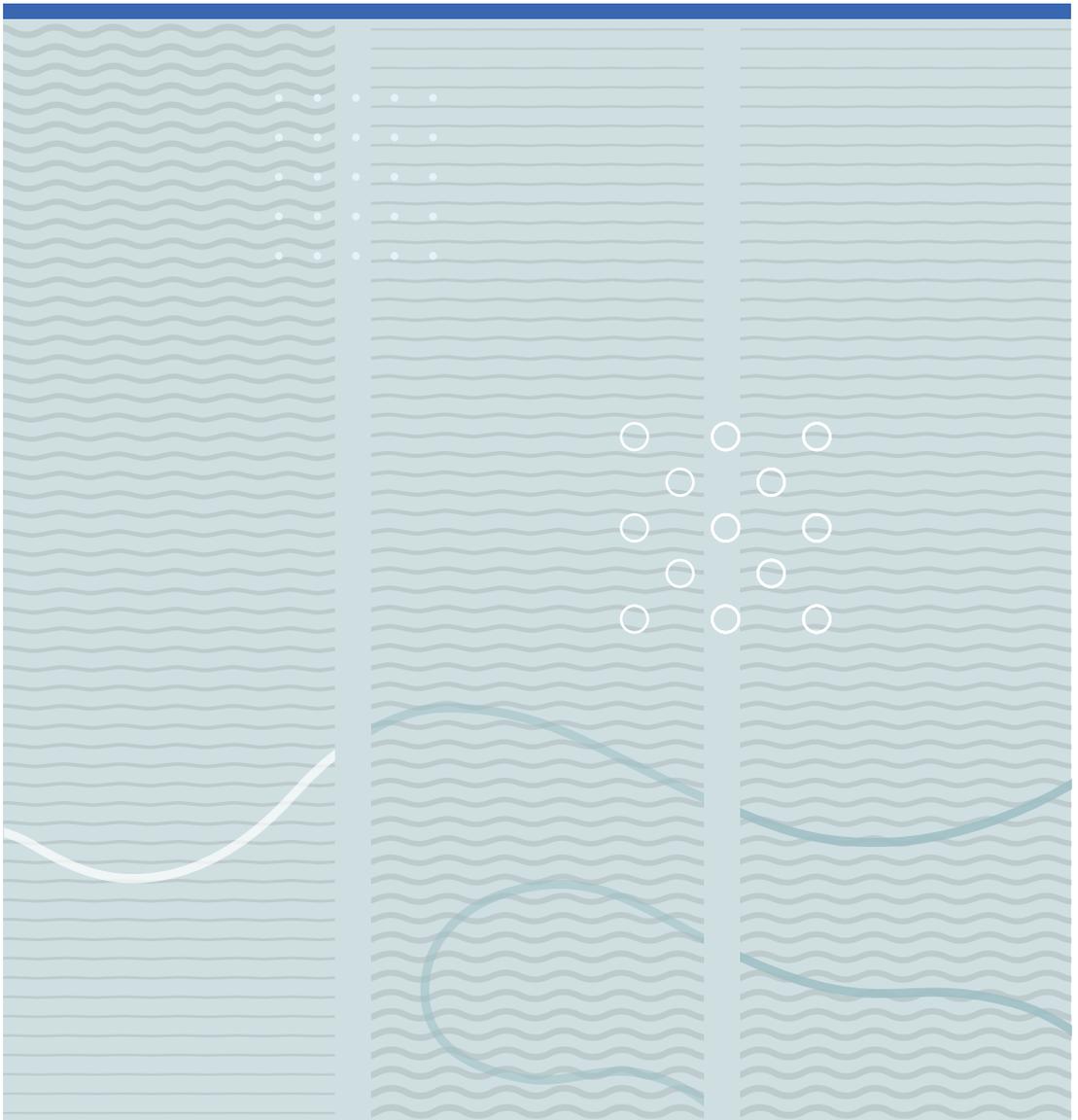Svein Arne Jervell Hansen

Evaluation of resolution enhancement in shifted superimposed projection displays: Simulations and experiments

USN

Svein Arne Jervell Hansen

Evaluation of resolution enhancement in shifted superimposed projection displays: Simulations and experiments

A PhD dissertation in
Applied Micro- and Nanosystems

Dedication

First of all, I would like to thank my main supervisor Professor Muhammad Nadeem Akram for his guidance, feedback and encouragement during this project. I also want to thank my co-supervisors Professor Jon Yngve Hardeberg and Dr. Øyvind Svensen for their support and discussions along the way.

Thanks to Barco and my colleagues in both Fredrikstad and Kortrijk for banter and support, and for all our technical discussions when I need somebody to spar with.

Finally I want to thank my wife, Mona, for always backing me up and to my children Leon and Cassandra for their curiosity and patience.

Preface

This doctoral thesis is submitted in partial fulfillment of the requirements for the degree of Philosophiae Doctor at the Faculty of Technology, Natural Sciences and Maritime Sciences at the University of South-Eastern Norway (USN), Norway, from 1st October 2015 to 1st October 2019.

The PhD work was carried out at the Department of Micro and Nanosystem Technology under the supervision of Professor Muhammad Nadeem Akram (USN), Professor Jon Yngve Hardeberg (NTNU) and Dr. Øyvind Svensen (Barco Fredrikstad).

This PhD work is a part of project HiLase 245569, Next Generation UHD/4K Projection based on Solid-state Illumination, founded by the Research Council of Norway in cooperation with Barco Fredrikstad.

Abstract

Spatial resolution is one of the key performance parameters of a projected display, and the projector industry continuously aims to increase it. Projector resolution is traditionally based on the resolution of the spatial light modulator (SLM), and increasing the projector resolution is typically done by increasing the SLM pixel count. However, increasing the amount of pixels on a single SLM is both cost intensive and technically challenging as not all SLM technologies support a sufficiently high pixel count. For this reason there has been a significant focus in recent years to develop resolution enhancing methods that allow projector resolution to exceed the native resolution of the projector SLM.

Shifted superimposition is one of the popular methods for increasing the resolution of the projected image. This is commonly done by shifting every other frame spatially on the projected screen with sub-pixel precision to form a new pixel grid with finer pixel pitch. Even though this is an established method, there are still an open question of how well this technique performs in comparison to native resolution, and how high the effective resolution gain really is.

In this PhD research we explore different aspects of resolution enhancement through shifted superimposition to gain more knowledge of how this method performs, and how to evaluate the actual resolution gain of this resolution enhancement method. We also review different ways the subframes may be generated, and what impact the shifting direction has on the resulting superimposed displayed image.

Through simulations and subjective observer experiments, we have established that the MSSSIM image quality metric is the most suitable metric to evaluate the simulated superimposed image since the results of this metric corresponds best with our subjective view of a good image. We have used this metric to evaluate different subframe generation techniques, and also as a tool to investigate the impact of the direction of the shift on the displayed image.

In this study we demonstrate the characteristics of the different categories of subframe generation methods, and that the subframe generation method used have a great impact on the overall quality of the superimposed image. Focusing on the shifting direction we establish that the quality of the superimposed image is almost invariant of the direction of the shift as long as the shift is not in horizontal or vertical direction.

The last part of this thesis explores the resolution gain of the shifted superimposition method, and examine the concept of resolution and quality in relation to each other. The research work involves simulations as well as measurements on a super-resolution projector. In this study we prove that the shifted superimposition method enhance the resolution approximately 40% above the native resolution given the desired source resolution.

Keywords: Superimposition, Display, Projector, Resolution enhancement, Image processing

Article list

(Due to publishers' restrictions, the articles are omitted from this version.)

Journal papers:

Article 1

Hansen, S.A.J., Akram, M.N., Hardeberg, J.Y. & Pedersen, M. (2018). Preferred image quality metric for shifted superimposition-based resolution-enhanced images, Journal of Electronic Imaging, 27(3), 1-13. doi:10.1117/1.JEI.27.3.033017

Article 2

Byanju, R., Hansen, S.A.J. & Akram, M.N. (2018). A comparative study of superimposition techniques for enhancing the projector resolution: Simulations and experiments, Displays, 55, 80-89. doi:10.1016/j.displa.2018.11.002

Article 3

Hansen, S.A.J., Akram, M.N. & Hardeberg, J.Y. (2019). Resolution enhancement through shifted superimposition: The influence of shift direction, Ready for submission.

Article 4

Hansen, S.A.J., Akram, M.N. & Hardeberg, J.Y. (2020). The effects of source resolution on resolution enhancement through shifted superimposition projection, Accepted for publication in "Journal of the Society for Information Display".

Conference papers:

Article 5

Hansen, S.A.J, Akram, M.N. & Hardeberg, J.Y. (2017). Resolution enhancement through superimposition of projected images – How to evaluate the quality?, Electronic Imaging, 1, 141-146. doi:10.2352/ISSN.2470-1173.2017.12.IQSP-231

Article 6

Hansen, S.A.J., Akram, M.N. & Hardeberg, J.Y. (2018). Pixel shifting in projected displays, Proceedings of the Image 2018 conference, 1, 62-67.

Article 7

Hansen, S.A.J., Akram, M.N. & Hardeberg, J.Y. (2019). Evaluation of Adaptive Shifted Superimposition Technique for Enhancing the Projector Resolution, Proceedings volume 11145 8th laser display and lightning conference (LDC 2019), doi:10.1117/12.2540345

List of Tables

List of Figures

Abbreviations

CCD	Charge-Coupled Device
CMOS	Complementary Metal–Oxide–Semiconductor
CRT	Cathode-Ray Tube
CSF	Contrast Sensitivity Function
CTF	Contrast Transfer Function
DCT	Discrete Cosine Transform
DLP	Digital Light Processing
DMD	Digital Micromirror Device
DSLR	Digital Single Lens Reflex
DSS	Different Subframe Shifting
ESSIM	Edge-based Structural Similarity
FSIM	Feature Similarity
GBP	Gaussian Belief Propagation
HOG	Histogram of Oriented Gradients
HVS	Human Visual System
ICDM	International Committee for Display Metrology
IDMS	Information Display Measurement Standard
IEEE	Institute of Electrical and Electronics Engineers
IQM	Image Quality Metric
ISO	International Organization for Standardization
JND	Just Noticeable Difference
LCD	Liquid Crystal Display
LCOS	Liquid Crystal on Silicon
LED	Light Emitting Diode
LR	Low Resolution
MSE	Mean Square Error
MSSSIM	Multi-Scale Structural Similarity
MTF	Modulation Transfer Function
PSNR	Peak Signal-to-Noise Ratio
RGB	Red Green Blue
SFR	Spatial Frequency Response
SID	Society for Information Display
SLM	Spatial Light Modulator
SNR	Signal-to-Noise Ratio
SR	Super Resolution
SR-SIM	Spectral Residual based Similarity
SSIM	Structural Similarity Index Metric

SSPOS	Shifted Superposition
SSS	Single Subframe Shifting
VIF	Visual Information Fidelity
VSNR	Visual Signal-to-Noise Ratio
WQXGA	Wide Quad eXtended Graphics Array (2560 x 1600 pixels)
WSNR	Weighted Signal-to-Noise Ratio
XGA	eXtended Graphics Array (1024 x 768 pixels)

Table of Contents

1. Introduction

1.1.   Motivation and focus of this PhD

Spatial resolution is one of the key performance parameters of a projector, and the projector industry continuously aims to increase it. In a projector, the spatial resolution is usually limited by the number of pixels in the spatial light modulator (SLM). If the projector is given a video signal containing a higher resolution than its native SLM resolution, the projector is forced to downscale the video signal and therefore also inevitably lose details in the image.

Shifted superimposition of projected images is a cost effective way of enhancing the resolution above the native resolution of the SLM in a projector [4]. Superimposition may be implemented either with a multi-projector setup as proposed by Takahashi [5] and Jaynes [6], or with an opto-mechanical wobulator within a single projector as introduced by Allen and Ulichney [7]. As long as superimposition consists of two or more images superimposed on one projected surface, the resulting image will be an additive function of the projected subimages.

Resolution enhancement currently has gained momentum because of the market drive for 4K images and video. Some SLM technologies still do not have cost efficient 4K modulators available, and for these modulator technologies it is necessary to have other means for reaching the 4K resolution. Resolution enhancement through shifted superimposition is currently the preferred method for enhancing the resolution above the native SLM resolution. 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 impact both the visual quality and the resolution measurements. The introduced artefacts 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.

Most projectors that utilize the superimposition method today have an optome-chanical actuator that spatially shifts every n^{th} frame with sub-pixel precision [8]. The two most common shift configurations are either half a pixel in one diagonal (two positions) as shown in Figure 1.1, or half a pixel in both diagonals (4 positions) [9].
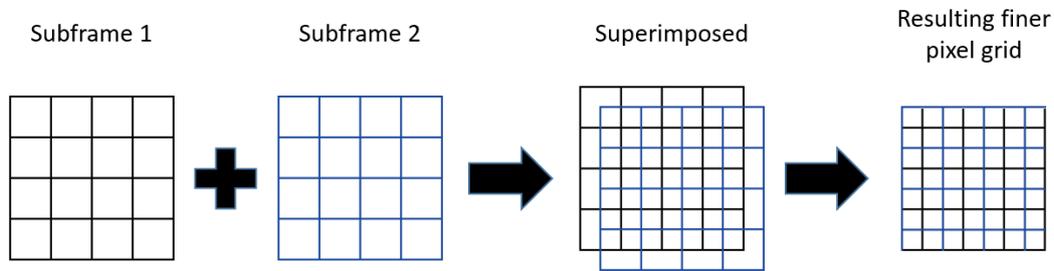
Figure 1.1: Subframe 1 and subframe 2 shifted half a pixel diagonally from each other. The overlap results in a finer sub-pixel grid consisting of approximately twice the amount of pixels in both horizontal and vertical direction. The resulting finer pixel grid is illustrated at the far right with the edges trimmed off. In this illustration we see that the new finer pixel grid has a pixel size of approximately a quarter of the original pixel size.

Since the frames are projected at higher frame rate than the flicker threshold, the two frames at different positions blend together in the human visual system (HVS) and appear as one frame with a new pixel grid. The new pixel grid consists of a larger number of distinct pixels, approximately doubled both horizontally and vertically, with a finer pixel pitch, meaning that the distinct pixels are smaller than those in the un-shifted case. The traditional way of determining resolution is to count the number of pixels at the screen. This pixel count would suggest that the resolution has doubled both horizontally and vertically, but since the overlap has introduced a dependency between certain neighbouring pixels, it is not as straightforward to claim the final effective resolution of the projected image.

This research aims to explore the different aspects of resolution enhancement through shifted superimposition, and to see in what degree this method enhances the resolution of the projected display. The following questions have been raised, and researched:

Q1: How are the projected subimages generated, and what impact do they have on the overall quality of the superimposed image?

Q2: What is the best metric to assess the overall quality of enhanced images? How does it match with subjective experiments with human observers?

Q3: Is the typical diagonal shift the ideal shifting direction?

Q4: How much resolution is gained with this method, and how does it impact the quality of the image?

1.2. Thesis structure

Section 1.1 starts with presenting the motivation for this PhD, and raising a number of questions to be answered. Section 1.3 goes through the relevant previous work published in this field of research, while Section 1.4 and Section 1.5 provide an overview of the basic theory that is necessary to understand the scope of the PhD.-project. Section 1.6 presents how the questions raised in Section 1.1 are answered and in what papers these answers are published.

Chapter 2 presents ways to evaluate the quality of a superimposed image in simulations, while Chapter 3 introduce a number of ways the subframes may be generated and how these subframes impact the overall quality of the projected image. Chapter 4 presents how the direction of the shift affects the superimposed image and also how the direction of the shift may interact with some subframe generation methods. Chapter 5 goes in depth on how the resolution of the image source affects the

resolution of the projected image, and how the resolution of the resulting projected image may be measured.

Chapter 6 discusses and concludes on the findings in the previous sections, and puts them in context of the questions raised in Section 1.1 to conclude upon the research presented.

1.3.   Literature review

The literature section give an overview of the work already published within this field, and the chronological presentation gives a outline of how the field has evolved.

Takahashi et al. [5] proposed a setup in 1995 with four 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 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. They used the Modulation 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, number of projectors and the projection lens and prism performance. Since this method relies upon both a cumbersome multi-projector setup and the low fill factor of the LCD projectors, it is not as relevant for current day designs. Therefore we have not explored the method from Takahashi et al. [5] further in this work.

Jaynes et al. [6] 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. 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. Since the method from Jaynes et al. [6] relies upon a cumbersome multi-projector setup it is not as relevant for current day designs. Therefore we have not explored the method further in this work.

Allen and Ulichney [7] 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 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. This work from Allen and Ulichney [7] is very interesting as it presents the basic form of shifted superimposition within a single projector. For this reason the method presented here is used as a baseline reference throughout the work in this thesis.

Majumder [10] explores the question if spatial super resolution is feasible using overlapping projectors in 2005. This analyse is based on multi projector displays with overlapping images from multiple projectors. The work presents a thorough

theoretical analysis to answer this question using signal processing and perturbation theory. Mujamders analysis is supported by results from a simulated overlapping projector display. This analysis shows that achieving spatial super-resolution using overlapping projectors is infeasible. The analysis in Majumders work [10] is an interesting approach, but since this analysis have been proven inaccurate by Said [11] and Damera-Venkata and Chang [13] we do not see this as relevant enough to elaborate upon in this thesis.

Said [11] 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 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. Parts of this work is presented in the papers [9] and [12]. The work presented from Said in 2006 is very relevant for current day techniques. Some of the methods we analyse and build upon is based on the naïve approach introduced by Said.

Damera-Venkata and Chang [13] 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 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. Damera-Venkata and Changs work [13] is interesting as it also proves that resolution enhancement through shifted superimposition is feasible, and thus that Majumders conclusion [10] was incorrect. Other than that this work do not present any relevant methods to include in this thesis.

6

Napoli et al. [14] describe a projection system that presents a 20 megapixel image using a single XGA SLM and time-division multiplexing, apertures and a piezoelectric mirror scanner. The system can be configured as a high-resolution 2-D display or a highly multi-view horizontal parallax display. They present a technique for characterizing the light transport function of the display and for precompensating the image for the measured transport function. The techniques can improve the effective quality of the display without modifying its optics. Precompensation is achieved by approximately solving a quadratic optimization problem. Compared to a linear filter, this technique is not limited by a fixed kernel size and can propagate image detail to all related pixels. Results of the algorithm are presented based on simulations of a display design. Simulated results of the characterization and precompensation process are presented. This method use a lenticular array and a aperture to decrease the fill grade of the pixels, so that when spatially shifting the projected frame the pixels will not overlap. Since such a method lose out too much on both framerate and light output, it is not deemed as a good enough approach for a high lumen state-of-the-art system. Thus we have not elaborated further upon this approach in this thesis.

Okatani et al. [15] explored the theory from Damera-Ventaka and Chang [13] 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. Okatani et al. [15] focus on multi projector systems which is not currently relevant, so the specific method is not included in this thesis. But the analysis showing the differences in white-on-black and black-on-white representation gave the idea for the priority based techniques described in Chapter 3.

Didyk et al. [16] propose in 2010 a novel method applied to moving images that takes into account the human visual system and leads to an improved perception of

such details. They display images rapidly varying over time along a given trajectory on a high refresh rate display. Due to the retinal integration time the information is fused and yields apparent super-resolution pixels on a conventional-resolution display. The work discusses how to find optimal temporal pixel variations based on linear eye-movement and image content and extend the solution to arbitrary trajectories. This step involves an efficient method to predict and successfully treat potentially visible flickering. Finally the resolution enhancement is evaluated in a perceptual study that shows that significant improvements can be achieved both for computer generated images and photographs. As the method proposed by Didyk et al. [16] focus on the motion of images and reduction of perceptible smear effects, this is not directly relatable to the case of shifted superimosition.

Sajadi et al. [17] 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 work presented by Sajadi et al. [17] is interesting and the edge enhancing approach could be currently relevant in commercial displays like home cinema. But since the unpredictable nature of the method do not comply with the pixel accuracy demands of state-of-the-art professional displays we have not included this optical pixel sharing approach in this thesis.

Berthouzoz and Fattal [18] present a method that makes use of the retinal integration time in the human visual system for increasing the resolution of displays. Given an input image with a resolution higher than the display resolution, they compute several images that match the display's native resolution, and then render these low-resolution images in a sequence that repeats itself on a high refreshrate display. The period of the sequence falls below the retinal integration time and therefore the eye integrates the images temporally and perceives them as one image. In order to achieve resolution enhancement they apply small-amplitude vibrations to the display panel and synchronize them with the screen refresh cycles. This approach achieves resolution enhancement without having to move the displayed content across the screen and hence offers a more practical solution than existing approaches. Moreover, they use their model to establish limitations on the amount of resolution enhancement achievable by such display systems. In this analysis they draw a formal connection between their display and super-resolution techniques and find that both methods share the same limitation, yet this limitation stems from different sources. The method presented by Berthouzoz and Fattal [18] is inspired by Didyk et al. [16], but places it closer to the concept of spatially shifted superimposition. While this work do show that it is possible to achieve the shifted superimposition concept with other display technologies than projected displays, it does not add any new methods to explore for projected displays or subframe generation.

The year after, Sajadi et al. [19] 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 [7], 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 SSIM [20] 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 Oriented Gradients (HOG) 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. The method proposed by Sajadi et al. [19] is targeting low-cost systems with less performing resolution enhancement, so this optical method is not discussed further in this thesis. But the analysis presented in the paper is current relevant, and is discussed and built upon both in Chapter 3 and in Chapter 4.

Heide et al. [21] 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 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. The multiplicative method presented in this work is very relevant, and should also be explored further in projected displays. This track was eventually excluded in this thesis because of prioritization of other topics.

The same year Heide et al. [22] released more work based on display architectures, exploring new optical device configurations and compressive computation. Previous research have shown how to improve the dynamic range of displays and facilitate high-quality light field or glasses-free 3D image synthesis. In this paper they introduce a new multi-mode compressive display architecture that supports switching between 3D and high dynamic range (HDR) modes as well as a new super-resolution mode. The proposed hardware uses readily-available components and is driven by an nontraditional splitting algorithm that computes the pixel states from a target high-resolution image. In effect, the display pixels present a compressed representation of the target image that is perceived as a single, high resolution image. As

in Heide et al. [21] this work use the concept of multiplicative superimpositioning. The multiplicative superimpositioning method is very relevant, and should also be explored further in projected displays. The multiplicative superimpositioning track was eventually excluded in this thesis because of prioritization of other topics.

Barshan et al. [23] 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 [7], but the generation of the sub-images are done independently instead of using the same sub-image for both positions. The quality improvement in this work is verified by visual inspection and by using the SSIM [20] metric as well. Barshan et al. used two computer generated test images and one natural image as test scenes in their work. The work presented by Barshan et al. [23] is relevant, and is included in some of the discussions in this thesis.

1.4. Quality evaluation

As seen in Section 1.3 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 Section 1.3, and also present other quality metrics that will be used in this work.

Since 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 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 while 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 [24]. 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 PSNR-HVS [25] and PSNR-HVSM [26] 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 use discrete cosine transform (DCT) to calculate contrast masking. By taking the contrast sensitivity function of the HVS into account the metric ignores the same contrast steps that the HVS also will 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 [20] 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 [27] (MSSSIM), ESSIM [28], SR-SIM [29], FeatureSIM [30] (FSIM), DC-Tex [31], VIF [32] and VSNR [33]. MSSSIM is a multiscale 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 are 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 sub-band and each local region contributes to the entire distortion independently. This assumption is reasonable since most typical distortions have few (linear) correlation both between the sub-bands and between the neighbourhoods at large spatial scales. The HVS contrast sensitivity function and texture mapping property are used to weight the contribution from the different sub-bands 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 luminance and visual angle (rather than on digital pixel values and pixel-based dimensions) to accommodate different viewing conditions.

1.5. Resolution

Resolution is a widely used term, but it turns out that the definition of resolution is highly dependant on the context. While many digital devices treat resolution as a mere pixel count, the analogue counterparts of the same devices often have definitions derived from the device's or the observer's ability to resolve details.

The resolution definition and limitations are also dependant on where in the imaging chain we refer to, as illustrated in Figure 1.2.

The real world scenes have unlimited resolution, as these scenes are continuous. Every stage after the real world scene have the possibility to introduce new limitations and constraints on the image, and by that also affect the resolution of the image. All

Figure 1.2: Imaging chain roughly illustrating the different steps.

of these steps have a different view-point and definition of resolution, as the concept of resolution is seen within their own realm. These examples show that the concept of resolution is imprecise, as the definition often is based on the context it is used in.

1.5.1. Resolution in the imaging chain

1.5.1.1. *Capturing device*

The device used to capture a real world scene is typically a camera, and such a device has in itself several components that may limit the resolution in itself, and therefore also have their own idea of what resolution is and how we should handle the concept.

1.5.1.2. *Optics*

The camera optics is prone to different aberrations and diffraction, which all have some impact on the final representation of the image. Focus shift and spherical aberrations both affect how the real world scene is transferred onto the sensor in the end, and may severely impact how different frequencies of the scene are represented at the sensor plane [2].

A type of target commonly used to test the performance of an optical system consists of a series of alternating light and dark bars of equal width. Several sets of patterns of different spacings are usually imaged by the system under test and the finest set in which the line structure can be discerned is considered to be the limit of resolution

of the system, which is expressed as a certain number of lines per millimetre. We express the contrast in the image as a modulation, given by the equation

$$modulation = \frac{max - min}{max + min} \qquad (1.1)$$

where max and min are the image illumination levels measured at the different linewidths. We can then plot the modulation as a function of the number of lines per millimetre in the image, and by adding a line representing the limiting resolution we get a modulation curve as shown in Figure 1.3 (a).
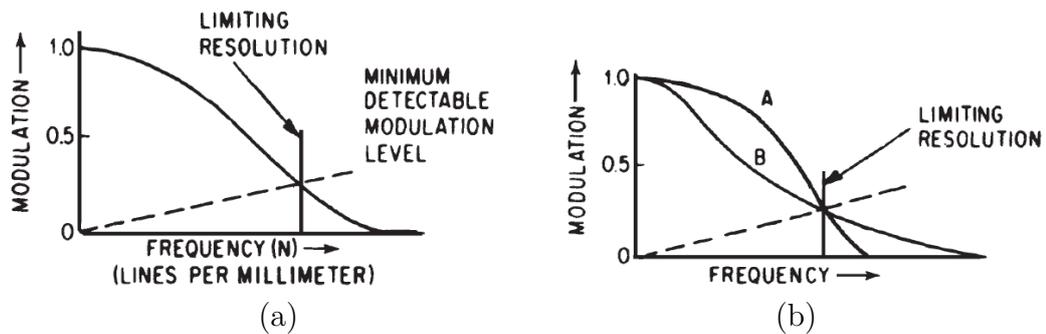
Figure 1.3: (a) The image modulation can be be plotted as a function of the frequency of the test pattern. When the modulation drops below the minimum that can be detected, the target is not resolved. (b) The system represented by A may produce a superior image, although both A and B have the same limiting resolution. [2]. The dotted line indicating the minimum detectable modulation level is often called an AIM curve, where the initials stand for the aerial image modulation required to produce a response in the system or sensor.

