

Viewing-Distance Aware Super-Resolution for High-Definition Display

Chih-Tsung Shen, *Member, IEEE*, Hung-Hsun Liu, Ming-Hsuan Yang, *Senior Member, IEEE*, Yi-Ping Hung, and Soo-Chang Pei, *Life Fellow, IEEE*

Abstract—In this paper, we propose a novel algorithm for high-definition displays to enlarge low-resolution images while maintaining perceptual constancy (i.e., the same field-of-view, perceptual blur radius, and the retinal image size in viewer's eyes). We model the relationship between a viewer and a display by considering two main aspects of visual perception, i.e., scaling factor and perceptual blur radius. As long as we enlarge an image while adjust its image blur levels on the display, we can maintain viewer's perceptual constancy. We show that the scaling factor should be set in proportion to the viewing distance and the blur levels on the display should be adjusted according to the focal length of a viewer. Toward this, we first refer to edge directions to interpolate a low-resolution image with the increasing of viewing distance and the scaling factor. After images are interpolated, we utilize a local contrast to estimate the spatially varying image blur levels of the interpolated image. We then further adjust the image blur levels using a parametric deblurring method, which combines L1 as well as L2 reconstruction errors, and Tikhonov with total variation regularization terms. By taking these factors into account, high-resolution images adaptive to viewing distance on a display can be generated. Experimental results on both natural image metric and user subjective studies across image scales demonstrate that the proposed super-resolution algorithm for high-definition displays performs favorably against the state-of-the-art methods.

Index Terms—Super-resolution, high-definition display, viewing distance, spatially-varying blur, and image deblurring.

I. INTRODUCTION

IN THE recent years, thanks to the advanced panel technologies, the resolutions of displays become higher and higher; however, the resolutions of contents remain degraded and limited. For these reasons, research on ultra high-definition

Manuscript received March 15, 2014; revised July 8, 2014 and October 1, 2014; accepted November 13, 2014. Date of publication November 25, 2014; date of current version December 22, 2014. This work was supported by the Ministry of Science and Technology, Taiwan, under Contract MOST 103-2221-E-002-102. The work of M.-H. Yang was supported in part by the National Science Foundation (NSF) CAREER under Award 1149783 and in part by the NSF Information & Intelligent Systems under Grant 1152576. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Sina Farsiu.

C.-T. Shen and Y.-P. Hung are with the Graduate Institute of Networking and Multimedia, National Taiwan University, Taipei 106, Taiwan, and also with the Institute of Information Science, Academia Sinica, Taipei 11529, Taiwan (e-mail: shenchihtsung@gmail.com; hung@csie.ntu.edu.tw).

H.-H. Liu is with Telecommunication Laboratories, Chunghwa Telecom Company, Ltd., Taoyuan 32601, Taiwan (e-mail: hliu@cht.com.tw).

M.-H. Yang is with the School of Engineering, University of California at Merced, Merced, CA 95343 USA (e-mail: mhyang@ucmerced.edu).

S.-C. Pei is with the Department of Electrical Engineering, National Taiwan University, Taipei 106, Taiwan (e-mail: pei@cc.ee.ntu.edu.tw).

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.2014.2375639



Fig. 1. The proposed super-resolution can maintain human perceptual constancy across image scales. From left to right, images are enlarged for phones, pad-like devices, and televisions by a factor of 2, 3, and 4, respectively. At the corresponding distance, the viewer can have the same visual experience among these images on different displays.

televisions (UHDTVs) has attracted much attention as it requires delicate algorithms to process, render and display high-quality images. In this paper, we propose a novel algorithm to generate high-resolution images for HD displays while maintaining perceptual constancy at different viewing distance. We model human eye as a digital single-lens reflex (DSLR) camera from which visual perception is approximated. Therefore, perceptual constancy can be approximated to the same amount of field-of-view, perceptual blur radius of circle-of-confusion (CoC) and retinal image size in human eyes.

From perceptual researches and user studies [1], [2], we find out that each person can find his own field-of-view and position to best view an image on a HD display. When an image is enlarged significantly, a viewer must move away from the display to see the enlarged image in order to maintain the best field-of-view and the original retinal image size at the corresponding viewing distance. Thus, it is necessary and important to take viewing distance into account for image super-resolution on HD Displays (Fig. 1).

The above-mentioned model alone does not work properly as the blur levels of super-resolution process do not grow with the viewing distance directly, but rather another factor related to the focal length of a viewer's eyes at different viewing distance (Figs. 2 and 3). As such, it is necessary to consider these two factors simultaneously when enlarging an image. In this work, we first deduce these two perceptual factors from image scale, viewing distance and optical geometry. Then, we propose a super-resolution algorithm to enlarge an image and adjust the image blur levels to maintain the same perceptual quality.

For super-resolution on HD Displays, we take the relationship between both viewers and displays into account. Similar to the formulation for compensating visual aberrations [3],

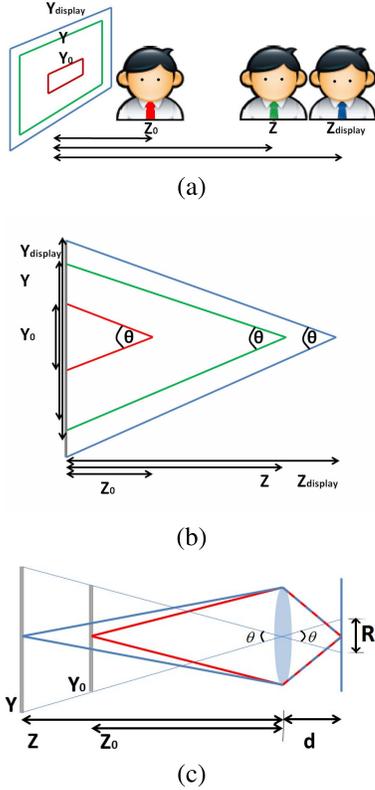


Fig. 2. (a) Relationship between viewing distance and enlarged image on the display where Z_0 is the viewing distance for a person to view the original image with the best field-of-view θ , and Y_0 is the size of the original image. (b) The field-of-view θ remains the same when the viewer moves away. (c) The two displays are at different viewing distances (left end). Both displays give the same response area (retinal image) R inside viewer's eye (right end). The lens represents the human crystalline lens.

we model human eye as a DSLR camera and analyze the relationship between perceptual factors and viewing distance. We show that image size increases with the scaling factor and viewing distance; while the perceptual blur (blur radius of CoC, σ_{eye}) grows with the focal length of human eyes. Since the viewing distance affects blur perception, it can be formulated as part of image formation process of the proposed super-resolution.

Based on our approximate model of visual perception for HD displays, we first calculate the perceptual factors: image scale and the corresponding image blur levels ($\sigma_{display}$) which can preserve human perceptual constancy. Then, we interpolate an input image based on its edge orientation to the desired size and solve the inverse problem of the super-resolution formulation. Since the interpolation operation usually enlarges the image blur levels as well, we have to estimate the blur levels of the interpolated image and remove an extra amount of blurriness so as to generate the final super-resolution image which can maintain perceptual constancy. While existing methods mainly deal with spatially invariant image blur, we estimate the spatially-varying blur map based on local contrast and then remove the extra image blur by spatially-varying deblurring [1], [2], [4]. For image deblurring, we propose to use a novel L1L2TTV term which integrates L1 and L2 reconstruction error terms, total variation as well as Tikhonov regularization to model the image formulation

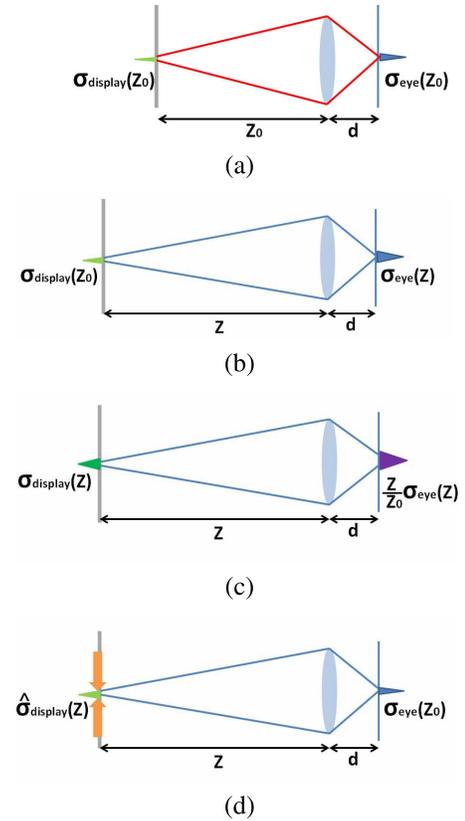


Fig. 3. Optical paths change with the viewing distance from the viewer (right end) to the display (left end). We adjust the image blur level on the display to have the same visual blur perception at distance Z_0 . (a) Optical path under the original resolution. (b) Perceptual blur radius changes at viewing distance Z . (c) Optical path when the image is enlarged with Z . (d) Adjust image blur levels to have perceptual constancy.

with image blur [5]–[14]. Based on the estimated blur map, we can adopt the pixel selection technique from the deblurred image stack and reconstruct the high-resolution output with which meets visual perceptual constancy. Experimental results on natural image metric and user subjective studies show that our proposed viewing-distance aware super-resolution algorithm performs favorably against the state-of-the-art methods qualitatively and quantitatively.

II. RELATED WORK

In this section, we revisit the previous super-resolution works and emphasize the contributions on our perceptual super-resolution. Different to the others, we not only pursue image quality which in terms of high image metric values but also preserve human visual perceptual constancy.

One of the early work on super-resolution is proposed by Tsai and Huang [15] where a high-resolution image is generated from multiple low-resolution frames based on image registration. Later on, the multi-frame reconstruction-based super-resolution methods [16], [17] are developed by considering the image formation process. Within this formation framework, images are formed in four steps: registration, downsampling, blurring and thermal noise estimation [18], [19]. The problem to construct a high-resolution image from multiple low-resolution images is

thus formulated as a series of inverse tasks. Farsiu et al. [6] propose a multi-frame super-resolution method using a L1 reconstruction term with bilateral total variation. Recently, Liu and Sun [20] adopt optical flow, estimate the spatially-invariant blur and noise level on their Bayesian adaptive video super-resolution scheme. It is true that multiple frames can bring more information to reconstruct the super-resolution image; however, the feature matching and optical flow in the registration stage is extremely expansive. Therefore, researchers in the field of image processing, vision and graphics try to seek the single-frame solution to super-resolution.

Numerous single-frame super-resolution algorithms have since been developed based on interpolation, reconstruction, learning, and integration approaches. Among these algorithms, interpolation can be regarded as the basic approach to enlarge an image. Lee et al. [21] propose an image resizing method based on an oblique rather than an conventional orthogonal projection operator for using faster and simpler B-spline algorithms. In [22], Li and Orchard present an edge-directed algorithm by estimating local covariance coefficients from a low-resolution image to adaptively interpolate a high-resolution image based on the corresponding geometric duality (i.e., consistency of geometric structure across resolutions). By grid filling and iterative correction of interpolated pixels, Giachetti and Asuni [23] develop an iterative curvature-based interpolation (ICBI) approach for super-resolution. Recently, Hung and Siu [24] use a weighted least squares regression function to address the mismatch issues with geometric duality for image interpolation. Generally, interpolation used to only combine the pixel information and edge orientation within a local region to fulfill the missing pixels (holes) on the grid. Therefore, we adopt the edge-aware interpolation as the foundation of our super-resolution algorithm.

