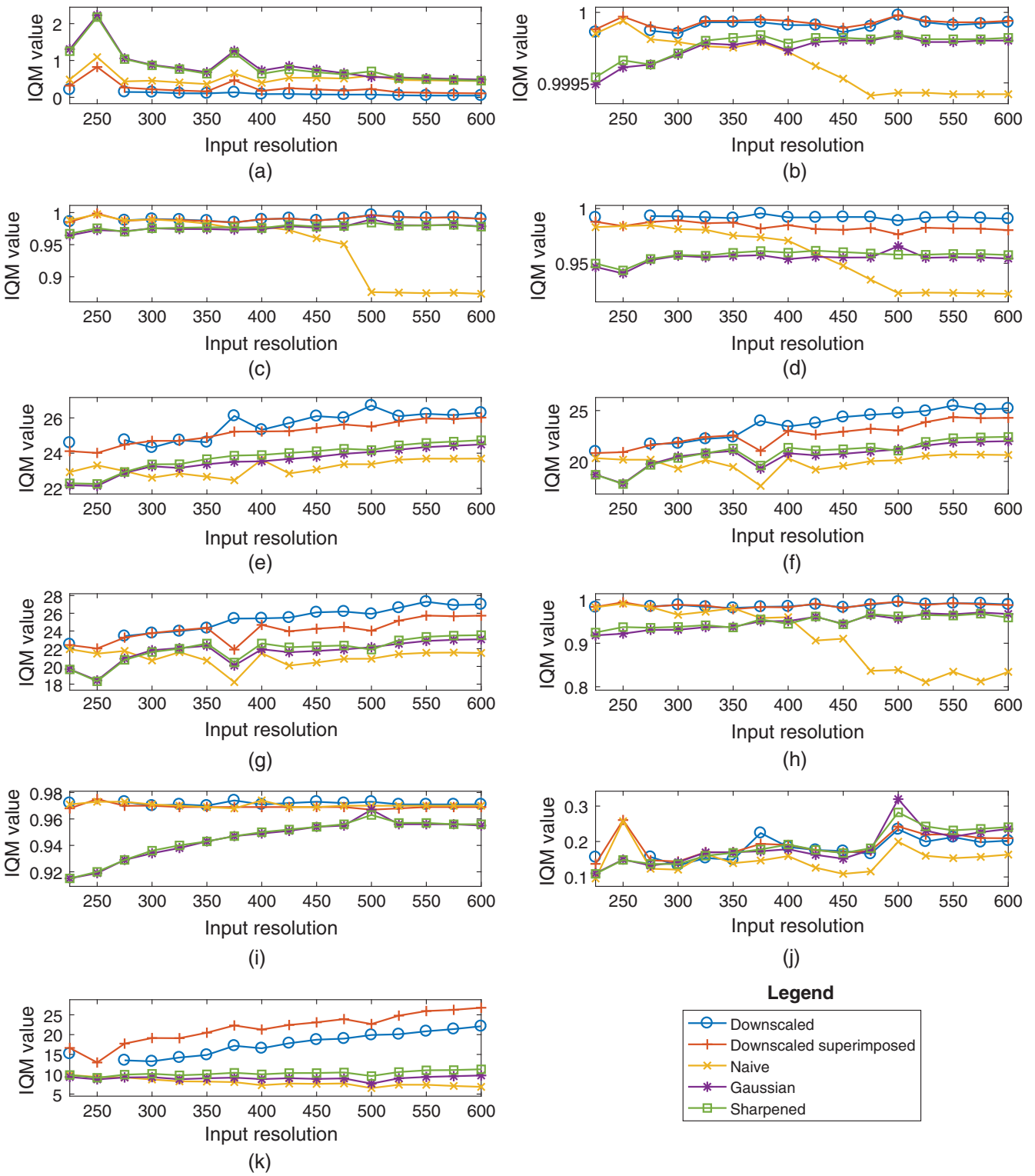

Fig. 5 Synthetic cross scene evaluated by all of the metrics. The *x*-axis represents the input resolution and the *y*-axis represents the IQM value. (a) Cross-DCTex, (b) cross-ESSIM, (c) cross-FeatureSIM, (d) cross-MSSSIM, (e) cross-PSNR, (f) cross-PSNR-HVS, (g) cross-PSNR-HVSM, (h) cross-SR-SIM, (i) cross-SSIM, (j) cross-VIF, and (k) cross-VSNR.

coefficients, the VIF IQM is the metric that performs best when rating according to the subjective view of the observers. Figure 10 shows how the VIF performs on all the images, with markers for different superimpositioning algorithms (markers) and the fitted linear regression curve (solid line).