The limiting resolution of an optical system is an effective way to see the system limitations, but it doesn't elaborate on the performance of the system up to that frequency. Figure 1.3 (b) shows us two different modulation curves with the same limiting resolution, but with very different performance up to this limiting resolution. For this reason, a measure of the area below the curve is sometimes used as a measure of the information a system may deliver, and therefore is also a measure of the available resolution [34].

A more universal form of performance measurement for the optical system is the Modulation Transfer Function(MTF). In this case the modulation is measured given sine wave instead of squares, and the MTF is the ratio of the modulation as a function of the frequency (cycles per unit of length) of the sine-wave pattern.

$$MTF = \frac{M_i}{M_o} \tag{1.2}$$

The MTF is now widely used as a performance measurement of optical systems, and the advantage of the MTF is that it can be cascaded by multiplying the MTFs of two or more sub-systems to obtain the MTF of the combination. MTF has thus been applied not only to lenses but to films, phosphors, scintillating plates, image tubes, the eye, and even to complete systems such as camera-carrying aircraft [2].

1.5.1.3. *Sensor*

Some decades ago the sensor part of the camera was photographic film, which was of analogue nature. These days a regular camera usually features an electronic sensor array that captures and stores the image electronically. Common for both these sensor types are that they are restricted by the sampling theorem as introduced by Nyquist and later expanded by Shannon [1]. They stated that if a time-varying function, v(t), contains no higher frequencies than f_{max}, then it is completely determined by giving its ordinates at a series spaced $0.5 * f_{max}$ sec apart. This means that the sampling interval must be

$$<= \frac{1}{2 * f_{max}} \tag{1.3}$$

to represent the function completely.

When Shannons sampling theorem is utilized on an image instead of a time-varying function, spatial coordinates is used instead of points in time. The theorem then states that the camera sensor is able to capture image frequencies up to f_{max}, deter-

mined by the spatial spacing between the sensor detector elements. This means that the highest detectable frequencies in the scene are determined by the pixel pitch of the camera sensor, given that the optics does not restrict the frequencies even before the scene reaches the sensor.

Since the real world scene contains frequencies above these limits, the captured image is prone to aliasing [1]. To prevent this unwanted aliasing, the image is filtered at the sampling frequency given by Equation 1.3, but since ideal filters are not possible to implement, these kind of filters usually restrict the frequency of the image even further.

Super-Resolution (SR)   Various techniques exist to enhance the resolution of the captured image above the resolution of the sensor, also known as super resolution [35]. Since the real world scene for all relevant purposes has unlimited resolution, the sensor is only capturing a fraction of this information in its fixed grid. Super resolution techniques are techniques that construct high resolution images from several low resolution images, thereby increasing the high frequency components. The basic idea behind SR is to combine the non-redundant information contained in multiple low resolution frames to generate a high resolution image. The non-redundant information contained in the these LR images is typically introduced by sub-pixel shifts between them. These sub-pixel shifts may occur due to uncontrolled motions between the imaging system and scene, e.g., movements of objects, or due to controlled motions such as shifting the sensor chip itself [35].

Since this technique involves image processing and the goal is to achieve a finer pixel grid than the sensor grid, many of these methods also include interpolation methods. Interpolation is a common way to upscale an image to a larger pixel grid, without adding additional information [36]. The need for more information is why the super resolution techniques utilize several low resolution frames to construct one high resolution image.

1.5.1.4.   *Electronics*

The electronics of the system may also contribute to limitations of the information available through the system, and thus also the resolution obtained. High data rate systems, such as high resolution imaging systems, are dependant on high electronic bandwidth to get all of the information from the image sensor to a storage unit, and from the storage unit over to the display.

Especially in low cost systems this can be a bottleneck, introducing the need for compressing the image. Compression may be done lossless, but if the bandwidth is very limited then lossy compression may be necessary.

Within the electronics the image is also commonly processed in one way or another. Usually the image is formatted into a standard format to transport or store the image in a way that other devices will understand. In this image processing step details in the image may be lost, which may impact the resolution and the quality of the image. The quantization noise due to finite A/D bits is also a source of quality degradation. This occurs when the image is captured by the image sensor and is converted from analogue to digital signals within the sensor.

1.5.1.5.   *Display device*

The display is a device that presents the image to the observer, and these devices come in many different categories. In this work we divide them into CRT based displays, digital flat screen displays and projected displays.

CRT-based display   Even though the video signal and electronics may be digitized, the CRT is an analog display device. The image is formed by one or more electron beams scanning the CRT faceplate covered with phosphor. The electron beam is scanning through a metal mask forming up dots in the phosphor, and the dot pitch in this mask is often referred to as the resolution of the CRT [1].

Since the nature of the phosphor makes the light scatter and spread in the CRT faceplate, the smallest individual details that can be made out are often larger than

the dot pitch. Therefore the shrinking raster method is often applied to CRT based displays to measure the observed resolution of the display [1]. With the shrinking raster method a large number of parallel lines are displayed with some spacing in between the lines. This spacing is then shrunk until a flat field condition is obtained, which means that the spacing between the lines can no longer be discerned. The shrunk raster resolution is then given as the number of lines per centimetre on the display. Under nominal viewing conditions, experienced observers may no longer perceive the raster when the luminance variation (ripple between the lines) are less than 5%.

Flat screen displays In flat screen displays, for instance LED, LCD and Plasma displays, the image is made up by discrete pixels, which are typically independently controlled [1]. Resolution in such displays with discrete pixels are usually given as a mere pixel count, where the separate pixels are counted in horizontal and in vertical direction.

In these displays each pixel is usually also addressable, so the addressable resolution is then the same as the physical resolution(number of pixels). This number only describes the number of image elements that the display is able to reproduce, and does not say anything about how well these image elements are reproduced.

Projected displays In digital projection, the image is made up by imager devices with discrete pixels, which are typically independently controlled in the same manner as the digital flat panel displays [34]. But in addition the image is projected through an internal optical system, making the image prone to the same optical artefacts as described in Section 1.5.1.2. So even though the mere pixel count is used as a definition for resolution in some projected applications, more professional markets utilize the MTF as a measure for the projected displays performance [34].

Capturing devices have different applications and may have a need for several different types of MTF curves. The lower curve in Figure 1.3 (b) may for instance be

more desirable in a camera for edge detection applications, like automatic number-plate recognition. But in a display application it is important how much modulation is available at mid-resolution. Detail and sharpness may generally be determined by the area under the MTF curve up to the limiting resolution [34]. The MTF curve shows the degradation that occurs through the electronic and optical system associated with the projector.

Different projector technologies favour different aspects of the projector, so some projector technologies (eg. DLP) preserve high pixel-to-pixel contrast, while LCD based projector technologies show a reduced contrast at the same resolution [34]. This gives a DLP projector a higher limiting resolution than and LCD projector with the same number of pixels and similar optics. Studies show that for a human observer the limiting resolution occurs when the modulation drops to 3-5% [34]. So the limiting resolution of the projector may be determined from the MTF curve by observing where the MTF curve crosses the 3-5% modulation level [34].

1.5.1.6. Observer

The observer is also an important part of the imaging chain when considering system resolution. The human visual system (HVS) introduces it's own limitations, and must be taken into account for system design.

1.5.1.7. HVS

The Human Visual System (HVS) use the eye as a detector, and the eye itself relies on an array of photoreceptors that sets the boundaries on the spatial detail available for neural representation [37]. As with any array of sensors this array is also limited by the sampling theorem [38] [1], which means that the highest spatial frequency that is detectable should be dependant on the spacing between the photoreceptors in the array. It is the foveal cone spacing in particular that is the basis of resolving details [37], and the distribution of the cones are highest in the center.

From the observer point of view the spatial observations are not discerned as individual sensor elements as pixels, but rather as an angular plane of receptive fields [38].

Receptive fields have a many-to-one relation to the photoreceptors in the eye, which means that even though we have a higher spatial acuity in the center of the eye, it is still not reaching the maximum frequency according to the sampling theorem [38]. From an observers position the perceived spatial resolution is given as degrees of visual angle, often referred to as line pair per degree of visual angle.

Perceptibility occurs when targets are larger than 1 arc minute, termed as one cycle per arc minute, as this corresponds to normal visual acuity of 20/20 [1]. An observer with poorer acuity will simply have to move closer to the screen while an observer with better acuity may move further from the screen without losing out on the resolution.

The angular representation of resolution is very accurate from an observer point of view, but has the restriction that the position of the observer is fixed in relation to the display. For regular display resolution this may not be suitable, but for specialised display systems where the observer is designed into the system, the angular representation may be very suitable. One example of such a system is military simulators where there is an established standard named the Johnsons Criteria. This criteria states the critical dimensions for an observer to detect, decide the orientation, recognize and identify a target [1]. The observer will in this case typically be a pilot within a simulator, so the observers position is fixed and the system will then be tuned at 1 cycle per arc minute.

Discrimination Level	Meaning	Cycles across target
Detection	An object is present	1.0 ± 0.025
Orientation	Orientation may be discerned	1.4 ± 0.35
Recognition	Class of object (tank, truck, etc...)	4.0 ± 0.8
Identification	Identify type (friend or foe)	6.4 ± 1.5

Table 1.1: Johnsons criteria [1].

It is important to note that since the HVS is a part of the imaging chain, it may also be the limiting factor of the chain. If the screens resolution is so high that the

display is displaying more than one detail per arc minute, the HVS will be limiting and the observer will not be able to discern all the details that are displayed.

In addition, the visual attention of the HVS limits what we are able to perceive as observers. While objects at approximately 1 arc minute may be discerned at our center of attention, our ability to select specific objects in a group of similar objects is poorer. To follow and track a single object in a group of similar objects HVS requires that the spacing between them is higher than three arc minutes [39].

1.5.1.8.   System resolution

All of the components in the imaging chain will have the possibility to affect the resulting resolution observed by the observer, and if the whole chain is known up front it is possible to design for optimal resolution in all parts of the chain. The optimal matching would be of all parts of the chain have the same limiting resolution, while still maintaining a good MTF for the frequencies that are important for the given application. The military simulators mentioned in Section 1.5.1.7 is such a system where it is possible to match all of the components and feed the observer with the information he/she needs without wasting resources on over-performing.

But in most cases the whole system is not known, making the different subsystems only relate to their own perception of resolution and the standards that interface the different parts of the chain. The camera captures the image or the film with as high resolution as it is capable of, and stores in a standardized format. The display will take the information stored in this format and present it to the observer. Typically the resolution given in the image or video standard is only a pixel count, and it is not certain that this parameter is enough to evaluate the system resolution.

1.5.2.   Resolution summary

Given that every part of the image chain employs a unique concept of resolution, it is apparent that resolution is an ambiguous expression. Even though some of these expressions for resolution vary, the general understanding is that resolution is a representation of the number of details represented at that stage in the image

chain. How the resolution is given and measured is dependant of the technology, and even different disciplines and industries that use the same technology may have different definitions of resolution.

Holst [1] lists up a number of these definitions in his book "CCD Arrays, Cameras, and Displays" and some of them are given in Table 1.2.

Subsystem	Resolution metric	Description
Optics	Rayleigh Criterion	Ability to distinguish two adjacent point sources
	Airy disk diameter	Diffraction-limited diameter produced by a point source
	Blur diameter	Actual minimum diameter produced by a point source
Detectors	Detector-angular-subtense	Angle subtended by one detector element
	Instantanous field-of-view	Angular region over which the detector senses radiation
	Effective-instantaneous field-of-view	One-half of the reciprocal of the object space spatial frequency at MTF equal 0.5
	Detector pitch	Center-to-center spacing
Electronics	Bandwith	Capacity to transfer data
Electronic Img.system	Limiting resolution	Spatial frequency at which MTF equals 2-10%
	Nyquist frequency	One-half of the sampling frequency
Displays	TV Limiting resolution	Number of resolvable lines per picture height
	Pixels, datels, disels	Number of image elements.
Observer	Ground resolved distance	The smallest test target (1 cycle) the one may distinguish
	Ground resolution	the limiting feature size one may distinguish

Table 1.2: Resolution metrics from literature [1].

The resolution metrics given in Table 1.2 is only a selection of the metrics found in the literature, but it illustrates the diverse usage of this definition. Most of these metrics are in some form describing the ability to discern a single element, so it seems that this is the ability that the resolution metrics have in common.

One of the more common fields that we use the term resolution today is within cameras and displays. Usually resolution refers to the number of pixels in these devices, but as we see in Section 1.5.1 this definition is incomplete. The pixel count may have information about the number of image elements in the device, but it does not say anything about the quality of the image presented or captured.

An image presented with higher MTF is generally judged by observers as having higher quality [1], and MTF also gives a very good description of the performance of the device in question. The Definitions and Standards Committee International Committee for Display Metrology (ICDM) have released the Information Display Measurement Standard (IDMS) where they propose different ways to measure and define display resolution, where several of these methods are based on MTF or Contrast Sensitivity Function (CTF) [40]. These methods are not widely adopted yet, but going in the direction of MTF-like curves will give much more information about the device at hand than a mere pixel count.

Super resolution techniques in cameras and computational displays like pixel-shifted projectors also press upon this issue. Both these types of devices are accessing resolutions well above their number of image elements, and we need a way of describing the performance of these devices.

1.6. PhD contributions

As stated in Section 1.1, this research aims to explore different aspects of resolution enhancement through shifted superimposition, and to see in what degree this method enhances the resolution of the projected display. Theoretical simulation work as well as practical implementation and detailed measurements on a shifted superimposition DMD-projector are done. The main contributions of this PhD work are:

Q1: How are the projected subimages generated, and what impact do they have on the overall quality of the superimposed image?
In the paper "A comparative study of superimposition techniques for enhancing

the projector resolution: Simulations and experiments" we have evaluated different existing and newly developed subframe generation algorithms and assessed the superimposed image quality by detailed simulations as well as by experiments. In this research we also evaluate a newly proposed category of subframe generation techniques, based on local minimum and maximum values. We also extended the Single-subframe iterative technique to Two-subframe iterative technique to evaluate its full potential. These methods were implemented, tested and compared on a shifted superimposed DMD projector for the first time to the authors knowledge. We do see that the subframe generation technique has a great impact on the quality of the superimposed image.

The paper "A comparative study of superimposition techniques for enhancing the projector resolution: Simulations and experiments" was published in the journal Displays in 2018, and the work in this paper is covered in Chapter 3 of this thesis.

Q2: What is the best metric to assess the overall quality of enhanced images? How does it match with subjective experiments with human observers? In the paper "Preferred image quality metric for shifted superimposition-based resolution enhanced images" we assess this quality impact with different kinds of Image Quality metric, and find that the MSSSIM image quality metric is well suited to evaluate and compare different subframe generation methods. This is supported by subjective measurements performed on a group of observers.

The paper "Preferred image quality metric for shifted superimposition-based resolution enhanced images" was published in the Journal of Electronic Imaging in 2018, and the work in this paper is covered in Chapter 2 of this thesis.

Q3: Is the typical diagonal shift the ideal shifting direction?
The simulations in "Resolution enhancement through shifted superimposition: The influence of shift direction" we show that a projection system with enough computational power to generate each subframe based on its spatial position will be

indifferent to the shifting direction. Both the gradient preservation in different directions and the quality of the image is equal when generating the different subframes according to the shift. The exception to this rule is shifting in horizontal or vertical direction, which gives lower quality and poor gradient preservation.

The paper "Resolution enhancement through shifted superimposition: The influence of shift direction" is in preparation, and the work in this paper is covered in Chapter 4 of this thesis.

Q4: How much resolution is gained with this method, and how does it impact the quality of the image?

In "The effects of source resolution on resolution enhancement through shifted superimposition projection" we show that the achieved resolution with the shifted superimposition technique does increase as we increase the source resolution. This is valid up to a certain threshold, where the shifted superimposition method reaches it's limit because of the physical size of the projected SLM pixels and the overlap of these pixels in different positions. The limit seems to be about 40% above the SLM resolution. To the authors knowledge this is the first time such quantitative measurements of a shifted superimposed projector is published, and also the first time the relationship between the source resolution and the measured resolution is investigated through measurements.

The paper "The effects of source resolution on resolution enhancement through shifted superimposition projection" have been accepted for publication in Journal of the Society for Information Display, and the work in this paper is covered in Chapter 5 of this thesis.

2. Image quality assessment of superimposed images

In this work we investigate 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 work 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 Section 1.3 provides an insight into the prior work done in the field of superimposition with respect to how quality was evaluated in these papers. Section 1.4 presents a set of relevant quality metrics used in this quality assessment research.

The rest of this chapter is organized as follows: Section 2.1 describes the experimental setup that is used to tests different metrics and superimposition methods. The simulated results are presented in Section 2.2 while the subjective experiment is presented in Section 2.3. Thereafter the findings are summarized and discussed in the discussion Section 2.4. Finally, Section 2.5 concludes on how to evaluate the resulting image quality for this particular application and the future work still to be done.

2.1. 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.

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-HVSM, ESSIM, FeatureSIM (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.

We have implemented four different ways of generating the superimposed image in this simulation framework, and are also comparing the superimposed methods to 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 [7, 9, 19] or methods that build further upon them. The Naïve and the filtered Naïve methods are quite close to actual methods used in products on the market. Several vendors use the idea of upscaling the input image to double the resolution of the SLM, and then do some image processing in that doubled resolution domain [3, 41]. 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 thereby preserving more details, before the subframe pixels are chosen among the resulting upscaled pixels.

In theory, iterative algorithms like the one presented by Sajadi [19] 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 since 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.

2.1.1. 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

$$OutImage = Resize(RefImage, SLMresolution), \qquad (2.1)$$

where $Resize$ in the equations of this chapter is referring to the Matlab function imresize using bicubic interpolation, and $SLMresolution$ is the resolution of the SLM in use. $RefImage$ is the original high resolution image.

2.1.2.   Downscaled superimposed

This method generates the sub-images as the Downscaled method, but then these sub-images are spatially shifted and superimposed on themselves. It is not an ideal method, but as seen in Section 2.3 it is a step up in perceived quality from the regular downscaled version in some instances. Allen and Ulichney [7] used this version to verify the superimpositioning in their wobulation paper. In this method, both subframes will be equal and given by

$$SubframeA = SubframeB = Resize(RefImage, SLMresolution). \tag{2.2}$$

2.1.3.   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 since we just select every other pixel.

$$IntermediateFrameNaive = Resize(RefImage, 2 * SLMresolution)$$
$$SubframeA(i, j) = IntermediateFrame(2 * i - 1, 2 * j - 1) \tag{2.3}$$
$$SubframeB(i, j) = IntermediateFrame(2 * i, 2 * j)$$

2.1.4.   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

$$IntermediateFrameGauss = imfilter(IntermediateFrameNaive,$$
$$fspecial('gaussian')),$$

(2.4)

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 $3x3$ filter with a standard deviation of 0.5.

2.1.5. 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

$$Sfilter = b * [0, a, 0; a, (-4) * a, a; 0, a, 0]$$
$$InputImageMask = imfilter(IntermediateFrameGauss, Sfilter)$$
$$IntermediateFrameSharpened = IntermediateFrameGauss+$$
$$k * InputImageMask$$

(2.5)

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

2.1.6. 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.

2.1.7.   Test images

We have used thirteen 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 content, ranging from buildings and architecture to random edge structure in waves. The natural image are presented in Figure 2.1, while the synthetic images are presented in Figure 2.2.

Natural images are in itself a very diverse group of images, and it is not given that one IQM will be optimal 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 the IQM analyses in the image. To accommodate these factors we have included a broad selection of natural images that include different features, but also that include details that makes 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. When resolution increase is done through superimposition, it acts as a lowpass filter, resulting in loss of detail in high frequency areas. The natural images presented in Figure 2.1 are selected because most of them feature objects with a high degree of detail that the different superimposition algorithms may affect in different ways. The images in Figure 2.1 a,b,c,d,e,g and h is from pixabay, image f is from wikipedia under the creative commons license, and image i and j are from the CID:IQ database [42].

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 horizontal direction and three line pairs in 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 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 remains 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. The natural images are presented in the subjective experiment section.

2.1.8. Test scenario

In our work to identify the objective imperfections in the synthetic images, we have defined an SLM with the resolution of 250x250 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 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 to above double of the native resolution.

For the natural images, we have set up a Subjective experiment as described in Section 2.3. In this experiment the input images is 512x512 pixels, while the SLM resolution is kept at 256x256 pixels.

Since we are simulating sub-pixel behaviour, each pixel of the reference image is quadrupled into four pixels in the sub-images. 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

Figure 2.1: 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 j) Church.

result both horizontally and vertically, bringing the simulated resulting image into the same resolution as the reference image.

2.2. Simulation results

For the synthetic images we have concrete symptoms to look for. The Line pair image have three distinguishable line pairs that will eventually fuse 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 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 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 2.3 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
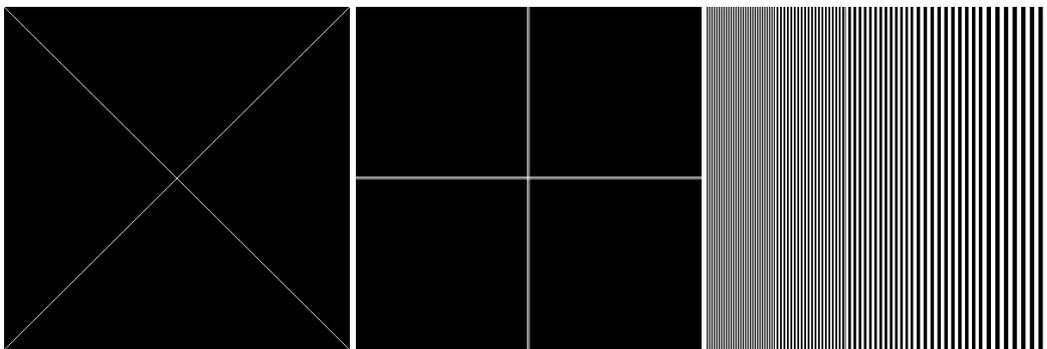
Figure 2.2: Synthetic test images used from left: Cross, Line pairs and H-frequency.

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 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 2.4, 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 2.5 illustrates 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.

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 Figure 2.6 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 2.6 shows how the different superimpositioning methods produce varying amount of aliasing.