In the recent years, prior knowledge becomes a very useful skill for image processing and computer vision. Researchers also introduce the prior technique to both reconstruction-based and learning-based super-resolution. To reconstruction, Fattal [25] presents an image up-sampling algorithm by using edges statistics learned from a set of natural images. Sun et al. develop a super-resolution method to generate high-resolution output where edges can be reconstructed well based on gradient prior learned from natural images [26]. Shan et al. pose the super-resolution problem within a feedback-control framework for up-sampling images and videos [9]. Most recently, Mallat and Yu [27] introduce a class of inverse problem estimators for super-resolution by mixing a family of linear functions based on different priors. Different to reconstruction, learning-based super-resolution algorithms have been developed by stitching best matched low-resolution patches directly from the input image or external datasets so as to meet the prior settings. Freeman et al. [28], [29] match low-resolution patches to find the corresponding high-frequency high-resolution ones from an external dataset and combine them using belief propagation. Baker and Kanade [30] use local features of low-resolution images to generate high-resolution face images while HaCohen et al. [31] use segments where each region is

matched against a database to generate a high-resolution image with similar texture. Several methods based on dictionary learning have been proposed [32]–[35] in which each pair of low-resolution and high-resolution patches is represented by one or multiple dictionaries, from which super-resolution images are constructed. However, underlying these approaches are the inherent difficulty to find good patch matches effectively and efficiently. As a result, learning-based super-resolution methods often generate noisy artifacts and require heavy computational loads.

Recently, methods that integrate different aspects of interpolation, reconstruction and learning have been proposed. Glasner et al. exploit patch self-similarity within and across image pyramid from a single image to construct a higher-resolution image [36]. Similarly, Freedman and Fattal use self-examples within local regions to reduce search time for images and video super-resolution [37]. Kim and Kwon develop a super-resolution method based on reconstruction and learning approaches using a regression function and learned image gradient prior [38].

After revisiting the above-mentioned super-resolution, we surprisingly observe that most works neglect human visual feeling at all. Accordingly, in this paper, we motivate ourselves to examine the relationship between viewers and HD displays. In the aspect of algorithm designs, our proposed single-frame super-resolution is an integrated approach which combines edge-aware interpolation and reconstruction with sparse priors.

III. VIEWER PERCEPTION AND DISPLAY

In this work, the super-resolution problem for HD Displays is formulated by taking both viewing distance and image formulation into account. We show that the scaling factor and image blur levels should be processed simultaneously in order to maintain perceptual constancy of a viewer.

A. Scaling Factor, Image Size and Viewing Distance

When viewing an image, each person has his best field-of-view and viewing distance to a display. In this work, we model visual perception in terms of the thin-lens model of a DSLR camera [3]. Fig. 2 shows an input image of size Y_0 on a HD Display is viewed by a person at distance Z_0 with field-of-view θ . The goal of this work is to enlarge an image with the same field-of-view θ and the same response area (retinal image R) as the viewer moves away the display. That is, the image size Y is enlarged in proportion to the viewing distance Z from the viewer to the display with the following relationship:

$$\tan\left(\frac{\theta}{2}\right) = \frac{Y_0}{2 \cdot Z_0}. \quad (1)$$

In order to maintain perceptual constancy, the field-of-view remains the same. Then, we obtain the relationship between the arbitrary image size Y and the corresponding viewing distance Z as follows:

$$Z = \frac{Y}{2 \cdot \tan\left(\frac{\theta}{2}\right)}. \quad (2)$$

Thus, the scaling factor $K_{Y,Z}$ is obtained by

$$K_{Y,Z} = \frac{Z}{Z_0} = \frac{Y}{Y_0}. \quad (3)$$

We can also denote $Z = Z_0 + \Delta Z$ and $Y = Y_0 + \Delta Y$ to obtain the scaling factor,

$$K_{Y,Z} = \frac{Z_0 + \Delta Z}{Z_0} = \frac{Y_0 + \Delta Y}{Y_0}. \quad (4)$$

When the image is enlarged to the size $Y_{display}$, we obtain the relationship between the $Z_{display}$ and $Y_{display}$ while maintaining the best field-of-view of an individual viewer by

$$Z_{display} = \frac{Y_{display}}{2 \cdot \tan(\frac{\theta}{2})}. \quad (5)$$

B. Relationship Between the Viewing Distance, Perceptual Blur Radius and Image Blur Levels

When an image is enlarged, the perceptual blur radius insides human eyes may not be changed in proportion to the scaling factor $K_{Y,Z}$. Fig. 3(a) shows the thin-lens model of human eyes of focal length f_0 with respect to the original image at distance Z_0 described by

$$\frac{1}{Z_0} + \frac{1}{d} = \frac{1}{f_0}, \quad (6)$$

where d represents the distance from the lens to the retina. As the viewer moves to distance Z and the image is enlarged to size Y , the viewer adaptively changes the focus point with length f by

$$\frac{1}{Z} + \frac{1}{d} = \frac{1}{f}. \quad (7)$$

We compute the blur radius of CoC (perceptual blur level) [39], [40] of human eyes $\sigma_{eye}(Z_0)$ and $\sigma_{eye}(Z)$ at the viewing distance Z_0 and Z , respectively. Since the lighting condition around the display is assumed to be the same, the aperture size N which represents human pupil remains unchanged. Thus, we have the perceptual blur radius $\sigma_{eye}(Z_0)$ and $\sigma_{eye}(Z)$, respectively [39], [40].

$$\sigma_{eye}(Z_0) = \frac{1}{2} \cdot \frac{(d - f_0) \cdot Z_0 - f_0 \cdot d}{N \cdot Z_0}, \quad (8)$$

$$\sigma_{eye}(Z) = \frac{1}{2} \cdot \frac{(d - f) \cdot Z - f \cdot d}{N \cdot Z}. \quad (9)$$

Based on (8) and (9), we derive the relationship between the perceptual blur radius of human eyes according to different viewing distance Z . The ratio between $\sigma_{eye}(Z)$ and $\sigma_{eye}(Z_0)$ changes with respect to focal length f and f_0 ,

$$\sigma_{eye}(Z) = \frac{f}{f_0} \cdot \sigma_{eye}(Z_0). \quad (10)$$

This shows that human perception of blur radius changes (CoC changes) with respect to focal length rather than the scaling factor $K_{Y,Z}$ directly.

The derivations are as follows:

$$\begin{aligned} \sigma_{eye}(Z) &= \frac{1}{2} \cdot \frac{(d - f) \cdot Z - f \cdot d}{N \cdot Z} \\ &= \frac{1}{2} \cdot \frac{(\frac{d}{f} - 1) \cdot Z - d}{N \cdot Z \cdot \frac{1}{f}} \\ &= \frac{1}{2} \cdot \frac{(\frac{d}{f_0} - \frac{d}{Z_0} + \frac{d}{Z} - 1) \cdot Z - d}{N \cdot Z \cdot \frac{1}{f}} \\ &= \frac{1}{2} \cdot \frac{(\frac{d}{f_0} - \frac{d}{Z_0} - 1)}{N \cdot \frac{1}{f}} \\ &= \frac{1}{2} \cdot \frac{N \cdot \frac{1}{f_0} \cdot (2 \cdot \sigma_{eye}(Z_0))}{N \cdot \frac{1}{f}} \\ &= \frac{f}{f_0} \cdot \sigma_{eye}(Z_0). \end{aligned}$$

The change of the blur radius of CoC inside human eyes is described by

$$\begin{aligned} \Delta \sigma_{eye}(Z) &= \sigma_{eye}(Z) - \sigma_{eye}(Z_0) \\ &= \sigma_{eye}(Z_0 + \Delta Z) - \sigma_{eye}(Z_0) \\ &= \frac{f_0 + \Delta f}{f_0} \cdot \sigma_{eye}(Z_0) - \sigma_{eye}(Z_0) \\ &= \frac{\Delta f}{f_0} \cdot \sigma_{eye}(Z_0). \end{aligned} \quad (11)$$

The image (blur) formation process of the thin-lens model can be approximated by a convolution with a low-pass Gaussian filter $G(\sigma_{display}(Z))$ with the standard deviation $\sigma_{display}(Z)$ as the image blur level on the display.

In addition, we denote the image formation process from a display to human eyes as a low-pass blurring function $H(\sigma(Z))$. Based on the central limit theorem, we model the image formation process $H(\cdot)$ as a convolution with a Gaussian kernel [1], [40]. Thus, the blurring function on human eyes is a Gaussian blurring function $G(\sigma_{eye}(Z))$ with its blur radius of CoC $\sigma_{eye}(Z)$. We express the image (blur) formation processes at distance Z_0 and Z , respectively (See Fig. 3(a) and Fig. 3(b)).

$$G(\sigma_{display}(Z_0)) \otimes H(Z_0) = G(\sigma_{eye}(Z_0)), \quad (12)$$

$$\begin{aligned} G(\sigma_{display}(Z_0)) \otimes H(Z) &= G(\sigma_{eye}(Z)) \\ &= G(\frac{f}{f_0} \cdot \sigma_{eye}(Z_0)), \end{aligned} \quad (13)$$

where \otimes denotes the convolution and the general form of the Gaussian blurring function is

$$G(x, y) = \frac{1}{2\pi\sigma^2} \cdot \exp(-\frac{x^2 + y^2}{2\sigma^2}), \quad (14)$$

with the following property,

$$G(\sqrt{(\sigma(a))^2 + (\sigma(b))^2}) = G(\sigma(a)) \otimes G(\sigma(b)). \quad (15)$$

After edge-directed interpolation [22], both image blur level on the display and the perceptual blur radius of CoC inside

human eyes increase. As shown in Fig. 3(c), we have

$$\begin{aligned} G(\sigma_{display}(Z)) \otimes H(Z) &= G\left(\frac{Z}{Z_0} \cdot \sigma_{eye}(Z)\right) \\ &= G\left(\frac{Z}{Z_0} \cdot \frac{f}{f_0} \cdot \sigma_{eye}(Z_0)\right). \end{aligned} \quad (16)$$

To obtain the same visual perception, we set the right-hand side to be $G(\sigma_{eye}(Z_0))$. Therefore, the image blur levels on the display $\hat{\sigma}_{display}(Z)$ is obtained by

$$G(\hat{\sigma}_{display}(Z)) \otimes H(Z) = G(\sigma_{eye}(Z_0)). \quad (17)$$

Using (16) and (17), we obtain

$$\hat{\sigma}_{display}(Z) = \frac{\sigma_{display}(Z)}{\frac{Z}{Z_0} \cdot \frac{f}{f_0}} = \frac{\sigma_{display}(Z_0)}{\frac{f}{f_0}}. \quad (18)$$

We adjust the image blur levels on the display such that viewers have the same visual perception as the original image Y_0 at distance Z_0 . After obtaining the perceptual-tuning factors f and $\frac{Z}{Z_0}$, we can adjust the image blur levels on the display by using spatially-varying deblurring method.

