Resolution enhancement through superimposition of projected images – How to evaluate the quality?

Svein Arne Jervell Hansen^{1,2}, Jon Yngve Hardeberg³, Muhammad Nadeem Akram¹

¹ University College of Southeast Norway, Borre, Norway

² Barco Fredrikstad AS, Fredrikstad, Norway

³ Norwegian University of Science and Technology, Gjøvik, Norway

Abstract

Increased demand for high-resolution projection displays makes the projector industry search for ways of enhancing the resolution above the native resolution of the projector's image panel. Resolution enhancement through superimposition is one method of enhancing the resolution that has gained popularity in the industry the last couple of years. This method consists of shifting every other projected frame spatially with sub-pixel 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 to the native resolution, and how high the effective resolution gain really is. To determine which application the superimposition method is best suited for, it is also interesting to look at how this method performs over different kinds of image and video content.

To help investigate these questions we have developed a simulator that simulates different superimposition methods over several classes of image content. The superimposed images are then evaluated by several image quality metrics with the goal of finding out which quality metrics are most applicable to the superimpositioning case.

We found that the MSSSIM metric is the most suitable to evaluate superimposed images. VIF also performs quite well, but MSSSIM performs slightly better. However, none of these metrics identifies all the artefacts introduced by the superimpositioning. More research is needed to develop an ideal metric.

Introduction

Resolution is one of the key performance parameters of a projector, and the projector industry continuously aims to increase the resolution. 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). This method is gaining momentum in the industry as an adopted method. Superimposition may be implemented either with a multi-projector 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.

Resolution enhancement currently has some momentum because of the market drive for 4K images and video. Not all SLM technologies have cost efficient 4K modulators available. For these kind 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 artefacts in the image. Since the optical overlap of superimposed images acts like a low-pass filter, some high frequency content is lost in the image. The spatial artefacts manifest as blurring in the image, and these artefacts impacts both the visual quality and the resolution measurements. The introduced artefacts raises the question if the resulting image on the wall really consist of 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 will investigate different methods of superimposition and explores how these methods compare to each other in quality. 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 and experimental setup" section describes the experimental setup that is used to tests different metrics and superimposition methods. The simulated results and the discussion of these are presented in the following section. Finally, the last section concludes on how to evaluate the resulting image quality for this particular application and the future work.

Related work

Takahashi et al. [1] proposed a setup in 1995 with four LCD projectors projecting on the same screen with an elaborate mirrorsetup. 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 vertically. This setup is very cumbersome and requires careful adjustment in the installation phase. Over time, the fill factor of LCD panels have also increased, leaving one of the main prerequisites of this method obsolete. Takahashi et al. use MTF as a main parameter to evaluate the resolution enhancement. The MTF is obtained in this case through optical simulations

Jaynes et. al. [2] proposed a system where several projectors project at the same screen, and then they are calibrated to determine the relative sub-pixel shift for each projector. The goal of this calibration is to derive an accurate mapping of each projectors 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. verify their work by displaying images showing the quality improvement.

Allen and Ulichney [3] made a breakthrough with their idea to keep the whole system within one projector unit, and instead include an opto-mechanical image shifter to shift every n'th 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 the 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 blurry limage. The authors present the gained image quality as visual results side-by-side, and they do not quantify the quality gain.

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 uses PSNR and visual representation for evaluating the quality.

Damera-Venkata and Chang [5] proposed the year after a method to produce superimposed images through multi-projector 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 visual inspection of the superimposed results.

Okatani et al. [6] explored the theory from Damera-Ventaka 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 visual inspection of the resulting images, and no quality metric is used.

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, but the quality evaluation is determined only by visual inspections of images taken from the test setup.

The year after, Sajadi et al [8] proposed a low-cost approach which shifts the whole image with sub-pixel 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. do not timemultiplex 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 cost-efficient. The quality gain of this method is quantified through the SSIM metric, and they use the CIELAB ΔE to check if the colors have drifted. Sajadi et al also evaluates the content preservation in the image by calculating Histogram of Gradients (HOG) for different combinations of pixelshift and numbers of superimposed frames.

Heide et al. [9] 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 sub-pixel 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 sub-images is added on top of each other. This method apparently provides good results, which is verified by visual inspection, PSNR, SSIM, and MTF analysis.

Barshan et al. [10] proposed their own superimposition scheme in 2015 named Shifted Superposition (SSPOS). This method is quite similar to the wobulation method proposed by Allen and Ulichney [3], but the generation of the sub-images are done in a more elaborate way. The quality improvement in this work is verified by visual inspection and by using the MSSIM metric as well.

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 images representing the visual gain of the superimpositioning, but this is a poor method for comparing different algorithms objectively. This section will look briefly into the different quality metrics mentioned in the previous section, and present some other quality metrics that may also be used.

