Study of image quality of superimposed projection using multiple projectors

著者	岡谷　貴之
journal or publication title	IEEE Transacation on Image Processing
volume	18
number	2
page range	424-429
year	2009
URL	http://hdl.handle.net/10097/46846

doi: 10.1109/TIP.2008.2008899

Correspondence

Study of Image Quality of Superimposed Projection Using Multiple Projectors

Takayuki Okatani, *Member, IEEE*, Mikio Wada, and Koichiro Deguchi, *Member, IEEE*

***Abstract*—In this correspondence, we discuss the quality of images realized by a recently proposed method of generating a high-resolution image by superimposing multiple low-resolution images projected by different projectors. We show several fundamental properties of this method: 1) the accuracy of the image realization (e.g., resolution of the realized image) is heavily affected by the structures of the images to be realized, and 2) there is a tradeoff between the image quality and the maximum brightness of the images to be realized. These are properties peculiar to multiprojector super-resolution and are in contrast with multicamera super-resolution. The results will be helpful in evaluating the usefulness of the method.**

***Index Terms*—Multiprojector system, projector-camera, super-resolution, superimposed projection.**

I. Introduction

Recently, it has been shown [1], [2] that super-resolution can be realized by the use of multiple projectors; a high-resolution image can be generated by superimposing multiple lower-resolution images projected by different projectors. A similar technology of superimposed projection using a single projector/display is known as Wobulation [3].

As compared to a standard tiled image projection, a superimposed projection enables more flexible scalability; an arbitrary number of projectors can be added to or subtracted from the system. Although it remains uncertain, it is possible that the brightness resolution could increase by the use of superimposed projection; this is not possible in the case of tiled projection. The difference in the layout of the projectors could also be a potential advantage.

The method of superimposed projection has also a few disadvantages as compared to that of tiled projection. First, it requires far more accurate calibration; for example, geometric calibration needs to be of sub-pixel precision for every pixel, i.e., approximately 1/10–1/100 of a pixel. Although this is not impossible, it entails extra costs. Second, unlike tiled projection, the image resolution does not increase in proportion to the number of projectors (rather, the spatial resolution appears to have an upper bound). Despite these disadvantages, the method of multiprojector superimposed projection is a promising technology [1], [2], [4], [5]; it could have applications not limited to ordinary image projection.

In this correspondence, we analyze the quality of the images realized by this method and show its several fundamental properties. More specifically, we consider the accuracy of image realization, i.e., the similarity between the image that we want to display and the image that is actually realized by superimposed projection. Assuming a particular point spread function (PSF) of the projectors, we derive several results.

First, we show that the realization accuracy is heavily affected by the structures of the images that we want to realize. We explain its mechanism by considering toy problems in the case of 1-D projectors. Moreover, we present experimental results to examine the phenomenon in detail, in which black-and-white text images are used as an extreme and realistic example.

Next, we show that there is a tradeoff between the quality and the maximum brightness of the realized images. When K projectors are used in a superimposed projection, one may expect that the realized image can potentially be K times as bright as the image of a single projector. However, the image quality tends to deteriorate with the increasing brightness of the image to be realized; thus, the maximum possible brightness in the superimposed projection is likely to be lower than the expected value of brightness when prioritizing realization accuracy. We also note that the extent of this tradeoff varies depending on the image structure. This is demonstrated by experimental results obtained for the same set of text images used in the experiment above.

A key problem in superimposed projection is the computation of the images fed to the component projectors in order to realize a desired image. Probably, the best solution in terms of image quality is to minimize the L_2 norm of the difference between the desired image and the realized image. It is shown in [1] and [2] that this minimization is then reduced to a quadratic programming (QP) problem. Since the cost function and constraints are both convex, a solution is always found [1]; however, this requires a large computational cost. In order to cope with this difficulty, a method based on a linear approximation of the QP optimization is proposed [1]. Once the impulse responses at all the pixels are determined in an off-line step, the online computation of the projector images can be performed in real-time by applying only linear filtering to the desired image. In this correspondence, we consider only the images realized by solving the QP problem. Nevertheless, the results that will be shown are valid for any other algorithm for computing the projector images including the above linear method, since the QP-based method is considered to achieve the maximum image quality among all algorithms.

