Superresolution in color videos acquired through turbulent media

B. Fishbain,[1,2,*] I. A. Ideses,[1] G. Shabat,[1] B. G. Salomon,[1] and L. P. Yaroslavsky[1]

[1]*Department of Physical Electronics, Faculty of Engineering, Tel Aviv University, Tel Aviv 69978, Israel*
[2]*Department of Industrial Engineering and Operational Research, University of California, Berkeley, Berkeley, California, 94720-177, USA*
**Corresponding author: barak@berkeley.edu*

Received July 24, 2008; revised January 14, 2009; accepted January 16, 2009; posted January 26, 2009 (Doc. ID 99387); published February 23, 2009

Color videos acquired with a single CCD through turbulent media can be enhanced in their resolution beyond the limit defined by the image sampling rate. We provide a mathematical justification for this claim, present an efficient superresolution algorithm and its experimental verification on a real-life video, and finally, discuss its potentials and limitations. © 2009 Optical Society of America

OCIS codes: 100.2000, 100.3020, 100.6640, 110.7050.

Videos acquired through turbulent media, or, for brevity, turbulent videos suffer from local image instabilities due to random fluctuations of the refraction index of the medium. In [1–3] the idea of making a profit from atmospheric turbulence-induced image geometrical spatial–temporal degradations to compensate image sampling artifacts and generate stabilized images of the stable scene with higher resolution than that defined by the camera sampling grid was advanced. This idea has been implemented to enhance the resolution of a single-channel (monochromatic) image. In this Letter, we extend this superresolution method to turbulent color videos and show that, while superresolution in color images is somehow limited, due to the color video acquisition mechanism, when compared to grayscale images, image resolution enhancement is achievable as well.

A color image can be represented by combining three separate monochromatic images. Ideally, each image pixel contains three data measurements, one for each of the three color bands, R, G, and B. In practice, commonly used digital camera with a single CCD array provides only one color measurement (red, green, or blue) per pixel. The detector array, in such cameras, is a grid of CCDs, each made sensitive to one color by placing a color filter array (CFA) in front of the CCD. The Bayer pattern, shown in Fig. 1, is a very common example of such a color filter arrangement. The values of missing color bands at every pixel are synthesized using some form of interpolation from neighboring pixel values. This process is referred to as color demosaicing.

Linear demosaicing methods can be represented by three sets of weights for neighborhoods with a R, G, or B pixel in the center. For the (x,y)-th pixel it is given by

$$I_{(x,y)}^{R,G,B} = \sum_{\Delta_1,\Delta_2 \in \Omega} w_{(\Delta_1,\Delta_2)}^{R,G,B} C_{(x+\Delta_1,y+\Delta_2)}^{R,G,B}, \tag{1}$$

where $I_{(x,y)}^{R,G,B}$ are either the R, G, or B interpolated values, $C_{(x+\Delta_1,y+\Delta_2)}^{R,G,B}$ are the R, G, and B sampled output values of CCD cells covered by the Bayer filter, Ω is a neighborhood centered at (x,y), and $w_{(\Delta_1,\Delta_2)}^{R,G,B}$ are the corresponding interpolation R, G, and B weights normalized so as

$$\sum_{\Delta_1,\Delta_2 \in \Omega} w_{(\Delta_1,\Delta_2)}^{R,G,B} = 1. \tag{2}$$

Applying the Fourier transform to Eq. (1) gives the modulation transfer function (MTF) of the interpolation procedure:

$$H_{(r,s)}^{R,G,B} = H_{(r,s)} \sum_{\Delta_1,\Delta_2} w_{(\Delta_1,\Delta_2)}^{R,G,B} \cdot e^{-2\pi i(r\cdot\Delta_1 + s\cdot\Delta_2)}, \tag{3}$$

where r and s are the image spatial frequencies and $H_{(r,s)}$ is the MTF of a, full cell, 2D rectangular pixel [1,4,5].