Since we do 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 Human Visual System (HVS) inspired metrics.

The error-based calculations are mathematical metrics based on error quantification between two images. They are popular since they are simple to understand, easy to use, and have a low computational cost. Typical examples of these metrics are Mean Square Error (MSE) and different versions of Signal to Noise Ratio (SNR). SNR and Peak SNR (PSNR) are based on the principle that the distorted image consists of the original image and a noise component in addition as an independent signal. SNR is defined as the ratio of average signal power to noise signal power. These raw mathematical error-based calculations have their limitations in that they often do not correlate well with subjective quality assessments.

To make up for these shortcomings, the Weighted SNR (WSNR) was developed to take the HVS contrast sensitivity function into account [11]. 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, since it is an error-based metric (SNR) modified slightly by using some of the HVS attributes. Other metrics like PSNRHVS [14] and PSNRHVSM [15] use the principles from PSNR and modifies this metric based on the frequency based contrast sensitivity of the HVS.

Pure HVS inspired metrics takes the attributes of the HVS into account and aims to measure specific image attributes that the HVS is particularly sensitive to. SSIM [12] 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 have fostered several quality metrics such as Multi scale SSIM [13] (MSSSIM), ESSIM [16], SRSIM [17], Feature-SIM [18] (FSIM), DCTex [19], VIF [20] and VSNR [21].

Methodology and experimental setup

In this research, we concentrate on verification through simulation, thus the entire setup is carried out within a simulation environment written in Matlab.

We have implemented five methods of displaying the resulting image in this simulator. *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. *Downscaled superimposed* – this method generates the sub-images as the downscaled method, but then these sub-images are spatially shifted and superimposed on itself. It is not an ideal method, but 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. *Naïve* – in this method we upscale the input image to the double horizontal and vertical resolution of the SLM, then we pick the pixels for the different sub-frames directly from the up-scaled frame. The Naïve method produces quite sharp images, but some details will be lost since we just pick every other pixel. Gaussian - this is the same method as the Naïve method, but in addition we have filtered the upscaled image with a Gaussian filter. By doing this we produce an image that is slightly more blurred, but we will not lose as many details as in the Naïve method. Gaussian Sharpened - this 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 adding a bit more randompatterned noise in the image. We are not aiming to develop the best method for superimposing images in this paper, so we have picked some methods that are distinguishable from each other, with different properties.

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 experimentations. We have not investigated into more than two positions or other techniques, but only additive superimpositioning in our experiments.

We have used five different test-images to test different image properties (see Figure 1). *Lenna* – this is a classical natural image to see how the metric performs on natural images. *Hair* – a natural



Figure 1 The test images used from upper left: Lenna, Hair, Cross, Linepairs and H-frequency.

image with lots of details. More high-frequency content than Lenna. Cross - a white cross on a black background with single pixel diagonals. This is included to see how the metrics pick up distortion of single pixel details. *Linepairs* – a synthetic image consisting of three line pairs in horizontal direction and three line pairs in vertical direction. 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 metrics included in this setup are the following; PSNR is one of the most widely used error calculation metrics. For metrics taking the human visual system into account, we have included the metrics PSNR-HVS, PSNR-HVS-M, ESSIM, Feature-SIM (FSIM), DCTex and VIF. In addition to these categories, we also have used metrics that are purely looking at the structure in the image, like SSIM, SRSIM and MSSSIM.

In our work, we have defined an SLM with the resolution of 250x250 pixels. We have chosen to keep the resolution low here for keeping the computational time down. We have then scaled the input resolution in 25 pixel steps from 225x225 to 600x600 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 and above double of the native resolution.

Simulated results

The natural images are clearly improved by the superimpositioning method, and we rate the image quality from different methods in the following order from best to worst: Gaussian Sharpened, Gaussian, Naïve, Downscaled Superimposed and Downscaled. This is based on a visual assessment from a small group of people. The downscaled superimposed and the downscaled methods are head-to-head sometimes, since the blur added by the superimpositioning is dominant in some cases. Several of the metrics rate the different methods in this order, like PSNR, SSIM, VSNR, MSSSIM and VIF. The rest of the metrics tend to weigh the methods a bit differently, and some of them, especially DCTex, appreciates the similarity of the downscaled image much more than we rate this image subjectively, and score the downscaled image quite high. Figure 2 shows the DCTex results for the natural image Lenna, illustrating how some methods rate the downscaled image very high.



Figure 2 DCTex results from the image Lenna, lower values better. Notice how DCTex rates different methods, compared with MSSSIM results in Figure 6.