As is discussed in [2], the PSF of projectors plays an essential role in the image formation by superimposed projection. In what follows, we mainly consider the box PSF having the same size as a pixel, which is considered to be a good approximation in the case of a DLP projector. However, our analysis does not require that the PSF is exactly a box function. Furthermore, we verify the validity of the assumption by implementing superimposed projection using a real multiprojector system and examining the resulting image quality.

II. Multiprojector Super-Resolution

We start by showing the basic formulation of superimposed projection along with the notations used in this correspondence.

A. Model of the Image Formation

For the purpose of simplicity, we consider only gray-level images throughout the correspondence. We assume that the projectors are photometrically calibrated and the brightness-radiance relation of the projector pixels is known. Let y_j^k be the brightness of pixel j of projector

Manuscript received March 12, 2008; revised August 22, 2008 and October 09, 2008. First published December 22, 2008; current version published January 09, 2009. This work was supported in part by KAKENHI(19680009). The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Stanley J. Reeves.

The authors are with the Graduate School of Information Sciences, Tohoku University, 980-8579 Sendai, Japan (e-mail: okatani@fractal.is.tohoku.ac.jp).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2008.2008899

k and $\hat{x}_i^k$ be the brightness of the image realized on the screen, which is resampled at the coordinates $\mathbf{u}_i$. Then, $\hat{x}_i^k$ is given by [1]

$$\hat{x}_i^k = \sum_j a_{ij}^k y_j^k \tag{1}$$

where a_{ij}^k is defined as

$$a_{ij}^k = \int p(\mathbf{u} - \mathbf{u}_i) r^k \left(\mathbf{f}^k(\mathbf{u}) - \mathbf{v}_j^k\right) d\mathbf{u} \tag{2}$$

where $p(\cdot)$ is a prefilter (anti-alias filter); $r^k(\cdot)$, the PSF reconstruction filter of the projector; $\mathbf{f}^k(\mathbf{u})$, a mapping function that maps a screen point $\mathbf{u}$ to its corresponding point $\mathbf{v}$ on the k-th projector image; and $\mathbf{v}_j^k$, the coordinates of the projector pixel j. By geometrically and photometrically calibrating the system, a_{ij}^k can be calculated.

Let $\mathbf{y}^k = [y_1^k, \ldots, y_n^k]^\top$ be the k-th projector image and $\hat{\mathbf{x}}^k = [\hat{x}_1^k, \ldots, \hat{x}_m^k]^\top$ be its projection on the screen. Further, let $\mathbf{A}^k$ be the $m \times n$ matrix containing a_{ij}^k as the (i, j) element. Then, (1) is rewritten as $\hat{\mathbf{x}}^k = \mathbf{A}^k \mathbf{y}^k$. When superimposing K projector images on the screen, the realized image $\hat{\mathbf{x}} = [\hat{x}_1, \ldots, \hat{x}_m]^\top$ is given by

$$\hat{\mathbf{x}} = \sum_{k=1}^{K} \mathbf{A}^k \mathbf{y}^k. \tag{3}$$

B. Computing the Projector Images

Given a desired image $\mathbf{x} = [x_1 \ldots, x_n]^\top$ that we wish to display, we first compute K projector images $\{\mathbf{y}^1, \ldots, \mathbf{y}^K\}$ such that the realized image $\hat{\mathbf{x}}$ is as close to $\mathbf{x}$ as possible. According to [2], it is theoretically shown that super-resolution, i.e., higher spatial resolution of $\hat{\mathbf{x}}$ than a single projector image, is achievable by finding the projector images $\{\mathbf{y}^1, \ldots, \mathbf{y}^K\}$ that minimize

$$J = |\mathbf{x} - \hat{\mathbf{x}}|^2. \tag{4}$$

This is subject to the lower and upper bounds of the projector pixel brightness. By applying some normalization, these bounds are represented as

$$0 \leq y_j^k \leq 1. \tag{5}$$

As is described in [1] and [2], the problem to be solved reduces to quadratic programming (QP) in which J of (4) is minimized subject to the inequality constraints (5). Since the solution of this QP problem requires a large computational cost because of the high dimensionality of the searchable space, a linear approximate method is proposed in [1]. As mentioned earlier, in what follows, we will consider only the images obtained by solving the above QP problem.