7 Discussion

To determine which metric is the best one to use, we must first decide what the metric should detect. The purpose of the superimpositioning is to increase the perceived resolution of the image above the native resolution of the SLM. This

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Electronic-Imaging on 22 Mar 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

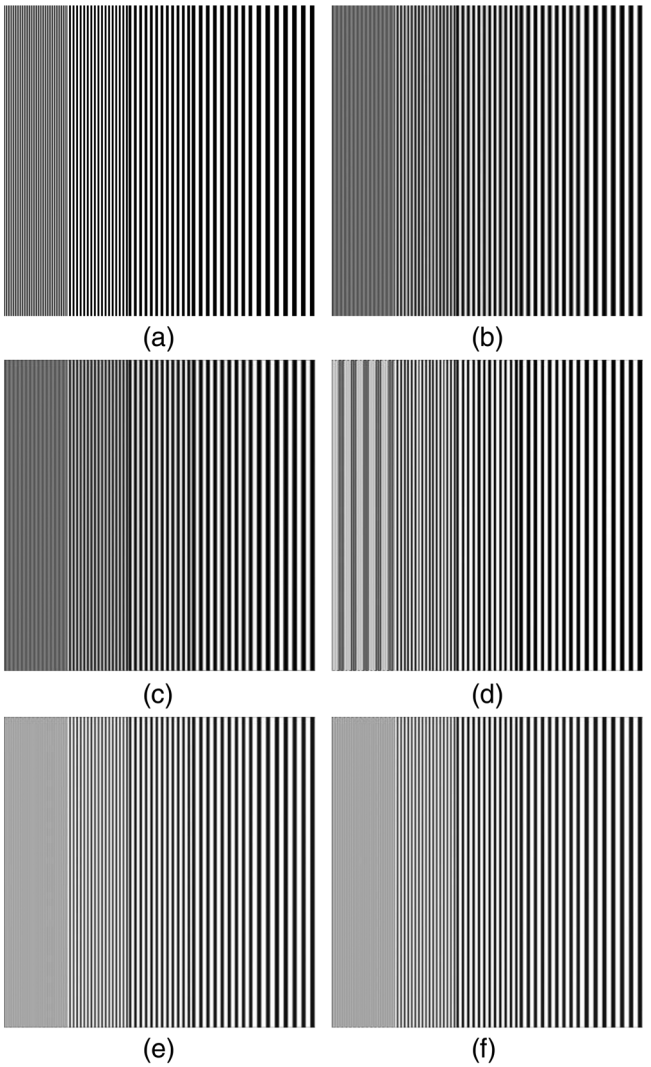

Fig. 6 Horizontal frequency image at 350 pixels input resolution. (a) Reference image, (b) downscaled, (c) downscaled superimposed, (d) Naïve, (e) Gaussian, and (f) Gaussian sharpened. Note how, for instance, the Naïve method (d) has severe aliasing. The methods containing the Gaussian filter have less aliasing as this method has filtered out some of the highest frequency components.

increased resolution should result in both an improved visual experience of the image, and preservation of more details from the input image. For this reason, we divide our investigation into two parts, objective detail preservation and subjective visual preference.

For detail preservation, we have generated three images provoking different types of image artifacts. The single pixel detail loss in the cross image is detected by the metrics ESSIM, SR-SIM, FeatureSIM, and MSSSIM. These metrics have been designed to analyze the structure in the image and evaluate the structural similarity between the original reference image and the target image. These algorithms should then be ideal for detecting the distortion in the objective detail preservation, and they perform well for the single pixel detail loss. The two other test images test pattern preservation and provoke errors that are more visible in the frequency domain, and that seems hard to detect in the spatial domain. All of the metrics fail to detect both the loss of line-pairs in the line pair image and the added aliasing in the H-frequency image. It is apparently not enough to analyze the structural similarity to detect the errors in this case. This may be because the local contrast goes down, even when we still may distinguish the different line-pairs in the line pair image. This contrast shift may trick the structural metrics. The same shift in contrast may also trick the metric when evaluating the H-frequency image. Here we look for aliasing, but one may advocate that the structure in the H-frequency image is quite similar with or without the aliasing.