For the synthetic images we have concrete parameters to look for. The *Line pairs* image have three distinguishable line pairs that will eventually melt together when the input/output ratio gets too high. The goal of the superimpositioning is to preserve the details in the image at frequencies above the spatial frequency of the SLM, so the superimpositioning methods should preserve the line pairs for higher resolutions than the downscaling method. In addition, we are looking for metrics that pick up when we lose line pairs in the different superimpositioning methods. The different methods perform as following; Downscaling and Downscaling superimposed both lose one line pair when the input resolution go 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. 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 picks up these details, and some of the metrics even rate the two worst methods as the two best ones. Again, we note the preference for the Downscaling and the downscaled Superimposed methods because they add less blur and preserve more local contrast in the image, even though they lose details in the image.



Figure 3 Zoomed in on the resulting line pairs at 300 pixels input resolution. A) Reference image. B) Downscaled. C) Downscaled superimposed. D) Naive. E) Gaussian. F) Gaussian sharpened.

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 non-shifted direction. This diagonal gets worse at higher resolutions and is completely lost at 500 pixels and above.

The loss of details is visualized in Figure 4, showing how the *Naïve* superimpositioning looks with and input resolution of 300x300, 400x400, 450x450 and 500x500 pixels. The SLM resolution is in this case kept at 250x250 pixels.



Figure 4 Results from Naive superimpositioning with A) 300, B) 400, C) 450 and D) 500 pixels input resolution.

Several of the metrics do not seem to care about this severe loss of details, but ESSIM, SR_SIM, FEATURESIM, VIF and MSSSIM picks this up. Figure 5 illustrates how SSIM metric misses the degradation of quality in the case of *Naive* superimpositioning. In this case, MSSSIM detects the loss of details in the *Naïve* method (Figure 6) but the SSIM method does not detect this defect.



Figure 5 SSIM results from the image Cross. Notice how SSIM do not detect the loss of detail in the Naïve method, compared with MSSSIM results in Figure 6.

The synthetic *H-frequency* image is hard for the superimpositioning methods to represent correctly when the input resolution increases, and it also introduces aliasing in some cases. We see in the visual results 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 metrics introduce quite severe aliasing in some instances.



Figure 6 MSSSIM results from all five test images.

Discussion

To determine which metric is the best metric 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 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, visual preference and detail preservation.

For the visual preference we see that several of the metrics follow our subjective preferences. PSNR, SSIM, VSNR and VIF correlate with our subjective assessment of the quality in the natural images. Most of the other metrics favor the downscaled version more than we do, and we may argue that the downscaled image is more similar to the original image since 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, since we are looking for a metric that appreciates the superimpositioning way of enhancing resolution and that differentiates the different ways of superimposing.

For detail preservation, we have generated three images provoking different type of image artefacts. The single pixel detail loss in the *Cross* image is picked up by the metrics ESSIM, SR_SIM, FEATURESIM, VIF and MSSSIM. The two other test images tests pattern preservation and provokes errors that are more visible in the frequency domain, and that seems hard to pick up in the spatial domain. All of the metrics fail to detect both the loss of line-pairs in the *Linepair* image and the added aliasing in the *Hfrequency* image.

Most image quality metrics have been designed to meet special requirements, for instance to detect degradation in specific elements of the image. The requirements we have to rate the image enhancement in different superimpositioning methods against each other is not covered entirely by the existing metrics. To cover both the visual resolution enhancement and the image detail preservations, the metric should look both at the spatial properties of the image, and in addition analyze the image in the frequency domain to look for line preservation and aliasing introduced.

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 the projected systems. Of the metrics tested, MSSSIM is the preferred one since this metric both rates the natural images in the preferred order, and detects loss in single-pixel details. VIF is also quite good, but MSSSIM is a bit better at detecting the single pixel defects. However, all of the metrics included in this survey fails in detecting loss of line-pairs and also fail in detecting aliasing introduced in high frequency patterns.

Different applications have different image features that is 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 should include analysis in the frequency domain to detect the pattern deviation.

We should also perform a full psycho-visual image quality study to verify the subjective ratings. The subjective ratings presented in this paper are performed with too few participants, with no statistical analysis and are not sufficiently trustworthy to make strong conclusions.

Acknowledgments

The image quality metrics in this paper are calculated in Matlab using code provided by Marius Pedersen at NTNU.