III. Dependency of Realization Accuracy on Image Structures

A. Analyzing Toy Problems

Let us consider the case of 1-D images. We assume the prefilter p and the reconstruction filter r^k to be box functions whose widths are equivalent to the pixel size of the target image and the projector image, respectively. In the real system of a projector, the PSF reconstruction filter is determined by several factors such as the optical lens of the projector, its imaging device (e.g., DLP and LCD), and the reflectance property of the screen surface onto which an image is projected. For LCD projectors, the pixel-size box function does not give a good approximation; a projected image tends to have a relatively large gap between pixels because of the structures of a LCD panel (i.e., the spread of the PSF tends to be smaller than the pixel size). For DLP projectors, it is considered to give a good approximation, because of the structure of micro-mirror arrays.

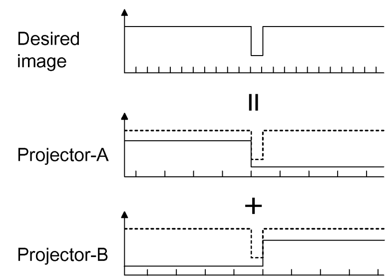

Fig. 1. Realization of an image of constant brightness with a single dark pixel. The ticks on the horizontal axes represent the boundaries of pixels. This configuration perfectly realizes the desired image.

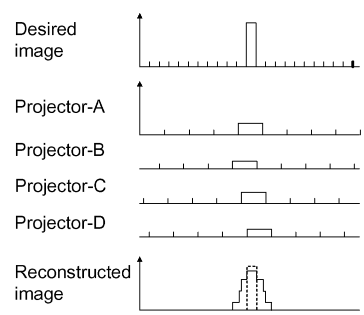

Fig. 2. Realization of an image with a single bright pixel. The shape cannot be generated even by an infinite number of projectors.

Suppose we wish to realize the image shown in Fig. 1, whose brightness is constant except for a single dark pixel. If there are a sufficient number of projectors, it is possible to choose two specific projectors from among them such that their pixel boundaries coincide with the boundaries of the dark pixel of the target image, as shown in the figure. The image is then realized in an exact manner by providing the appropriate component images to the two chosen projectors.

In reality, it is highly improbable to find two projectors having exactly matched pixel boundaries; the implication of this somewhat special example has generality because of the continuous relation between the pixel geometry and the final image quality: the closer the nearest pixel boundary is to the target pixel boundary, the more accurate is the realized image. [recall that the realization accuracy is given by the cost (4)]. Thus, for this single-dark-pixel image, the chance of accurately realizing the desired image is very high; the best result could even be its perfect realization. We may conclude that this image structure is considered to be the most suitable for the superimposed projection.

On the other hand, suppose we want to realize the image shown in Fig. 2 in which the brightness is consistently zero (or small) except for a single bright pixel. Unlike the above image, it is easily seen that the image will never be perfectly realized, even when there are an infinite number of projectors. Thus, although it is only a reversal of the first image (Fig. 1), the realization accuracy for this image is considered to be more limited (by a considerable amount in some cases).

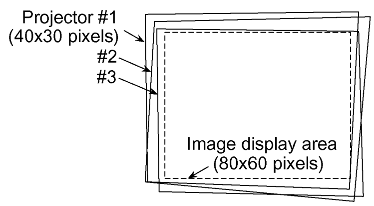

Fig. 3. Configuration of the projectors used in the experiments.

This asymmetric nature of the realization accuracy for these two image structures stems from a fundamental constraint of the superimposed projection: the image brightness can only be added and cannot be subtracted.

B. Experimental Results

In order to examine the above analysis in greater detail, we conducted experiments that simulated a superimposed projection. In the experiment, the text images shown in the top row of Fig. 4 are used as target images. Since letters comprise of thin lines and small dots, the left image (i.e., text in black letters on a white background) has a nature similar to the image of a single dark pixel (Fig. 1), whereas the right image (i.e., text in white letters on a black background) has a nature similar to the single-bright-pixel image (Fig. 2). These are extreme and realistic image examples.