2.3. Subjective experiments

The subjective experiments are designed to see which of the IQMs 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
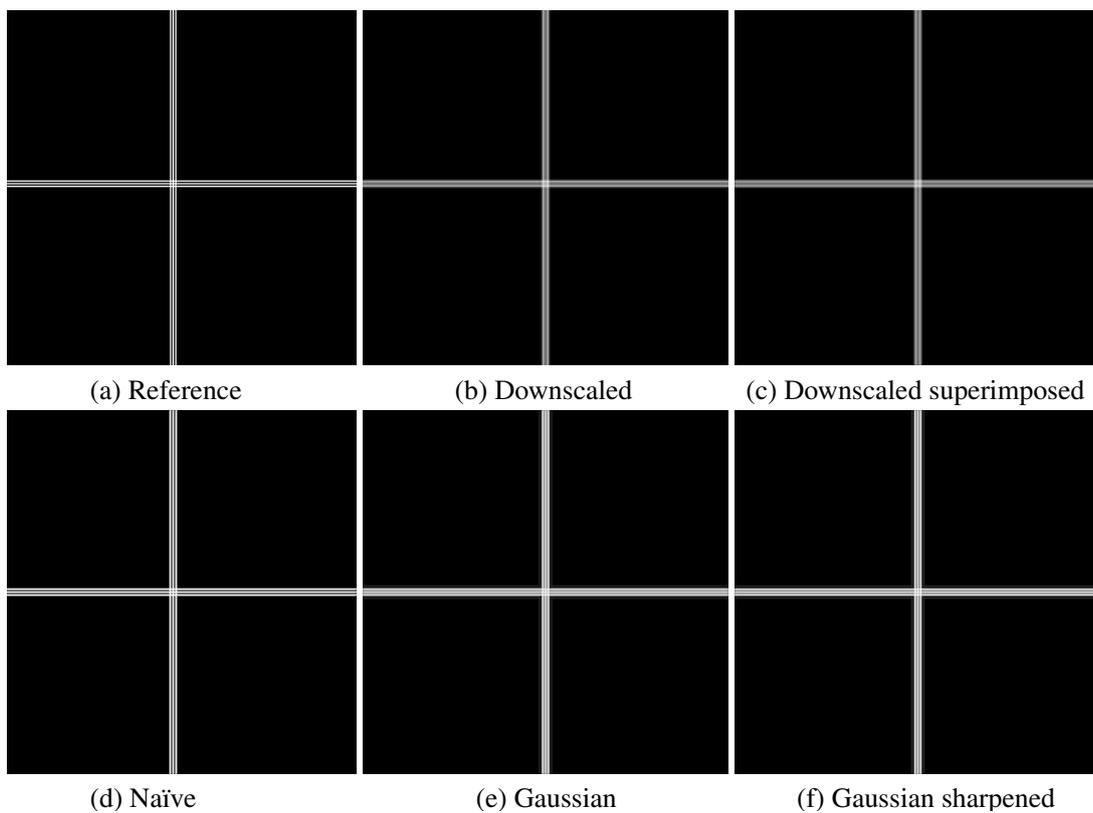
Figure 2.3: Zoomed in on the resulting line pairs at 275 pixels input resolution. a) Reference image b) Downscaled c) Downscaled superimposed d) Naïve e) Gaussian f) Gaussian sharpened. Note how b) and c) have lost one line pair.
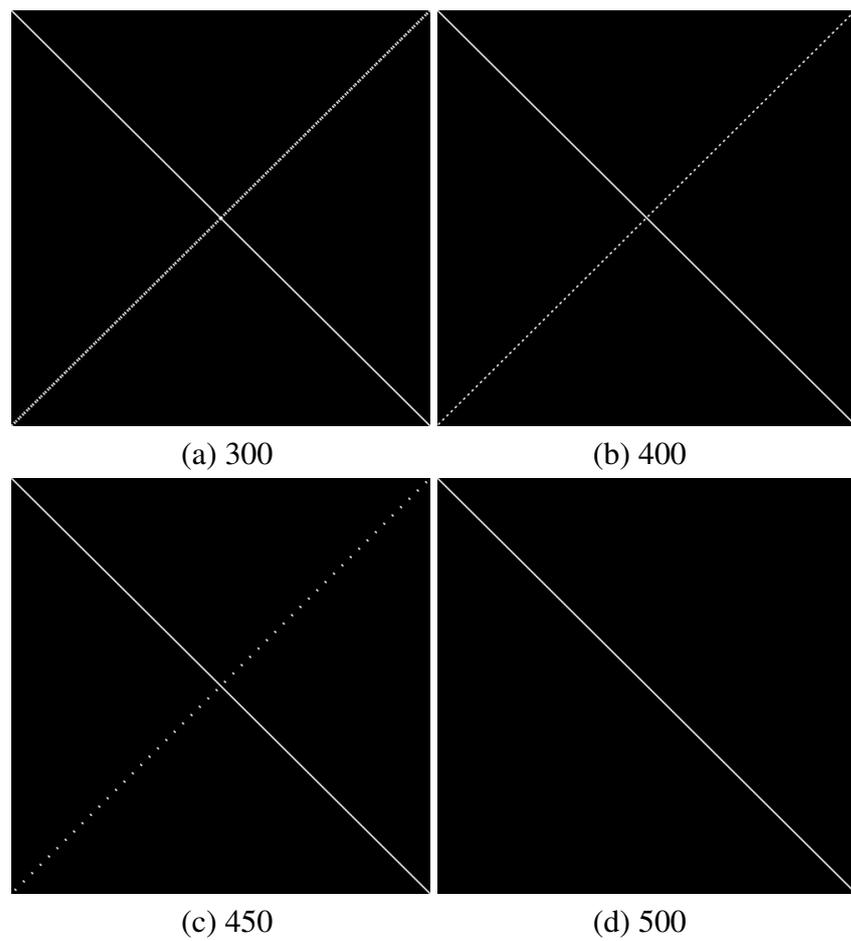
Figure 2.4: Results from Naïve superimpositioning with a) 300, b) 400, c) 450 and d) 500 pixels input resolution.
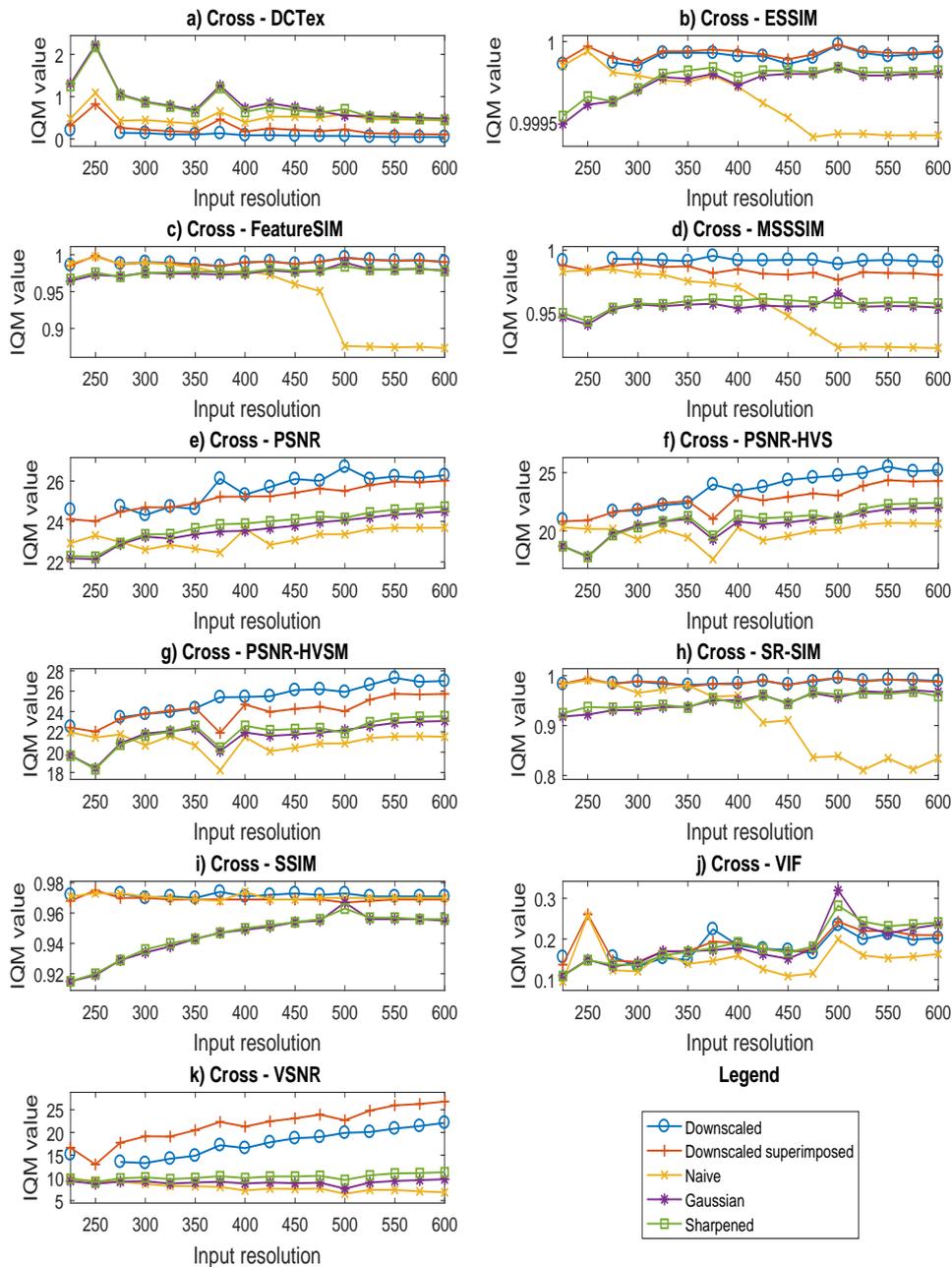
Figure 2.5: Synthetic Cross scene evaluated by all of the metrics. The x-axes represents the input resolution and the y-axes represents the IQM value.

(a) Reference   (b) Downscaled

(c) Downscaled superimposed   (d) Naïve

(e) Gaussian   (f) Gaussian sharpened
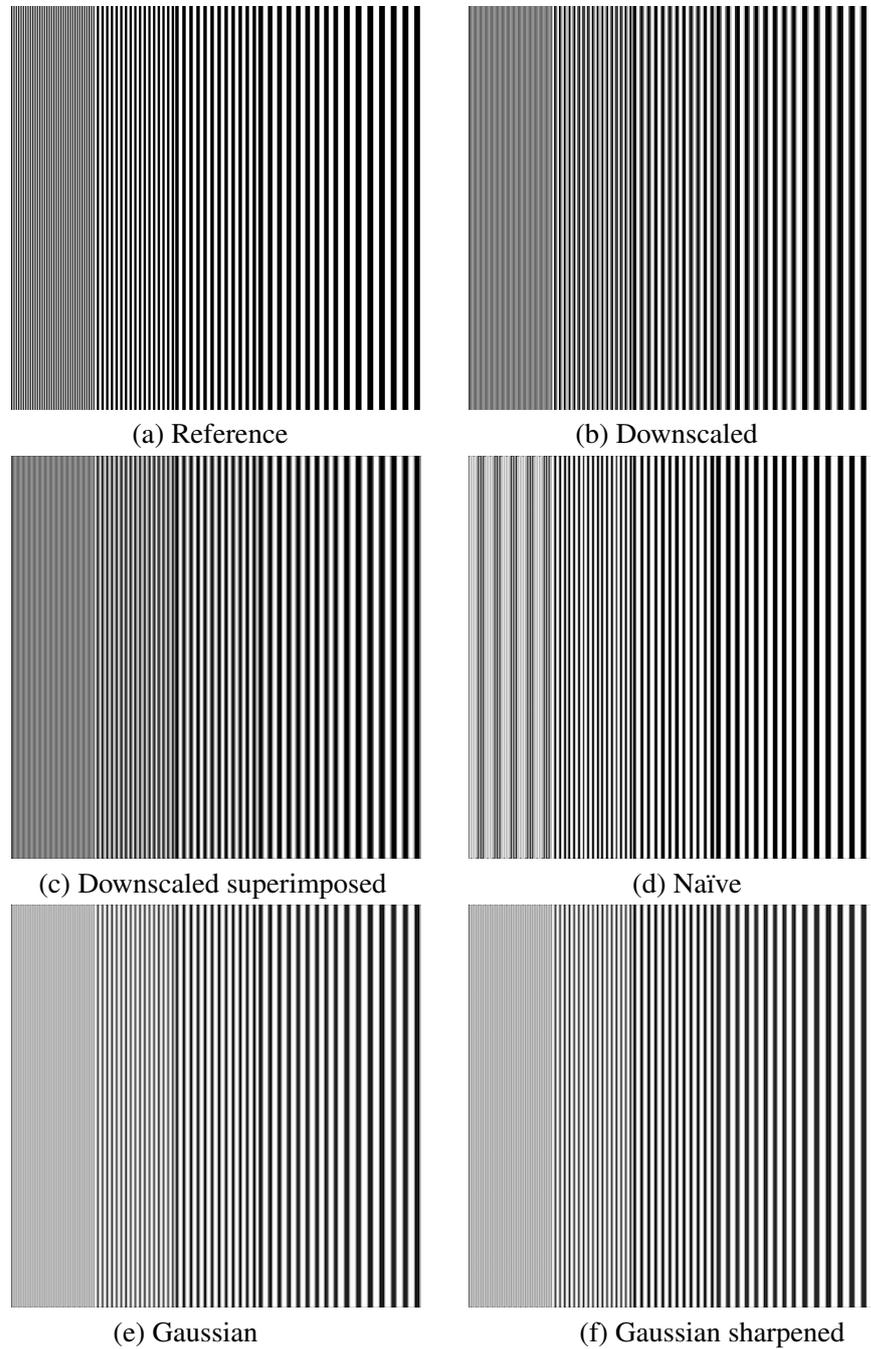
Figure 2.6: Horizontal frequency image at 350 pixels input resolution. a) Reference image b) Downscaled c) Downscaled superimposed d) Naïve e) Gaussian f) Gaussian sharpened. Note how for instance the Naïve mathod (d) have severe aliasing. The methods containing the gaussian filter have less aliasing since this method have filtered out some of the highest frequency components.

variety of content and structure, and we have included images of persons, texture, nature, buildings, vehicles and text. Please see Figure 2.1 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 ten images. We have kept the input resolution over SLM resolution ratio at a factor of 2x, meaning that an input image of 512x512 is simulated with an SLM resolution at 256x256. 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 [43]. The participants were told to select the visually preferred image in each image pair. Since 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. Since this experiment tests the perceived resolution and spatial quality enhancement of the image, this condition was deemed to be good enough. The z-scores [44] from this test are visualized in Figure 2.7. The general trend in Figure 2.7 shows that the methods where each superimposition subframe is calculated individually gives a much better perceived image than the Downscaled and Downscaled superimposed methods.

The correlation of each IQM towards these z-scores was then calculated for both Pearson and Spearman coefficients, and the results from these calculations is presented in Table 2.1 and Table 2.2, and these results are visualized in Figure 2.8 and Figure 2.9. 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 Figure 2.8 and Figure 2.9.
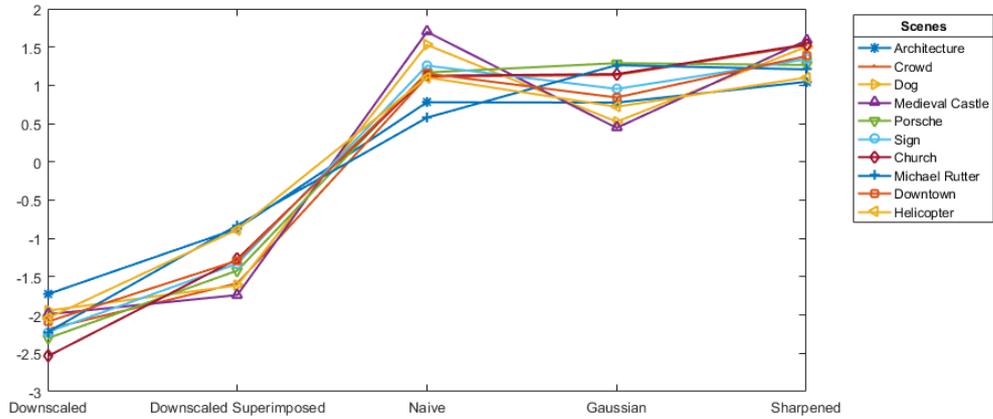
Figure 2.7: Z-scores from the subjective test of the superimposed images. Notice how different algorithms gives better results for different images. There is no universal best algorithm for all images.
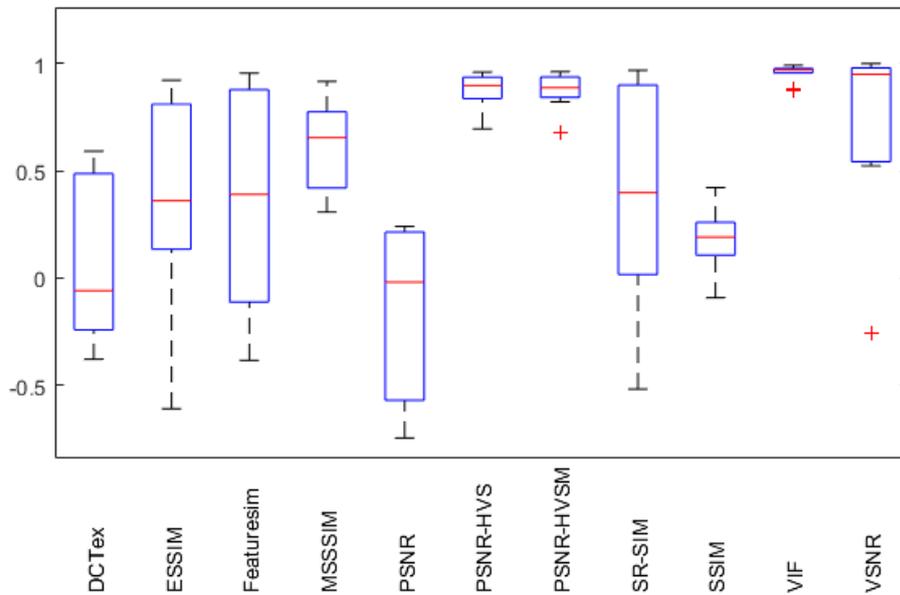
Figure 2.8: Pearson correlation coefficient for the different IQMs and the z-score.

Scene	DCTex	ESSIM	Feature Sim	MS-SSIM	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 2.1: Pearson correlation coefficients.

Several of the IQMs perform well in the subjective experiment. Based on the Pearson and Spearman correlation coefficients, the VIF IQM is the metric that performs best when rating according to the subjective view of the observers. Figure 2.10 shows how the VIF performs on all the images, with markers for different superimpositioning algorithms (markers) and the fitted linear regression curve (solid line).

2.4. 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 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 type of image artefacts. 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 analyse the structure in the image and evaluate the structural similarity
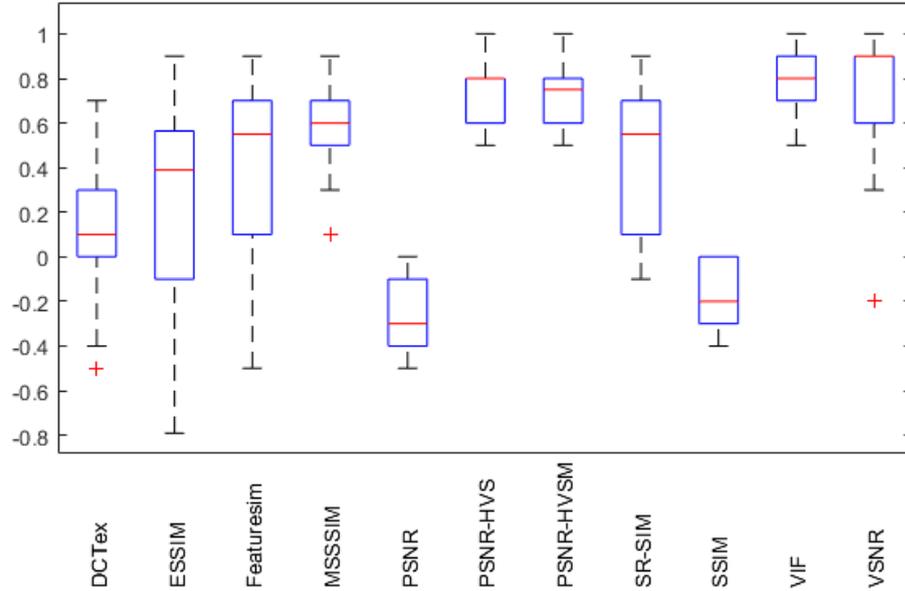
Figure 2.9: Spearman correlation coefficient for the different IQMs and the z-score.
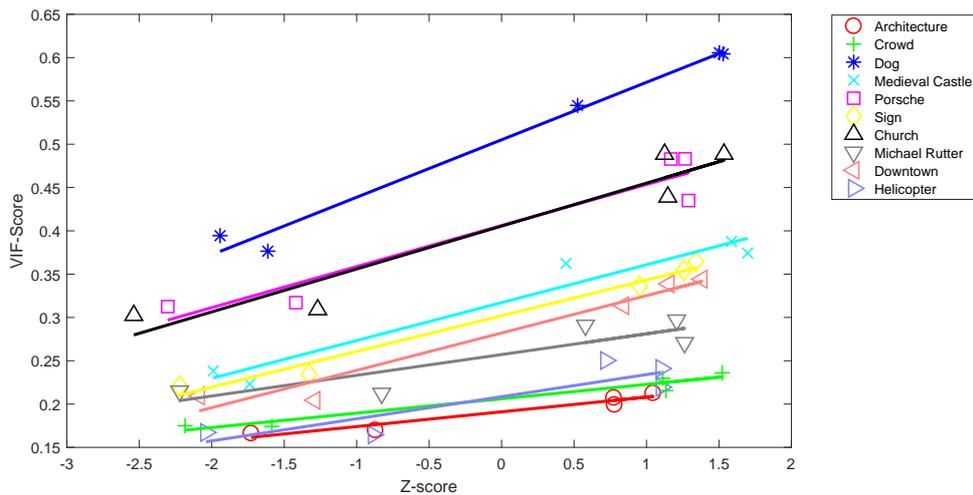
Figure 2.10: VIF score for each image plotted against the respective Z-scores. Solid lines indicate the fitted linear regression curve.

Scene	DCTex	ESSIM	Feature SIM	MS-SSIM	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

Table 2.2: Spearman correlation coefficients.

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 remaining two test images are used to establish pattern preservation and provoke errors that are more visible in the frequency domain than 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 analyse 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 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

evaluates the superimpositioning way of enhancing resolution and that differentiates different ways of superimposing.

As seen in the Z-scores [44] in Figure 2.7, 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. 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 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 those that generate different content for the two spatially shifted positions. This shows that to really utilize the potential of the shifted superimpositioning one should have different content 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 selects pixels in a way that ensures sharpness in the image, but also results in significant loss of information. Images with sharp details without straight geometric lines like the stray hairs on the Dog in Figure 2.1 (b), the random shaped stones in the Medieval castle in Figure 2.1 (g) and the braking waves on the Helicopter image in Figure 2.1 (h), benefit from this technique. But when looking at images with geometric structures like the buildings in Architecture in Figure 2.1 (d) and the text in the Michael Rutter image in Figure 2.1 (f) and Porsche in Figure 2.1 (e), the missing details that your mind still knows are there are probably influencing our view of these images. These images are then visually better enhanced with the Gaussian method, that is low-pass filtering the high-resolution image slightly 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 has the highest overall score, but it is not the best algorithm for all content. 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 many criteria to look for when matching against subjective quality.

From Figure 2.10 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 indicate a very similar quality, as seen in the graph for the Dog image in Figure 2.10.

Most IQMs have been designed to meet special requirements, for instance to detect degradation in specific elements of the image. The requirements 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 regards to visual preference take the HVS into account, while the metrics that are better at picking up detail preservation analyse 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.

2.5. Conclusion

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 these fully cover 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, while 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 more important. For applications where 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 first standardise the oversampling factor on the captured image and define 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 aberrations in the image, and it is also crucial that physical factors that influence the human visual system, like brightness, should also be taken into account.

3. Subframe generation methods

In additive superimposition, it is not possible to define each pixel on the enhanced resolution grid accurately. The aim of the subframe generation techniques developed to date is to produce subframes that when superimposed, will result in an image with as many pixels with similar values to the input high resolution image as possible. In this work we investigate and compare different categories of such methods, and introduce two new techniques that have been developed with the aim of prioritising certain pixels so that they are addressed perfectly.

The work starts with a brief description of all the techniques studied. Based on their principles, the techniques have been divided into three categories. The first category is based on Naïve approach. The second one is an iterative approach based on the single-subframe technique proposed by Sajadi et al. [19]. This technique has been further extended in this work to two independent subframes to evaluate its full potential. The third category describes two new priority-based techniques which were developed during this work. A simulation framework has been developed to generate the subframes and superimpose them to obtain the final image and assess its quality as compared to the original high resolution image. A measurement setup has also been used to sequentially project the subframes on the wall and access the image quality by capturing the individual subframes with a high-speed camera and the superimposed final image as well. Detailed simulation and measurements results are also presented .

3.1. Brief description of the techniques

3.1.1. Techniques based on the Naïve approach

Naïve technique was described by Said in 2006 for mathematical analysis of the subframe generation techniques [9]. The rest of the techniques described under this category are based on simple intuitive modifications to the Naïve technique intended for better results.

Technique	Hardware type	Time multiplexing	Subframe generation technique
Damera-Venkata and Chang, 2007 [13]	Multiple projectors	No	Iterative/ fast filter banks
Allen and Ulichney, 2005 [7]	Opto-mechanical actuator	Yes	Downscaled superimposed
Berthouzoz and Fattal, 2012 [18]	Opto-mechanical actuator	Yes	Iterative with multiple subframes
Sajadi et al., 2012 [17]	Multiple SLM with optical pixel sharing	Yes	Special technique based on edge detection
Sajadi et al., 2013 [19]	System of two lens producing optical shift and overlay	No	Single subframe iterative technique
Barshan et al., 2015 [23]	Opto-mechanical actuator	Yes	Iterative/ fast filter banks similar to those proposed by Damera-Venkata and Chang

Table 3.1: Different subframe generation techniques and hardware types

3.1.1.1. Naïve technique

Naïve technique considers pixels which are row wise odd numbered in the input image in one subframe while even numbered in the other. In Figure 3.1, subframe 1 would be P1 (odd) and Subframe 2 would be P4 (even).

3.1.1.2. Pick Mean technique

The aim of Pick Mean technique is to increase the dependency on all four pixels in the input image. This technique considers the mean value of the four pixels. That is Subframe 1 = Mean (P1, P2, P3, P4) and Subframe 2 = Mean (P4, P5, P6, P7).

3.1.1.3. Pick Minimum and Maximum technique

Pick Minimum and Maximum (Pick Min. and Max.) also considers one out of four possibilities but it takes exact values from the original image and considers darkest and brightest pixel out of the four possibilities. This technique is much more content dependent than the basic Naïve technique. This technique considers the least value in one subframe while the maximum value in the other subframe. That is Subframe 1 = Minimum (P1, P2, P3, P4) and Subframe 2 = Maximum (P4, P5, P6, P7).

3.1.1.4. Techniques using gaussian filter

Gaussian and Gaussian sharpened technique are described briefly by in Chapter 2. Although both Gaussian and Gaussian sharpened techniques have been used to generate subframes in this work, results from only the Gaussian sharpened technique are presented since the difference between them is marginal and Gaussian sharpened technique is considered an improvement.
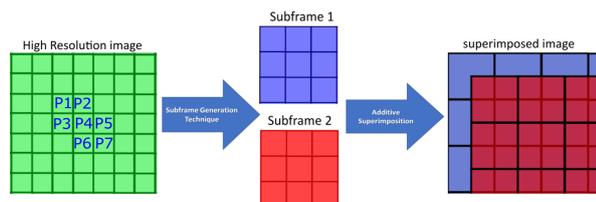
Figure 3.1: The overall process of additive superimposition

3.1.2.   Iterative techniques

3.1.2.1.   Single-subframe iterative approach

Sajadi et al. proposed an iterative technique to develop subframes but used only single subframe to minimize the computation time and resources. In this approach, the problem is formulated as a constraint linear least square problem and the aim is to reduce the error between the resulting superimposed image and the original high-resolution image [19].

The problem is seen as a system of linear equations, see Equation 3.1.

$$AI = I_T \qquad (3.1)$$

Where I is the subframe which is unknown, A is a sparse matrix which defines the linear equations and I_T is the given high-resolution image. The paper readdresses the problem as quadratic programming problem using the bound $0 \le I \le 1$, and then proceeds to solve it using Gaussian Belief Propagation (GBP) solver, which does not guarantee convergence but is faster than other algorithms such as Jacobi or Gauss-Seidel [45]. The limit $0 \le I \le 1$ is enforced using Lagrange multipliers updated using gradient ascent at each iteration. The objective of the iterations is to minimize Equation 3.2:

$$min \frac{1}{2} I^T A^T A I - I^T A^T I_T \qquad (3.2)$$

After considering the Lagrange multipliers and defining $J = A^T A$ and $h = A^T I^T$, the objective function can be redefined as Equation 3.3:

$$\frac{1}{2} I^T J I - I^T h - I^T \gamma + (I - 1)^T \gamma \qquad (3.3)$$

This technique does not use time multiplexing, hence to adjust the brightness of the scene, either the bound must be changed from $0 \leq I \leq 1$ to $0 \leq I \leq 0.5$ or Equation 3.1 should be rewritten as:

$$AI = 2I_T \tag{3.4}$$

3.1.2.2.   Two-subframe iterative approach

The aim of extending the previous technique to two independent subframes was to examine the full potential of the superimposition setup, since single subframe iterative technique was found to be the best among the techniques we explored as presented in later sections.

The same approach as mentioned above can be extended to create two independent subframes:

$$I_1 + I_2 = I \tag{3.5}$$

$$AI_1 + BI_2 = I_T \tag{3.6}$$

$$A'I' = I_T \tag{3.7}$$

A and B are two sparse matrices which depend on the size of the image. Solving Equation 3.5 just like in the case of the single subframe method using GBP solver, the two subframes can be calculated. Similar adjustment for time multiplexing has to be implemented here as well.

3.1.3.   Priority-based techniques

Since all the pixels in the high resolution image cannot be addressed perfectly in the superimposition setup, it is interesting to have a technique which is able to prioritise

which pixels to translate perfectly. Iterative technique tries to fit all the pixels perfectly while the rest of techniques examined here do not give any importance to the content of the image. The algorithms described in this section were developed with the view to prioritise certain pixels over others.