Using (15), we have

$$\hat{\sigma}_{display}(Z)^2 + \sigma_{display,de}(Z)^2 = \sigma_{display}(Z)^2, \quad (19)$$

where $\sigma_{display,de}(Z)$ is the adjusted image deblur level on the display.

In our formulation, while the image size is enlarged by $\frac{Z}{Z_0}$, the blur radius of CoC inside human eyes $\sigma_{eye}(Z_0)$ is enlarged by $\frac{Z}{Z_0} \cdot \frac{f}{f_0}$. As long as we adjust the image blur levels on the display $\hat{\sigma}_{display}(Z)$, we can maintain perceptual constancy. However, since we enlarge an image in the edge-directed interpolation step, the image blur levels on the display are also enlarged with the scaling factor. Thus, we use deblurring operations such that the resulting image blur levels are $\hat{\sigma}_{display}(Z)$. According to different viewing distance Z , we remove the extra blur amount by de-convoluting a Gaussian blur function with its standard deviation $\sigma_{display,de}(Z)$.

We first derive the equations in the continuous domain and then apply to the digital domain. Using (18) and (19), we obtain the approximated image deblur levels $\sigma_{display,de}(Z)$ at viewing distance Z ,

$$\begin{aligned} \sigma_{display,de}(Z) &= \sqrt{1 - \frac{1}{\left(\frac{Z}{Z_0}\right)^2 \cdot \left(\frac{f}{f_0}\right)^2}} \cdot \sigma_{display}(Z) \\ &= \sqrt{\left(\frac{Z}{Z_0}\right)^2 - \left(\frac{f_0}{f}\right)^2} \cdot \sigma_{display}(Z_0). \end{aligned} \quad (20)$$

Moreover, since each person has one's best field-of-view to see the original image at the corresponding viewing distance Z_0 , Z_0 can be fixed. In addition, since we model human eye as a DSLR, the image distance d can be regarded as the distance from crystalline lens to retina and set to a fixed value in this work. Thus, according to (6) and (7), the ratio between f and f_0 is only related to the viewing distance Z which corresponds to its scaling factor $K_{Y,Z}$ and facilitates the deblur adjustment,

$$\frac{f}{f_0} = \frac{1 + \frac{d}{Z_0}}{1 + \frac{d}{Z}} = \frac{1 + \frac{d}{Z_0}}{1 + \frac{d}{K_{Y,Z} \cdot Z_0}}. \quad (21)$$

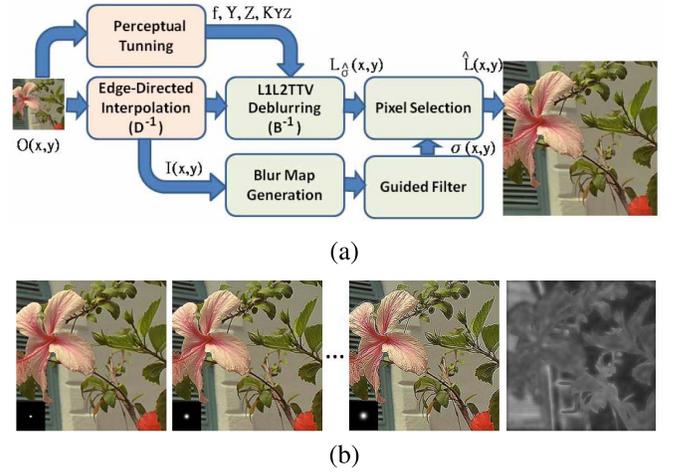


Fig. 4. Proposed super-resolution algorithm consisting of four modules: edge-directed interpolation, blur map generation, L1L2TTV deblurring with perceptual-tuning parameters and pixel selection. Different scaling factors lead to different parameters for perceptual constancy. (a) Pipeline. (b) Image stacks and blur map.

When an image is enlarged by the scaling factor $K_{Y,Z}$ in proportion to the viewing distance Z until $Z = Z_{display}$, the human perceived blur radius are increased by a factor of $\frac{f}{f_0}$.

IV. PROPOSED SUPER-RESOLUTION ALGORITHM

Fig. 4 shows the steps of the proposed super-resolution algorithm¹ which consists of four modules: edge-directed interpolation, blur map generation, L1L2TTV deblurring with perceptual-tuning parameters and pixel selection.

A. Image Formation Model for Super-Resolution

We propose a direct approach based on an image formation model for super-resolution. Let $O(x, y)$, $B(x, y)$, and $L(x, y)$ denote the observed low-resolution image, blurring function and the latent high-resolution image, respectively. The observed image is modeled by

$$O(x, y) = D\{B(x, y) \otimes L(x, y)\} + n(x, y), \quad (22)$$

where D is the downsampling operator and $n(x, y)$ denotes noise. To reconstruct the latent high-resolution image $L(x, y)$ from an observed input $O(x, y)$, we need to estimate noise, interpolate the image, and deblur the image. The inverse process is expressed as follows:

$$\hat{L}(x, y) = B^{-1}\{D^{-1}\{O(x, y) - n(x, y)\}\}. \quad (23)$$

We first use a median filter to remove noise while maintain the blurry condition. Second, we use an edge-directed interpolation method to upsample the image to the desired size according to the scaling factor. Third, different from previous works, we estimate a spatially-varying blur map for future blur adjustment. To reconstruct a high-quality latent image, we propose a L1L2TTV deblurring algorithm to remove the spatially-varying image blur for maintaining perceptual constancy at different viewing distance.

¹Preliminary results of this work are presented in [41].

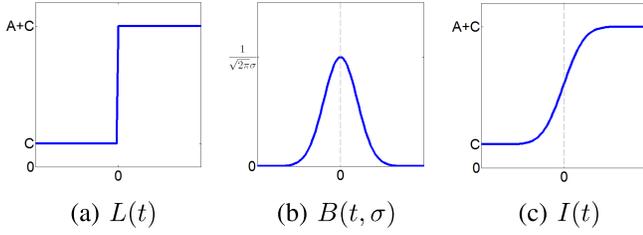


Fig. 5. From left to right, (a) the latent signal, (b) Gaussian blur function, and (c) the blurred signal, respectively. Here A is the local edge amplitude, C is the offset, and σ is the blur level of the Gaussian function.

B. Estimation of Spatially-Varying Image Blur

After image noise is removed by a median filter and edge-directed interpolation [22] is carried out, the spatially-varying image blur (i.e., B in (23)) is estimated from local contrast in spirit similar to [42]–[45]. Considering a 1D latent signal $L(t)$, we can model it as a linear combination of the unit step function $u(t)$ by

$$L(t) = A \cdot u(t) + C, \quad (24)$$

where A is the local edge amplitude, and C is the corresponding offset as shown in Fig. 5. The blurred signal $I(t)$ is modeled as a convolution with a blurring function $B(t, \sigma)$:

$$I(t) = L(t) \otimes B(t, \sigma). \quad (25)$$

Although the proposed super-resolution algorithm accommodates various blur kernels, we use a Gaussian function in this work due to the central limit theorem. The 1D blurring function is expressed by

$$B(t, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{t^2}{2\sigma^2}\right), \quad (26)$$

where σ represents the blur level. From (25) and (26), we have

$$\frac{\partial}{\partial t} I(t) = \frac{\partial}{\partial t} (L(t) \otimes B(t, \sigma)) = A \cdot B(t, \sigma). \quad (27)$$

The above formulation can be extended to 2D signals [44], [45]. When we compute the normalized gradient of the 2D blurred signal $I(x, y)$ inside a window (x', y') , due to $\max |\exp(0)| = 1$, the spatially-varying blur map (which represents all blur levels of an image), is obtained as follows,

$$\begin{aligned} & \frac{\max_{(x,y) \in (x',y')} \sqrt{\left(\frac{\partial}{\partial x} I(x,y)\right)^2 + \left(\frac{\partial}{\partial y} I(x,y)\right)^2}}{\max_{(x,y) \in (x',y')} I(x,y) - \min_{(x,y) \in (x',y')} I(x,y)} \\ & \approx \frac{A \cdot \max |B(t, \sigma)|}{A + C - C} \\ & = \frac{A}{A + C - C} \cdot \left| \frac{1}{\sqrt{2\pi}\sigma(x,y)} \cdot \exp(0) \right| \\ & = \frac{1}{\sqrt{2\pi}\sigma(x,y)}. \end{aligned} \quad (28)$$

In a general form, the spatially-varying blur map is computed by

$$\sigma(x,y) \approx \frac{1}{\sqrt{2\pi}} \cdot \frac{\max_{(x,y) \in (x',y')} I(x,y) - \min_{(x,y) \in (x',y')} I(x,y)}{\max_{(x,y) \in (x',y')} \sqrt{\left(\frac{\partial}{\partial x} I(x,y)\right)^2 + \left(\frac{\partial}{\partial y} I(x,y)\right)^2}}. \quad (29)$$

In addition, we use a guided filter [46] to remove ambiguous regions around strong edges and $\sigma(x, y)$ can be regarded as the 2D representation of $\sigma_{display}(Z)$ discussed in Section III-B.

C. L1L2TTV Deblurring

Once the spatially-varying blur map is estimated, the high-resolution image $L(x, y)$ can be recovered via a deconvolution operator $B^{-1}\{\cdot\}$. Let the aforementioned $I(x, y)$ denote the current interpolated image:

$$\begin{aligned} I(x, y) &= D^{-1}\{(O(x, y) - n(x, y))\} \\ &= B(x, y) \otimes L(x, y) + \epsilon(x, y), \end{aligned} \quad (30)$$

where $\epsilon(x, y)$ is the residual error and

$$B(x, y) = \frac{1}{2\pi\sigma^2} \cdot \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right). \quad (31)$$