The system configuration is as follows. The target image has 80×60 pixels, and the projector images have 40×30 pixels. The geometric configuration of the projector images is randomly chosen, as shown in Fig. 3. A planar homography gives the mapping from each projector image onto the screen. We choose a box prefilter p of the same size as the target image pixels and also a box reconstruction filter r^k of the same size as the projector image pixels. Then, the matrix $\mathbf{A}^k$ of the image formation $\hat{\mathbf{x}} = \mathbf{A}^k\mathbf{y}^k$ is calculated. This configuration is used in all the simulation-based experiments described in the remainder of the correspondence.

Fig. 4 shows the original images and their optimal realizations obtained using 1, 3, and 10 projectors, respectively. These images are obtained by solving the QP problem. It is observed from the results that the quality of the realized image is clearly improved by using superimposed projection. The texts, which are unreadable in the single projector image, becomes sufficiently clear in the superimposed projection. It is also observed that a better image quality is realized when more projectors are used. These findings confirm the result stated in [1].

More importantly, it is observed that the apparent quality of the realized images differs between the black-on-white and white-on-black images. Specifically, the lettering tends to be sharper in the black-on-white image than in the white-on-black image; for example, the heads of the letter "t" are realized more accurately in the former. The realized images appear to be more blurred in the white-on-black image. This tendency can be observed more clearly in the plot of the residue of the cost function (4) shown in Fig. 5; the residue for the white-on-black image is smaller as compared to the black-on-white image.

Thus, the earlier analysis is proven to be correct for realistic image examples. We make the following conclusion.

Result 1. *The accuracy of image realization by superimposed projection is affected by the target image itself. A typical example is that black-on-white text images tend to be realized more accurately than white-on-black text images.*

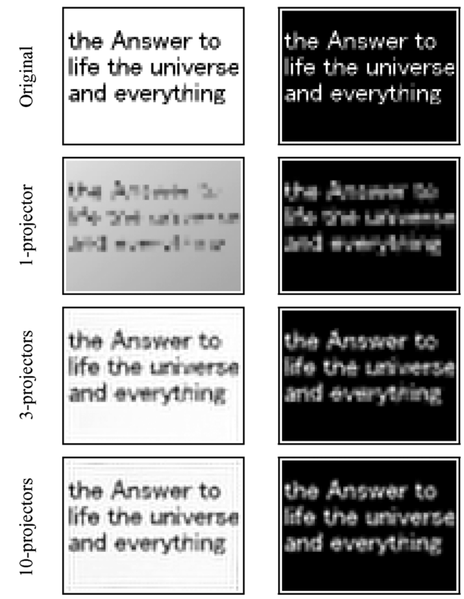

Fig. 4. Realizations of text images (80×60 pixels) for the case of 1, 3, and 10 projectors.

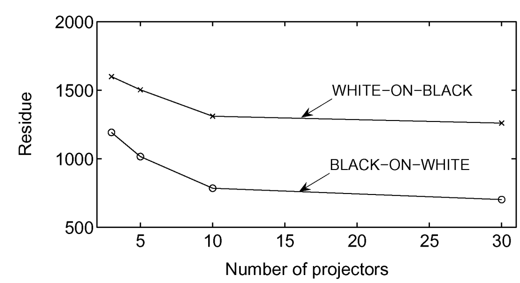

Fig. 5. Residue of the difference between the image to be realized and the realized image versus number of projectors.

IV. Tradeoff Between Realization Accuracy and Image Brightness

A. Degrees of Freedom in the Projector Image Optimization

When superimposing the images of K projectors of the same performance, the maximum brightness of the final image will theoretically be K times the brightness due to a single projector. Maximum brightness is considered to be an important factor in projection-based-display. In simple terms, a larger value of maximum brightness is more preferable. Thus, we now discuss the relation between image quality and maximum brightness.

When an image is optimally realized by superimposed projection, the brightness of a single pixel of the realized image is the contribution of multiple pixels of multiple projectors. When the target pixel is less bright, none of these projector pixels may reach the upper bound of the pixel brightness. However, when the target pixel has brightness exceeding a certain level, as a result of the QP optimization, some of these projector pixels will have reached the upper bound of brightness; only the rest of the projector pixels are permitted to have intermediate brightness values.