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

References

- Y. Takahashi, T. Nomura, and S. Sakai, "Optical System and Characteristics of an LCD Projector with Interleaved Pixels Using Four LCD Projectors," Transactions on Circuits and Systems for Video Technology, IEEE, vol. 5, no. 1, pp. 41-47, Feb 1995.
- [2] C. Jaynes and D. Ramakrishnan, "Super-resolution composition in multi-projector displays" in Proc. IEEE International Work- shop on Projector-Camera Systems (ProCams), Nice, France, Oct 2003.
- [3] W. Allen and R. Ulichney, "Wobulation: Doubling the ad- dressed resolution of projection displays," in Proc. SID Sympo- sium Digest of Technical Papers (SID), Boston, MA, pp. 1514–1517, May 2005.
- [4] A. Said, "Analysis of Systems for Superimposing Projected Images," HP Tech. Rep. HPL-2006-129, Oct 2006.
- [5] N. Damera-Venkata and N. L. Chang, "Realizing Super-Resolution with Superimposed Projection," IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, pp. 1-8, 2007
- [6] T. Okatani, M. Wada, K. Deguchi, 2009 "Study of Image Quality of Superimposed Projection Using Multiple Projectors" IEEE Transactions on image processing, vol. 18, no. 2, pp. 424-429 Feb 2009
- [7] B. Sajadi, M. Gopi, and A. Majumder. "Edge-guided resolution enhancement in projectors using optical pixel sharing." ACM TOG, 2012.
- [8] B. Sajadi, D. Qoc-Lai, A. Ihler, M. Gopi, A. Majumder, "Image enhancement in projectors via optical pixel shift and overlay" In IEEE International Conference on Computational Photography (ICCP), pp. 1–10. 2013
- [9] F. Heide, D. Lanman, D. Reddy, J. Kautz, K. Pulli, D. Luebke, "Cascaded Displays: Spatiotemporal Superresolution using Offset Pixel Layers" ACM Transactions on Graphics (Proceedings of SIGGRAPH), Aug 2014
- [10] E. Barshan, M. Lamm, C. Scharfenberger, P. Fieguth "Resolution Enhancement Based on Shifted Superposition" SID Symposium Digest of Technical Papers, 2015
- [11] N. Damera-Venkata, T. D. Kite, W. S. Geisler, B. L. Evans, and A. C. Bovik, "Image quality assessment based on a degradation model," IEEE Trans. Image Processing, vol. 9, pp. 636–650, 2000.
- [12] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity" IEEE Trans. Image Processing, vol. 13, no. 4, pp. 600–612, 2004.
- [13] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in Asilomar Conference on Signals, Systems and Computers, IEEE, Pacific Grove, CA, vol. 2, pp. 1398–1402, Nov 2003.
- [14] K. Egiazarian, J. Astola, N. Ponomarenko, V. Lukin, F. Battisti, M. Carli, New full-reference quality metrics based on HVS, CD-ROM Proceedings of the Second International Workshop on Video Processing and Quality Metrics, Scottsdale, USA, 2006.

- [15] N. Ponomarenko, F. Silvestri, K. Egiazarian, M. Carli, J. Astola, V. Lukin, "On between-coefficient contrast masking of DCT basis functions", CD-ROM Proceedings of the Third International Workshop on Video Processing and Quality Metrics for Consumer Electronics VPQM-07, Scottsdale, Arizona, USA, Jan 2007.
- [16] X. Zhang, X. Feng, W. Wang, and W. Xue, "Edge strength similarity for image quality assessment," IEEE Signal Processing Lett, vol. 20, no. 4, pp. 319–322, Apr 2013.
- [17] L. Zhang and H. Li, "SR-SIM: A fast and high performance iqa index based on spectral residual," in International Conference on Image Processing, pp.1473–1476, Sep 2012,
- [18] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," IEEE Trans. Image Processing, 2011.
- [19] F. Zhang, L. Ma, S. Li, and K. N. Ngan, "Practical image quality metric applied to image coding," IEEE Trans. On Multimedia, vol. 13, no. 4, pp. 615–624, Aug 2011.
- [20] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," IEEE Trans. Image Processing, vol. 15, no. 2, pp. 430–444, 2006.
- [21] D. M. Chandler and S. S. Hemami, "VSNR: A waveletbased visual signal-to-noise ratio for natural images," IEEE Trans. Image Processing, vol. 16, no. 9, pp. 2284–2298, Sep 2007.

Author Biography

Svein Arne Hansen received his Master in Electronics in 2007 at the Norwegian University of Science and Technology. Currently he is a PhD student at the University College of Southeast Norway, and he is in addition to this working at Barco Fredrikstad in Norway. His research is focusing on resolution enhancement of projected displays.

Jon Yngve Hardeberg is Professor at the Norwegian Color and Visual Computing Laboratory at the Faculty of Computer Science and Media Technology at NTNU in Gjøvik. He has a MSc in Signal Processing from NTNU, and a PhD in Signal and Image Processing from the Ecole Nationale Supérieure des Télécommunications in Paris, France. His current research interests include multispectral colour imaging, print and image quality, colorimetric device characterisation, colour management, medical imaging, and cultural heritage imaging, and he has co-authored more than 200 scientific publications.

Muhammad Nadeem Akram is a professor at the University College of Southeast Norway. His background is in optics and photonics.