Since L1 and L2 reconstruction errors have been shown to be effective for image deblurring [7], [8], [12], they are used in the proposed L1L2TTV deblurring algorithm with the Tikhonov and total variation regularization terms:

$$\begin{aligned} & \min_{L(x,y)} \gamma \sum_{(x,y)} |B(x,y) \otimes L(x,y) - I(x,y)| \\ & + \frac{\mu}{2} \sum_{(x,y)} |B(x,y) \otimes L(x,y) - I(x,y)|^2 \\ & + s \sum_{(x,y)} \sqrt{\left(\frac{\partial}{\partial x} L(x,y)\right)^2 + \left(\frac{\partial}{\partial y} L(x,y)\right)^2} \\ & + (1-s) \sum_{(x,y)} \left(\left(\frac{\partial}{\partial x} L(x,y)\right)^2 + \left(\frac{\partial}{\partial y} L(x,y)\right)^2 \right), \end{aligned} \quad (32)$$

where γ , μ and s are weights; $\frac{\partial}{\partial x} L(x, y)$ and $\frac{\partial}{\partial y} L(x, y)$ denote the gradients of $L(x, y)$ in the x and y directions. The third term of (32) is the total variation of $L(x, y)$ and the last term is the Tikhonov regularizer. We use variable-splitting and penalty techniques for the proposed L1L2TTV optimization in a way similar to [7], [12], and [47]. At each pixel, we introduce an auxiliary variable $w(x, y)$ to transfer $\frac{\partial}{\partial x} L(x, y)$ and $\frac{\partial}{\partial y} L(x, y)$ outside the non-differentiable term and penalize the difference between $w(x, y)$, $\frac{\partial}{\partial x} L(x, y)$ and $\frac{\partial}{\partial y} L(x, y)$. In addition, we introduce another variable $v(x, y)$ to substitute the L1 reconstruction term. Thus, the L1L2TTV optimization problem is solved by

$$\begin{aligned} & \min_{L,v,w} \frac{\gamma\alpha}{2} \sum_{(x,y)} |B(x,y) \otimes L(x,y) - I(x,y) - v(x,y)|^2 \\ & + \gamma \sum_{(x,y)} |v(x,y)| \\ & + \frac{\mu}{2} \sum_{(x,y)} |B(x,y) \otimes L(x,y) - I(x,y)|^2 \\ & + s \sum_{(x,y)} |w(x,y)| \\ & + \frac{s\beta}{2} \sum_{(x,y)} |w(x,y) - \sqrt{\left(\frac{\partial}{\partial x} L(x,y)\right)^2 + \left(\frac{\partial}{\partial y} L(x,y)\right)^2}|^2 \end{aligned}$$

$$+ (1-s) \sum_{(x,y)} \left(\left(\frac{\partial}{\partial x} L(x,y) \right)^2 + \left(\frac{\partial}{\partial y} L(x,y) \right)^2 \right), \quad (33)$$

where β and γ are balancing parameters such that $w(x,y) = \sum_{(x,y)} \sqrt{\left(\frac{\partial}{\partial x} L(x,y)\right)^2 + \left(\frac{\partial}{\partial y} L(x,y)\right)^2}$ and $v(x,y) = B(x,y) \otimes L(x,y) - I(x,y)$. With this approximation, the minimization is modified as a penalty function in terms of $L(x,y)$, $v(x,y)$, and $w(x,y)$. We use an alternating approach to minimize the penalty function with respect to $L(x,y)$, $v(x,y)$ and $w(x,y)$. For a fixed $L(x,y)$, (33) is equivalent to

$$\begin{aligned} \min_{v,w} & \frac{\gamma\alpha}{2} \sum_{(x,y)} |B(x,y) \otimes L(x,y) - I(x,y) - v(x,y)|^2 \\ & + \gamma \sum_{(x,y)} |v(x,y)| + s \sum_{(x,y)} |w(x,y)| \\ & + \frac{s\beta}{2} \sum_{(x,y)} \left| w(x,y) - \sqrt{\left(\frac{\partial}{\partial x} L(x,y)\right)^2 + \left(\frac{\partial}{\partial y} L(x,y)\right)^2} \right|^2. \end{aligned} \quad (34)$$

Since $v(x,y)$ and $w(x,y)$ are independent, we solve them as follows,

$$\begin{aligned} v(x,y) &= \max\{|B(x,y) \otimes L(x,y) - I(x,y)| - \frac{1}{\alpha}, 0\} \\ & \cdot \frac{B(x,y) \otimes L(x,y) - I(x,y)}{|B(x,y) \otimes L(x,y) - I(x,y)|}, \end{aligned} \quad (35)$$

and

$$\begin{aligned} w_x(x,y) &= \max \left\{ \sqrt{\left(\frac{\partial}{\partial x} L(x,y)\right)^2 + \left(\frac{\partial}{\partial y} L(x,y)\right)^2} - \frac{1}{\beta}, 0 \right\} \\ & \cdot \frac{\frac{\partial}{\partial x} L(x,y)}{\sqrt{\left(\frac{\partial}{\partial x} L(x,y)\right)^2 + \left(\frac{\partial}{\partial y} L(x,y)\right)^2}}, \end{aligned} \quad (36)$$

$$\begin{aligned} w_y(x,y) &= \max \left\{ \sqrt{\left(\frac{\partial}{\partial x} L(x,y)\right)^2 + \left(\frac{\partial}{\partial y} L(x,y)\right)^2} - \frac{1}{\beta}, 0 \right\} \\ & \cdot \frac{\frac{\partial}{\partial y} L(x,y)}{\sqrt{\left(\frac{\partial}{\partial x} L(x,y)\right)^2 + \left(\frac{\partial}{\partial y} L(x,y)\right)^2}}, \end{aligned} \quad (37)$$

where $w(x,y) = w_x(x,y) + w_y(x,y)$ and the convention $0 \cdot \frac{0}{0} = 0$ [7]. Alternatively, we fix both $v(x,y)$ and $w(x,y)$, and minimize $L(x,y)$ by solving the quadratic equation:

$$\begin{aligned} \min_L & \frac{\gamma\alpha}{2} \sum_{(x,y)} |B(x,y) \otimes L(x,y) - I(x,y) - v(x,y)|^2 \\ & + \frac{\mu}{2} |B(x,y) \otimes L(x,y) - I(x,y)|^2 \\ & + \frac{s\beta}{2} \sum_{(x,y)} |w_x(x,y) - \frac{\partial}{\partial x} L(x,y)|^2 \\ & + \frac{s\beta}{2} \sum_{(x,y)} |w_y(x,y) - \frac{\partial}{\partial y} L(x,y)|^2 \end{aligned}$$

$$+ (1-s) \sum_{(x,y)} \left(\left(\frac{\partial}{\partial x} L(x,y) \right)^2 + \left(\frac{\partial}{\partial y} L(x,y) \right)^2 \right), \quad (38)$$

which has a closed-form solution using the techniques proposed in [7] and [12]. Assuming that $L(x,y)$ is under the periodic boundary condition, we replace the time-consuming spatial convolution operator using the 2D discrete Fast Fourier Transforms by

$$L(x,y) = F^{-1} \left(\frac{s\beta F^*(\Delta)F(w) + (\mu + \gamma\alpha) F^*(B)F(I) + \gamma\alpha F^*(B)F(v)}{(s\beta + 2 - 2s)F^*(\Delta)F(\Delta) + (\mu + \gamma\alpha)F^*(B)F(B)} \right) \quad (39)$$

where $F(\cdot)$ represents the Fast Fourier Transforms, $*$ is the complex conjugacy, $F^*(\Delta)F(w) = F^*(\partial_x)F(w_x) + F^*(\partial_y)F(w_y)$, $F^*(\Delta)F(\Delta) = F^*(\partial_x)F(\partial_x) + F^*(\partial_y)F(\partial_y)$, and the multiplication and the division are both component-wise. Since the Fast Fourier Transforms have been developed numerous speed-up versions, upsampled images can thus be computed efficiently by our L1L2TTV image deblurring algorithm.

The alternating minimization problem is summarized as follows. Given $I(x,y)$ and $B(x,y)$ with a given quantized $\hat{\sigma}$, **Step 1:** Initialize $L(x,y) = I(x,y)$; **Step 2:** Iteratively compute $v(x,y)$ and $w(x,y)$ according to (35), (36), and (37) for fixed $L(x,y)$, and compute $L(x,y)$ according to (39) for fixed $v(x,y)$ and $w(x,y)$; **Step 3:** Until the minimization process converges, we obtain the final deblurred image $L_{\hat{\sigma}}(x,y)$ with the given $\hat{\sigma}$.

D. Pixel Selection

Once the blur map is estimated, a latent high-resolution image $\hat{L}(x,y)$ can be computed. Using blur map as the index, we deal with 9 different image blur levels $\hat{\sigma}$ and generate the final super-resolution image by selecting pixels from the deblurred interpolated image stack (9 images deblurred by different image blur levels $\hat{\sigma}$).

$$\hat{L}(x,y) = \{L_{\hat{\sigma}_{display}(x,y)}(x,y)\}, \quad (40)$$

where $\hat{\sigma}_{display}(x,y)$ is the largest quantized image blur level close to the blur map $\sigma(x,y)$ for reducing ringing artifacts [48]. Since we use local contrast to estimate the blur levels of the edge-directed interpolated $I(x,y)$ (which is equal to $\sigma_{display}(Z) = K_{Y,Z} \cdot \sigma(Z_0)$ as discussed in Section III-B), given $K_{Y,Z}$, d and the initial viewing distance Z_0 , we can remove the extra amount $\sigma_{display,de}(x,y)$ to maintain visual perceptual constancy by

$$\sigma_{display,de}(x,y) = \frac{\sqrt{(K_{Y,Z})^2 - \left(\frac{f_0}{f}\right)^2}}{K_{Y,Z}} \cdot \sigma(x,y). \quad (41)$$

V. EXPERIMENTAL RESULTS

To examine our proposed viewing-distance aware super-resolution, we design several experiments on both image quality and human mean-of-score across image scales. In this section, we first introduce the settings of our visual



Fig. 6. From left to right, input image, images upsampled by a factor of 2, 3 and 4, respectively. To view the original image clearly, one needs to be very close to the display. However, the upsampled images can be viewed clearly at different distance. This experiment can be carried out by selecting a viewing distance for the leftmost image and then move away from the display. The upsampled images by the proposed algorithm facilitates a viewer to have the same visual perception at different distance.

perceptual model between viewers and the displays. Since our super-resolution adopt image deconvolution operations, we propose a method to learn the weight parameters for our L1L2TTV deblurring. Later on, we show our quantitative and visual results across scales, and we also compare our super-resolution to the state-of-art methods. To ensure that our proposed viewing-distance aware super-resolution is feasible, we set up two subjective examinations and invite viewers to examine our super-resolution results on the HD displays. In the final, based on these experimental results, we analyze the limitation of the proposed super-resolution.

A. Experimental Settings

1) *Human Perceptual Model*: In this work, human eye is considered as a DSLR camera such that visual perception at varying distance is modeled. As long as we adjust the image blurs and image size on the display, we can perceive the same visual feeling in our eyes. Therefore, we set the image distance of this virtual camera d to 24 mini-meter which is similar to the distance from the crystalline lens to retina in human eyes. Although each viewer has his own best field-of-view and initial viewing distance Z_0 to see the original image, we set Z_0 to 120 mini-meter to generate our super-resolution results. According to our perceptual model, the initial focal length f_0 and the varying ratio $\frac{f}{f_0}$ to the corresponding viewing distance can be computed with these settings. As discussed in previous sections, the image size grows in proportion to the scaling factor and the viewing distance, i.e., $Y = K_{Y,Z} \cdot Y_0 = \frac{Z}{Z_0} \cdot Y_0$. On the other hand, we carefully adjust the image blur levels on the display according to the perceptual-tuning ratio $\frac{f_0}{f}$ to maintain perceptual constancy irrespective of different viewing distance.