Thus, it can be predicted that *as the image to be realized becomes brighter, the degrees of freedom of the projector images decrease, which could cause a deterioration in the realization accuracy*. If this holds true, a natural question is—how large is the extent of the deterioration that occurs and what its its tendency. Based on the results of the previous section, we predict the following: *the deterioration in realization accuracy is more significant for images having an image structure similar to the single-dark-pixel image shown in* Fig. 1, *whereas it is less significant for images having a structure similar to the single-bright-pixel image shown in* Fig. 2.

The reason for this prediction is as follows. For the image structure of the single-dark-pixel image (Fig. 1), when the overall brightness of the image is small, its optimal realization will be given by a small number of projectors; only a limited number of projectors contribute to the final image brightness, as is seen in the previous section. Therefore, if the target image brightness exceeds the ability of a single projector, it is more likely that such highly selective use of projectors is not possible. This tendency becomes more significant with the increasing brightness of the target image.

As is discussed previously, for the image structure of the single-bright-pixel image (Fig. 2), multiple projectors tend to have an equal contribution in shaping the single bright pixel, as compared to the former case. Although a deterioration in realization accuracy will emerge when the target image becomes brighter, it is expected to be smaller and less critical than the deterioration in the former case, owing to the equal contribution of the projectors.

B. Experimental Results

Experiments were conducted to examine the above analyses. In the experiments, the five projectors that are used are configured in the same way as the earlier experiments. Moreover, the same target images (Fig. 4) are used. The range of the brightness value of every projector pixel is set to [0:1]. Then, the superimposed projection is simulated while varying the brightness multiple α of the target images.

Fig. 6 shows the results of the QP optimization. The brightness of the realized images is appropriately normalized for the sake of comparison. It is observed that as the brightness multiple α increases, the image quality monotonically deteriorates for both target images, which supports one of the above predictions. Figs. 7 and 8 show the component projector images for the black-on-white image in the case of $\alpha = 1$ and $\alpha = 4$, respectively. As discussed above, it is observed that the degrees of freedom in the determination of the projector images differs for the two cases ($\alpha = 1$ and 4), which must have affected the realization accuracy. In the case of $\alpha = 1$, the projector images are generally dark and are a mixture of extremely bright pixels and dark pixels; the projectors are used in a highly selective manner. On the other hand, in the case of $\alpha = 4$, it is seen that the projector images are generally brighter and similar to each other; the projectors are used in a more equal manner. These results agree with the above prediction.

A more detailed inspection of Fig. 6 reveals that the deterioration is more significant for the black-on-white image than for the white-on-black image. In the case of the black-on-white image, the realized image appears to be the sharpest for $\alpha = 1$, and the sharpness tends to decrease with the increase in α. In the case of the white-on-black image, the realized image already appears to be blurred for $\alpha = 1$, and the extent of the blur only mildly increases with α. These observations are more clearly confirmed by plotting the residue of the cost function (4) versus the brightness multiple α; see Fig. 9. For the black-on-white image, the residue rapidly increases with α, whereas for the white-on-black image, the residue increases more slowly. This result agrees with the second prediction. Unlike the case of the black-on-white image, the projector images for the white-on-black image are more similar to each other for different α. The projectors are used less selectively even for the case of $\alpha = 1$, which is the case with the largest degrees of freedom. This leads to the conclusion that realization accuracy is less sensitive to the degrees of freedom.

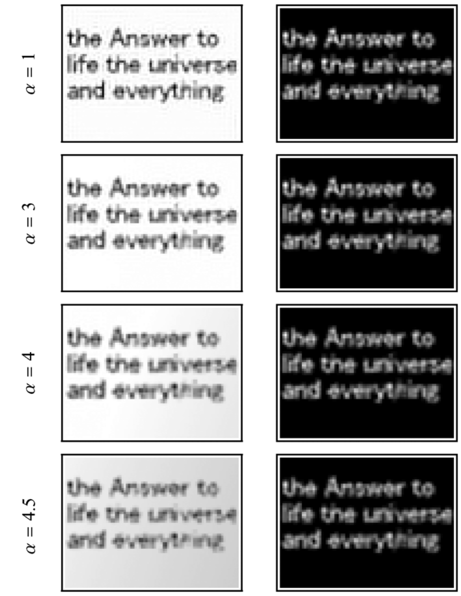

Fig. 6. Dependency of the realized image on the maximum brightness of the target image; α is the brightness multiple. Experiments conducted with five projectors.