For the monochrome case, it has been shown, that the larger fill factor is, the heavier unrecoverable resolution losses are, and, therefore, superresolution methods are potentially more efficient for images acquired with cameras with a small fill factor [1,5]. Equation (3) shows that the color MTF for R, G, or B pixels is a weighted average of monochrome MTFs. This implies that, like in monochrome videos, larger fill factors limit the achievable superresolution results in color images.

In commonly used color video cameras, the Moving Picture Experts Group and JPEG image compression standard is used in which color images are first transformed from raw red-green-blue-representation to the YCbCr representation [6]. In this representation, most of the energy is concentrated in the luminance (Y) component, with little energy in the chrominance (Cb,Cr) components [7]. Therefore, for processing color turbulent videos one can apply a

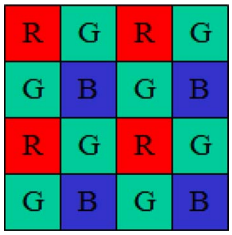


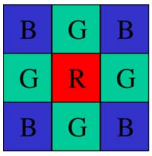
Fig. 1. (Color online) Bayer CFA.

0146-9592/09/050587-3/$15.00

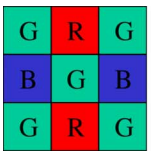
monochrome superresolution algorithm to only the Y component and use simpler processing of the Cb and Cr components. In this way computational efficiency is achieved, and color artifacts, which may appear when superresolving each of the R, G, B channels separately, are avoided. In our proposed method, the Y component is superresolved through elastic registration, with subpixel accuracy, followed by resampling of the frames according to the registration results [1]. The Cb and Cr components are computed using a pixelwise temporal (framewise) median filtering, which was shown to be a good estimation of the stable scene, [2,8]. For JPEG encoded images, the Y component is given by [9]

$$\mathrm{Y} = 0.299\ \mathrm{R} + 0.587\ \mathrm{G} + 0.114\ \mathrm{B}. \tag{4}$$

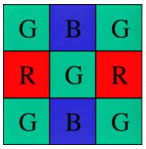
The Y component's MTF depends on the interpolation method. In JPEG encoding standard, the bilinear interpolation is used, and each color channel is interpolated independently. Substitution of Eq. (4) into Eq. (3) yields the Y components' MTFs for the cases when the central pixel is red, green, and blue, respectively,



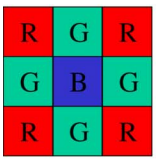
$$Y^{R}_{(r,s)} = H_{(r,s)} \cdot \left[\begin{array}{c} 0.299 + 0.2935\left[\cos\left(2\pi\frac{r}{N}\right) + \cos\left(2\pi\frac{s}{M}\right)\right] \\ +0.114\cos\left(2\pi\frac{r}{N}\right)\cdot\cos\left(2\pi\frac{s}{M}\right) \end{array} \right], \tag{5}$$



$$Y^{G_{\text{Red vertical}}}_{(r,s)} = H_{(r,s)} \cdot \left[0.299\cos\left(2\pi\frac{r}{N}\right) + 0.587 + 0.057\cos\left(2\pi\frac{s}{M}\right)\right], \tag{6}$$



$$Y^{G_{\text{Red Horizontal}}}_{(r,s)} = H_{(r,s)} \cdot \left[0.299\cos\left(2\pi\frac{s}{M}\right) + 0.587 + 0.057\cos\left(2\pi\frac{r}{N}\right)\right], \tag{7}$$



$$Y^{B}_{(r,s)} = H_{(r,s)} \cdot \left[\begin{array}{c} 0.299\cos\left(2\pi\frac{r}{N}\right)\cdot\cos\left(2\pi\frac{s}{M}\right) \\ +0.2935\left[\cos\left(2\pi\frac{r}{N}\right) + \cos\left(2\pi\frac{s}{M}\right)\right] + 0.114 \end{array} \right], \tag{8}$$

where for each equation the corresponding Bayer scheme setup is shown and N and M are the image dimensions on each axis.