To upsample color images, each one is first transformed to the $YCbCr$ color space. We only apply our super-resolution on the gray-scale because human perception is more sensitive to the intensity information, and the chromatic images in other two channels are only interpolated. To estimate image blur levels on HD displays, we compute local contrast within a window of 11×11 pixels to obtain the blur map for an edge-directed interpolated image. To handle image noise within each local window, we average the fifth highest/lowest pixels to obtain the maximum $I(x, y)$ and minimum $I(x, y)$. After computing the 2D gradients of the image, we obtain the blur

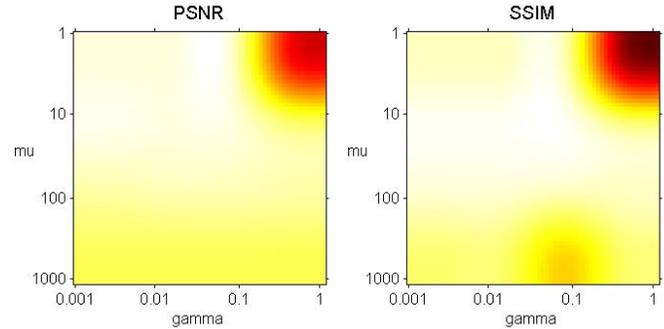


Fig. 7. Deblurring results with different γ and μ where γ is changed from 0.001 to 1, and μ is varied from 1 to 1000. Brighter pixels represent higher PSNR and SSIM values. We set $\alpha = 1$, γ to 0.5 and μ to 20 for L1L2TTV deblurring since the PSNR and SSIM values of the deblurred images are high.

map $\sigma(x, y)$. To process images with a set of discrete blur levels, we quantize $\sigma(x, y)$ with the step size of 0.5 pixels. The blur levels are increased linearly from 0.5 to 4.5 pixels. According to (19), (40) and (41), 9 corresponding $\hat{\sigma}$ values are used to generate deblurred interpolated image stack from which pixels are selected to form the resulting super-resolution image as shown in Fig. 4.

2) *Weight Parameters for L1L2TTV Deblurring*: Since there are a few parameters in the proposed L1L2TTV deblurring algorithm, we use a set of natural images to determine the weights of the reconstruction and regularization terms. These parameters are fixed for evaluation on a different dataset for super-resolution. As the total variation term tends to preserve the edges of resulting images, we set s in (33) to 0.99 to balance the total variation and Tikhonov regularization terms, and β is the update parameter which is increased by two in each iteration as suggested in [7]. In addition, according to (40), we can further set α to 1 and analyze the effects of μ and γ from numerous spatially-invariant deblurring experiments using peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) index with different image blur levels.

We examine the performance of our L1L2TTV deblurring method using images ('aerial', 'boat', 'bridge', 'couple', 'elaine' and 'man') from the USC SIPI database sipi.usc.edu/DATABASE/database.php. We blur each test image with a given image blur level (i.e., $\sigma = 2, 4, 6, \text{ and } 8$), and then deblur each blurred image to com-

TABLE I
PSNR AND SSIM VALUES USING THE USC SIPI DATASET WITH GAUSSIAN BLUR KERNEL

Factor	Bicubic	Shan [9]	Glasner [36]	Freedman [37]	Kim [38]	Yang [32]	Mallat [27]	Ours (s=0.8)	Ours (s=0.99)	ICBI
2	25.26	25.87	25.64	25.52	25.46	25.43	25.16	26.10	26.26	24.61
3	23.50	23.71	23.97	23.78	23.66	23.69	23.29	24.64	24.77	23.20
4	22.41	22.56	22.80	22.75	22.57	22.60	22.34	23.76	23.90	22.26
2	0.847	0.852	0.858	0.852	0.844	0.847	0.836	0.897	0.905	0.813
3	0.731	0.731	0.759	0.747	0.736	0.740	0.716	0.820	0.832	0.714
4	0.650	0.647	0.672	0.669	0.654	0.658	0.642	0.756	0.767	0.639

pute the overall PSNR and SSIM values across scales and test images. Since $s = 0.99$, we analyze the response of μ and γ at different discrete values with high PSNR and SSIM values (See Fig. 7), and empirically determine the values, i.e., $\alpha = 1$, $\gamma = 0.5$, and $\mu = 20$.

B. Quantitative and Visual Results on Super-Resolution

1) *Quantitive Results on Simulated Blurry Images:* To examine these weight parameters for our viewer-display perceptual model are set properly or not, we evaluate the deblurred results using the synthesized blurry images from the USC SIPI dataset. We adopt ‘aerial’, ‘boat’, ‘bridge’, ‘couple’, ‘elaine’, and ‘man’ as the high-resolution images, and then we blur each high-resolution image by a Gaussian blurring function with different kernel width (i.e., $\sigma = 2, 3, 4$) and downsample the blurred images by a scaling factor of 2, 3, and 4 to generate low-resolution images, respectively. We reconstruct the spatially-invariant super-resolution images with comparisons to those generated by the state-of-the-art approaches [9], [27], [32], [36]–[38]. (To upsample an image to 3-times larger, we adopt ICBI to upsample to 4-time larger and then downsample since ICBI can only deal with the power of two.) Table I shows that our spatially-invariant super-resolution algorithm achieves higher PSNR and SSIM values in the Gaussian blurring kernel cases which meet the perceptual model setting. Since the proposed model is viewing-distance aware, we mainly consider the Gaussian blurring kernels which connect the relationship between the viewer and the display. Our L1L2TTV deblurring is also designed for blur adjustment. Since the viewers move backwards to see the enlarged image, we mainly deal with large kernel widths (i.e., 32×32). Therefore, our proposed L1L2TTV deblurring cannot deal with bicubic kernel well due to the limited kernel size (i.e., 4×4) and also the negative weights of the bicubic kernel (Table II). The limited 4×4 bicubic kernel is not proper to the viewing system model while the negative weights are unfit the positive pixel value assumption and the optical phenomena in physics. In other word, different to the conventional super-resolution methods which mostly pursue image quality, our super-resolution is designed to maintain viewer’s visual experience across scales. Therefore, we deal with large kernel width and enhance edges and textures by deblurring. However, these operations may degrade the image quality, but maintain visual feeling and preserve perceptual quality.

We further evaluate the proposed super-resolution algorithm quantitatively using twenty four images from the well-known dataset [36] (available at www.wisdom.

weizmann.ac.il/~vision/SingleImageSR.html). As the PSNR or SSIM metric is not effective for assessing image quality of super-resolution results for visual perception [49], [50], and it is difficult to deal with the spatially-varying cases and the alignment problems of enlarged images, we adopt a blind spatial image quality estimator (BRISQUE) [51] to judge the natural image quality. BRISQUE uses scene statistics of locally normalized luminance coefficients to quantify the naturalness loss. High scores by BRISQUE suggest large amount of distortion in images. In other words, lower BRISQUE score refers to better image quality. As we can see in Table III that our proposed methods can generate better natural image quality for viewers according to the scales given in the dataset.

2) *Visual Results Across Scales:* We shows super-resolution images generated by the proposed algorithm with different scaling factors from [36] for readers to examine our viewing-distance aware super-resolution. As human perception is considered in the proposed algorithm, these super-resolution images can be viewed with similar visual perception at the corresponding viewing distance. While one can clearly see the original image at a distance very close to the display, one can view the rightmost upsampled image clearly at a distance four times away. Furthermore, Fig.6 and Figs. 8-10 show the super-resolution images by the proposed algorithm and the state-of-the-art methods.

The super-resolution images by the proposed algorithm can be perceived similarly at different scaling factors and the corresponding viewing distance whereas some details (e.g., beard, eye corners and wrinkles) are missing in the results generated by the other two methods. In addition, it is also clear in Fig. 9 that visually pleasing textures are reconstructed by the proposed algorithm while mid-frequency textures are missing by the other methods. The proposed method generates sharper results that are similar to the original image observed at the best viewer distance. In Fig. 10, the super-resolution results generated by the proposed algorithm are sharper than those by the other methods especially in the background and head regions.

Generally, compared to the spatially-invariant super-resolution algorithm (which includes a high-quality image deconvolution [52]), the proposed super-resolution algorithm can spatially-varying adjust the scale/blur relationships to maintain the same visual feeling. Since the learning-based methods used to adopt the iterative backward projection with only a given blurring function with a single blur level, the background regions remain blurry. In addition, the learning-based methods used to look for the similar patches and then



Fig. 8. From left to right, original image (*old man*), images enlarged by a factor of 2, 3 and 4, respectively. To observe the original image clearly, one needs to be close to the display. However, the enlarged images can be viewed clearly at different distance. This experiment can be carried out by selecting a viewing distance for the leftmost image and then move away from the display. The enlarged images by the proposed algorithm facilitates a viewer to have the same visual perception at different distance. In addition, we show regions of images upsampled by a factor of 4. (a) Freedman [37]. (b) Kim [38]. (c) Proposed. (d) Bicubic. (e) Shan [9]. (f) Glasner [36]. (g) Freedman [37]. (h) Kim [38]. (i) Yang [32]. (j) Mallat [27]. (k) Proposed.

stitch them together so as to generate the blocking artifacts as shown in Yang et al.'s result.

C. Subjective Examinations

To evaluate how the proposed algorithm preserves perceptual constancy, we invite 15 persons and carry out human

subject studies using the dataset from [36] with scaling factors of 2, 3, and 4, respectively. (Inviting 15 subjects for a subject experiment meets the requirement of Recommendation ITU-R BT.500-13.)

As it is time-consuming to examine all the images against all the methods, the subjects are asked to compare the results



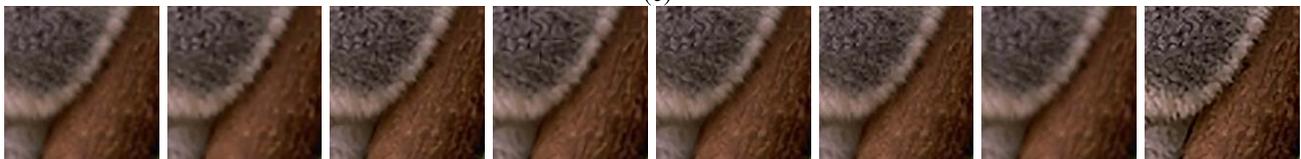
(a)



(b)



(c)



(d) (e) (f) (g) (h) (i) (j) (k)

Fig. 9. From left to right, original image (*koala*), images enlarged by a factor of 2, 3 and 4, respectively. We also show the regions of images enlarged by a factor of 3. (a) Freedman [37]. (b) Kim [38]. (c) Proposed. (d) Bicubic. (e) Shan [9]. (f) Glasner [36]. (g) Freedman [37]. (h) Kim [38]. (i) Yang [32]. (j) Mallat [27]. (k) Proposed.

generated by the proposed algorithm and the representative exemplar-based method proposed by Yang et al. [32], and the representative reconstruction-based method proposed by

Mallat and Yu [27]. In addition, subjects evaluate the results with scores ranging from 1 to 5. Score 1 represents the lowest perceptual constancy between the original low-resolution