Many of the other metrics favor the downscaled version more when evaluating the synthetic images, and we may argue that the downscaled image is more similar to the original image as the superimpositioning is adding some noise and blur in the image. However, the metrics that rate the downscaled image higher are not suitable in the superimpositioning case as we are looking for a metric that evaluates the superimpositioning way of enhancing resolution and that differentiates different ways of superimposing.

As seen in the Z-scores[27] in Fig. 7, there are some noticeable differences in the subframe generation methods. The downscaled method is always the worst rated, showing that all of the superimpositioning methods are increasing the perceived quality in all of the images. The next algorithm is the downscaled superimposed. The fact that this algorithm is better rated than the regular downscaled algorithm shows that superimposing two equal images and by that removing

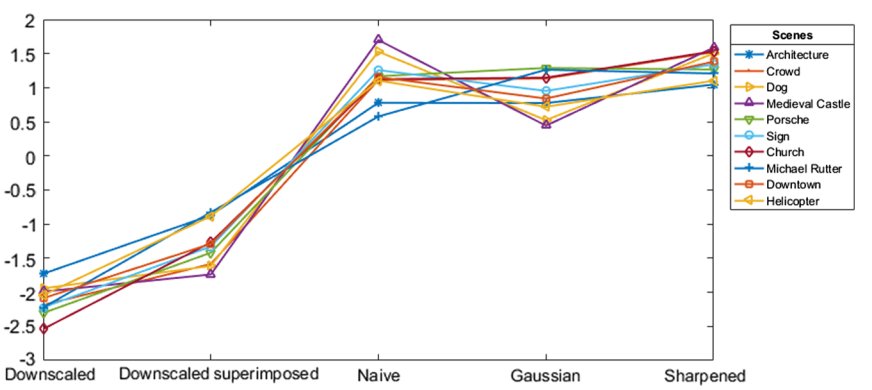

Fig. 7 Z-scores from the subjective test of the superimposed images. Notice how different algorithms give better results for different images. There is no universal best algorithm for all images.

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Electronic-Imaging on 22 Mar 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Table 1 Pearson correlation coefficients.

Scene	DCTex	ESSIM	Feature Sim	MSSSIM	PSNR	PSNR–HVS	PSNR–HVSM	SR–SIM	SSIM	VIF	VSNR
Architecture	0.59	−0.61	−0.38	0.39	−0.75	0.90	0.89	−0.52	−0.09	0.97	0.52
Crowd	0.49	−0.25	−0.11	0.54	−0.57	0.90	0.88	−0.10	0.12	0.97	−0.25
Dog	−0.38	0.86	0.89	0.92	0.22	0.84	0.86	0.90	0.11	0.99	1.00
Medieval castle	−0.36	0.81	0.88	0.78	0.24	0.95	0.96	0.94	0.42	0.97	1.00
Porsche	0.06	0.68	0.82	0.78	−0.01	0.86	0.84	0.71	0.25	0.96	0.89
Sign	−0.20	0.23	0.37	0.70	−0.03	0.96	0.95	0.40	0.26	0.99	0.98
Church	−0.17	0.49	0.41	0.73	0.20	0.93	0.92	0.39	0.21	0.96	0.94
Michael Rutter	0.54	0.17	0.15	0.31	−0.59	0.70	0.68	0.06	0.01	0.88	0.54
Downtown	−0.24	0.92	0.96	0.61	0.21	0.94	0.94	0.97	0.37	0.98	0.97
Helicopter	0.16	0.13	−0.32	0.42	−0.55	0.83	0.82	0.02	0.17	0.88	0.96
Average correlation	0.05	0.34	0.37	0.62	−0.16	0.88	0.88	0.38	0.18	0.95	0.75

Table 2 Spearman correlation coefficients.