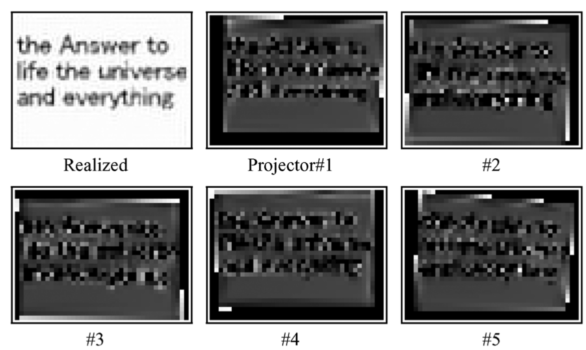

Fig. 7. Realized images and five component projector images for the cases of $\alpha = 1$.

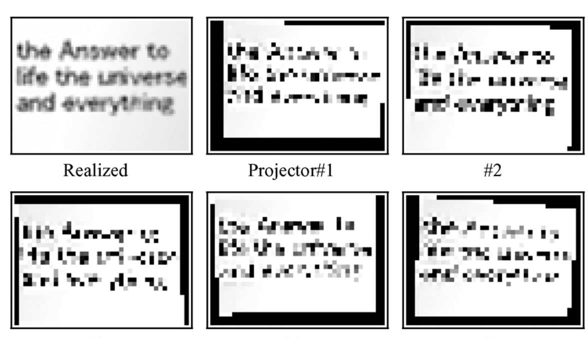

Fig. 8. Realized images and five component projector images for the cases of $\alpha = 4$.

We conclude this section as follows.

Result 2. *There is a general tendency that the realization accuracy monotonically deteriorates as the brightness of the target image increases. This tendency is more significant for the image structures similar to the single-dark-pixel image shown in Fig. 1 than for those similar to the single-bright-pixel image shown in Fig. 2.*

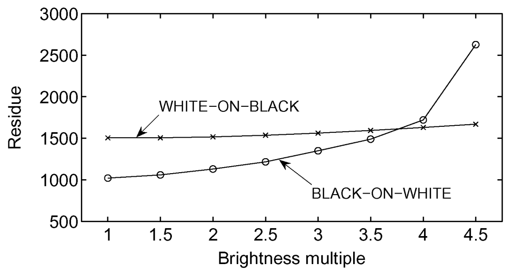

Fig. 9. Residue versus maximum brightness (α) of the target image.

Fig. 10. Experimental system. Left: The layout of eight projectors and a screen. Right: A video camera (left) for the calibration and a still-image camera (right) for verifying the image quality of the realized images.

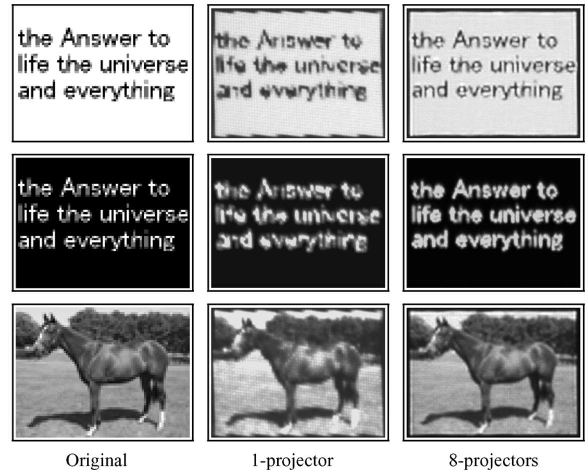

Fig. 11. Original images and the images realized by superimposed projection using one and eight projectors, respectively.

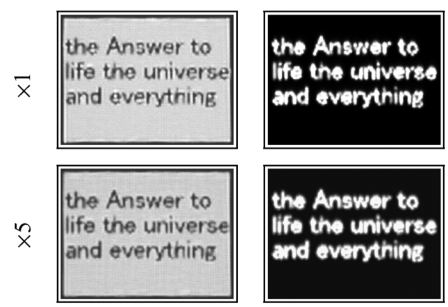

Fig. 12. Images realized by superimposed projection of eight projectors when the brightness of the target image is multiplied by one and five.

V. Demonstration by a Real System

In order to verify the results obtained thus far, we constructed a real system and conducted several experiments.