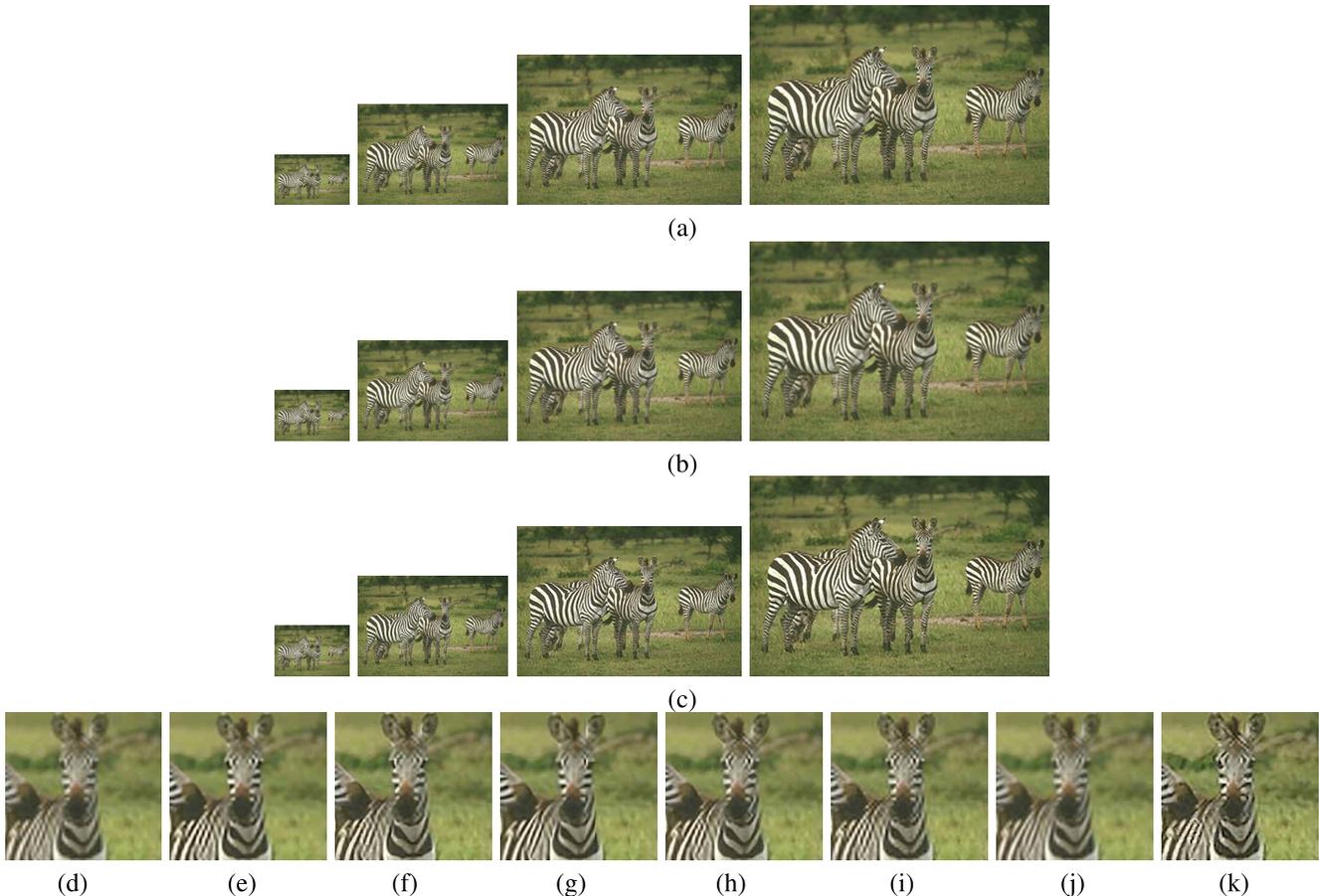


Fig. 10. From left to right, original image (*zebra*), images enlarged by a factor of 2, 3 and 4, respectively. The enlarged images by the proposed algorithm facilitates a viewer to have the same visual perception at different distance. In addition, we show regions of images enlarged by a factor of 3. (a) Yang et al. [32]. (b) Mallat and Yu [27]. (c) Proposed. (d) Bicubic. (e) Shan [9]. (f) Glasner [36]. (g) Freedman [37]. (h) Kim [38]. (i) Yang [32]. (j) Mallat [27]. (k) Proposed.

TABLE II
PSNR AND SSIM VALUES USING THE USC SIPI DATASET WITH BICUBIC BLUR KERNEL

Factor	Bicubic	Shan [9]	Glasner [36]	Freedman [37]	Kim [38]	Yang [32]	Mallat [27]	Ours (s=0.8)	Ours (s=0.99)	ICBI
2	29.56	28.72	28.80	29.47	30.87	30.77	28.73	24.57	24.52	28.41
3	25.67	27.67	27.55	24.48	28.07	25.76	25.73	24.18	24.17	25.82
4	25.58	26.14	26.24	25.47	26.48	26.06	24.60	21.96	21.95	24.72
2	0.984	0.969	0.974	0.981	0.994	0.995	0.971	0.852	0.852	0.965
3	0.882	0.923	0.940	0.857	0.942	0.895	0.888	0.830	0.830	0.891
4	0.857	0.867	0.894	0.870	0.889	0.886	0.826	0.691	0.691	0.830

image and the super-resolution image while Score 5 represents the highest perceptual constancy between the original low-resolution image and the super-resolution image. Score 5 also means that the super-resolution image at the corresponding viewing distance Z can give the exact similar feeling to the viewer as seeing the original low-resolution image at Z_0 .

In each run of our conducted experiments, we pick up one image from the dataset [36], and then we fix one scale to subjectively evaluate the enlarged images. The images enlarged by bicubic interpolation, Yang et al.'s exemplar-based algorithm [32], Mallat and Yu's inverse estimators [27], and our viewing-distance aware method are shown by random order. We set up two subjective experiments. One is to deal

with only one display to simulate the digital home theater environments. In this case, subjects have to see the original image, move backwards to the corresponding viewing distance to see the enlarged image, and evaluate the perceptual constancy among the comparisons. Another one is to deal with two displays simultaneously for side-by-side examination. Subjects have to see the original image on one display and also see the enlarged image on another display placed at the corresponding viewing distance. The viewing distance is ensured by Bosch GLM 7000 laser rangefinder.

We carry out two subject experiments to evaluate the super-resolution images. Moreover, we also adopt the blind image spatial quality evaluator (BRISQUE) to evaluate the quality across image scales.

TABLE III
BRISQUE EVALUATION

Image	Factor	Bicubic	Shan [9]	Glasner [36]	Freedman [37]	Kim [38]	Yang [32]	Mallat [27]	Ours (s=0.8)	Ours (s=0.99)
8143	3	48.136	54.410	43.541	46.045	50.227	51.276	48.557	45.046	39.012
23084	3	51.965	50.105	46.278	46.156	41.331	52.7353	44.995	42.915	36.627
35070	3	69.118	72.863	56.901	57.633	63.340	54.124	59.999	48.850	39.034
56028	3	48.842	50.934	48.916	49.748	50.613	55.021	48.191	46.746	45.843
69040	3	55.513	56.879	49.606	52.712	50.160	57.058	53.898	46.545	42.229
78004	3	57.733	57.275	53.436	57.572	52.062	56.923	54.588	49.317	46.162
86000	3	30.073	21.622	19.696	19.232	20.578	37.656	27.444	28.105	22.792
103041	3	55.036	53.973	44.358	50.090	41.510	47.085	54.564	46.144	39.596
198054	3	49.786	51.102	42.287	47.325	45.486	55.550	48.794	46.891	38.592
253027	3	47.733	39.876	43.317	41.482	42.061	58.782	40.600	43.148	39.326
auditorium	3	49.806	47.977	39.734	43.227	41.232	54.083	46.890	43.084	37.831
baby	3	43.932	41.806	28.273	32.449	28.928	36.497	38.966	30.723	22.930
butterfly	3	46.960	39.970	22.306	26.973	23.801	38.939	42.613	35.454	25.251
child	4	53.492	53.500	45.625	49.276	41.954	45.084	57.410	45.483	38.473
chip	4	66.252	77.384	62.755	64.314	59.622	56.025	64.497	55.100	46.905
colorblind	3	50.267	62.830	39.266	51.403	57.473	50.289	53.097	47.056	42.455
flowerfield	3	50.055	48.058	39.402	42.605	35.838	45.188	48.873	42.175	35.089
freckles2	3	49.491	46.224	40.129	39.624	33.675	43.235	45.959	41.995	36.259
kitchen	3	49.419	48.098	49.076	47.927	43.196	52.547	48.871	44.936	39.107
kitchen2	3	53.327	57.401	58.157	55.740	54.902	59.506	48.771	47.032	43.357
koala	3	57.187	56.310	50.552	54.021	46.904	54.575	56.433	53.110	47.289
oldman	4	58.062	57.001	51.457	53.854	45.733	49.340	61.119	50.303	43.298
sculpture	4	63.826	64.105	52.505	60.175	53.898	50.620	66.079	52.449	43.619
street3	3	47.282	51.087	48.658	47.103	47.637	58.437	44.389	50.292	45.871
Avg.		52.221	52.533	44.843	45.365	42.874	50.857	50.233	45.121	39.040

TABLE IV
SUBJECT EVALUATION (SETUP 1)

Factor	Bicubic	Yang [32]	Mallat [27]	Ours (s=0.8)	Ours (s=0.99)
2	2.403	3.708	2.544	3.283	3.797
3	2.072	3.236	2.217	2.994	3.503
4	1.947	2.944	2.083	2.875	3.344

1) *Examination Setup 1*: The first subject experiment is to evaluate super-resolution images on one HD display at different viewing distance to simulate the digital home theater environments (Fig. 2). We use the ASUSTek VS278Q LED display panel with a digital HDMI connection and an Intel HD graphics card to render the super-resolution images. The subject first best view the original image on the display. We record the viewing distance Z_0 and ask the subject viewer to move backwards to the corresponding viewing distances (i.e., $2Z_0$, $3Z_0$ and $4Z_0$). Later on, we randomly show the enlarged images proposed by bicubic interpolation, Yang et al.’s exemplar-based result [32], Mallat and Yu’s reconstructed result [27], and our super-resolution result. To each test image, subject viewers report the perceptual opinion score to compare the enlarged image viewed at the corresponding viewing distance with the original image viewed at Z_0 . We analyze the opinion scores according to three scaling factors and five super-resolution scheme. As we can observe in Table IV, human subjects prefer the super-resolution images generated by the proposed algorithm among different scaling factors under this setup 1. The subjects also point out the fact that bicubic interpolation and sparse mixing estimators may generate blurry results. Therefore, according to subjective examination, we can also know that human perception tends to notice sharp edges and rich textures. Our proposed super-resolution with L1L2TTV deblurring is suitable to fit human visual perception.

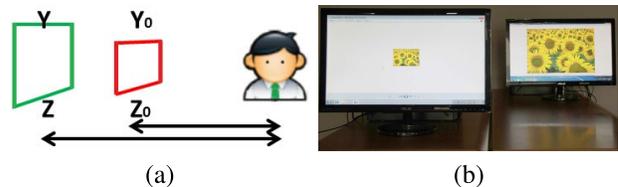


Fig. 11. (a) The second setup to examine the relationship between viewing distance and enlarged image on the display. (b) The testing environment. The viewer can see the two displays simultaneously. The display near to the viewer shows the original image while the display far to the viewer shows the enlarged images according to the viewing distances.