Scene	DCTex	ESSIM	Feature SIM	MSSSIM	PSNR	PSNR–HVS	PSNR–HVSM	SR–SIM	SSIM	VIF	VSNR
Architecture	0.1	−0.79	0.1	0.7	−0.5	0.9	0.9	0.1	0	0.9	0.9
Crowd	0.5	0.05	0.3	0.3	−0.1	0.8	0.6	0.3	−0.3	0.8	−0.2
Dog	0.1	0.78	0.8	0.87	−0.3	0.8	0.8	0.8	−0.3	0.8	0.9
Medieval castle	0.1	0.56	0.6	0.6	−0.4	0.8	0.8	0.6	−0.1	0.8	0.9
Porsche	0.3	0.37	0.6	0.5	−0.3	0.5	0.5	0.6	−0.4	0.7	0.6
Sign	−0.4	0.56	0.7	0.9	−0.1	1	1	0.7	0	1	1
Church	−0.5	0.41	0.5	0.6	−0.1	0.6	0.7	0.5	−0.3	0.9	0.7
Michael Rutter	0.7	−0.1	−0.1	0.1	−0.3	0.5	0.5	−0.1	−0.4	0.6	0.3
Downtown	0	0.9	0.9	0.6	0	0.8	0.8	0.9	0	0.9	0.9
Helicopter	0.1	−0.1	−0.5	0.6	−0.4	0.8	0.6	−0.1	−0.1	0.5	0.9
Average correlation	0.1	0.27	0.39	0.58	−0.25	0.75	0.72	0.43	−0.19	0.79	0.69

both the screen-door effect and blurring out some of the sharp and jaggy edges may produce a visually more pleasing image. The three best algorithms are the algorithms that generate different contents for the two spatially shifted positions. This shows that to really utilize the potential of the shifted superimpositioning one should have different contents for the two positions, and consider the spatial shift in the algorithm itself when generating the subframes.

The top three subframe generation methods also have some independent differences. The Naïve method is picking pixels in a way that ensures sharpness in the image, but it also lose a lot of information. Images with sharp details without straight geometric lines like the stray hairs on the dog in Fig. 1(b), the random-shaped stones in the Medieval castle in Fig. 1(g), and the braking waves on the Helicopter image in Fig. 1(h) are benefiting from this technique. But when looking at images with geometric structures like the buildings in architecture in Fig. 1(d) and the text in the Michael Rutter image in Fig. 1(f) and Porsche in Fig. 1(e), the missing details that your mind still knows are there probably influence our view of these images. These images are then visually better enhanced with the Gaussian method that

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Electronic-Imaging on 22 Mar 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

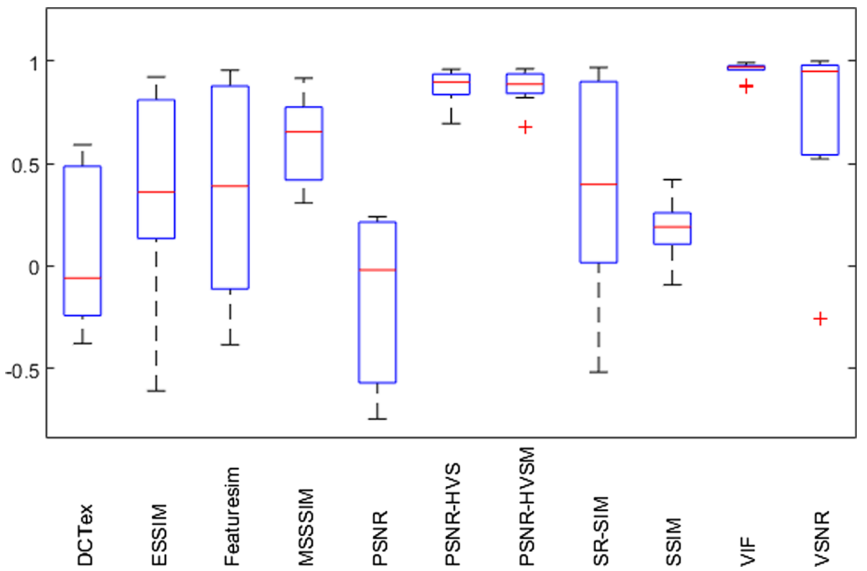

Fig. 8 Pearson correlation coefficient for the different IQMs and the *z*-score.