A. Experimental System

Fig. 10 shows our experimental system. It consists of eight Toshiba TDP-FF1A projectors. Their image size is 800 $\times$ 600 pixels. These projectors are DLP-based and have a LED-based light source. They are stored in two vertical racks, and the racks are horizontally placed about 1 m from the screen, as shown in the figure. The camera for the calibration of the system is a Sony HDR-HC7 video camera, which is placed between the two racks. The camera for capturing superimposed images for their verification is a Nikon D80 camera equipped with a Sigma 150-mm macro lens.

B. Experimental Results

Using the above system, we conducted several experiments. Before image projection, we performed geometric and photometric calibrations. For the geometric calibration, we used the phase-shift method, which is widely used in range finders, to establish the pixel-to-pixel correspondences between the images of the projector and the calibration camera. Using a sinusoidal brightness pattern for projection, we estimates the phase of the pattern from the brightness changes at each camera pixel when shifting the phase of the sinusoidal pattern. We have confirmed that the calibration can be performed with subpixel accuracy (order of 1/100 pixel size). Then, the QP optimization [(4) with the inequality constraint (5)] was performed to determine the component projector images. We used a few target images and they are all 80 $\times$ 60 pixels. The resolution of the associated projection area on the screen is about 60 $\times$ 45 pixels for each projector. Thus, the magnification ratio in the image resolution is 1.33.

Fig. 11 shows the results of superimposed projection by one and eight projectors for three different target images. Note that these images are captured by the camera different from the calibration camera and then trimmed. It is observed from the results that the image resolution increases for the case of eight projectors as compared with the case of a single projector, which confirms the effectiveness of superimposed projection. Moreover, it is observed that the black-on-white text image tends to be more precisely realized than the white-on-black image. The difference is significant in the thickness of the lines of their letters; in the black-on-white image, the lines are thin and their thickness is close to that of the target image, whereas in the white-on-black image, they are thicker than that of the target image.

Fig. 12 shows the realized images for different overall brightnesses of the target images; eight projectors are used and the brightness of the target images is multiplied by one and five. Comparing the results for different brightness multiple ($\times 1$ and $\times 5$), it is observed that for both target images, their realized images lose sharpness in the case of $\times 5$. The difference is more significant for the white-on-black image; comparing the appearance of the white regions enclosed by the lines of "e" and "g", it is observed that their areas are smaller in the case of $\times 5$ than in the case of $\times 1$. There is no such significant difference in the case of the white-on-black image. (It is noted that the difference is small even in the the case of the black-on-white image, as compared to the results (Fig. 6) of the synthetic experiments. This may be because many other factors than those considered in the theoretical analyses are involved in the real experiments, such as the difference between the real and ideal projector PSFs, the reflecting property of the screen material, the PSF of the camera taking the image projected on the screen, etc.)

These results verify our theoretical results with respect to the image quality of superimposed projection; they hold in the case of real systems and images.

VI. Summary and Discussions

In this correspondence, we have discussed the image quality of the multiprojector superimposed projection—the method of generating a high-resolution image by the use of multiple projectors. The main results are as follows: 1) the realization accuracy depends on the structure of the target image to be realized, and 2) it is also affected by the overall brightness of the target image, and, moreover, the extent of the affection is dependent on the structure of the target image. The underlying mechanisms behind these phenomena, which are related to the fundamental constraints of superimposed projection, are also shown.

Note that these underlying mechanisms are universal, that is, the above results hold for general images, although we have used two particular images for the purpose of explanation. Note also that the above results describe fundamental properties of the multiprojector superimposed projection. That is, the above results are valid for any algorithm for computing projector images, although we have considered only the QP optimization. This is because the QP optimization is considered to achieve the upper bound of image quality of the multiprojector superimposed projection; for any other algorithm, which should be *suboptimal* as compared with the QP optimization, its performance will be more or less governed by the same mechanisms. The method of (multi)projector superimposed projection is itself a novel imaging technology and could have many potential applications not limited to ordinary image projection. The results of this correspondence will be useful also in this context.