3.1.3.1.   Dark priority technique

In this technique, the image is first converted from RGB to YCbCr format and the operations are done only on the luminescence channel (Y). This helps to decrease the computation time and helps isolate brightness of the image better. The values for the other two channels are picked while the pixels for the luminescence channel is picked. High resolution image is first analysed and all the pixels are sorted from dark to bright along with their positions in the image. Then taking the darkest pixel, equivalent subframe pixels are calculated before being placed in the superimposed grid, thereby keeping half of the pixel value in each subframe:

$$Subframe1 = DarkestPixelValue/2 \qquad (3.8)$$

$$Subframe2 = DarkestPixelValue/2 \qquad (3.9)$$

such that

$$DarkestPixelValue = subframe1 + subframe2 \qquad (3.10)$$

Meanwhile, the values for chrominance channels (Cb and Cr) are picked from the same position as the pixels for the Y channel and are not prioritised separately. The flow of this method is illustrated in Figure 3.2.

3.1.3.2.   Bright priority technique

This technique is similar to Dark Priority technique, but brighter pixels are prioritised instead of the darker ones. The pixels are sorted from brighter to darker values
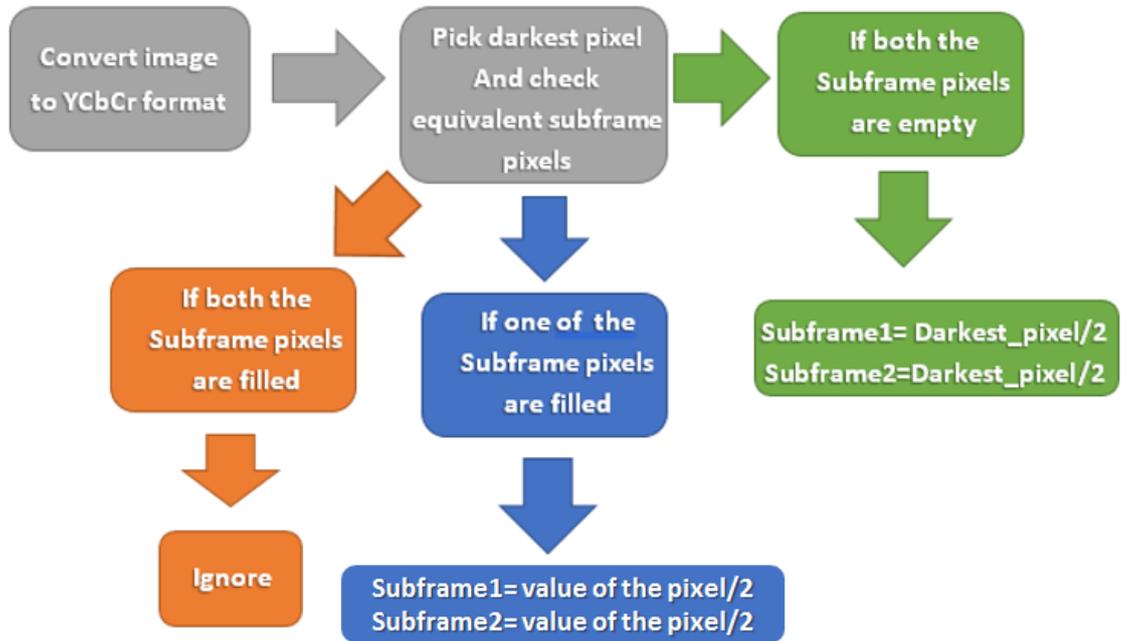
Figure 3.2: Flow chart showing Dark Priority technique.

and the values for other pixels are picked from same positions as in luminescence channel.

3.1.3.3. *Dark and bright priority technique*

The idea of combining the results from prioritizing dark pixels with bright pixels came after observing the results from both (shown for test image 2 in Figure 3.5 (j)). They both act as filters and one has details which are missing in the other and vice versa. Dark Priority acts as a filter which passes only darker details while Bright Priority passes only brighter details. So combining both seemed a good idea to preserve the details in all the areas. Therefore, first subframes are generated using Dark Priority algorithm and then using Bright Priority algorithm. Then the mean value of the two subframes are taken as subframes for the superimposition. The flow of this method is illustrated in Figure 3.3.
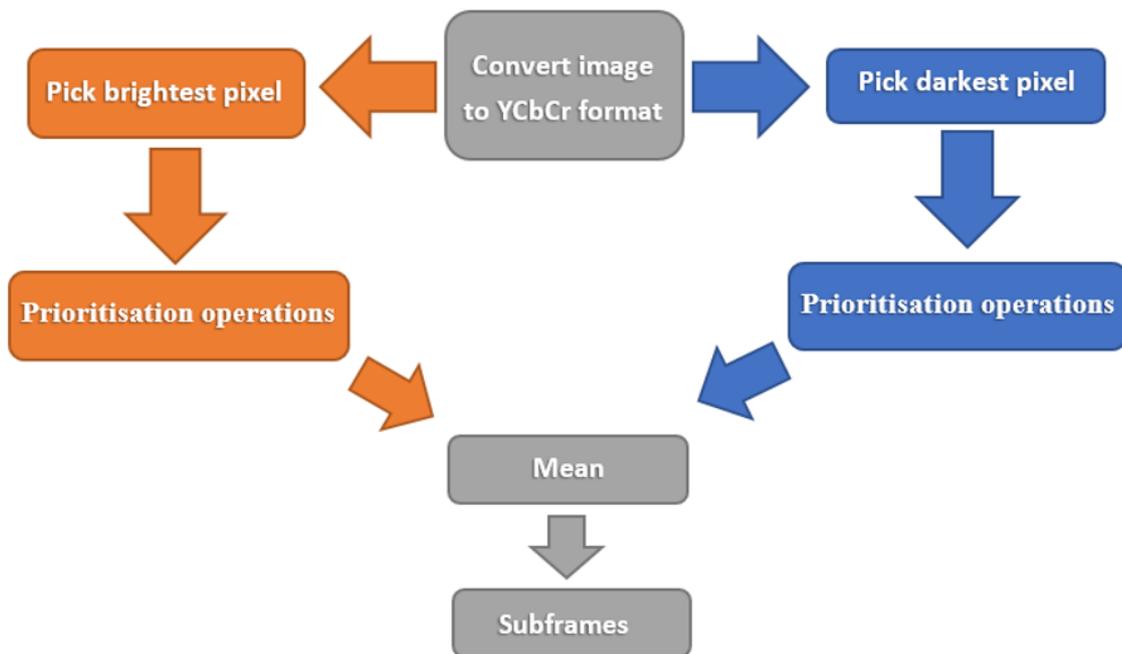
Figure 3.3: Flow chart showing Dark and Bright Priority technique.

3.2.   Results from simulation

Simulations were done using MATLAB with three test images. The subframes were generated using the techniques described above. The subframes were then scaled to mimic the high resolution image grid by repeating each pixel four times in a 2x2 block such that each block can be considered as a pixel when projected. One of the subframe was then shifted diagonally by one pixel on the high resolution image grid and superimposed on the other. The resulting images were inspected both visually as well as measured using the Image Quality Metric MSSSIM. Visual comparison is made between the downscaled version of the image at the native resolution of the projector and the high resolution version of the image.

Figures 3.4, 3.5 and 3.6 show the superimposed image results, the original high resolution image and the downscaled version of the test image 1, test image 2 and test image 3, respectively.
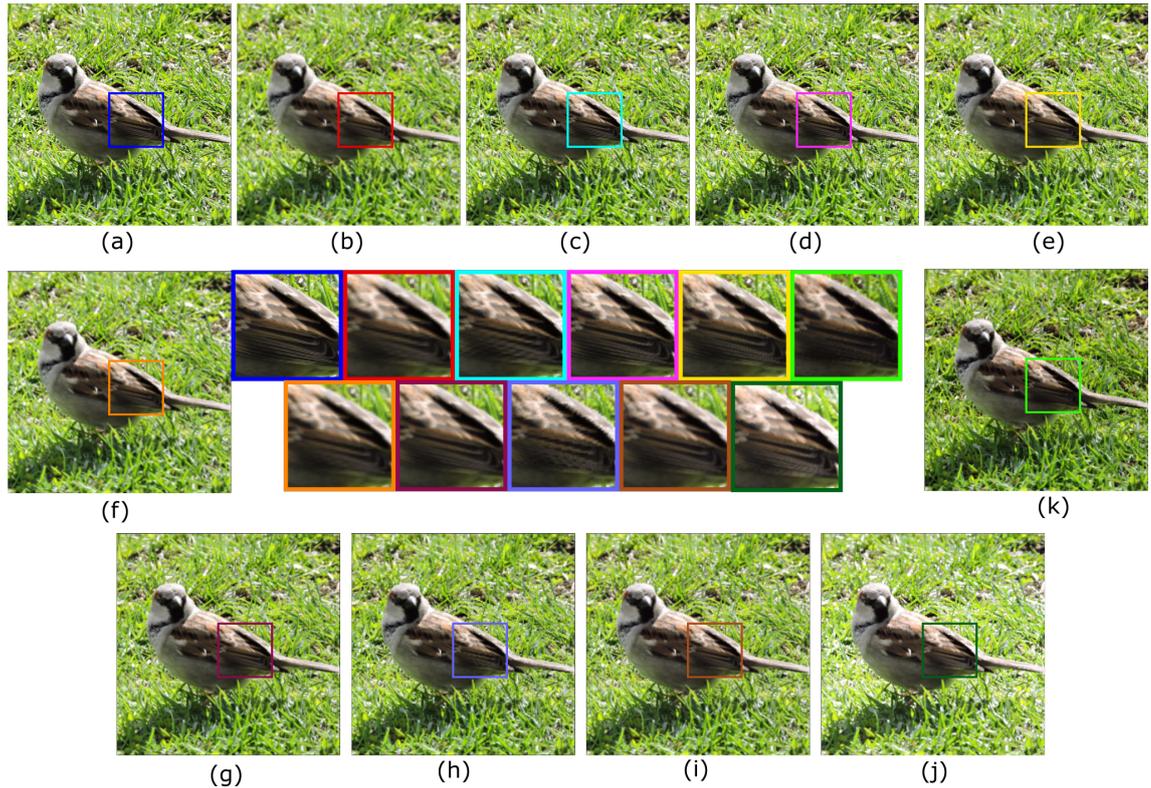
Figure 3.4: Simulation results for test image 1 for (a) original high resolution image (400x400 pixels) (b) Downscaled version of the image (200x200 pixels) (c) Single-subframe iterative technique (d) Two-subframe iterative technique (e) Naïve technique (f) Gaussian sharpened technique (g) Pick Mean technique (h) Pick Min. and Max. technique (i) Dark Priority technique (j) Bright Priority technique (k) Dark and Bright Priority technique
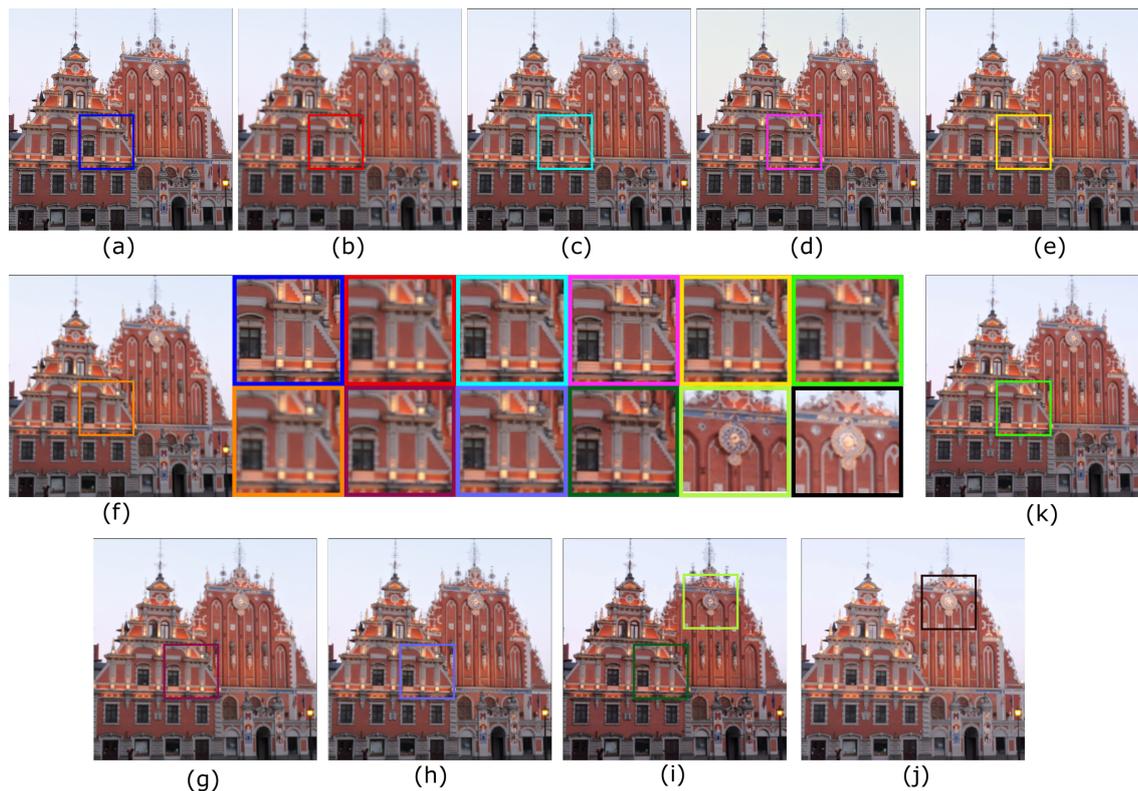
Figure 3.5: Simulation results for test image 2 for (a) original high resolution image (400x400 pixels) (b) Downscaled version of the image (200x200 pixels) (c) Single-subframe iterative technique (d) Two-subframe iterative technique (e) Naïve technique (f) Gaussian sharpened technique (g) Pick Mean technique (h) Pick Min. and Max. technique (i) Dark Priority technique (j) Bright Priority technique (k) Dark and Bright Priority technique
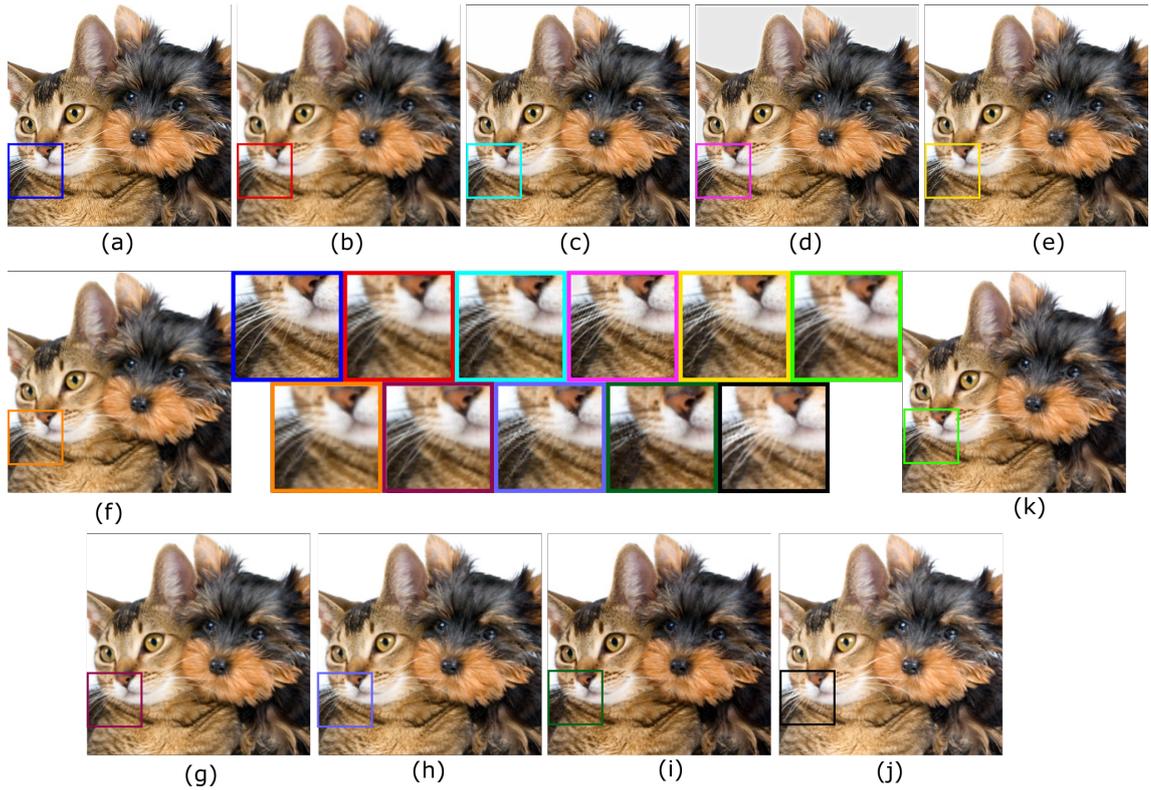
Figure 3.6: Simulation results for test image 3 for (a) original high resolution image (400x400 pixels) (b) Downscaled version of the image (200x200 pixels) (c) Single-subframe iterative technique (d) Two-subframe iterative technique (e) Naïve technique (f) Gaussian sharpened technique (g) Pick Mean technique (h) Pick Min. and Max. technique (i) Dark Priority technique (j) Bright Priority technique (k) Dark and Bright Priority technique

3.2.1. Techniques based on the naïve approach

3.2.1.1. Naïve technique

The results from the superimposition of subframes generated using the Naïve techniques show more details which are missing in the downscaled version of the image. Figures 3.4 (b), 3.5 (b) and 3.6 (b) show downscaled version of the image. There are definitely more details in the resulting image as shown in figures 3.4 (e), 3.5 (e) and 3.6 (e). The magnified part of Figure 3.5 (e) show the edge of the roof is detailed in much better way in the resultant image than in the downscaled version; however resultant images are blurrier than the high resolution image shown in 3.4 (a), 3.5 (a) and 3.6 (a) in all the test images.

3.2.1.2. Pick minimum and maximum technique and pick mean technique

Picking actual values from the input image, as seen in the Naïve approach or Pick Minimum and Maximum approach, produces visually better results than the Pick Mean value approach, although the MSSSIM value is higher for the later. Figures 3.4 (g), 3.5 (g) and 3.6 (g) show results from the superimposition of subframes generated by the Pick Mean technique while figures 3.4 (h), 3.5 (h) and 3.6 (h) show results from the Pick Min. and Max. technique.

Clearly, visual inspection suggests superior results for Pick Min. and Max. technique. This shows that in the image data, it is important to get the exact values for the pixels from the input image to achieve better visual experience, and the MSSIM metric misses this in the assessment. In comparing with the downscaled image which would be what a native projector would project, these techniques still produce better results.

3.2.1.3. Gaussian technique and Gaussian sharpened technique

Figures 3.4 (f), 3.5 (f) and 3.6 (f) show results from the Gaussian sharpened techniques. The use of Gaussian filter further decreases the noise in the Naïve result but introduces blurriness and we see an increase in the average MSSSIM value from

0.95 to 0.96. Gaussian sharpened technique makes the image slightly sharper in the visual inspection but does not improve the MSSSIM value.

3.2.2. Iterative methods

3.2.2.1. Single-subframe method

As shown in the magnified part of the Figures 3.4 (c), 3.5 (c) and 3.6 (c), the superimposed images are sharper than the downscaled image in the case of single-subframe iterative technique and they look almost as sharp as the original image. There are additional details present in the superimposed image which are clearly missing in the downscaled image. Such as in Figure 3.4 (c), the feathers in the lower left part are completely missing in the downscaled version, while they are present in the superimposed image.

3.2.2.2. Two-subframe method

The superimposed images observed are close to the original image apart from the noise which is shown in the magnified parts of figures 3.4 (d), 3.5 (d) and 3.6 (d). There are small tile-like rectangular artefacts which are the main differences between the original and the superimposed image.

The quality of the iterative techniques is also dependent on the number of iteration and the convergence tolerance selected in the GBP algorithm. Hence, another experiment was conducted where the quality of the superimposed image was measured using the MSSSIM metric after every iteration. The aim of this experiment was to see if it is possible to achieve good result with less iterations and to determine if it can be practical to use this technique. The experiment was conducted on all three images and the number of iterations was taken from 1 to 20. Results are based on 1 to 20 iterations and are shown in Figure 3.7. It is evident from the graph that the five iterations are sufficient with this technique as the MSSSIM values saturate to their peak after five iterations. Still it takes significant time to do five iterations. The result after 5th iteration and 20th iteration for the test image 2 are shown in Figure 3.8.
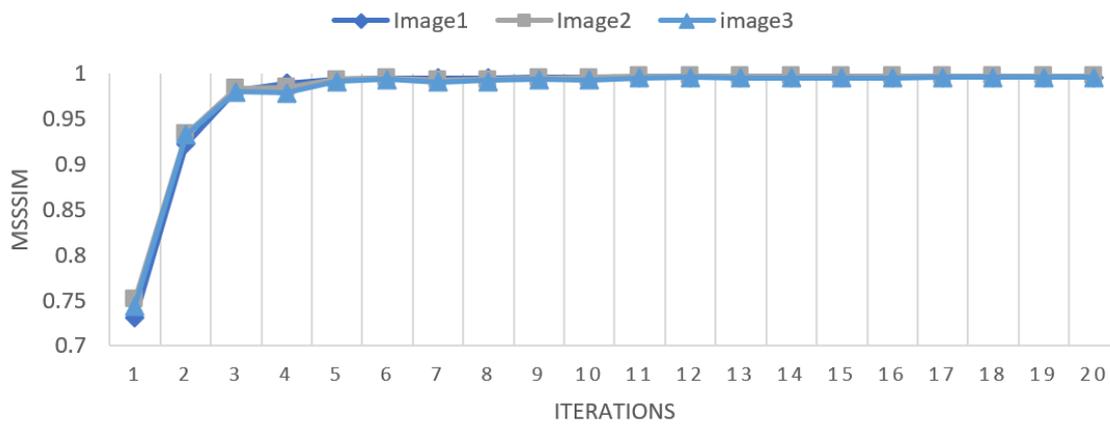
Figure 3.7: Chart showing the result of iteration test
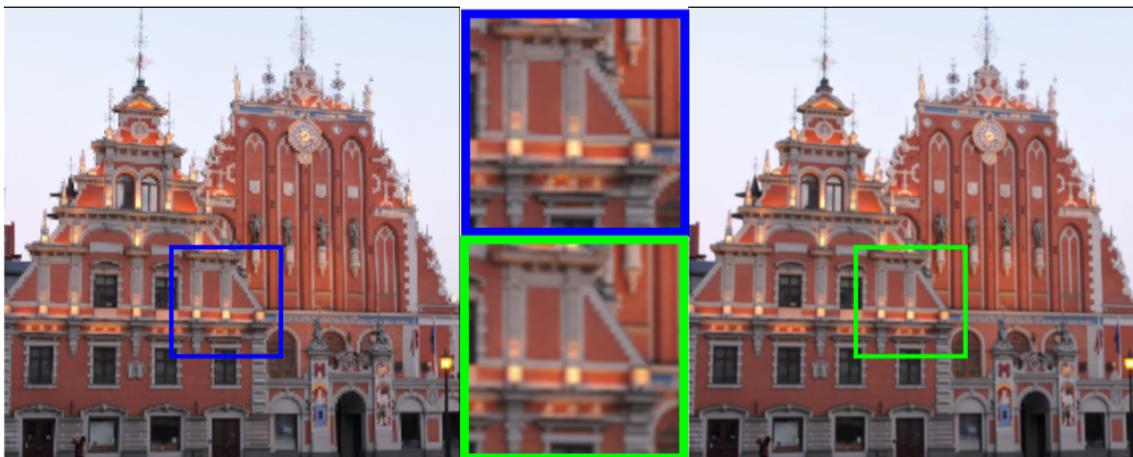
Figure 3.8: Comparison between 5th (left) and 20th iteration (right) for the Two-subframe iterative technique

3.2.3.   Priority techniques

3.2.3.1.   Dark priority technique

Figures 3.4 (i), 3.5 (i) and 3.6 (i) show the superimposed image result from the Dark Priority technique. The resultant images are much better when they have less bright pixels than in the original image. This can be seen in Figure 3.4 (i) where magnified part shows feathers of the bird with much better clarity. The resultant image however misses brighter details which are prominently seen in the test image 3 as shown in Figure 3.5 (i), where the white whiskers of the cat are missing. The simulation result of this technique showed that while dark details appear faithfully in the result as expected, brighter details are missing.

3.2.3.2.   Bright priority technique

Figures 3.4 (j), 3.5 (j) and 3.6 (j) show the result from Bright Priority technique. The images seem brighter than the original image as bright details seem to glow. This however results in the absence of darker details in some parts of the image.

3.2.3.3.   Dark and bright priority technique

Figures 3.4 (k), 3.5 (k) and 3.6 (k) show the results from Dark and Bright Priority technique. The resulting superimposed images have details which are represented better than the downscaled version. The magnified part of the Figure 3.5 (k) shows that details of the roof area are better presented in the resulting image than the downscaled version. The images are however blurrier than the original image. While the Dark Priority and Bright Priority techniques brings out darker and brighter details respectively, they are both incomplete as other details are missed out. Dark and Bright Priority technique tries to preserve features from both the individual techniques such that none of the details are missed out in the resultant image.

3.2.4.   Image quality assessment

Figure 3.9 shows the average MSSSIM values of the three test images for all the techniques discussed here. According to the MSSSIM values, the best technique is

the two-subframe iterative technique. This is the expected result, and in line with the visual inspection. It is followed by Pick Mean technique which is not in agreement with the visual assessment. This technique produced blurry details when we look at the resulting image. This result highlights the weakness of the MSSSIM metric. Even though MSSSIM metric does much more than comparing the error between the two images, it fails to recognize the blurriness observed visually. Next best technique is Single-subframe iterative technique followed closely by Dark and Bright Priority technique. Among the Naïve based techniques, even though Naïve technique itself scored average in the MSSSIM metric, the visual appearance was much sharper and most details were reproduced, albeit with some degree of distortion.

3.3.   Measurement setup

Barco F70 4K/UHD projector was used for the experiments. A block diagram for the projector is shown in Figure 3.10. The Barco F70 4K/UHD projector has an optomechanical actuator in its optical path. The optical actuator shifts the projected image by half pixel in up-right direction with the frequency of 120Hz, giving a superimposed frame rate of 60Hz. The projector can also be used without the optical shift activated.