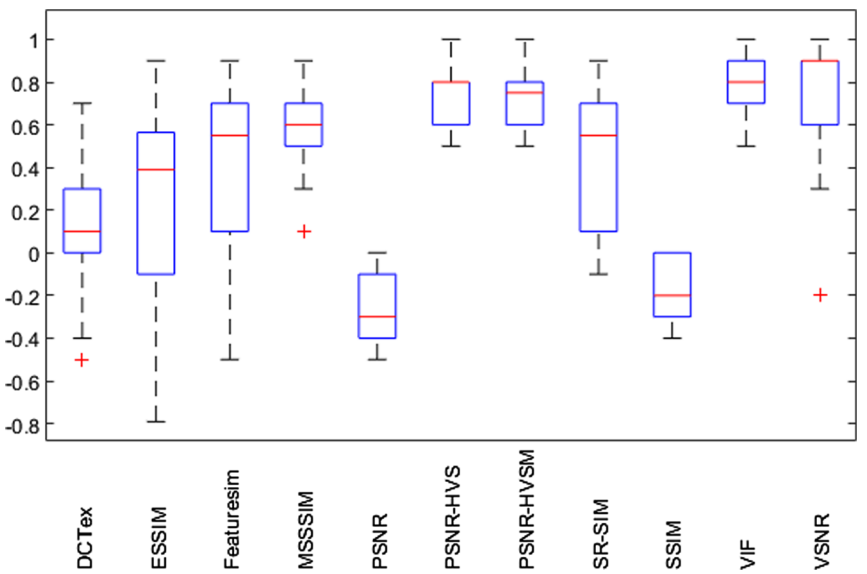

Fig. 9 Spearman correlation coefficient for the different IQMs and the *z*-score.

is low-pass filtering the high-resolution image a bit just to ensure that the details in the missing pixels are not completely lost. The Gaussian-sharpened algorithm is combining these two approaches by sharpening the edges in the Gaussian method. The Gaussian-sharpened algorithm is the method that has the highest overall score, but it is not the best algorithm for all contents. It seems that the most suitable algorithm is dependent on the structure of the image, and on what type of geometry the details is made up from.

For the visual preference, we see that several of the metrics do well. VIF have the best results in this experiment, with PSNR-HVS and PSNR-HVSM close behind. All of these metrics have been designed with the HVS in mind, and they correlate best with our subjective assessment of the quality in the natural images. Most of the structural metrics do not perform as well as the HVS metrics. This may be because the HVS is a complex system, and the sensibility for structural similarity is just one of the many criteria to look for when matching against subjective quality.

From Fig. 10, we see that the VIF metric fits very well with the observer ratings of the different test images. We also notice that for most of the images, both the VIF metric and the observers rate the three methods based upon the Naïve approach close to each other. Especially the last two methods are close to each other in the quality, like we see, for instance, in the graph for the dog image in Fig. 10.

Most IQMs have been designed to meet special requirements, for instance, to detect degradation in specific elements of the image. The requirement we have in this work is to rate the image enhancement of different superimpositioning methods against each other. We see that the metrics that are performing well in rating the algorithms with regard to visual preference take the HVS into account,

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Electronic-Imaging on 22 Mar 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

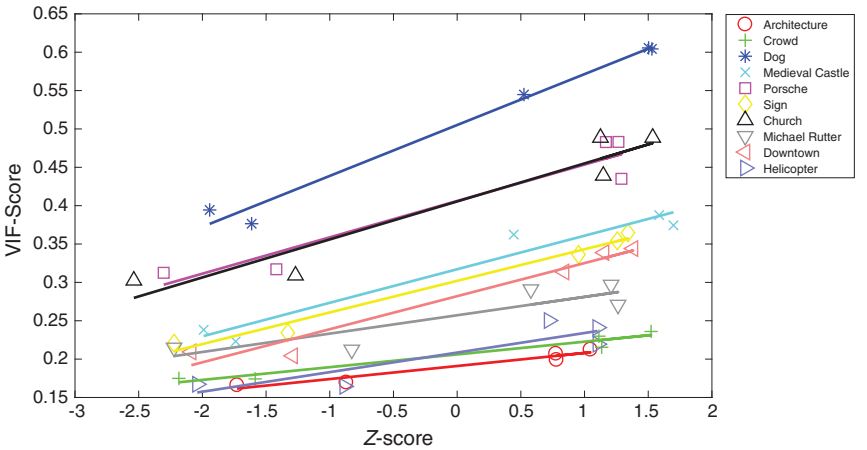

Fig. 10 VIF score for each image plotted against the respective *Z*-scores. Solid lines indicate the fitted linear regression curve.