TABLE V
SUBJECT EVALUATION (SETUP 2)

Factor	Bicubic	Yang [32]	Mallat [27]	Ours (s=0.8)	Ours (s=0.99)
2	2.628	3.900	2.867	3.689	3.947
3	2.481	3.683	2.642	3.711	4.042
4	2.444	2.606	2.531	3.806	4.039

2) *Examination Setup 2*: In the second subject experiment, we display the images on two HD display for viewer to examine perceptual constancy at a fixed viewing distance. The display close to the viewer shows the original low-resolution image while the other display at a distance shows the upscaled images according to the scaling factors (Fig. 11). We use two ASUSTek VS278Q displays with digital HDMI connections, and two NVIDIA Geforce GTX650 graphics cards to render both original image and the super-resolution images at different distances, respectively. Table V shows that human subjects also prefer the super-resolution images generated by the proposed algorithm at different scaling factors under this setup 2. We observe that the visual quality degrades with the scale when applying Yang et al.’s exemplar-based method [32] or Mallat and Yu’s

TABLE VI
BRISQUE EVALUATION

Factor	Bicubic	Yang [32]	Mallat [27]	Ours ($s=0.8$)	Ours ($s=0.99$)
2	38.05	36.88	41.07	36.49	30.52
3	51.55	50.49	48.58	43.90	37.79
4	59.87	50.33	59.84	50.10	42.57

reconstruction-based method [27]. Since our proposed method can deal with spatially-varying blurry condition and adjust the scale/blur relationship, our proposed method can give the similar visual quality for viewers to maintain perceptual constancy. In addition, these results also show that subjects prefer strong edges and clear object boundaries (i.e., s is set to 0.99).

3) *Numerical Examination on BRISQUE*: As we can observe in Table VI that our super-resolution results can give better natural quality at each scale. It is also reported that enlarging an image may degrade the natural image quality. However, our proposed super-resolution can still preserve more natural feature such as sharp edges and rich textures so as to maintain better natural quality for viewers to see on the HD display.

D. Limitation

One main limitation of the proposed algorithm is that the largest size of an enlarged image is constrained to the dimension of the display (Fig. 2). In such cases, the proposed algorithm can only improve image sharpness as the scaling factor is constrained to the maximal value of the display. Another limitation is the scenario when the viewing distance and the screen size are both fixed and non-movable (eg. the seats in the digital cinema). To the cinema cases, our proposed algorithm super-resolute the image to the highest resolution on the screen and report the comfort zone in terms of the seats which can give the viewers to maintain perceptual constancy. It is also true that our proposed super-resolution may also enhance the background. To solve this problem, we can adopt the background extraction and only enhance the objects in the foreground. Furthermore, we can also adopt the content analyses with the map technique which is widely used in computer graphics and the industry.

VI. CONCLUDING REMARKS

We propose a super-resolution algorithm based on human perception by considering the relationship between image size, scaling factor and viewing distance as shown in Fig. 2 and Fig. 3. Based on the proposed model, we estimate the corresponding blur levels of upsampled images at different viewing distance. To preserve perceptual constancy, we propose the L1L2TTV deblurring algorithm to adjust the spatially-varying image blur while maintain the perceptual blur radius in human eyes. Experimental results show that the proposed algorithm performs favorably against the state-of-the-art methods qualitatively, quantitatively, and subjectively at different viewing distance and scaling factors.

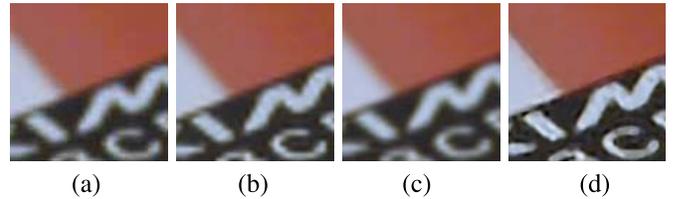


Fig. 12. Failure examples. The scaling factor is 4. (a) Bicubic. (b) Yang [32]. (c) Mallat [27]. (d) Proposed.

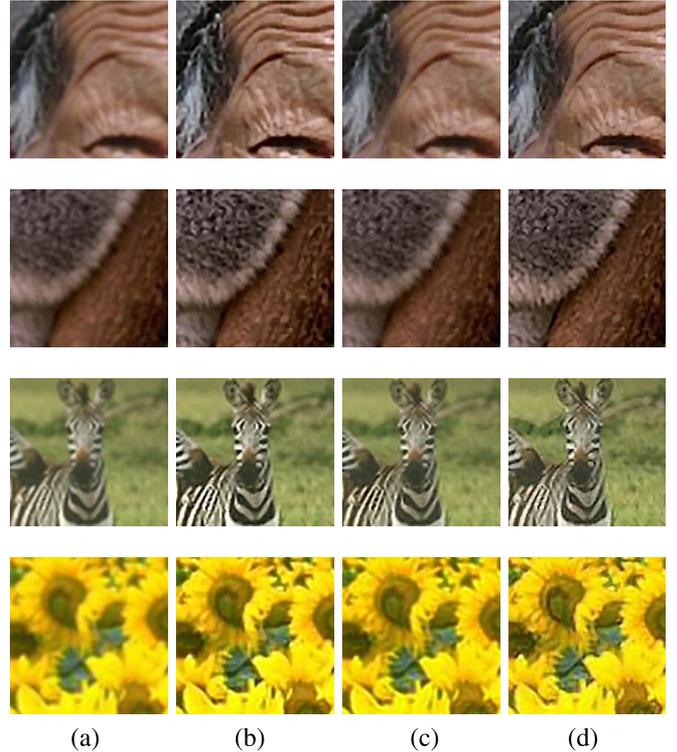


Fig. 13. Non-uniform deblurring interpolated images. The regions of image *flowfield* are enlarged by a factor of 3. (a) ICBI [23]. (b) Krishnan [56]. (c) Whyte [55]. (d) Proposed.

In the future, we plan to extend the proposed algorithm to upsample images for multiple viewers in the cinemas as well as video super-resolution at different viewing distance. In addition, we also plan to integrate our viewing-distance aware super-resolution with the depth estimation techniques on UHDTV. With the knowledge of viewer's distance, our algorithm can auto-tune the parameters and provide pleasant visual experiences. Furthermore, we plan to combine the proposed super-resolution with the novel kernel regression techniques [53], [54] to deal with arbitrary blurring functions.

APPENDIX

A. Some Failure Cases

We also show the failure cases in Fig. 12. Since our super-resolution currently adopt Gaussian blurring function to connect viewer's perception and the image blur levels on the HD displays, we may have some artifacts around edges. Although these artifacts can compensate in part by human lens, we still need more precise blurring kernels (arbitrary-shaped PSFs) to improve image quality.

B. Compared to Other Spatially-Varying Deconvolution

We compare the proposed viewing-distance aware super-resolution with the other integrations of interpolation and spatially-varying (non-uniform) deconvolution. The input degraded images are first interpolated to the desired size by the ICBI method [23]. Then, we apply the state-of-the-art blind deconvolution on these interpolated images. Among numerous spatially-varying deconvolution methods, we adopt two most representative deconvolution methods proposed by Whyte *et al.* [55] and Krishnan *et al.* [56], respectively.

Whyte *et al.* propose a parametrized geometric blurring model in terms of the rotation velocity of the camera during exposure time. With their estimated rotational kernel, they can reconstruct the final latent images by the modified Richardson-Lucy algorithm proposed by Yuan *et al.* [57]. Furthermore, Krishnan *et al.* design an image regularization to ensure the scale-invariant sparsity and also extend it to deal with 3D rotation camera shakes. Fig. 13 shows that our spatially-varying deconvolution scheme generates sharper results as compared to Whyte *et al.*'s approach. Although Krishnan *et al.*'s deconvolution can always generate high quality results, our scheme can reconstruct more details in the background regions of the test images *zebra* and *flowerfield*. Different from these two methods that reconstruct the all-in-focus images, we do not remove the entire blur condition. Instead, we adjust the blur levels inside an image and reconstruct it to meet the perceptual constancy according to the viewing distance.