The experimental setup is shown in Figure 3.11. The two cameras (Phantom M310 high-speed camera and Nikon D5100 DSLR camera) are placed on either sides in front of the screen to ensure that they do not block the area where the subframes are projected. A DSLR camera is used to capture the superimposed image in higher resolution and with better quality than the high-speed camera. The resolution of Nikon D5100 is 4928x3264 pixels. It was controlled by the software provided by the manufacturer called "Camera control pro 2". The high-speed camera can capture images up to 3200 frames per second with 1280x800 pixel resolution. It is possible to increase the frame rate by decreasing the resolution. This camera was controlled by Phantom CV 2.8 control software provided by the manufacturer [19]. All the camera settings, such as the required resolution and the image capture speed, could
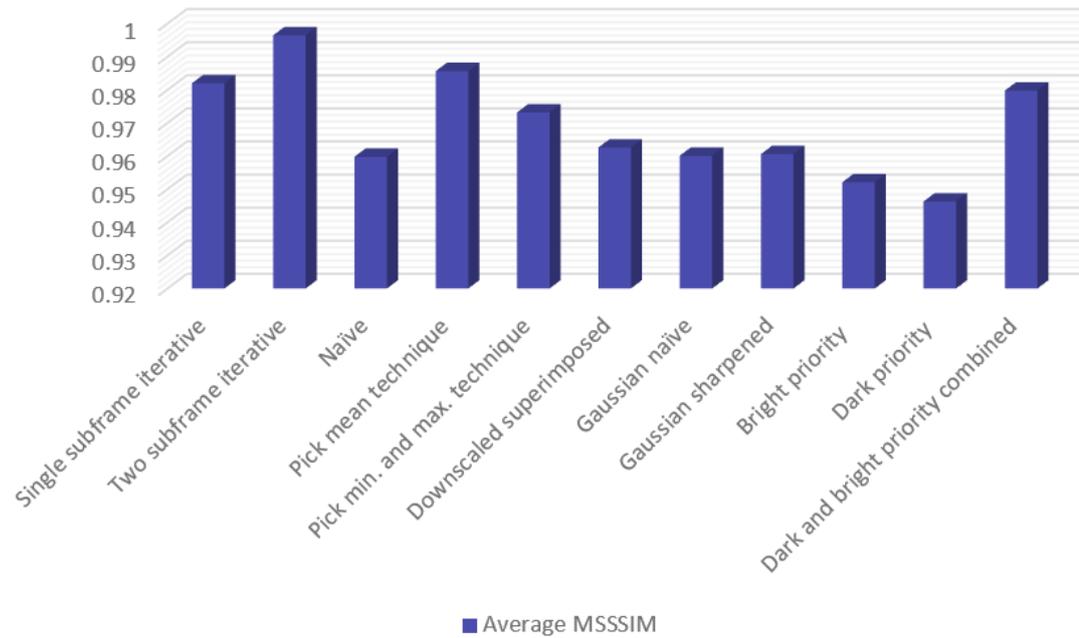
Figure 3.9: Chart showing the average MSSSIM values of three test images from simulation of different techniques
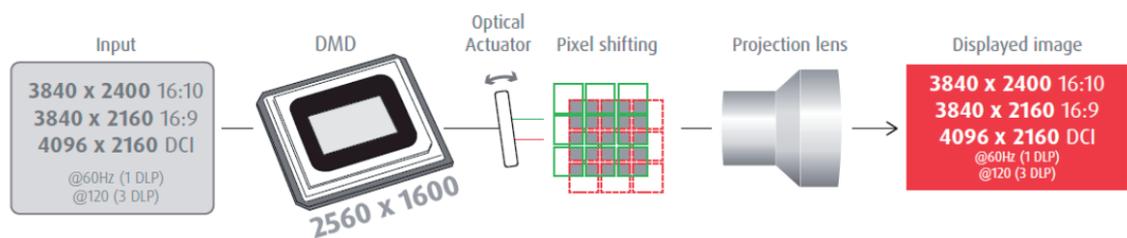
Figure 3.10: Schematic of Barco F70 4K/UHD projector with optomechanical shifter [3]

be controlled using the camera control software. It is essential to synchronize the high-speed camera with the projector actuator to ensure that the subframes are captured at the right instant. If not synchronized, the captured image can represent the transition between the two subframes which would have different brightness and color. This synchronization was enabled using a trigger input in the camera which was connected to the trigger output of the projector.

The experiment was done with test image 2 and test image 3 for all the techniques discussed earlier. The original images generated in MATLAB were 400 x 400 but framed with black background of 2560x1600 to fit the native resolution of the projector. The two subframes were sent as inputs from two different display ports to the projector.

3.4.   Results from the measurements

Figure 3.12 shows two individual subframes captured by the high-speed camera and the result from the simulation (top) and actual setup (bottom) for Two-subframe iterative technique. This illustrates the difference observed in the actual measurement due to noise and difference in brightness. They are visually similar apart from the noise and the brightness level of the two images. The noise is prominent due to the low resolution of the high-speed camera and limited brightness of the projected image. The results were similar for all the other techniques and the simulated and the captured images closely resembled each other.

The screen-door effect (the gap between the individual mirrors in the DLP chip which appears as dark grid on the screen) which was present in the native projector was absent in the superimposed versions. This shows the advantage of superimposition over the native projector. Downscaled superimposed image was used to observe the screen-door effect. Image was first projected without shifting the optical actuator and then projected with the optical actuator wobulating. Both results were captured with the Nikon D5100 camera. The screen-door effect can be seen in Figure 3.13, where the left side shows the image without shift and superimposition.
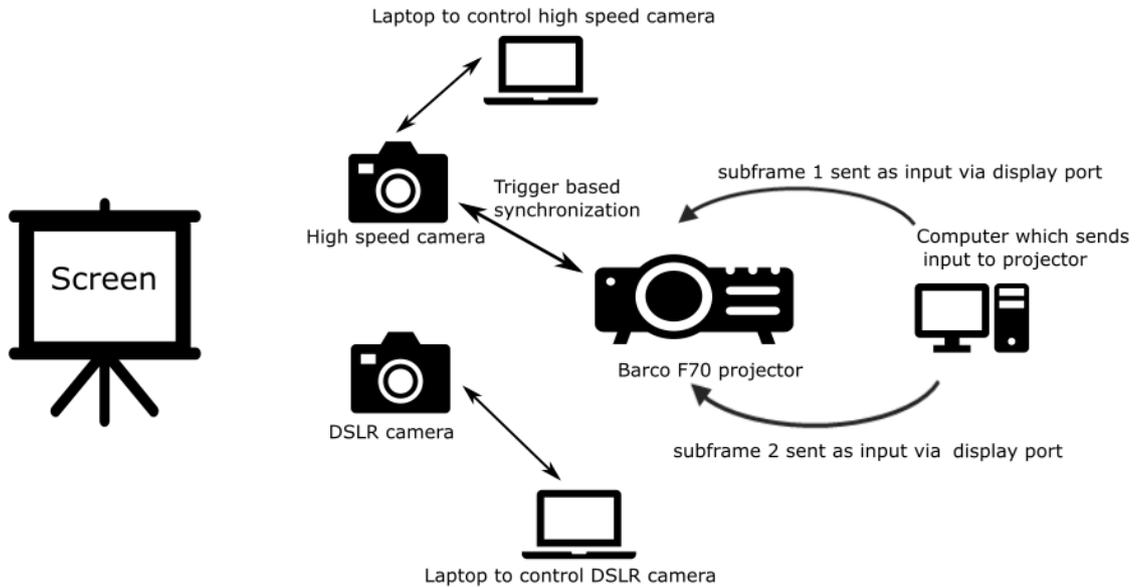
Figure 3.11: Schematic diagram of the measurement setup.

The magnified part shows small black boxes which are there due to the screen-door effect. The right-side image is the same image when the optical actuator is shifting and superimposing the image. Here the screen-door effect is not present making picture visually smoother.

3.4.1. Techniques based on the Naïve method

The measurement results were similar to the simulations, the blurriness of Pick Mean method was prominent when compared to the Naïve and the Pick Minimum and Maximum techniques. It was difficult to find a clear advantage of using the Gaussian filter and sharpening filter due to the difference being marginal and the presence of noise.

3.4.2. Iteration-based techniques

Unlike results from the simulation, the difference is not remarkable between the two iterative techniques, but the Two-subframe technique still shows superior results as compared to the Single-subframe technique. The results are shown in figures 3.14 (a),(b) and 3.15 (a),(b). The ability of the Two-subframe technique to produce fine
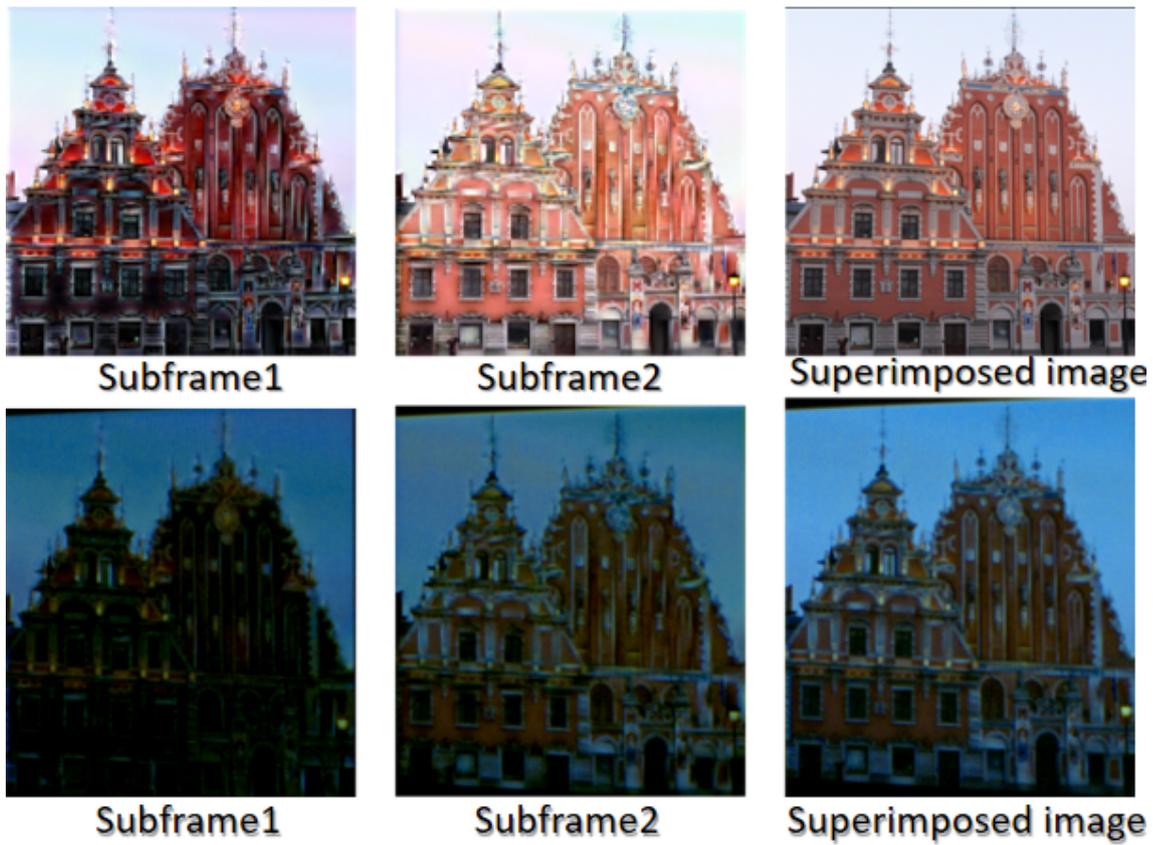
Figure 3.12: Superimposed image from simulation (top) and superimposed image from experiment (bottom) by Two-subframe iterative approach
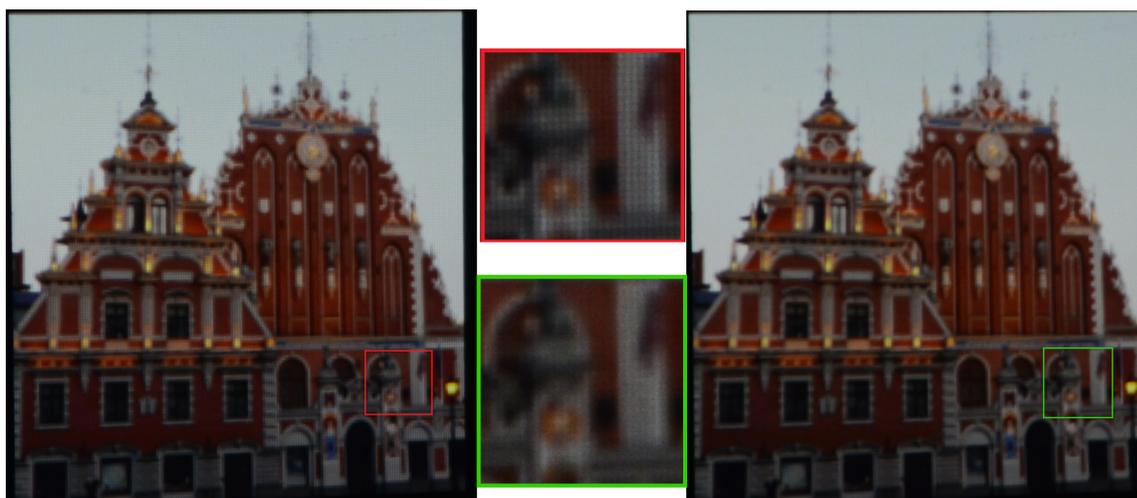
Figure 3.13: Screen-door effect observed for test image 2, with (right) and without (left) superimpositioning enabled

detail is hard to notice due to the presence of noise. However, more prominent artefacts seen in the simulation of Single-subframe technique are still present in the measurement. For example, in the Figure 3.15 (b), the noise seen in the whisker of cat is present in the measurement results as well. This part is smooth in the Two-subframe technique (Figure 3.15 (a)), showing the superiority of the Two-subframe technique.

3.4.3. Priority-based techniques

The measurement results are similar to what was seen in the simulations. Darker details are present and brighter details are missing in case of Dark Priority (shown in figures 3.14 (g) and 3.15 (g)). On the other hand, much more details are present in case of Dark and Bright Priority technique as shown in figures 3.14 (h) and 3.15 (h).

3.5. Discussion

As a baseline, we can take the downscaled superimposed image which is simply a downscaled version of the high resolution image superimposed on itself. Techniques that score higher than this method in terms of MSSSIM metric and/or visual assessment can be said to have improved the result. Although MSSSIM metric itself is not a conclusive evidence for usefulness of the technique, as demonstrated by the visual inspection result of Pick Min. and Max. technique and Pick Mean technique, it is still a good indicator. The simulations and measurements indicate that the iterative techniques produce the best results among the techniques studied in this work. Both iterative techniques (Single-subframe and Two-subframe) scored greater than 0.98 in the MSSSIM metric evaluation. The major downside of these techniques is the extra computation time needed to calculate the subframes. Among the techniques based on the Naïve approach, Pick Mean (0.984) and Pick Min. and Max. (0.97) are the two techniques that produce better results than the downscaled superimposed method according to the MSSSIM metric. This indicates that other three techniques (Naïve, Gaussian and Gaussian sharpened) do not do much better than
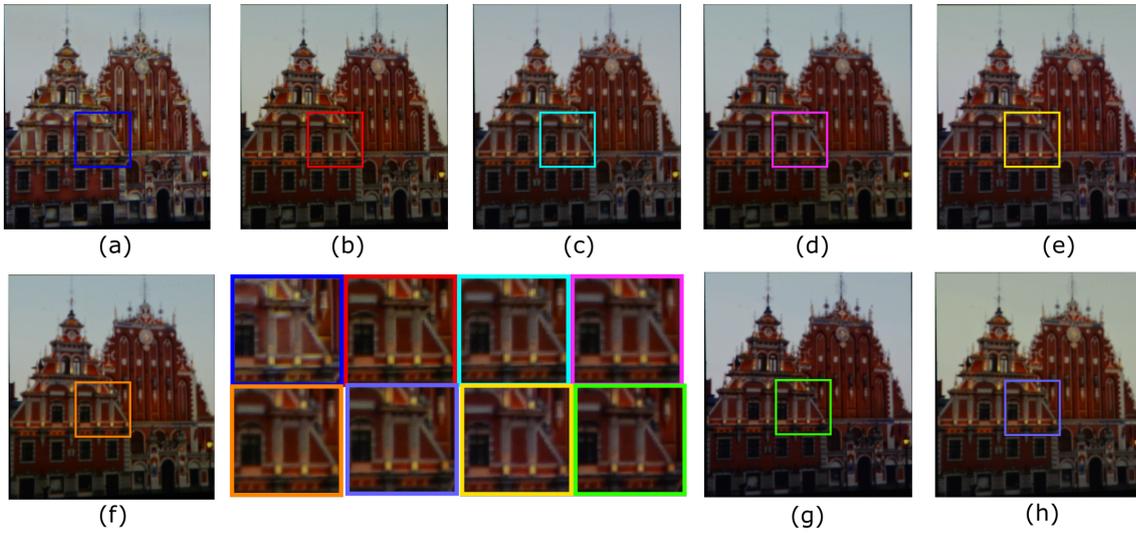
Figure 3.14: Results from measurements (a) Iterative technique (Two-subframe) (b) Iterative technique (single subframe) (c) Naïve technique (d) Pick Mean technique (e) Pick Min. and Max. technique (f) Gaussian sharpened technique (g) Dark Priority technique (h) Dark and Bright Priority technique
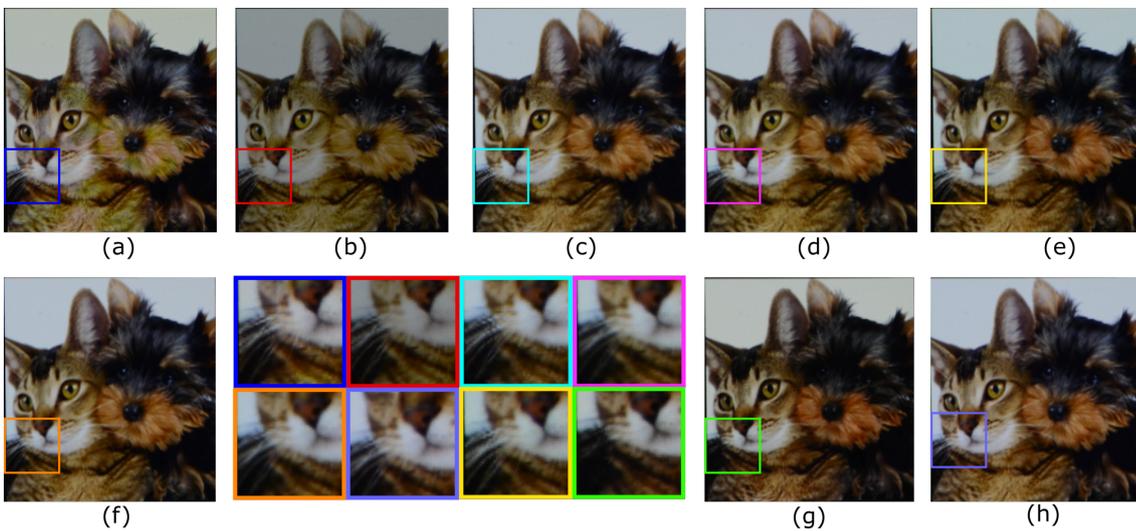
Figure 3.15: Results from measurements (a) Iterative technique (Two-subframe) (b) Iterative technique (single subframe) (c) Naïve technique (d) Pick Mean technique (e) Pick Min. and Max. technique (f) Gaussian sharpened technique (g) Dark Priority technique (h) Dark and Bright Priority technique

superimposing the downscaled version of the high-resolution image. Even though Pick Mean technique scores well in the MSSSIM metric, visual inspection indicates otherwise. The blurriness observed in the resulting images suggests that Pick Min. and Max. technique is the best among the Naïve based techniques. The Dark Priority and Bright Priority techniques bring out darker and brighter details well, but both of them score lower than the downscaled superimposed image. The Dark and Bright Priority technique however gives higher value (0.978) in the MSSSIM metric evaluation as well as in the visual inspection. It does not miss out on any details although it does have some blurriness in the resulting image produced due to the averaging operation. Even though only one property (brightness) has been studied in this work, it is possible to extend the techniques for other properties such as certain frequency range in the image or even certain part of the image. This technique is also computationally expensive due to the need to sort the image matrix according to its brightness values.

Comparing the results of the simulations and the experiments, some of the artefacts seen in the simulations seem to become less prominent in the actual measurement because of the noise and low resolution of the high speed camera. A significant case for this is the Iterative techniques which did not stand out as much as they did in the simulations, due to noise which took away possibilities to see fine details which iterative techniques are excellent at preserving. Even so, the results were devoid of blurriness and artefacts as seen in other techniques, which made this technique best among the studied techniques.

3.6. Conclusion

In this work we have evaluated different existing and newly developed subframe generation algorithms and assessed the superimposed image quality by detailed simulations as well as by experiments. Two-subframe iterative technique gave the best results but is not practical due to high computational resources needed. Among the Naïve techniques, Pick Min. and Max. technique produced the best results

and it is simpler to implement. The results from the Pick Mean technique showed that the metrics such as MSSSIM alone cannot be conclusive in determining the quality of the subframe generation technique. Priority-based techniques based on the brightness of the scene were developed and tested for the first time. Dark and Bright Priority-based method showed the best results among the priority-based techniques. However, these techniques are also computationally expensive. In the future, the priority-based techniques can be extended to cover other properties, such as frequency content of the image. Finally, experiments on a projector with an optomechanical shifter showed similar results to the simulations, apart from the noise present in the captured images due to the low resolution of the high-speed camera and low brightness of the projected scenes.

4. Shifting direction: Effect on the image quality

Most projectors that utilize the superimposition method today have an opto- mechanical actuator that spatially shifts every n^{th} frame with sub-pixel precision [8]. The two most common shift configurations are either half a pixel in one diagonal (two positions) as shown in Figure 4.1 (a), or half a pixel in both diagonals (4 positions) as shown in Figure 4.1 (b).
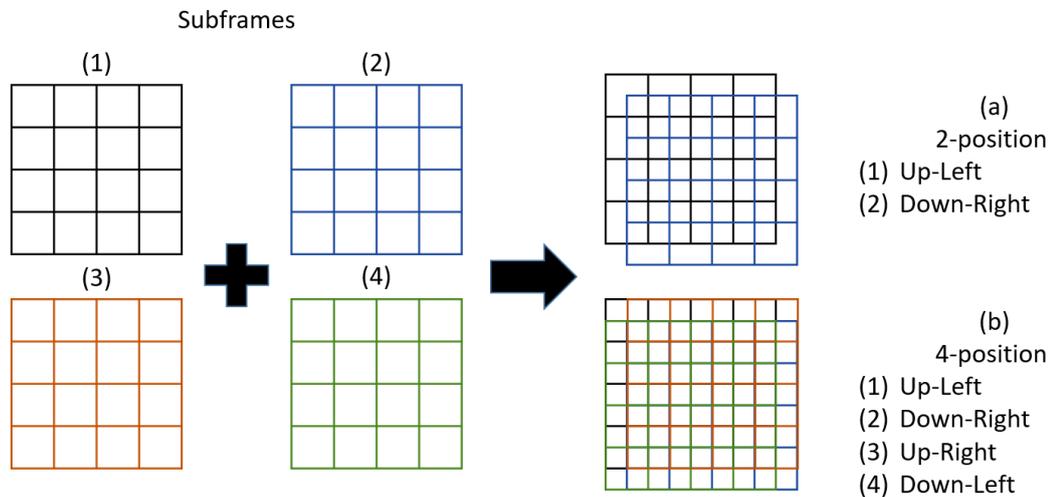
Figure 4.1: Illustration on how subframes may be superimposed in a two position system and in a four position system. In the two position system either diagonal may be used, and in the four position system both diagonals are used.

As shown in the bottom row of Figure 4.2, different spatial shifts of the overlaid pixel grid give different frequency responses for the overlapping pixels. The non-uniformity of the frequency response indicates that there are some directions that are better represented than others in the frequency domain. With a diagonal shift the orthogonal diagonal is the preferred direction, while the shifted diagonal will lose some high frequency content. This is seen by the width of the frequency response in a given direction in Figure 4.2. The wider frequency band in the orthogonal direction, indicates that higher frequency components are being reproduced in the image in that direction. When shifting in the diagonal directions, the shift produces

a new grid of overlapping sub-pixels where all pixels are of equal size and shape. This equality makes both the horizontal and vertical direction have equal priority. The frequency response of different pixel shifts is also described further by Sajadi *et al.* [19].

The frames that are projected in the different spatial locations are called subframes, and the total quality and resolution enhancement gain of the final superimposed image is affected by how these subframes are generated. One approach is to use the same subframe in both positions, and calculate the optimal subframe for overlapping on itself. This approach is computationally cost effective, but often results in sub-optimal quality. Another approach is to generate different subframes for each spatial position. Calculating a new subframe for each position is computationally expensive, but it also results in a better quality superimposed image.

Most projection systems that use the shifted superimposition method shift in one or two fixed diagonals. In this work we investigate if this fixed diagonal configuration is optimal, or if we should rather shift in directions other than the diagonal(s). In Section 4.1 we describe our approach to this subject and describe the method used in our analysis. In Section 4.2 the results from the analysis are presented and in Section 4.3 we discuss these results. Section 4.4 concludes on the subject.

4.1.   Our approach

In this work we explore the impact of the shift direction, with the objective to exploit the non-uniform frequency response of the overlapping pixels. Such understanding of shifted superimposition is important to achieve optimum image quality and design of the optimal system using this resolution enhancement technique.

As part of this, we explore how the high frequency content from a high resolution image source is preserved when displayed with a projector with a lower native SLM resolution than the image source. We also look at the overall image quality in
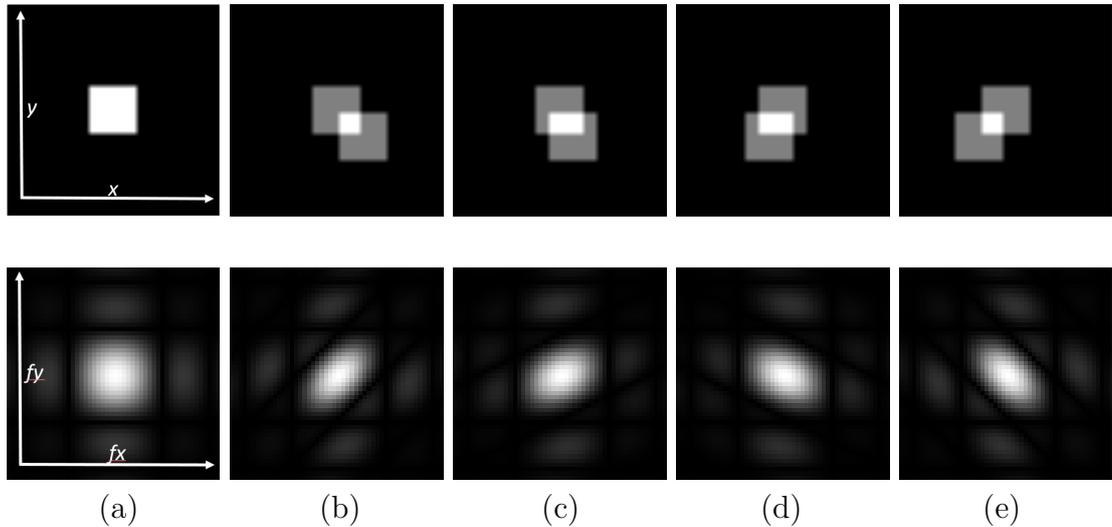
Figure 4.2: Overlapping pixels in different configurations. Spatial representation in top row. Frequency response represented by the magnitude of the Fourier transform in the bottom row.

the reproduced superimposed image through the use of the Image Quality Metric MSSSIM [27].