whereas the metrics that are better at picking up detail preservation analyze the structure in the image. It is important to pick a metric from the correct category to evaluate the case you are looking at. As shown in the results section, metrics that are good at rating the visual preference may not be as good to evaluate the objective distortions and vice versa. Using the wrong metric may very well lead to false results.

8 Conclusion and Further Work

We have evaluated several image quality metrics to assess which metric is most suitable to evaluate different methods of generating superimposed images for enhancing the resolution in projector systems. Of the metrics tested, none of the metrics is covering all of the criteria. But when partitioning the problem into finding a metric to evaluate objective distortion in synthetic images and a different metric to rate natural scenes subjectively, we find that VIF correlates well with our subjective preferences, whereas some of the structural metrics are good at picking up single pixel defects in synthetic images. However, all of the metrics included in this survey fail in detecting loss of line-pairs and also fail in detecting aliasing introduced in high-frequency patterns.

Different applications have different image features that are most important. For the application where the detail preservation in line pairs and high-frequency content is crucial, we should develop new methods for evaluating the image. These methods may include analysis in the frequency domain to detect the pattern deviation.

It would also be valuable to find a way to utilize these IQMs in practical applications and real setups. This introduces some challenges as we need to standardize the oversampling factor on the captured image and on how to get the reference image and the captured image into the same resolution or pixel domain for comparison. Resizing this in software may introduce some deviances, and also other physical factors that influence the HVS, such as brightness, should also be taken into account.

Another domain to explore is video sequences instead of still images. It would be valuable to investigate how the shifted superimposition method performs on moving objects and how to evaluate this performance.

Acknowledgments

The Research Council of Norway, in cooperation with Barco Fredrikstad, funds the research presented in this paper (project HiLase 245569).

References

1. Y. Takahashi, T. Nomura, and S. Sakai, "Optical system and characteristics of an LCD projector with interleaved pixels using four LCD projectors," *IEEE Trans. Circuits Syst. Video Technol.* **5**, 41–47 (1995).
2. C. Jaynes and D. Ramakrishnan, "Super-resolution composition in multi-projector displays," in *Proc. IEEE Int. Workshop on Projector-Camera Systems (ProCams)* (2003).
3. W. Allen and R. Ulichney, "47.4: invited paper: wobulation: doubling the addressed resolution of projection displays," *SID Symp. Dig. Tech. Pap.* **36**(1), 1514–1517 (2005).
4. A. Said, "Analysis of systems for superimposing projected images," Tech. Rep. HPL-2006-129, Hewlett-Packard (2006).
5. N. Damera-Venkata and N. L. Chang, "Realizing super-resolution with superimposed projection," in *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–8 (2007).
6. T. Okatani, M. Wada, and K. Deguchi, "Study of image quality of superimposed projection using multiple projectors," *IEEE Trans. Image Process.* **18**, 424–429 (2009).
7. B. Sajadi, M. Gopi, and A. Majumder, "Edge-guided resolution enhancement in projectors via optical pixel sharing," *ACM Trans. Graphics* **31**, 1–122 (2012).
8. B. Sajadi et al., "Image enhancement in projectors via optical pixel shift and overlay," in *IEEE Int. Conf. on Computational Photography (ICCP)*, pp. 1–10 (2013).
9. Z. Wang et al., "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.* **13**, 600–612 (2004).
10. F. Heide et al., "Cascaded displays: spatiotemporal superresolution using offset pixel layers," *ACM Trans. Graphics* **33**, 60:1–60:11 (2014).
11. E. Barshan et al., "35.3: resolution enhancement based on shifted superposition," *SID Symp. Dig. Tech. Pap.* **46**(1), 514–517 (2015).
12. S. A. J. Hansen, J. Y. Hardeberg, and M. N. Akram, "Resolution enhancement through superimposition of projected images—how to evaluate the quality?" in *Electronic Imaging, Image Quality and System Performance XIV*, pp. 141–146, Society for Imaging Science and Technology (2017).
13. K. Egiazarian et al., "A new full-reference quality metrics based on HVS," in *Proc. of the Second Int. Workshop on Video Processing and Quality Metrics* (2006).
14. N. Ponomarenko et al., "On between-coefficient contrast masking of DCT basis functions," in *Proc. of the Third Int. Workshop on Video Processing and Quality Metrics* (2007).
15. X. Zhang et al., "Edge strength similarity for image quality assessment," *IEEE Signal Process. Lett.* **20**, 319–322 (2013).
16. L. Zhang et al., "FSIM: a feature similarity index for image quality assessment," *IEEE Trans. Image Process.* **20**, 2378–2386 (2011).
17. F. Zhang et al., "Practical image quality metric applied to image coding," *IEEE Trans. Multimedia* **13**, 615–624 (2011).
18. H. R. Sheikh and A. C. Bovik, "Image information and visual quality," in *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Vol. 3, pp. 709–712 (2004).