REFERENCES

- [1] M. A. Webster, M. A. Georgeson, and S. M. Webster, "Neural adjustments to image blur," *Natural Neurosci.*, vol. 5, no. 9, pp. 839–840, 2002.
- [2] M. Trentacoste, R. Mantiuk, and W. Heidrich, "Blur-aware image downsampling," *Comput. Graph. Forum*, vol. 30, no. 2, pp. 573–582, 2011.
- [3] V. F. Pamplona, M. M. Oliveira, D. G. Aliaga, and R. Raskar, "Tailored displays to compensate for visual aberrations," *ACM Trans. Graph.*, vol. 31, no. 4, 2012, Art. ID 81.
- [4] J. Kopf, A. Shamir, and P. Peers, "Content-adaptive image downscaling," *ACM Trans. Graph.*, vol. 32, no. 6, 2013, Art. ID 173.
- [5] P. Milanfar, *Super-Resolution Imaging*. Boca Raton, FL, USA: CRC Press, 2011.
- [6] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super resolution," *IEEE Trans. Image Process.*, vol. 13, no. 10, pp. 1327–1344, Oct. 2004.
- [7] Y. Wang, W. Yin, J. Yang, and Y. Zhan, "A new alternating minimization algorithm for total variation image reconstruction," *SIAM J. Imag. Sci.*, vol. 1, no. 3, pp. 248–272, 2008.
- [8] W. Yin, S. Osher, D. Goldfarb, and J. Darbon, "Bregman iterative algorithms for ℓ_1 -minimization with applications to compressed sensing," *SIAM J. Imag. Sci.*, vol. 1, no. 1, pp. 143–168, 2008.
- [9] Q. Shan, Z. Li, J. Jia, and C.-K. Tang, "Fast image/video upsampling," *ACM Trans. Graph.*, vol. 27, no. 5, 2008, Art. ID 153.
- [10] Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," *ACM Trans. Graph.*, vol. 27, no. 3, 2008, Art. ID 73.
- [11] M. Zibulevsky and M. Elad, "L1-L2 optimization in signal and image processing," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 76–88, May 2010.
- [12] L. Xu and J. Jia, "Two-phase kernel estimation for robust motion deblurring," in *Proc. 11th Eur. Conf. Comput. Vis.*, 2010, pp. 157–170.
- [13] Y. Dong, M. Hintermüller, and M. Rincón-Camacho, "A multi-scale vectorial L^T -TV framework for color image restoration," *Int. J. Comput. Vis.*, vol. 92, no. 3, pp. 296–307, 2011.
- [14] S. H. Chan, R. Khoshabeh, K. B. Gibson, P. E. Gill, and T. Q. Nguyen, "An augmented Lagrangian method for total variation video restoration," *IEEE Trans. Image Process.*, vol. 20, no. 11, pp. 3097–3111, Nov. 2011.
- [15] R.-Y. Tsai and T.-S. Huang, "Multiframe image restoration and registration," *Adv. Comput. Vis. Image Process.*, vol. 1, no. 2, pp. 317–339, 1984.
- [16] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 21–36, May 2003.
- [17] D. Capel and A. Zisserman, "Computer vision applied to super resolution," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 75–86, May 2003.
- [18] B. Basclé, A. Blake, and A. Zisserman, "Motion deblurring and super-resolution from an image sequence," in *Proc. 4th Eur. Conf. Comput. Vis.*, 1996, pp. 571–582.
- [19] M.-S. Lee, M.-Y. Shen, C.-C. J. Kuo, and A. Yoneyama, "Techniques for flexible image/video resolution conversion with heterogeneous terminals," *IEEE Commun. Mag.*, vol. 45, no. 1, pp. 61–67, Jan. 2007.
- [20] C. Liu and D. Sun, "On Bayesian adaptive video super resolution," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 2, pp. 346–360, Feb. 2014.
- [21] C. Lee, M. Eden, and M. Unser, "High-quality image resizing using oblique projection operators," *IEEE Trans. Image Process.*, vol. 7, no. 5, pp. 679–692, May 1998.
- [22] X. Li and M. T. Orchard, "New edge-directed interpolation," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1521–1527, Oct. 2001.
- [23] A. Giachetti and N. Asuni, "Real-time artifact-free image upscaling," *IEEE Trans. Image Process.*, vol. 20, no. 10, pp. 2760–2768, Oct. 2011.
- [24] K.-W. Hung and W.-C. Siu, "Robust soft-decision interpolation using weighted least squares," *IEEE Trans. Image Process.*, vol. 21, no. 3, pp. 1061–1069, Mar. 2012.
- [25] R. Fattal, "Image upsampling via imposed edge statistics," *ACM Trans. Graph.*, vol. 26, no. 3, 2007, Art. ID 95.
- [26] J. Sun, J. Sun, Z. Xu, and H.-Y. Shum, "Image super-resolution using gradient profile prior," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.
- [27] S. Mallat and G. Yu, "Super-resolution with sparse mixing estimators," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2889–2900, Nov. 2010.
- [28] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael, "Learning low-level vision," *Int. J. Comput. Vis.*, vol. 40, no. 1, pp. 25–47, 2000.
- [29] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," *IEEE Comput. Graph. Appl.*, vol. 22, no. 2, pp. 56–65, Mar./Apr. 2002.
- [30] 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.
- [31] Y. HaCohen, R. Fattal, and D. Lischinski, "Image upsampling via texture hallucination," in *Proc. IEEE Int. Conf. Comput. Photograph.*, Mar. 2010, pp. 1–8.
- [32] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010.
- [33] K. Zhang, X. Gao, D. Tao, and X. Li, "Multi-scale dictionary for single image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 1114–1121.
- [34] S. Wang, D. Zhang, Y. Liang, and Q. Pan, "Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch synthesis," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 2216–2223.
- [35] W. Dong, L. Zhang, G. Shi, and X. Li, "Nonlocally centralized sparse representation for image restoration," *IEEE Trans. Image Process.*, vol. 22, no. 4, pp. 1620–1630, Apr. 2013.
- [36] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in *Proc. Int. Conf. Comput. Vis.*, 2009, pp. 349–356.
- [37] G. Freedman and R. Fattal, "Image and video upscaling from local self-examples," *ACM Trans. Graph.*, vol. 30, no. 2, pp. 1–10, 2011.
- [38] K. I. Kim and Y. Kwon, "Single-image super-resolution using sparse regression and natural image prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 6, pp. 1127–1133, Jun. 2010.
- [39] P. Green, W. Sun, W. Matusik, and F. Durand, "Multi-aperture photography," *ACM Trans. Graph.*, vol. 26, no. 3, 2007, Art. ID 68.
- [40] D. A. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*. Upper Saddle River, NJ, USA: Prentice-Hall, 2011.
- [41] C.-T. Shen, H.-H. Liu, M.-S. Lee, Y.-P. Hung, and S.-C. Pei, "Spatially-varying super-resolution for HDTV," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2013, pp. 1195–1198.
- [42] S. Bae and F. Durand, "Defocus magnification," *Comput. Graph. Forum*, vol. 26, no. 3, pp. 571–579, 2007.
- [43] N. Joshi, R. Szeliski, and D. Kriegman, "PSF estimation using sharp edge prediction," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.

- [44] Y.-W. Tai and M. S. Brown, "Single image defocus map estimation using local contrast prior," in *Proc. 16th IEEE Int. Conf. Image Process.*, Nov. 2009, pp. 1797–1800.
- [45] S. Zhuo and T. Sim, "Defocus map estimation from a single image," *Pattern Recognit.*, vol. 44, no. 9, pp. 1852–1858, 2011.
- [46] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 6, pp. 1397–1409, Jun. 2013.
- [47] C.-T. Shen, F.-J. Chang, Y.-P. Hung, and S.-C. Pei, "Edge-preserving image decomposition using L1 fidelity with L0 gradient," in *Proc. SIGGRAPH Asia Tech. Briefs*, 2012, Art. ID 6.
- [48] T. F. Chan and C.-K. Wong, "Total variation blind deconvolution," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 370–375, Mar. 1998.
- [49] T. S. Cho, C. L. Zitnick, N. Joshi, S. B. Kang, R. Szeliski, and W. T. Freeman, "Image restoration by matching gradient distributions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 4, pp. 683–694, Apr. 2012.
- [50] Y.-N. Liu, Y.-C. Lin, Y.-L. Huang, and S.-Y. Chien, "Algorithm and architecture design of high-quality video upscaling using database-free texture synthesis," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 7, pp. 1221–1234, Jul. 2014.
- [51] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4695–4708, Dec. 2012.
- [52] Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," *ACM Trans. Graph.*, vol. 27, no. 3, 2008, Art. ID 73.
- [53] H. Takeda, S. Farsiu, and P. Milanfar, "Deblurring using regularized locally adaptive kernel regression," *IEEE Trans. Image Process.*, vol. 17, no. 4, pp. 550–563, Apr. 2008.
- [54] H. Zhang, J. Yang, Y. Zhang, and T. S. Huang, "Image and video restorations via nonlocal kernel regression," *IEEE Trans. Cybern.*, vol. 43, no. 3, pp. 1035–1046, Jun. 2013.
- [55] O. Whyte, J. Sivic, A. Zisserman, and J. Ponce, "Non-uniform deblurring for shaken images," *Int. J. Comput. Vis.*, vol. 98, no. 2, pp. 168–186, 2012.
- [56] D. Krishnan, T. Tay, and R. Fergus, "Blind deconvolution using a normalized sparsity measure," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2011, pp. 233–240.
- [57] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum, "Image deblurring with blurred/noisy image pairs," *ACM Trans. Graph.*, vol. 26, no. 3, 2007, Art. ID 1.



photography, optimization, inverse problems, and multimedia stream processing.

Chih-Tsung Shen received the B.S. and M.S. degrees in electrical engineering from National Taiwan University, Taipei, Taiwan. He was a Research Assistant with the Institute of Information Science, Academia Sinica, Taipei, from 2008 to 2012, and an Algorithm Engineer with Himax Technologies, Inc., Tainan, Taiwan, from 2012 to 2014. He is currently pursuing the Ph.D. degree with the Graduate Institute of Networking and Multimedia, National Taiwan University.

His research interests include computational

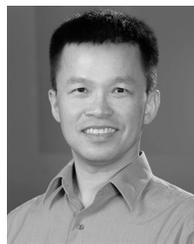


Data Communication Business Group, Chunghwa Telecom Company, Ltd. His research interests include 3DTV and digital rights management.

Hung-Hsun Liu received the B.S. and M.S. degrees in computer science from National Tsing Hua University, Hsinchu, Taiwan, in 2002 and 2004, respectively. He is currently pursuing the Ph.D. degree with National Taiwan University, Taipei, Taiwan. He was a Researcher with Telecommunication Laboratories, Chunghwa Telecom Company, Ltd., Taoyuan, Taipei, and a Visiting Scholar with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA. He is also a Software Engineer with the



Ming-Hsuan Yang (SM'06) received the Ph.D. degree in computer science from the University of Illinois at Urbana-Champaign, Champaign, IL, USA, in 2000. He is currently an Associate Professor of Electrical Engineering and Computer Science with the University of California at Merced (UC Merced), Merced CA, USA. He serves as the Program Chair of the 2014 Asian Conference on Computer Vision, and the Area Chair of the IEEE International Conference on Computer Vision, the IEEE Conference on Computer Vision and Pattern Recognition, the European Conference on Computer Vision, and the Asian Conference on Computer. He served as an Associate Editor of the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE from 2007 to 2011, and is currently an Associate Editor of the *International Journal of Computer Vision, Image and Vision Computing*, and the *Journal of Artificial Intelligence Research*. He was a recipient of the NSF CAREER Award in 2012, the Senate Award for Distinguished Early Career Research at UC Merced in 2011, and the Google Faculty Award in 2009. He is a Senior Member of the Association for Computing Machinery. His research interests include computer vision, pattern recognition, artificial intelligence, robotics, and machine learning.



Yi-Ping Hung received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, and the master's and Ph.D. degrees from Brown University, Providence, RI, USA. He is currently a Professor with the Graduate Institute of Networking and Multimedia and the Department of Computer Science and Information Engineering, National Taiwan University. Before 2002, he was with the Institute of Information Science, Academia Sinica, Taipei, where he received the Young Researcher Publication Award from Academia Sinica. He has served as the Deputy Director of the Institute of Information Science with Academia Sinica, and is the Director of the Graduate Institute of Networking and Multimedia with National Taiwan University. He was the Program Co-Chair of ACCV00 and ICAT00, and the Workshop Co-Chair of ICCV03. He has served in the Editorial Board of the *International Journal of Computer Vision* since 2004. He was a keynote speaker for PSIVT07, DeSForM09, and ACE10. His current research interests include computer vision, pattern recognition, image processing, virtual reality, multimedia, and human-computer interaction.



Soo-Chang Pei (SM'89–F'00–LF'15) received the B.S. degree from National Taiwan University, Taipei, Taiwan, in 1970, and the M.S. and Ph.D. degrees from the University of California at Santa Barbara, Santa Barbara, CA, USA, in 1972 and 1975, respectively, all in electrical engineering.

He was an Engineering Officer with Chinese Navy Shipyard, Varna, China, from 1970 to 1971, and a Research Assistant with the University of California at Santa Barbara, from 1971 to 1975. He was a Professor and the Chairman of the Department

of Electrical Engineering with the Tatung Institute of Technology, Taipei, and National Taiwan University, from 1981 to 1983 and 1995 to 1998, respectively. He is currently a Professor with the Department of Electrical Engineering, National Taiwan University. His research interests include digital signal processing, image processing, optical information processing, and laser holography.

Dr. Pei was a recipient of the National Sun Yet-sen Academic Achievement Award in Engineering in 1984, the Distinguished Research Award from the National Science Council from 1990 to 1998, the Outstanding Electrical Engineering Professor Award from the Chinese Institute of Electrical Engineering in 1998, the Academic Achievement Award in Engineering from the Ministry of Education in 1998, the Pan Wen-Yuan Distinguished Research Award in 2002, and the National Chair Professor Award from the Ministry of Education in 2002 and 2008. He was the President of the Chinese Image Processing and Pattern Recognition Society in Taiwan from 1996 to 1998.

Dr. Pei is a member of Eta Kappa Nu and the Optical Society of America.