Super-resolution by multiprojector superimposed projection is similar in concept to more familiar multicamera super-resolution. In fact, the former is sometimes referred to as the dual of the latter [1], [5]. However, there is an interesting difference between the two, apart from the apparent difference that one is image projection and the other is image acquisition. As shown in this correspondence, the quality of images realized by the multiprojector super-resolution depends on the images themselves to be realized. Specifically, the accuracy of the image realization can be asymmetric with respect to image brightness; it could differ for a positive image and a negative image, as is seen in the case of a single-bright-pixel image and a single-dark-pixel image. From a viewpoint of image resolution, this indicates that image resolution effectively varies depending on the structures of images. In the case of multicamera super-resolution in its narrowest sense, (i.e., the method of increasing image resolution by repeated resampling), such a behavior does not appear. In the multicamera super-resolution, the upper limit on the image resolution is simply constrained by the PSF of the imaging system [6], [7]. In the multiprojector super-resolution, the resolution limit is determined not only by the PSF of the projectors, as discussed in this correspondence.

Acknowledgment

The authors would like to thank anonymous reviewers for their helpful comments and suggestions.

References

[1] N. Damera-Venkata and N. L. Chang, "Realizing super-resolution with superimposed projection," presented at the IEEE Int. Workshop on Projector-Camera Systems, 2007.
[2] N. Damera-Venkata and N. L. Chang, "On the resolution limits of superimposed projection," presented at the IEEE Int. Conf. Image Processing, 2007.
[3] W. Allen and R. Ulichney, "Wobulation: Doubling the addressed resolution of projection displays," in *Proc. SID Symp. Dig. Technical Papers (SID)*, 2005, pp. 1514–1517.
[4] C. Jaynes and D. Ramakrishnan, "Super-resolution composition in multi-projector displays," presented at the Proc. IEEE Int. Workshop on Projector-Camera Systems, 2003.
[5] A. Majumder, "Is spatial super-resolution feasible using overlapping projectors?," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, 2005, pp. 209–212.
[6] S. Baker and T. Kanade, "Limits on super-resolution and how to break them," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 9, pp. 1167–1183, Sep. 2002.
[7] H. Lin and H. Shum, "Fundamental limits of reconstruction-based superresolution algorithms under local translation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 1, pp. 83–97, Jan. 2004.

High-Fidelity Data Embedding for Image Annotation

Shan He, *Member, IEEE*, Darko Kirovski, *Member, IEEE*, and Min Wu, *Senior Member, IEEE*

***Abstract*—High fidelity is a demanding requirement for data hiding, especially for images with artistic or medical value. This correspondence proposes a high-fidelity image watermarking for annotation with robustness to moderate distortion. To achieve the high fidelity of the embedded image, we introduce a visual perception model that aims at quantifying the local tolerance to noise for arbitrary imagery. Based on this model, we embed two kinds of watermarks: a pilot watermark that indicates the existence of the watermark and an information watermark that conveys a payload of several dozen bits. The objective is to embed 32 bits of metadata into a single image in such a way that it is robust to JPEG compression and cropping. We demonstrate the effectiveness of the visual model and the application of the proposed annotation technology using a database of challenging photographic and medical images that contain a large amount of smooth regions.**

***Index Terms*—High fidelity, human visual model, image annotation, image watermarking.**

I. Introduction

Data embedding provides a means to seamlessly associate annotation data with host images. One of the key requirements of data embedding based annotation is high fidelity, in order to preserve the artistic or medical value of host images. A proper visual model is important to control the distortion introduced by embedded information. There have been a number of human visual models proposed and employed in the watermarking literature. Cox *et al.* employed a simplified frequency scaling model in their spread-spectrum watermarking paper [1]. More explicit masking models, including frequency masking and spatial masking, were developed by Swanson *et al.* [2] and Podilchuk and Zeng [3]. Based on the work of [3], Wu *et al.* have proposed a refined visual model to reduce ringing artifacts along edges [4].

The stringent requirement in high-fidelity applications calls for more research into further reducing perceptual artifacts in data embedding

Manuscript received April 23, 2008; revised September 18, 2008. First published December 22, 2008; current version published January 09, 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Bruno Carpentieri.

S. He is with Thomson Corporate Research, Princeton, NJ 08540 USA (e-mail: shan.he@thomson.net).

D. Kirovski is with Microsoft Research, Redmond, WA 98052 USA (e-mail: darkok@microsoft.com).

M. Wu is with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA (e-mail: minwu@eng.umd.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2008.2008733