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Electronic-Imaging on 22 Mar 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

19. L. Zhang and H. Li, "SR-SIM: a fast and high performance IQA index based on spectral residual," in *19th IEEE Int. Conf. on Image Processing*, pp. 1473–1476 (2012).
20. Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Thirty-Seventh Asilomar Conf. on Signals, Systems Computers*, Vol. 2, pp. 1398–1402 (2003).
21. N. Damera-Venkata et al., "Image quality assessment based on a degradation model," *IEEE Trans. Image Process.* **9**, 636–650 (2000).
22. D. M. Chandler and S. S. Hemami, "VSNR: a wavelet-based visual signal-to-noise ratio for natural images," *IEEE Trans. Image Process.* **16**, 2284–2298 (2007).
23. BARCO, "White paper: 4k UHD explained," Tech. Rep., BARCO, https://www.barco.com/en/page/2016/entertainment/downloads/wp4kuhd/downloadpage (20 February 2018).
24. K. Bylsma, "Can you explain in detail how e-shift works?" 2014, http://usjvc.com/faq/index.php?action=artikel&cat=20&id=268 (20 February 2018).
25. X. Liu, M. Pedersen, and J. Hardeberg, "CID:IQ—a new image quality database," *Lect. Notes Comput. Sci.* **8509**, 193–202 (2014).
26. K. V. Ngo et al., "QuickEval: a web application for psychometric scaling experiments," *Proc. SPIE* **9396**, 93960O (2015).
27. P. Engeldrum, *Psychometric Scaling: A Toolkit for Imaging Systems Development*, Imcotek Press, Winchester, Massachusetts (2000).

Svein Arne Jervell Hansen received his master's degree in electronics in 2007 at the Norwegian University of Science and Technology. Currently, he is a PhD student at the University of South-Eastern Norway, and he is in addition to this working at Barco Fredrikstad in Norway. His research is focusing on resolution enhancement of projected displays.

Muhammad Nadeem Akram received his PhD in photonics from Royal Institute of Technology, Stockholm, Sweden, in 2005. He is a professor at the University of South-Eastern Norway. His research interests are semiconductor optoelectronics, vacuum electronics, imaging optics, speckle reduction, laser projectors, and human visual system modeling. He has published more than 100 articles in scientific journals and conferences.

Jon Yngve Hardeberg received his MSc degree in signal processing from NTNU, and his PhD in signal and image processing from Ecole Nationale Supérieure des Télécommunications in Paris, France. He is a professor at the Norwegian Color and Visual Computing Laboratory at the faculty of computer science and media technology at NTNU in Gjøvik. His current research interests include multispectral color imaging, print and image quality, colorimetric device characterization, color management, medical imaging, and cultural heritage imaging, and he has coauthored more than 200 scientific publications.

Marius Pedersen received his PhD in color imaging from the University of Oslo, Norway, in 2011. He is a professor at the Norwegian University of Science and Technology (NTNU), Norway. His work is centered on image quality assessment; he has more than 60 publications in this field. Currently, he is the head of the computer science group in Gjøvik in the Department of Computer Science, as well as the head of the Norwegian Colour and Visual Computing Laboratory, both at NTNU.

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Electronic-Imaging on 22 Mar 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use