One important conclusion derived from Eqs. (5)–(8) is that the MTF of the interpolation process is shift variant. The Y MTFs are depicted in Fig. 2, where the Y MTFs for neighborhoods centered in red, green, and blue pixels are given in solid, dotted, and dashed curves, respectively, and correspond to Eqs. (5)–(8). As one can see, for neighborhoods centered in blue or red pixels, some frequencies of the Y component (marked with arrows) are completely lost. For the sake of clear presentation, the elementwise square-root of the MTFs are presented. In the case at hand, the acquisition is held under turbulent conditions; therefore the same information is likely to be acquired by adjacent, but different, pixels in subsequent frames. This means that frequencies totally suppressed in a certain pixel, are likely to be recovered when the same information is acquired by a neighboring pixel with different color filter (red, green, or blue), in a subsequent frame [10].

The entire processing was experimentally tested on real-life turbulent degraded color videos acquired in a setup that simulates turbulence. The experimental setup is illustrated in Fig. 3. Colored tiles are laid on the bottom of a water tub, filled with water to the height of 30 cm, while the random motion of the water medium is caused by the filling water. The light propagates through the water and is acquired by the single-CCD camera, which has a Bayer filter on its sensor. The results are shown for a sequence where each frame is 200×200 pixels. The test sequence's average turbulent interframe motion magnitude is 2.5, and standard deviation is 1.3 pixels.

Figure 4 illustrates, through images' spectra, the efficiency of the superresolution process with aperture correction. Figure 4(a) presents the estimation of the stable scene of a turbulent degraded frame, interpolated to twice its original size and filtered for

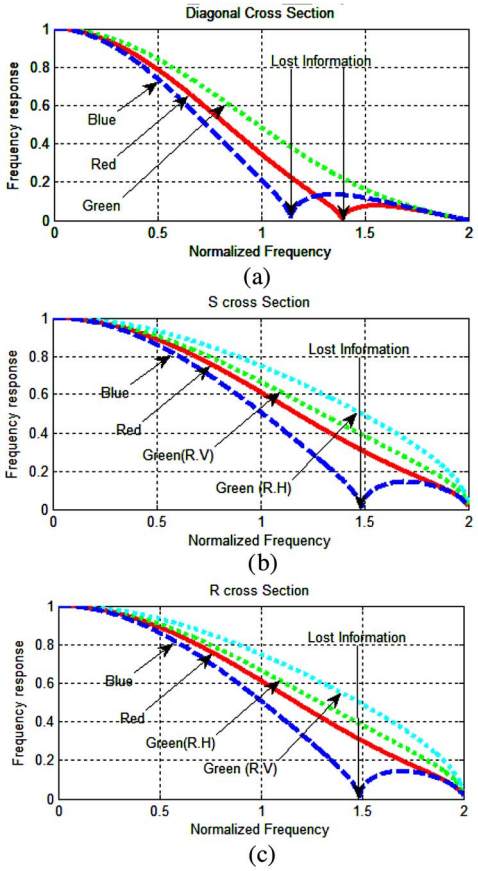


Fig. 2. (Color online) Sensor frequency response for fill factor of 1—cross section views. (a), (b), and (c) are, correspondingly, the diagonal, R, and the S cross sections of the 2D Y frequency response of a single sensor. The dotted, solid, and dashed curves represent the Y component frequency response for the green, red, and blue pixels in the Bayer scheme. The two greens, as defined in Eqs. (6) and (7), are marked by red horizontal (R.H) and red vertical (R.V). The arrows mark frequencies where for the blue or red channel all luminance information is lost.

aperture correction through unsharp masking. The corresponding superresolved image, after aperture correction, is depicted in Fig. 4(b). Figures 4(c) and 4(d) show corresponding image spectra intensities. It is evident that information in higher frequencies outside the base band of the original images is reconstructed in the process.