Shifted superimposition is a resolution enhancement method to enable the display to reproduce higher frequencies than the display would be capable of at the native SLM resolution. The first part of the analysis determines the preservation of high frequency content in the image. We analyze each image with a histogram of oriented gradients (HOG) analysis [46]. The HOG is divided into predefined bins to quantify the general high frequency information in different directions.

We do not discriminate whether the gradient is on an upward or downward slope, so both 0 and 180 degree angles will be placed in the same bin. The 180 degree angle space is then divided into 8 bins, giving us bins for 0-22.5, 22.5-45, 45-67.5, 67.5-90, 90-112.5, 112.5-135, 135-157.5 and 157.5-180 degrees.

The quality of the enhanced images is evaluated using the MSSSIM metric. Chapter 2 showed that this metric is suitable to evaluate the quality in resolution enhanced superimposed images.

There are several ways of generating the subframes for the shifted superimposition, and some methods were described in Chapter 3. In this work we include one single subframe method where the same subframe is used in both positions, and one different subframes method where the two different subframes are calculated from the high resolution source image based on their spatial shifted position. We have also included *downscaling to native resolution* as a reference to the projector's performance without any resolution enhancement methods applied.

4.1.1. Downscaling to native SLM resolution

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 and is taken as the reference image for comparison. The resulting output image will then be given by

$$OutImage = Resize(RefImage, SLMresolution), \tag{4.1}$$

where $Resize$ is the corresponding Matlab function and $SLMresolution$ is the resolution of the SLM in use. $RefImage$ is the original high resolution image.

4.1.2. Single subframe

It is cost efficient to use the same subframe in both positions, as this also keeps the complexity of the system down to a minimum. This scenario may also be accomplished by different optical configurations as shown by Sajadi *et al.* [19], but in a static configuration it will be difficult to change the direction of the shift from frame to frame. A system using single subframe for both positions is perceived as producing better quality images than the native resolution, and we aim to investigate how the direction of the shift influences this quality gain. Allen and Ulichney used a version of this method to verify the concept of superimpositioning in their original paper on this subject [7].

The subframe generation method used in this experiment is the downscaled super-imposed method described in Chapter 2. In the downscaled superimposed method, we downscale the high resolution frame to the SLM resolution and superimpose the same frame on itself. Since the overlapping pixels will be of same value and strength, this method should follow the frequency response in Figure 4.2 (b)-(e).

The subframes in the single subframe method in this work are given by

$$SubframeA = SubframeB = Resize(RefImage, SLMresolution). \qquad (4.2)$$

4.1.3. Different subframes

Using different subframes based on the spatial position of the shift is expected to be optimal concerning the image quality. In this configuration the two different subframes take different details in the high resolution input image into account to ensure higher detail preservation. The subframe generation method used in this experiment is the Gaussian Sharpened method described in Chapter 2. In the Gaussian Sharpened method, we upscale the input image to a larger pixel grid before we filter this up-scaled image with a Gaussian filter. Then we apply a sharpening filter to remove some of the blur added by the Gaussian filtering, but with the possibility of

amplifying noise in the image. The parameters and filters are described further in Chapter 2.

$$
\begin{aligned}
IntermediateResolution = {} & (m, n) * SLM resolution \\
IntermediateFrameNaive = {} & Resize(RefImage, IntermediateResolution) \\
IntermediateFrameGauss = {} & imfilter(IntermediateFrameNaive, \\
& fspecial('gaussian')) \\
Sfilter = {} & b * [0, a, 0; a, (-4) * a, a; 0, a, 0] \\
InputImageMask = {} & imfilter(IntermediateFrameGauss, Sfilter) \\
IntermediateFrameSharpened = {} & IntermediateFrameGauss+ \\
& k * InputImageMask \\
SubframeA(i, j) = {} & IntermediateFrame(m * i - 1, n * j - 1) \\
SubframeB(i, j) = {} & IntermediateFrame(m * i, n * j)
\end{aligned}
$$

$$(4.3)$$

where m and n are the rescaling factor in horizontal and vertical direction to stretch the pixel grid according the direction of the shift. The constants in the sharpening filter are set to $a = -1$, $b = 0.25$ and $k = 0.5$.

The example above generates subframes for a 45 degree diagonal shift when $m = n = 2$. By shifting in different directions the grid size of the IntermediateResolution is adjusted and the values for SubframeA and SubframeB is then picked from the appropriate coordinates of IntermediateFrame.

In the experiments in this paper the length of the pixel shift is always set to half the pixel width in the direction of the shift. This shift length is chosen to generate the pixel grid as uniform as possible.

4.2. Results

We developed a Matlab simulator to explore the different subframe generating methods and how they perform within shifted superimposition with varying shifting direction. The resolution enhancement performance is evaluated by calculating the HOG preservation, and the quality of the superimposed image is evaluated through the use of the MSSSIM metric.

For determining the HOG preservation we calculate the HOG for the high resolution source image, and then we calculate the HOG for the resulting superimposed image. The ratio between gradients in the original image and gradients in the superimposed image gives us the gradient preservation for each histogram bin. A preservation of 1.0 means that all gradients in that direction are preserved, while a preservation of 0.4 will mean that 40 percent of the gradients are preserved and 60 percent of the gradients in that particular direction are lost. Based on the non-uniformity of the frequency response we expect to see that gradients in some directions are better preserved than in other directions.

Figure 4.3 shows a cutout of a scene where we have the original high resolution source image at Figure 4.3 (a) and the native SLM resolution version at Figure 4.3 (b). When a display receives an image with a resolution higher than its number of pixels, the display needs to downscale the received image to the native resolution of the display as mentioned before. This is shown by the SLM resolution example in Figure 4.3 (b). Figure 4.3 (c) through Figure 4.3 (f) shows the same scene shifted with the single subframe method shifted in various angles. In Figure 4.3 (c) we see that the horizontal shift introduces horizontal lines in the image and the vertical lines (which have gradients in the horizontal direction) are blurred out. This corresponds with the poor frequency response in the shifted direction as observed at Figure 4.2 (b)-(e). The vertical shift in Figure 4.3 (e) shows us the same behaviour, only that this time it is the horizontal lines (vertical gradients) which are blurred out.

(a) Reference (b) Downsampled (c) SSS 0º

(d) SSS 45º (e) SSS 90º (f) SSS 135º

(g) DSS 0º (h) DSS 45º (i) DSS 90º

(j) DSS 135º

Figure 4.3: Cutout of a forest cabin image. a) reference image, b) downsampled to SLM resolution c) SSS 0º, d) SSS 45º, e) SSS 90º, f) SSS 135º, g) DSS 0º, h) DSS 45º, i) DSS 90º, j) DSS 135º. SSS notes the Single Subframe Shifting while DSS notes the Different Subframe Shifting. Notice the image artefacts appearing in different directions according to the direction of the shift.

In the 45 degree (Figure 4.3 (d), up-left) and 135 degree (Figure 4.3 (f), up-right) cases, we also see the same behaviour, especially at the end of the pitched roof. At the roof's edge, we see that the method favours the direction orthogonal to the shift as suggested by the frequency response.

When using two independent subframes generated according to the direction of the shift, as shown in Figure 4.3 (g) through Figure 4.3 (j), we see that the same artefacts apply in the 0 degree and the 90 degree shift. The respectively horizontal or vertical blocking artefacts of the shift are very visible. Unlike the single subframe method, the different subframe method favours both diagonals when the direction of the shift is at 45 or 135 degrees. This behaviour is visible at the end of the pitched roof in Figure 4.3 (h) and Figure 4.3 (j).

As shown in Figure 4.4 the method that use a single frame for both positions in the superimposition has a very high directivity in its frequency response and therefore also preserve more gradients in the orthogonal direction of the shift. This is because all of the pixels are overlapping themselves and are generating a frequency response as shown in Figure 4.2 (a)-(e). The downside of this approach is that we are losing some details in the direction of the shift. This behaviour is visible in Figure 4.3 where we see that the gradients in the shifting direction are less distinct, which is also according to the frequency response in Figure 4.2 where we see that the bandwidth is more narrow in the shifting direction.

Figure 4.4 shows that both the diagonals have preserved most of their gradients in the different subframe cases. This result is also visible in Figure 4.3 where we see that both sides of the roof are displayed with high quality and that the blocking artefacts are not as dominant in these examples.

The quality of the resulting shifted superimposed images is calculated using the MSSSIM metric. The results of these calculations are given in Figure 4.5 where we present the graphs for both the single subframe and the different subframe case. Figure 4.5 shows that the different subframe method produces better quality re-
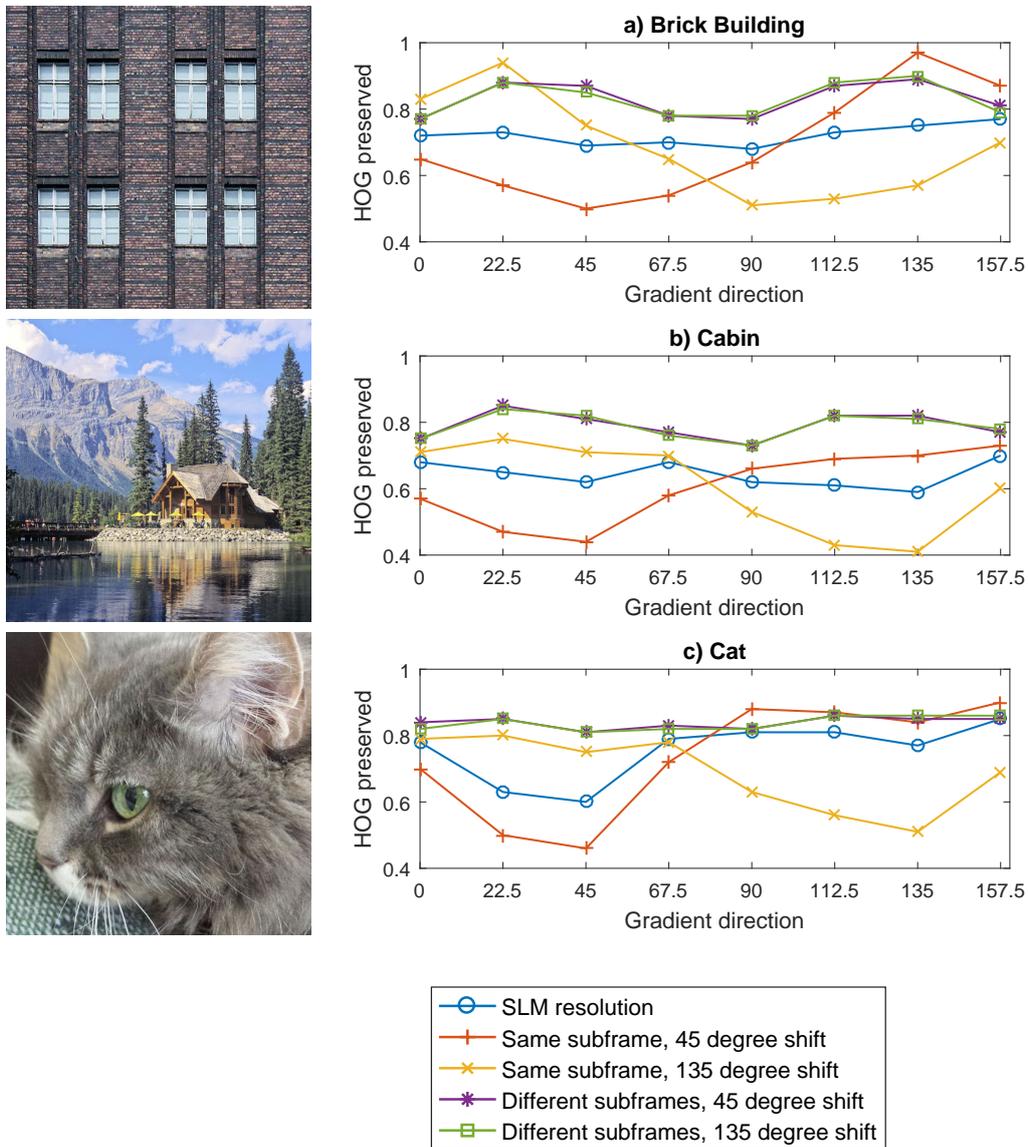
Figure 4.4: Gradient preservation shown as HOG analysis for three selected test scenes. In this example the SLM resolution is half the source resolution in both horizontal and vertical direction. The gradient direction starts at zero degrees in the positive x-axis, and increases counter-clockwise.

sults than the single subframe method. We can also see that the 0 and 90 degree shift (horizontal and vertical) produce the worst quality results within both of the methods.

One interesting observation is that even though the HOG preservation in Figure 4.4 shows that the single subframe method is very directional in reproducing gradients, the MSSSIM results in Figure 4.5 show that the directionality itself does not affect the quality.

4.3. Discussion

The enhanced resolution is a direct consequence of the decreased pixel pitch of the resulting pixel grid. The smaller pixels have the capability to display smaller image details, and the increased number of pixels are able to produce a larger number of details than the native SLM resolution. Thus it is logical that the 0 degree and 90 degree shift cause blocking artefacts in one direction since these shifts only decrease pixel pitch in one given direction. And since objects of interest often include a significant amount of gradients in the horizontal and the vertical direction, these artefacts becomes very visible.

These artefacts indicates that shifting in the horizontal or the vertical direction is not a good solution, since this shift only enhances the resolution in the vertical or the horizontal direction. Figure 4.3 illustrate these artefacts where we can see how the horizontal shift in Figure 4.3 (c) creates the illusion of horizontal stripes, and the vertical shift in Figure 4.3 (e) creates the illusion of vertical stripes in the image. The shift around the diagonals are more visually balanced.

The single subframe method is interesting for low cost systems since this method demands less calculation power. Chapter 2 showed that the downscaled superimposed method is preferred over regular downscaling to the SLM resolution. In their study they used 45 degree diagonal shift for all test scenes, but we see from the MSSSIM measurements in Figure 4.5 that the pure diagonal shift may not be the
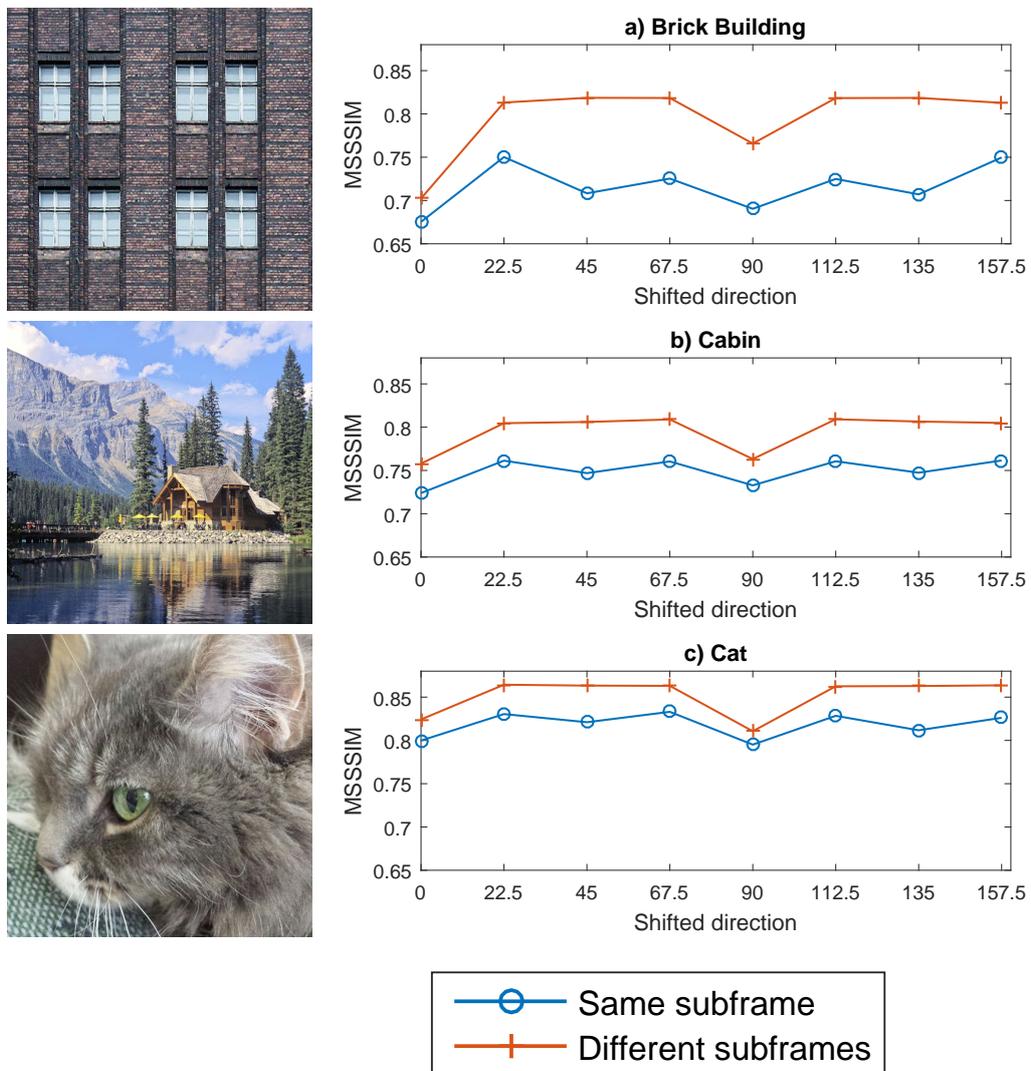
Figure 4.5: MSSSIM calculated for each directional shift for both the single subframe and the different subframe case. The test scenes are the same as referred to in Figure 4.4.

optimal shifted direction for the single subframe method. Potentially the single subframe method may score even higher in comparison studies such as those shown in Chapter 2 and Chapter 3 if a more optimal shifting direction is chosen.

Figure 4.2 shows the spatial overlap between pixels shifted in different directions, and how this overlap affects the frequency response of the overlapping pixels. One shortcoming of the analysis drawn from Figure 4.2 is that the overlapping pixels must be of equal strength and shifted in the same direction throughout the whole image for this response to be valid globally. This criteria is met for the single subframe method, and we see that the HOG preservation for the single subframe method in Figure 4.4 corresponds to the magnitude of the Fourier transform shown in the bottom row of Figure 4.2.

When different subframes are being used in each position, the overlapping pixels are no longer of equal strength. As shown in Figure 4.6 it is possible with the different subframe method to locally mimic different shifting directions by changing the pixel patterns locally in different parts of the image.

All of Figure 4.6 (a)-(d) is shifted in the same 45 degree diagonal direction. Figure 4.6 (c) and Figure 4.6 (d) illustrates how the different diagonal shifts may be imitated by using different subframes where different pixels are highlighted.

In Figure 4.6 (c) the right pixel is active in both positions while the left pixel is dark. This forms a pattern of a single pixel overlapping itself when shifted in the up-left diagonal direction. But in Figure 4.6 (d) the left pixel is active in the lower position (first subframe), and when the pixels are shifted in the up-left diagonal direction (second subframe) the left pixel goes dark while the right pixel becomes active. The resulting overlapping pattern is the same pattern as if a single active pixel is shifted in the up-right direction, which in this case is the orthogonal of the shifted direction. This property is what makes the different subframe cases in Figure 4.4 preserve both 45 degree and 135 degree quite equally. The two subframes are calculated to take into account the local gradients in all parts of the image,

and are therefore independent of the direction of the diagonal shift compared to the same subframe case.

The effect of adapting the frequency response locally instead of globally across the image, is observed in Figure 4.3 (d) and Figure 4.3 (f). In this Figure we see that both diagonals are displayed at a higher percieved resolution, where the low-resolution jaggedness of the cabin roof is enhanced in both directions.

The different subframe method will generate optimal pixel values based on the spatial position of the shift, and will therefore also shape the frequency response in a desired manner based on the image content even without adjusting the direction of the shift. In Figure 4.5 we observe that exact behaviour when noticing that apart from the 0 and 90 degree shift, most shifts produce approximately the same quality result with the different subframe method.

Another observation from the MSSSIM results in Figure 4.5 is that the diagonals are not necessarily the preferred shifting direction for the single subframe method. This also corresponds to the findings by Barshan *et al.* where they discover that the common half a pixel shift is not the shifting length that gives the best quality, but in several cases reducing the shift to 0.3 pixels may yield a better quality result [23]. The 22.5 degrees direction gives the best quality result at Figure 4.5 (a), and we see in general that the diagonals are local minima with the single subframe method for all the scenes in Figure 4.5.

4.4. Conclusion

The simulations in this work show that a projection system with enough computational power to generate each subframe based on its spatial position will be indifferent to the shifting direction. Both the gradient preservation in different directions and the quality of the image is equal when generating the different subframes according to the shift. The exception to this rule is shifting in horizontal or vertical direction, which gives lower quality and poor gradient preservation.

| (a) | (b) | (c) | (d) |
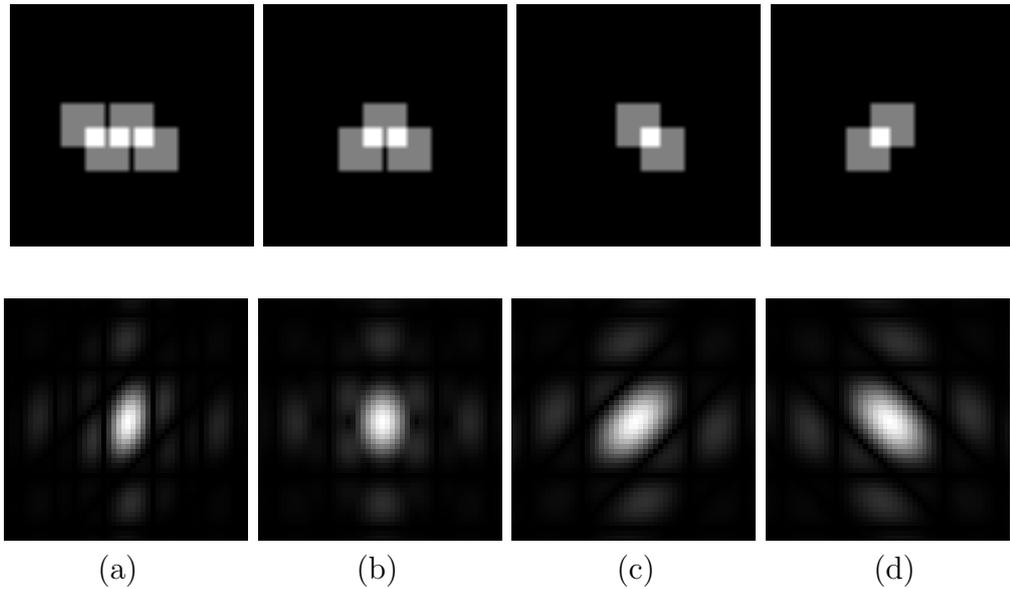
Figure 4.6: These illustrations are generated with two position 45 degree shift using the different subframe method. The figure shows that different weight to different pixels in the two subframes shapes the local frequency response of the overlapping pixels differently.

In a system with less computational power, it is a valid solution to use the same subframe for all positions since it does give better picture quality than the native resolution. Such a solution will make both the gradient preservation and the quality of the image dependant on the direction of the shift. It is important to take this directionality into account when designing such a system.

The different subframe superimposition method with a static shift along any of the diagonals will still be superior in quality to the single subframe method, but at a higher cost in complexity.

5. Influence of source resolution and resolution measurements

Even though the actual pixel count on the projected screen will increase with shifted superimposition, the resolution enhancement 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 [9]. The spatial artefacts manifest as blurring in the image, and these artefacts impacts both the visual quality and the resolution measurements. The introduced artefacts raise the question of how high resolution is actually achieved, and which factors impact the resulting resolution.

Pixel shifted displays challenge the traditional sense of resolution as this is a computational display and not a traditional display [47]. Each resulting sub-pixel is made up of the sum of the SLM pixels illuminating the resulting pixel, and each subframe is contributing to this. Because different subframes may include different details from the high-resolution source image, there will be more details present on the screen than the native resolution would be able to represent without the pixel shifting technology.

For the different subframes to be able to display different details from the source image, the source image must be of higher resolution and therefore contain more details than the native SLM resolution. This means that the resolution of the source image will affect the amount of available details for the projector to generate subframe details, and will therefore also affect the perceived resolution of the resulting superimposed image. In this work we experimentally investigate the relationship between source resolution and resulting superimposed resolution, and also how the resulting superimposed resolution is to be measured.

5.1.   Resolution measurement

Display resolution is in its simplest form defined as the number of pixels in the display available to form a image. This definition applies to all traditional forms of displays, and in a projected display the resolution would be the number of pixels in the SLM. There are atleast two aspects of the display that challenge this simple definition of resolution: The quality of the pixels and the physical build of the pixels.

The quality of the pixels may be seen as the pixels ability to represent different details. In most displays, the pixels are not completely independent of each other, and will in some form affect the pixels in near proximity with their own value.

This may for instance be as non-perfect optics in projected displays, backlight bleeding in LCD monitors, or as fringe field effects in LCD and LCOS displays [48]. All of these effects make the value of a pixel interfere with the appearance of the neighbour pixels. Shifted superimposition as a resolution enhancement method will add dependencies between some neighbour pixels because of the optical overlap of the pixels, and also because the SLM pixels illuminate a larger area than one resulting pixel in the superimposed pixelgrid.

The physical build and geometry of the pixels may also affect the perceived resolution of the displayed image. Projected displays usually have a uniform pixelgrid where the colors are superimposed on each other within the same pixelgrid. But other displays may have different pixel geometry, for instance flat panels with sub-pixel rendering where colors are adjacent to one another and arranged in a specific pattern. Even though each pixel in Figure 5.1 [49] is made up of a red, green and blue sub-pixel, those colored sub-pixels may be individually controlled to form different pixelpairs to increase the apparent resolution when needed.

These aspects illustrate the point of having a more thorough resolution definition than just the pixel-count. For a given display it will give direct information to mea-
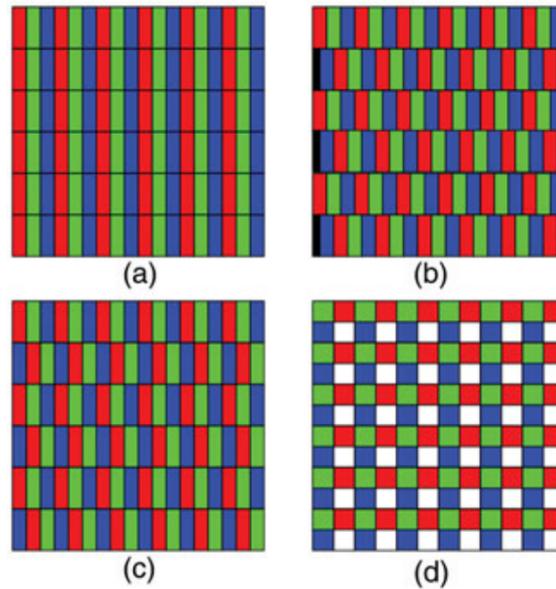
Figure 5.1: Examples of different pixel geometries in flat panel displays, and their sub-pixel placements. (a) RGB vertical stripe display, (b) RGB delta, (c) VPX (with 3 sub-pixel/pixel), and (d) VPW (with 4 sub-pixel/pixel).

sure the different specifications of the display, thereby also the resolution. By using the same measurement methods it will then also be possible to compare the performance of different displays. There is currently no measurement standard that is agreed upon by all display manufacturers in all markets, but there have been several measurement standards proposed. The International Committee for Display Metrology (ICDM) have included several proposals for spatial resolution measurement in the Information Display Measurements Standard (IDMS) [40].

5.1.1. Least resolvable line pairs

The least resolvable line pairs measurement is an established method for determining resolution based on measuring a limited number of line pairs. This is done by displaying an object in 3D-space consisting of a predetermined amount of line pairs. The object is then pushed back in 3D space further and further away, apparently shrinking the line pairs. When the line pairs are at the limit where they are nearly not resolvable anymore, the physical size of the line pairs is measured. The size of the whole projected screen is then divided by the size of a single line pair. By doing

these measurements both horizontally and vertically the total resolvable resolution in the system may be obtained.

One major drawback with this method is that it is very prone to subjective biases and measuring errors. The action of establishing when the line pairs are at their resolvable limit is in itself very subjective. In addition to that, any inaccuracy when measuring the physical size of the line pairs may have a major impact on the total resolution number. Another trait of this method is that it is measuring the system performance of the image generator, the projected display and the screen on at the same time. This fact may be a feature for a system integrator, but for measuring the display itself it introduces uncertainties.

Another drawback is that the results of this method will be dependant on the image generator. This could make this as a characterization method hard to compare objectively over different test sites and technologies. On the other hand it is a good system-level resolution measurement for fixed installations with in-system image generators. For these reasons it is a method often applied in the simulation industry where the display is a fixed component in the system.

5.1.2. Grille contrast modulation

The grille contrast modulation method is to measure the Michelson contrast with grille patterns (alternating black and white stripes) with different widths. The intermediate values between these points is extracted from linear interpolation between neighbour points as shown in Figure 5.2. This approach was proposed by The International Committee for Display Metrology (ICDM) within the Society for Information Displays (SID) in 2012 when they released the Information Display Measurements Standard (IDMS) [40].

The contrast vs line pair curve gives an indication on what contrast the display is able to reproduce for different detail sizes. In this sense, the curve gives not only a resolution measurement, but also a quality factor to that given resolution. It is debatable what contrast factor that is really needed for different display applications.
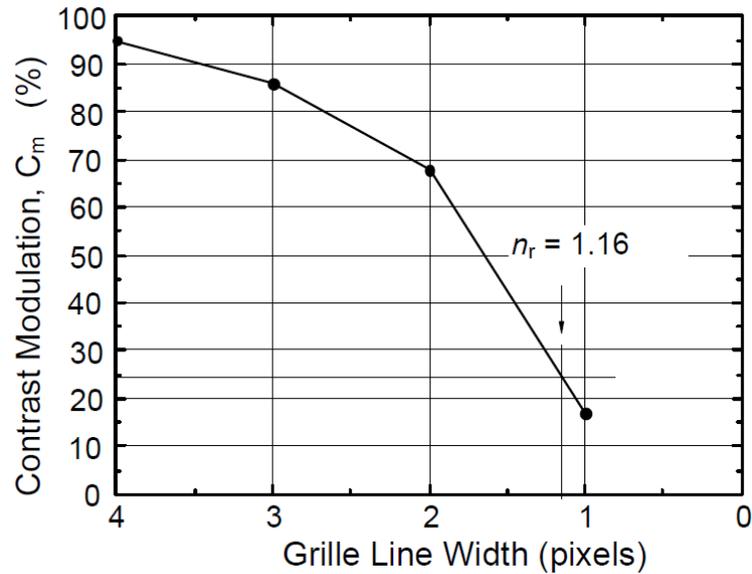
Figure 5.2: Contrast modulation example as shown in the IDMS standard. Measurements are done with Grilles of 1,2,3 and 4 pixels width, while linear interpolation makes up the intermediate values.

IDMS proposes that the display itself should have 50% contrast modulation for text applications and 25% contrast modulation for image applications. This means that the resolvable resolution at 25% contrast modulation is

$$ResolvableResolution = DisplayedResolution/n_r \qquad (5.1)$$

In the example given in Figure 5.2 the calculated grill line width n_r is 1.16, making the resolvable resolution lower than the input resolution to the display with a factor of 1.16.

5.1.3. Slanted edge measurement

The slanted edge method is to measure the spatial frequency response (SFR) as an approximation of the modulation transfer function (MTF). This method has become widely used within fields of optics, and has been adopted by multiple international standards within optics, including ISO and IEEE. The slanted edge approach for displays is described in the IDMS Chapter 7.7 [40].

In this method we measure the luminance of vertical or horizontal step patterns on the display with a slightly tilted camera so that the sampled image captured by the CCD camera oversamples the slanted edge. The SFR is then calculated as described in Chapter 7.7 of the IDMS to obtain the $M_D(f)$ curve. An example of such a curve is shown in Figure 5.3.

The SFR obtained by the slanted edge method give a continuous spectrum without the linear approximation as introduced by the Grille Contrast Modulation method. This means that the resolvable resolution numbers at 25% and 50% contrast modulation are calculated more precisely than those methods that use linear approximation between measured points.

For both the Slanted edge method and the Grille contrast method, there is an on-going debate regarding the appropriate contrast modulation required for different applications. In the IDMS, 25% and 50% for images and text respectively is suggested as points of interest in the slanted edge method also.

5.2. Experimental setup

In our experimental setup we used a Barco F70-4K pixel shifted projector equipped with a Barco EN41 lens. This projector has a native WQXGA (2560 by 1600 pixels) DMD as an SLM, and a pixel shifting mode where every other frame is shifted half a pixel diagonally. To capture the test scenes we used a Nikon D5100 SLM camera with a 23.6 x 15.6 mm 16.2 megapixel CMOS sensor. The captured images was stored in Nikons uncompressed raw format NEF. The setup is shown in Figure 5.4.

One purpose of these tests is to explore the relationship between the source resolution and the resulting measured resolution. To this end, all of these experiments are performed for different source resolutions, starting at the native WQXGA resolution. Then the experiment is redone several times with steps of 10% increase in the source resolution all the way up to 120% over the native resolution.
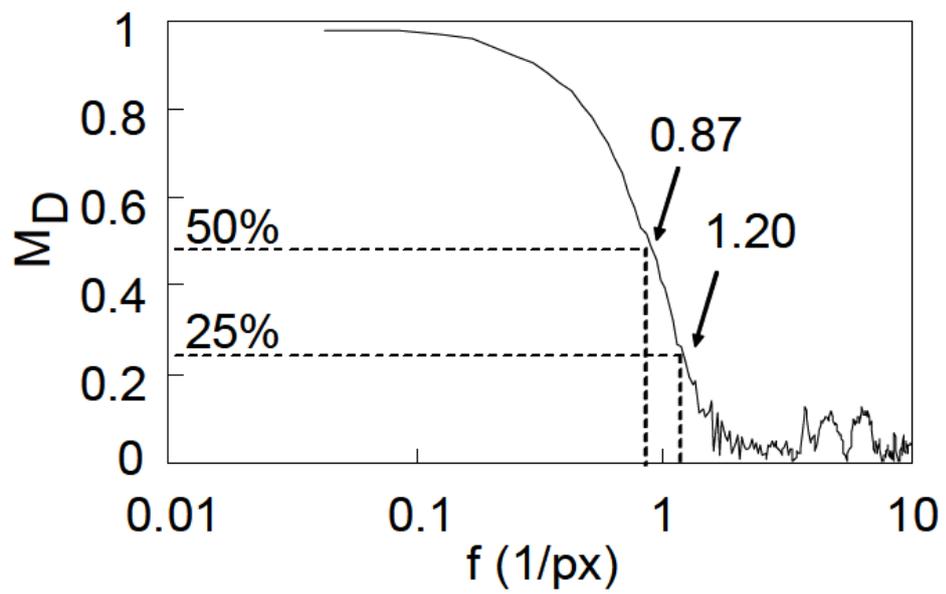
Figure 5.3:   Spatial Frequency response example curve from IDMS. MD is the frequency modulation for the display
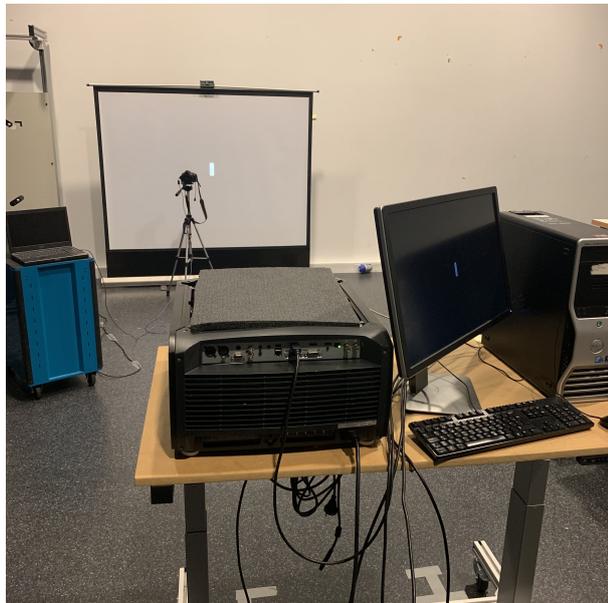
Figure 5.4:   Lab setup measuring the projected contrast with a camera.
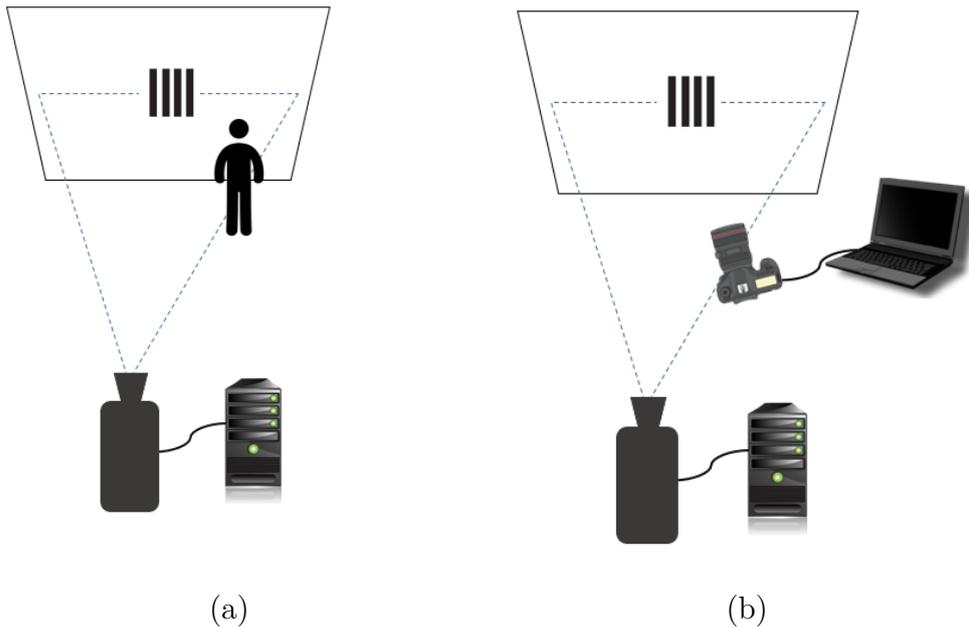
Figure 5.5: a) Test setup for the least resolvable line pair method. b) Test setup for the grille contrast and the slanted edge methods.

5.2.1. Line pair measurements

To perform the least resolvable line pair measurements we used a regular PC as the source connected to the projector. On this PC we had a program rendering a 3D object consisting of 3 black/white line pairs. This object could be moved back in 3D space to increase the distance between the viewpoint and the object, thus also decreasing the size of the object on the screen. This will be viewed as the line pairs is shrinking, all the way down to not being distinguishable anymore. The measurement setup was as illustrated in Figure 5.5 (a) where there is an observer who determines if the line pairs are resolvable or not.

The experiment itself was performed as described in Section 5.1.1. The object with the line pairs was moved away from the viewpoint until the three line pairs was not resolvable anymore. Then the object was moved towards the viewpoint one step again to make them resolvable. At this point the physical size of the line pairs on the screen was measured. The total size of the projected surface was then divided upon the line pair size to calculate the number of line pairs that would fit within

the projected area. This number is the measured resolution for this method at this given source resolution.

The least resolvable line pair measurement is redone for source resolutions ranging from 100% native SLM resolution up to 120% of the native SLM resolution.

5.2.2. Grille contrast measurements

The grille contrast measurement experiment setup was per Section 5.1.2. On the PC we had a program rendering 3 black/white line pairs of controllable width, and we did the measurements with 1,2,3 and 4 pixels wide lines. Measurements were done using the camera and DCRAW was used to convert raw image files to TIFF format. The images were then filtered and analysed in Matlab as described in the IDMS Chapter 7.2.

The measurement setup was as illustrated in Figure 5.5 (b) where the camera is capturing the test scene.

5.2.3. Slanted edge measurements

The grille contrast measurement experiment was setup as per Section 5.1.3. On the PC we had a program rendering a black/white step pattern. The measurements was done with the camera that we tilted 5 degrees relative to the edge of the step pattern. The raw image files were converted to TIFF images with DCRAW before they were filtered and analysed in Matlab as described in the IDMS Chapter 7.7.

The measurement setup is the same as for the Grille Contrast Modulation method as illustrated in Figure 5.5 (b), but with a different test image displayed and with the camera tilted 5 degrees as described in Chapter 7.7 of the IDMS .

There are several commercial software solutions available for calculating the slanted edge response of an image, but we did not find any that suited this use case. The nature of the DMD makes the projected pixels very distinctive with dark gaps in between, and the available software solutions required a solid line edge to calculate the response. For this reason, we developed our own solution that took the distance

between the pixels into account, and sampled one line for each pixel in the area of interest. Even though the shifted superimposition reduces the screen-door effect significantly, Figure 5.12 and Figure 5.13 illustrates that the resulting sub-pixels also are distinguishable. So for calculating the slanted edge of the shifted superimposed scenes, we took the resulting pixelsize into account, which in this case is half the size of the projected DMD pixel of the native image in Figure 5.13 (a).

5.3. Results

5.3.1. Least resolvable line pairs

The least resolvable line pair measurement was executed as described in Section 5.1.1. In these measurements the source resolution is set to WQXGA, the native resolution of the projector, and then increased in steps of 10% up to 120% over the native resolution.

Figure 5.6 illustrates four different measurements, where a) and b) show examples on resolvable line pairs, while c) and d) are not resolvable since two of the line pairs are melted together.

The results from the least resolvable line pairs measurements on all source resolutions is plotted in Figure 5.7. At the native resolution the measured resolution is below the native resolution, and the measured resolution increase steady until the measured resolution reaches about 40% above the native resolution.

(a)   (b)   (c)   (d)
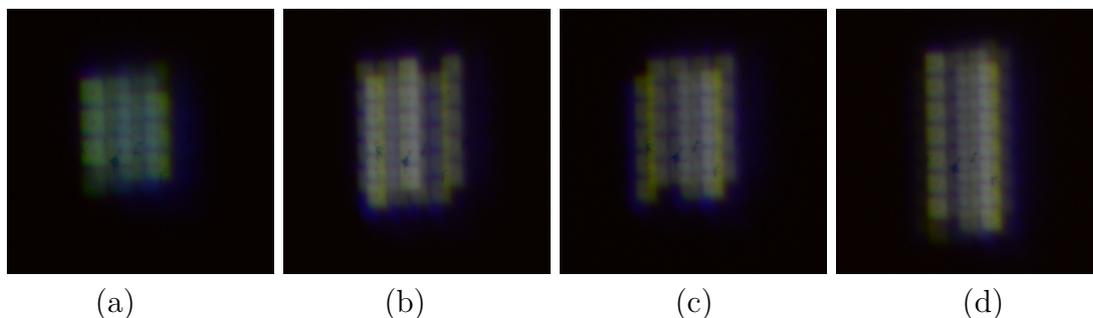
Figure 5.6:   Results from least resolvable line pair test. Horizontal source resolution is a) 4352, b) 5120, c) 5376, d) 5632.

When the source resolution reaches and goes beyond double the native resolution, the measured resolution drops down below the maximum measured resolution.

5.3.2. Grille contrast

The grille contrast measurement was executed as described in Section 5.2.2. In these measurements the source resolution is set to WQXGA, the native resolution of the projector, and then increased in steps of 10% up to 120% over native resolution.

Figure 5.8 shows the grille contrast measurement for 3328x2080 (30% over native resolution) and such a measurement was obtained for all the given source resolutions. The usage of such a curve is to look at the intersection point between the desired contrast value and the contrast curve of the display.

According to the IDMS a typical desired contrast value for displaying images is 25%, and extracting the measured resolution for 25% contrast value gives us the measured resolution curve shown in Figure 5.9.

The measured curve is monotonically increasing with input resolution equal the source resolution as long as the desired contrast is less than the measured contrast at grille1 (single pixel width line pairs). When the measured contrast at grille1 goes below the desired contrast we start to go up the contrast curve and n_r in Equation 5.1 goes up. This leads to the measured resolution dropping at that point such as at source resolution of 3580 in Figure 5.9.

Since the source resolution grid and the SLM resolution grid do not match each other in even numbers, the source pixels may be represented by an uneven number of overlapping pixels in the projected image. This means that the line pairs will sometimes be unevenly represented, which again leads to uneven contrast measurement between the different line pairs. When this occurs the resolution measurement is open for interpretation, since the different line pairs have different contrast ratio.

Figure 5.10 illustrates an example of this occurrence where Figure 5.10 (b) have different contrast measurements for the line pairs. Figure 5.9 have used the biggest
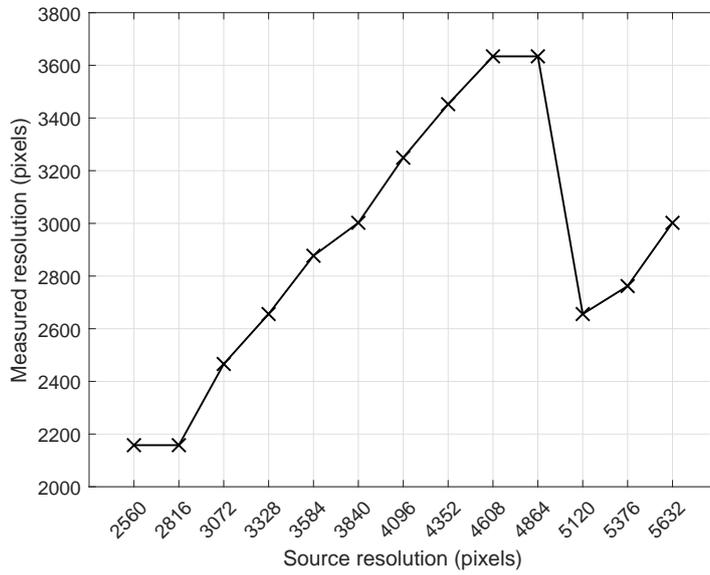
Figure 5.7: Horizontal measurement results from the least resolvable line pairs experiment on a pixel shifted projector with WQXGA (2560) native resolution.
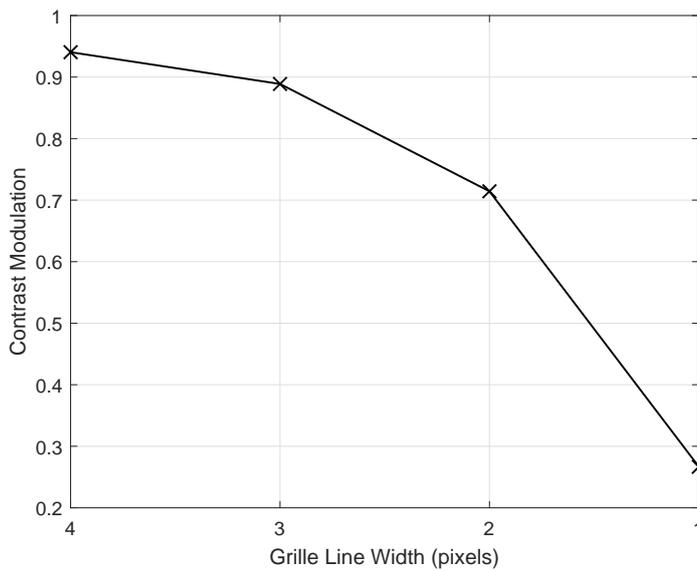
Figure 5.8: Grille contrast measurements for the source resolution 3328x2080.
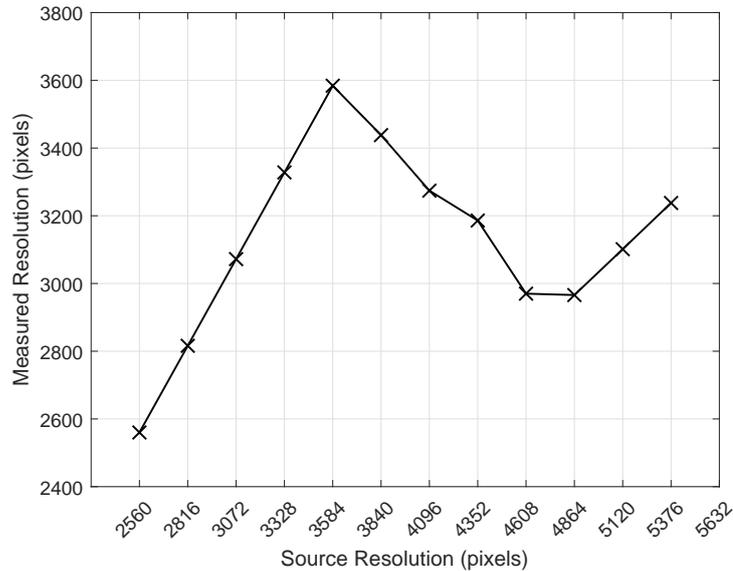
Figure 5.9: Resolution measurements versus source resolution, given 25% contrast.

local line pair contrast while Figure 5.11 shows how the resolution measurements also is when using the smallest local line pair contrast instead.

5.3.3. Slanted edge

The slanted edge measurement was executed as described in Section 5.2.3. In these measurements the source resolution is set to WQXGA, the native resolution of the projector, and then increased in steps of 10% up to 120% over native resolution.

As seen in Figure 5.13 the shifted edges are blurred because of the overlapping pixels; Figure 5.14 shows that the frequency response of the shifted images is lower than the native unshifted edge.

The slanted edge MTF curve of each measurement is calculated from the camera captured scenes, and the results of these is presented in Figure 5.14.