For quantitative evaluation of the resolution improvement we used a method for numerical evaluation of image effective bandwidth (IEBW) suggested in [11]. The IEBW measure ranges from 0 to 1, where higher values correspond to the presence of larger

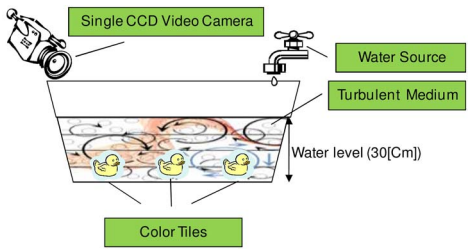


Fig. 3. (Color online) Acquisitioning setup.

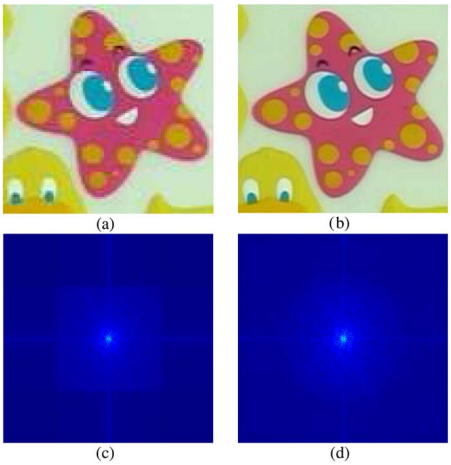


Fig. 4. (Color online) Superresolution through turbulent water currents. (a) is the stable reference frame interpolated to twice its original size with aperture correction applied. (b) is the superresolved image with aperture correction, using the same filter applied on (a). (c) and (d) show corresponding image spectra intensities (for display purpose raised to the power of 0.4) of (a) and (b).

amounts of energy in the higher frequencies of the image. The IEBW factors for the images shown in Fig. 4 are 0.1459 and 0.1814 for the interpolated [Fig. 4(a)] and superresolved [Fig. 4(b)] images, respectively. This suggests that the superresolved image does have higher effective bandwidths, hence more information in higher frequencies.

In conclusion, one can state that the presented results confirm that color sequences subjected to turbulence and acquired by a commonly used camera with a single CCD array can be considerably enhanced in their resolution using the superresolution process described in [1].

This work was partially supported by The Domestic Nuclear Detection Office, Department of Homeland Security (grant CBET-0736232), and the Ran Naor Highway-Safety-Research-Center-Foundation at the Technion, Israel.

References

1. L. P. Yaroslavsky, B. Fishbain, G. Shabat, and I. Ideses, Opt. Lett. **32**, 3038 (2007).
2. B. Fishbain, L. P. Yaroslavsky, and I. Ideses, Adv. Opt. Technol. **2008**, 546808 (2008).
3. Z. Zalevsky, Sh. Rozental, and M. Meller, Opt. Lett. **32**, 1072 (2007).
4. O. Hadar and G. D. Boreman, Opt. Eng. **38**, 782 (1999).
5. L. P. Yaroslavsky, G. Shabat, B. Fishbain, and I. A. Ideses, Proc. SPIE **6812**, 681205 (2008).
6. W. Wharton and D. Howorth, *Principles of Television Reception* (Pitman, 1971), p. 161.
7. M. Irani and S. Peleg, CVGIP Graph. Models Image Process. **53**, 231 (1991).
8. B. Fishbain, L. P. Yaroslavsky, and I. A. Ideses, J. Real-Time Image Process. **2**, 11 (2007).
9. E. Hamilton, JPEG File Interchange Format Version 1.02, September 1992.
10. L. P. Yaroslavsky and H. J. Caulfield, Appl. Opt. **33**, 2157 (1994).
11. B. Fishbain, L. P. Yaroslavsky, I. A. Ideses, and F. Roffet-Crété, Proc. SPIE **6808**, 68080X (2008).