5.4. Discussion

Increasing source resolution gives more details in the source image to include in the subframes that makes up the resulting projected image on the screen. There are several different ways the subframes may be generated, but the subframes are always
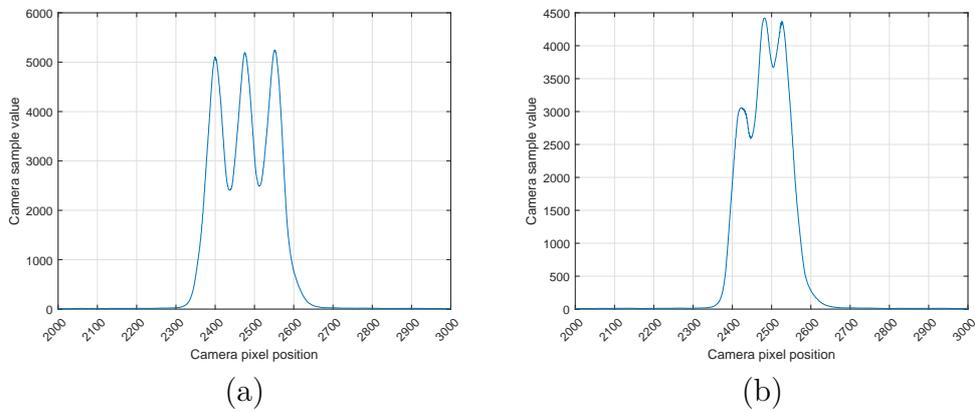
Figure 5.10:   Grille plots of three different source resolution a) 2560, c) 3584
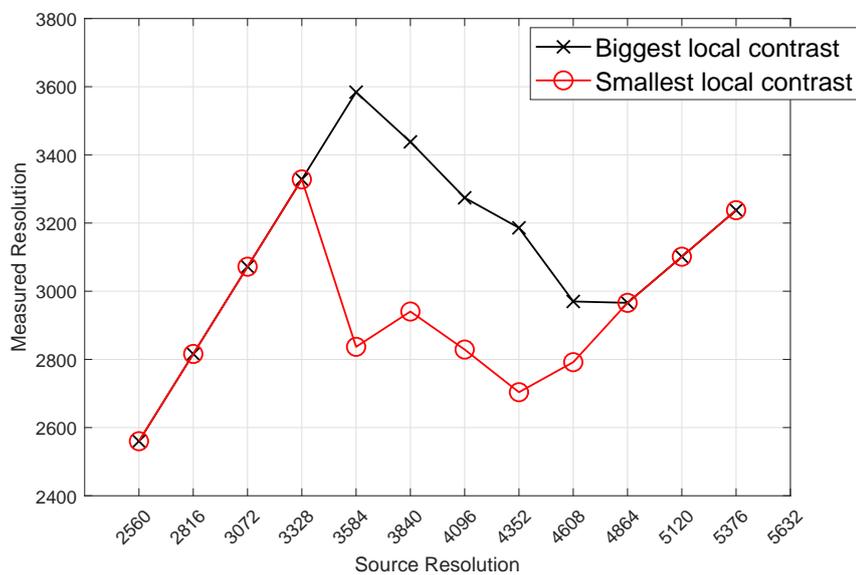
Figure 5.11:   Resolution measurements versus source resolution, given 25% contrast. This figure illustrates both the best case measurement and the worst case measurement.
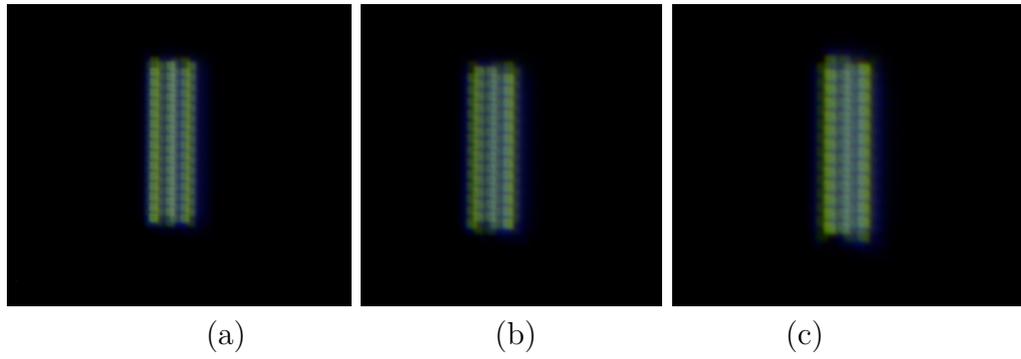
Figure 5.12:   Images taken of the grille11 measurements for a)3328 b)3840 c)4096
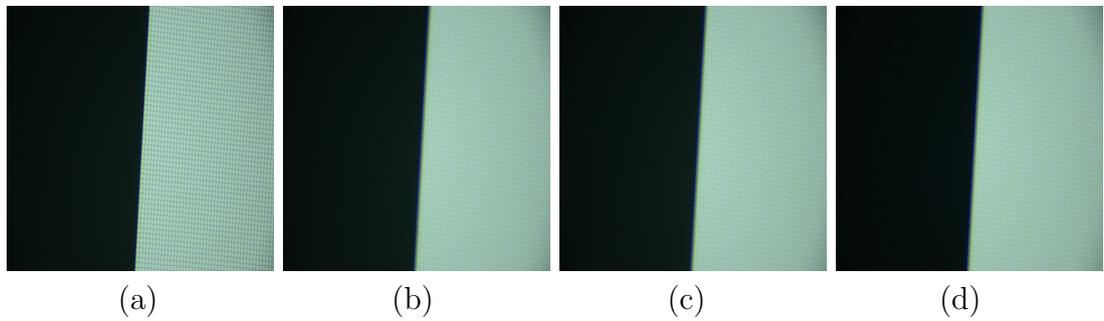
Figure 5.13:   Slanted edge measurements at horizontal source resolutions a) unshifted 2560, b) 3584, c) 4352, d) 5376.
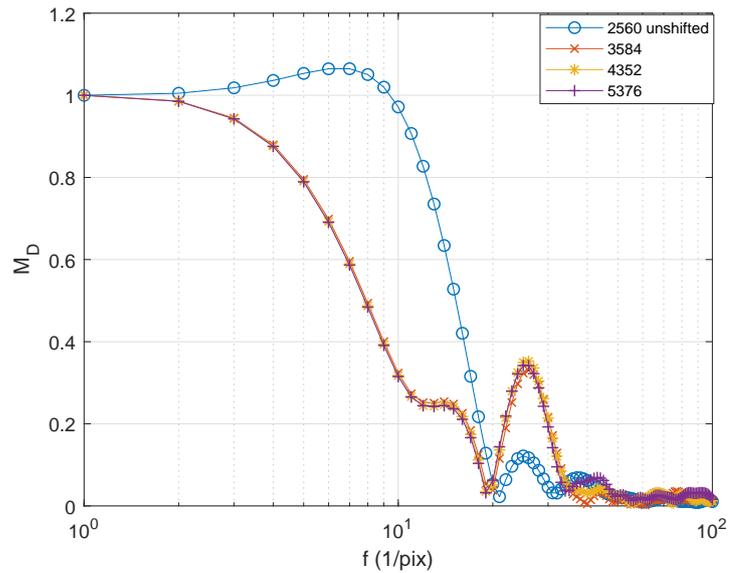
Figure 5.14:   Slanted edge MTF calculations.

more than one frame and are always generated at the SLM resolution. So for the different subframes to have different information in them, the source resolution needs to be higher than the SLM resolution to provide enough details and information. Therefore it is intuitive that higher source resolution also results in higher measured resolution.

But the shifted superimposition technique also has some physical limitations. The optical overlap of the pixels as shown in Figure 1.1 makes up the new and finer pixel grid, but it also illustrates that these new pixels are not independent of each other. Each resulting finer pixel in Figure 1.1 are made up of two overlapping SLM pixels from different subframes, and each of these SLM pixels are also influencing three other resulting pixels in the finer pixel grid. This dependency makes the optical overlap function as a low-pass filter, attenuating the highest frequencies of the resulting image.

These physical limitations ensure that even though the resulting resolution in the shifted superimposed image is increasing with increasing source resolution, there must be some limitations in how high resolution that may be obtained. When counting the resulting pixels in the new overlapping pixelgrid shown in Figure 1.1 we see that the number of separable pixels have doubled in both horizontal and in vertical directions. But because of the inter-pixel dependency each of these new pixels are not independently controllable, and the low-pass filter behaviour of the optical overlap will attenuate the highest frequencies. These aspects affect the resulting resolution so that the ideal double resolution will not be fully achieved.

The high frequency attenuation should be measurable, so the method we use to measure the resolution will also have an impact on the measured result. The least resolvable line pair method makes use of subjective observations to see when the line pairs is at the resolvable limit. The idea here is that the resolution represents the amount of separate distinguishable details, and to find this resolution number we need to see how small details the display is able to reproduce. This procedure

is well known in the industry and is widely adopted in some professional markets of projected display using the Johnson's criteria to design and verify their display systems [1]. The least resolvable line pair method is often used as a system resolution measurement rather than a display resolution measurement, but there is no reason not to use this method also as a pure display measurement. The results are dependant on the performance of the source and the projected screen, but the same can be said for the results from the other measurement methods.

The grille contrast modulation measurements is straightforward and not as open to interpretation when the measurement results follow the criteria in the IDMS. But this method will be open for interpretation when the measured data behaves as shown in Figure 5.10 (b) and (c). The problem with the data in Figure 5.10 (b) is that the line pairs have different contrast ratio, and the measurements will be heavily dependant on which of these contrast ratios that is used. The reason for this difference in contrast is that the pixel grid of the resulting pixels shown in Figure 1.1 do not necessarily correlate to the pixel grid of the source resolution. In these instances, rows and columns of the source image will be represented by different geometric compositions in the resulting pixel grid, as evidenced by the different widths of the line pairs shown in Figure 5.12. In Figure 5.12 (b) and (c) we see that the line pairs have different widths, so the contrast measurements of these examples will be of the nature in Figure 5.10 (b). In these cases it is not given which contrast to use, and the difference between using the best and the worst line pair contrast is illustrated in Figure 5.11. There is a significant difference in these resolution numbers, and it must be defined in such cases how to interpret the contrast measurements when the geometry of the line pairs are not consistent.

Figure 5.10 (c) shows another interesting phenomena. When the source resolution is double the SLM resolution (and beyond) one may end up losing a whole line. This is because the source resolution goes above the resolution of the resulting pixel grid caused by the pixel overlap, which is double the SLM resolution in both horizontal and vertical direction. When the source resolution goes above this limit, there are

more details represented in the source resolution than we have pixel elements on the projected screen. So some of these details will then be lost, and it is therefore possible to lose whole line pairs. The least resolvable line pair test will disqualify these results as one of the line pairs will be lost. But the grille contrast modulation method will still calculate a contrast ratio based on the remaining line pairs, and as we see in the measured resolution in Figure 5.11, the measured resolution actually goes up again at the higher resolutions even though we are starting to lose a line pair in the Grille measurements.

In Figure 5.12 we see three line pairs in both a), b) and c) so all of these examples would pass the least resolvable line pair test. In the grille contrast measurement however both Figure 5.12 (b) and (c) fall below 25% contrast and would therefore not pass that measurement. This raises the question if 25% really is a wise choice or if the target contrast should be lower. In older applications, like CRT monitors for the professional market, the target contrast limit was set between 2% and 10% [1] so the 25% target contrast may seem a bit too strict to measure the general resolution of a display. It is good to have a target contrast if you have a specific application that needs a given contrast to perform satisfactorily. But as a measurement for general resolution this method disqualifies details that are perfectly distinguishable just because the contrast doesn't reach this tests desired contrast levels.

The slanted edge measurement is a very good measurement for optical performance, but as we see from Figure 5.13 and Figure 5.14, the shifted slanted edge does not change significantly as a function of the input resolution. This is logical since the source image in this case is a step pattern that will be interpreted the same way in all of these resolutions. This makes the slanted edge measurement method an unsuitable method to measure the effect of input resolution on a shifted superimposed display.

The least resolvable line pairs and the grille contrast method both show that the measured resolution increase with the source resolution up until approximately 40%

over the native resolution. The differences in the measurement method give some differences in what source resolution that gives the best measured resolution, but they both indicate that the maximum gain of this resolution enhancement method is around 40%. Approximately 40% resolution increase also matches the number of pixels actually projected on the screen, since we are using two different subframes in two different positions for this measurements. With our WQXGA projector this gives us 2560 * 1600 * 2 number of pixels on the screen. Keeping the aspect ratio, this will equal the number of pixels in a 3620 * 2262 image which is 41% (square root of two) above the native resolution.

5.5.   Conclusion

The achieved resolution with the shifted superimposition technique does increase as we increase the source resolution. This is valid up to a certain threshold, where the shifted superimposition method reaches it's limit because of the physical size of the projected SLM pixels and the overlap of these pixels in different positions. The resolution enhancement limit seems to be about 40% above the SLM resolution. There are still open questions on how to measure this resolution increase in the best way. We have utilized the least resolvable line pair test, the grille contrast modulation method and the slanted edge method in this work. All these methods have their shortcomings, but the two best methods for this use case both measure a maximum resolution increase at 40% while the slanted edge method is found to be unsuitable for this measurement. The least resolvable line pair method seems to be better suited to measure the achieved resolution increase, since the grille contrast modulation method takes too many assumptions on what attributes the measurements should have to be defined as resolution.

6. Discussion, Conclusion and Future work

This research was undertaken to provide more insight in the concept of resolution enhancement through shifted superimposition, and several topics have been explored to provide understanding of the whole chain from subframe generation to evaluation of the final superimposed image on the projected screen. The evaluation and analysis presented in this thesis are based on simulations and on experimental work performed on a shifted superimposition DLP projector.

The goal of this work has not been to design the optimal subframe generation method, but rather to evaluate the different classes of subframe generation methods and provide tools to evaluate them. For this reason we have chosen to use methods that are easy to understand and analyse instead of optimizing the subframe generation methods in most of this research. This is explained in Chapter 2 where we first find the optimal image quality metric used to evaluate the subframe generation methods before we explore the different classes of subframe generation methods in Chapter 3. These chapters brings more insight in how important the subframe generation methods are for the quality of the superimposed image, and we also see that choosing the most suitable subframe generation method for a given system is a tradeoff between quality achieved on the screen and resource usage.

The original idea behind the work presented in Chapter 4 was that since the frequency response of a shifted and superimposed pixel is non-uniform, such a shift will influence the frequency content in the shifted image differently in different directions. If that is the case, then it should be possible to find a preferred shifting direction based on the content to be shifted. But our research show that as long as you use different subframes for the different positions and also take the direction and magnitude of the shift into account, the resulting superimposed image is practically invariant of the shifting direction. We planned to build a proof-of-concept projector with a 2-dimensional optical actuator, but since the simulation results did not show much improvement with this concept we did not build such a system.

In Chapter 5 we explore how much quantifiable resolution is gained from the shifted superimposition method. We do find that both the grille contrast modulation and the least resolvable line pair method are suitable for measuring the resolution enhanced image, and both these images give a resolution gain of about 40% above the native resolution.

All of this work explores the gained resolution and quality from the shifted superimposition method. This work also raises the question of resolution vs quality, and the question of what resolution actually is. Section 1.5 show that resolution is an ambiguous term, and that resolution and quality are tightly coupled. An image of higher resolution will also often be perceived as an image of higher quality, and many of the resolution definitions have some artificial limitation set to generate a quantifiable resolution number. One example of such an artificial limitation is the resolution measurement method grille contrast modulation, where one must choose a contrast limit to extract a resolution number. The notion that resolution need to be given by a single number seems to be very limiting.

The modulation plots explained in Section 1.5 show that different systems with the same limiting resolution may have very different characteristics. As long as we see resolution as a number given by the limiting resolution, these kind of systems will have the same resolution specifications even though they may perform very differently when inspected visually. This implies that in addition to the limiting resolution number there must be an indication of the quality this resolution is given by. For instance a pixel count on a flat panel will not give enough resolution if these pixels are not individually addressable by the system and distinguishable for the observer.

Since different applications may have different specifications for limiting resolution, modulation plots seems to be the best basis for a parameter for the systems resolution and quality capabilities. Such plots will give the user of the display system the opportunity to extract the resolution numbers they need for themselves based

on their specific application. The disadvantage with the modulation plot as a specification is that this plot is very advanced and not intuitive for a non-professional user. These users will probably still prefer to be given resolution as a single number rather than a curve. These numbers should still be derived from measurements instead of simple pixel counts, and also be accompanied by a parameter indicating the quality of the resolution. The area under the curve in the modulation plot may be a candidate for such a quality parameter.

As seen in Chapter 2 and Chapter 3, the gained resolution and quality of the shifted superimposition method is highly dependant of the subframe generation method. So choosing the correct method and allocating the needed calculation resources is one of the very important decisions when designing such a system.

Since the direction of the shift does not have the same impact on the resulting quality, this shifting direction is not as crucial. The horizontal and vertical direction will only expand the pixel grid in one direction and are the only directions that do not provide sufficient quality gain. For this reason all shifting directions are valid, except the angles close to the horizontal and vertical direction, and this gives a large degree of freedom for shifted superimposition systems.

For future work it would be very interesting to explore how the shifted superimposition method interacts with moving images. Does moving objects have any impact on the perceived quality of the image, and are there any correlation between the movement of the object and the movement of the subimages? In this context it is also interesting to see the impact of deriving the different subimages from different points in time. In such a system the moving objects should be able to move at the projectors native frame rate, but this temporal resolution will be traded off by the perceived spatial resolution of the moving objects. It would be very valuable to determine and quantify this trade-off and to explore how this impacts our overall perception of the presented quality.

7. Article summary

Article 1:

Preferred image quality metric for shifted superimposition-based resolution-enhanced images In this work we have developed a framework for simulating different superimposition methods over different image content, 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 VIF 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. MSSSIM metric which is based on the analysis of image structure is better at detecting this distortion.

My contribution: Defining the work and process, developed simulation framework, analysis, subjective experiments, manuscript presentation

Article 2:

A comparative study of superimposition techniques for enhancing the projector resolution: Simulations and experiments Among the techniques developed to enhance the resolution of a projector beyond its native resolution, a number of techniques use sequential superimposition of low resolution images with sub-pixel shift to produce a resulting high-resolution image on the wall. This work investigates different low-resolution subframe image generation techniques used for this purpose. Along with that, two new subframe generation techniques have been developed with the aim to prioritize darker and/or bright pixels in an image. We have also extended a Single-subframe iterative technique to Two-subframe iterative technique to evaluate its full potential. Detailed simulations, visual image quality analysis, image quality metric assessment and measurement results of the existing and the newly developed methods have been carried out. A comparative study of the techniques suggests that iterative subframe generation techniques give the

best overall image quality, but their high computational cost could make them less practical.

My contribution: Defining the work and problem, simulation framwork built upon my framework, supervising and assisting lab experiments and manuscript presentation

Article 3:

Resolution enhancement through shifted superimposition: The influence of shift direction Shifted superimposition is commonly done by shifting every other frame spatially on the projected screen with subpixel precision to form a new pixel grid with finer pixel pitch. By shifting every other frame diagonally, the frequency response of a pixel will be non-uniform with respect to different directions. This non-uniformity implies that frequency in some directions may be better represented than in other directions. In this work, we explore the possible benefits and attributes of shifting in different specific directions. We simulate the spatial shifting in 8 different spatial directions, using two different subframe generation methods. We find that as long as we do not shift in the horizontal or vertical direction, the direction of the shift does not influence the projected picture quality if the system is using different subframes individually optimized for each position. But when the same subframe is used for both positions, the direction of the shifts affects both the directionality of the frequency preservation and the quality of the projected image.

My contribution: Defining the work and process, developed simulation framework, analysis, manuscript presentation

Article 4:

The effects of source resolution on resolution enhancement through shifted superimposition projection Resolution in a projected display is traditionally defined by the number of pixels in the projectors spatial light modulator (SLM).

In the pixel shifting technology the display physically shifts every nth frame on the projected screen and the overlapping pixel grids forms a new subpixel grid with a higher pixelcount. There is still an open question how much this method increase the resolution and how to quantify it. The system will not be able to reproduce the source resolution to full extent, but the higher source resolution the system is provided with, the more details the system should be able to reproduce. In this work we experimentally investigate how the projector performs with resolution enhancement through pixel shifting, and how this method relates to the source resolution. We also investigate some known methods of resolution measurement, and evaluate how these perform for the shifted superimposition method. We find that the resolution enhancement through shifted superimposition enhance the resolution to about 40% over native resolution, and we also find two different measurement methods that is relevant for measuring resolution within such systems.

My contribution: Defining the work and process, executed lab experiments, analysis, manuscript presentation

Bibliography

[1] G. C. Holst. *CCD Arrays, Cameras and Displays.* JCD Publishing, 1998.

[2] W. J. Smith. *"Modern Optical Engineering".* SPIE Press, 2000.

[3] Barco. White paper: 4k uhd explained. Technical report, Barco, https://www.barco.com/en/page/2016/entertainment/downloads/ wp4kuhd/downloadpage, retrieved on 2018-02-20, 2016.

[4] N. Damera-Venkata and N. L. Chang. On the resolution limits of superimposed projection. In *2007 IEEE International Conference on Image Processing*, volume 5, pages 373–376, Sep. 2007.

[5] Y. Takahashi, T. Nomura, and S. Sakai. Optical system and characteristics of an LCD projector with interleaved pixels using four LCD projectors. *IEEE Transactions on Circuits and Systems for Video Technology*, 5(1):41–47, Feb 1995.

[6] C. Jaynes and D. Ramakrishnan. Super-resolution composition in multi-projector displays. *Proc. IEEE International Workshop on Projector-Camera Systems (ProCams)*, Oct 2003.

[7] W. Allen and R. Ulichney. 47.4: Invited paper: Wobulation: Doubling the addressed resolution of projection displays. *SID Symposium Digest of Technical Papers*, 36(1):1514–1517, 2005.

[8] M. N. Sing, T. A. Bartlett, W. C. McDonald, and J. M. Kempf. Super resolution projection: leveraging the mems speed to double or quadruple the resolution. volume 10932, 2019.

[9] A. Said. Analysis of multi-rate filters and signal design for projected image superimposition. In *2006 Fortieth Asilomar Conference on Signals, Systems and Computers*, pages 869–872, 01 2006.

[10] A. Majumder. Is spatial super-resolution feasible using overlapping projectors? *Proceedings. (ICASSP '05). IEEE int. conf. on acoustics, speech, and signal process.,*, pages 209–212, 2005.

[11] A. Said. Analysis of systems for superimposing projected images. Technical Report HPL-2006-129, Hewlett-Packard, 2006.

[12] A. Said. Analysis of subframe generation for superimposed images. *IEEE int. conf. on Image process., 2006*, pages 401–404, 2006.

[13] N. Damera-Venkata and N. L. Chang. Realizing super-resolution with superimposed projection. In *2007 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2007)*, pages 1–8, 2007.

[14] J. Napoli, S. R. Dey, S. Stutsman, O. S. Cossairt, T. J. Purtell, S. L. Hill, and G. E. Favalora. Imaging artifact precompensation for spatially multiplexed 3-D displays displays. *Stereoscopic Displays and Applications XIX, Proceedings of SPIE-IS&T Electronic Imaging, SPIE*, 6803, 2008.

[15] T. Okatani, M. Wada, and K. Deguchi. Study of image quality of superimposed projection using multiple projectors. *IEEE Transactions on Image Processing*, 18(2):424–429, Feb 2009.

[16] P. Didyk, T. Ritschel, K. Myszkowski, and T. Paristech. Apparent display resolution enhancement for moving images. *ACM trans. on graphics*, 29, 2010.

[17] B. Sajadi, M. Gopi, and A. Majumder. Edge-guided resolution enhancement in projectors via optical pixel sharing. *ACM Trans. Graph.*, 31(4):79:1–79:122, July 2012.

[18] F. Berthouzoz and R. Fattal. Resolution enhancement by vibrating displays. *ACM Trans. Graph.*, 31(2):15:1–15:14, April 2012.

[19] B. Sajadi, D. Qoc-Lai, A. H. Ihler, M. Gopi, and A. Majumder. Image enhancement in projectors via optical pixel shift and overlay. In *IEEE International Conference on Computational Photography (ICCP)*, pages 1–10, April 2013.

[20] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, April 2004.

[21] F. Heide, D. Lanman, D. Reddyand, J. Kautz, K. Pulli, and D. Luebke. Cascaded displays: Spatiotemporal superresolution using offset pixel layers. *ACM Trans. Graph.*, 33(4):60:1–60:11, July 2014.

[22] F. Heide, J. Gregson, G. Wetzstein, R. Raskar, and W. Heidrich. Compressive multi-mode superresolution display. *Optical society of America*, 22(12):1514–1517, 2014.

[23] E. Barshan, M. Lamm, C. Scharfenberger, and P. Fieguth. 35.3: Resolution enhancement based on shifted superposition. *SID Symposium Digest of Technical Papers*, 46(1):514–517, 2015.

[24] 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 Transactions on Image Processing*, 9(4):636–650, Apr 2000.

[25] K. Egiazarian, J. Astola, N. Ponomarenko, V. Lukin, F. Battisti, and M. Carli. A new full-reference quality metrics based on HVS. In *Proc. of the Second International Workshop on Video Processing and Quality Metrics*, jan 2006.

[26] N. Ponomarenko, F. Silvestri, K. Egiazarian, M. Carli, J. Astola, and V. Lukin. On between-coefficient contrast masking of DCT basis functions. In *Proc. of the Third International Workshop on Video Processing and Quality Metrics*, Jan 2007.

[27] Z. Wang, E. P. Simoncelli, and A. C. Bovik. Multiscale structural similarity for image quality assessment. In *The Thirty-Seventh Asilomar Conference on Signals, Systems Computers, 2003*, volume 2, pages 1398–1402 Vol.2, Nov 2003.

[28] X. Zhang, X. Feng, W. Wang, and W. Xue. Edge strength similarity for image quality assessment. *IEEE Signal Processing Letters*, 20(4):319–322, April 2013.

[29] L. Zhang and H. Li. SR-SIM: A fast and high performance IQA index based on spectral residual. In *2012 19th IEEE International Conference on Image Processing*, pages 1473–1476, Sept 2012.

[30] L. Zhang, L. Zhang, X. Mou, and D. Zhang. FSIM: A feature similarity index for image quality assessment. *IEEE Transactions on Image Processing*, 20(8):2378–2386, Aug 2011.

[31] F. Zhang, L. Ma, S. Li, and K. N. Ngan. Practical image quality metric applied to image coding. *IEEE Transactions on Multimedia*, 13(4):615–624, Aug 2011.

[32] H. R. Sheikh and A. C. Bovik. Image information and visual quality. In *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 3, pages 709–712, May 2004.

[33] D. M. Chandler and S. S. Hemami. Vsnr: A wavelet-based visual signal-to-noise ratio for natural images. *IEEE Transactions on Image Processing*, 16(9):2284–2298, Sept 2007.

[34] C. S. Swartz. *Understanding Digital Cinema - A Professional Handbook*. Focal Press, 2005.

[35] L. Yue, H. Shen, J. Li, Q. Yuan, H. Zhang, and L. Zhang. "image super-resolution: The techniques, applications, and future". *Signal Processing*, 128, 2016.

[36] D. Han. Comparison of commonly used image interpolation methods. *Proceedings of the 2nd International Conference on Computer Science and Electronics Engineering*, 2013.

[37] J. Hirsch and C. A. Curcio. "the spatial resolution capacity of human foveal retina". *Vision Research*, 29, 1989.

[38] H. Barrett and K. Myers. *Foundations of Image Science*. Wiley, 2004.

[39] J. Intriligator and P. Cavanagh. "the spatial resolution of visual attention". *Cognitive Psychology*, 43, 2001.

[40] International Committee for Display Metrology. Information display measurement standard. Technical report, SID, Society for information display, 2012.

[41] K. Bylsma. Can you explain in detail how e-shift works?, 2014. http://usjvc.com/faq/index.php?action=artikel&cat=20&id=268, retrieved on 2018-02-20.

[42] X. Liu, M. Pedersen, and J. Y. Hardeberg. Cid:iq – a new image quality database. volume 8509, 2014.

[43] K. V. Ngo, J. Storvik, C. A. Dokkeberg, I. Farup, and M. Pedersen. Quickeval: a web application for psychometric scaling experiments. volume 9396, pages 9396 – 9396 – 13, 2015.

[44] P.G. Engeldrum. *Psychometric Scaling: A Toolkit for Imaging Systems Development*. Imcotek Press, 2000.

[45] D. Bickson, O. Shental, D. Dolev, P. H. Siegel, and J. K. Wolf. Gaussian belief propagation solver for systems of linear equations. *int. symp. on inf. theory*, 2008.

[46] R. K. McConnell. Method of and apparatus for pattern recognition, 1986. U. S. Patent No. 4,567,610.

[47] B. Masia, G. Wetzstein, P. Didyk, and D. Gutierrez. A survey on computational displays: Pushing the boundaries of optics, computation, and perception. *Computers and Graphics*, 37(8):1012 – 1038, 2013.

[48] Y. Huang, E. Liao, R. Chen, and S. Wu. Liquid-crystal-on-silicon for augmented reality displays. *Applied Sciences*, 8(12), 2018.

[49] L. Fang, O. Au, K. Tang, and X. Wen. Increasing image resolution on portable displays by subpixel rendering - a systematic overview. *APSIPA Transactions on Signal and Information Processing*, 1, 12 2012.

**Evaluation of resolution enhancement in shifted superimposed
projection displays: Simulations and experiments**
Dissertation for the degree of PhD

—
Svein Arne Jervell Hansen

—

usn.no