Image-adaptive Color Super-resolution

Umamahesh Srinivas[1], *Xuan Mo*[1], *Manu Parmar*[2], *Vishal Monga*[1]
[1] *Pennsylvania State University, University Park, PA;* [2] *Qualcomm MEMS Technologies, San Jose, CA*

Abstract

Image super-resolution is the problem of recovering a high resolution (hi-res) image from multiple low resolution (lo-res) acquisitions of a scene. The main focus and the most significant contributions of research in this area have been on the problem of super-resolving single channel (grayscale) images. Multi-channel (color) image super-resolution is often treated as an extension to grayscale super-resolution by simply considering the luminance component of the image more carefully than the chrominance components. In this paper we address explicitly the problem of color image super-resolution by formulating an optimization problem that leads to convergence guarantees. The key contribution of this work is the inclusion of a color regularizer that effectively accounts for both luminance and chrominance geometry in images. We show results demonstrating substantial image quality improvement over the state of the art, especially for images with significant chrominance geometry.

Introduction

The resolution of an imager is limited by the resolution of its image sensor and the quality of its optics. In several imaging applications it is useful to have the ability to recover an image with resolution higher than that permitted by the capabilities of the imager. Image super-resolution fills this need by recovering a high resolution (hi-res) image from multiple low resolution (lo-res) acquisitions of a scene - provided, of course, that the different low-res images capture different (at sub-pixel level) views of the scene. There are several imaging applications where super-resolution finds use - for instance, medical imaging applications that use images for computer vision tasks benefit from high-res images. Another application where super-resolution is particularly appropriate is in surveillance applications where a video stream, which can provide input lo-res frames to the super-resolution algorithm, is continuously acquired.

In simple terms, the super-resolution problem is addressed by describing first the several lo-res images on a grid finer than the resolution of single images (an image registration problem) followed by filling in values for missing pixels (an image interpolation problem). There has been significant advance in super-resolution research in recent years. Park et al. [1] give an overview of the problem and describe early advances. Most performance improvements come from solutions to the image registration problem with better motion estimation techniques.

A common thread in most work is the focus on grayscale image super-resolution. Color image super-resolution is often treated simply by assuming that the luminance component of the image carries its spatial features. Algorithms that consider the chrominance components will only use them to improve image registration by better motion estimation [2, 3]. Very few researchers consider explicitly the relationship between the color channels in the interpolation problem. When they do, a common approach is to assume that spatial high-frequency components across the color channels are strongly correlated. In other words, if an edge (or a feature) is sensed in one channel, it implies that the edge (or feature) exists in all channels. Farsiu et al. [4] use this approach in a joint demosaicking and super-resolution problem formulation with good results.

We note that the assumption about strong interchannel correlation in high frequency components is akin to assuming that most spatial features (edges and texture) appear in some luminance-type component - found either with decomposition to a standard luminance-chrominance space like YCbCr, or with the PCA technique for decorrelating the color components. This assumption is clearly untrue for images with strong chrominance geometry - images in which edges and textures are not a result of ambient illumination but due to edges between objects with different chrominance.

In this work we consider explicitly the problem of color image super-resolution. The key contribution of this work is the inclusion of a color regularizer that effectively accounts for both luminance and chrominance geometry in images. We propose an optimization framework that is separably convex, leading to convergence guarantees, along with the enforcement of constraints consistent with real-world imaging physics. We show results demonstrating substantial image quality improvement over the state of the art, especially for images with significant chrominance edge features.

Image-adaptive color super-resolution

We first present the mathematical formulation of our color SR framework. We use the camera imaging model [1]:

$$\mathbf{y}_k = \mathbf{DBT}_k\mathbf{x}+\mathbf{n}_k, \quad 1 \le k \le K, \tag{1}$$

where $\mathbf{x} = [\mathbf{x}_r^T \ \mathbf{x}_g^T \ \mathbf{x}_b^T]^T \in \mathbb{R}^{3n}$ is the unknown (vectorized) hi-res image that we seek to reconstruct (subscripts r, g, b indicate the red, green and blue color channels respectively), $\mathbf{y}_k \in \mathbb{R}^m$ represents the k-th observed lo-res image, $\mathbf{T}_k \in \mathbb{R}^{3n\times 3n}$ is the k-th geometric warping matrix, $\mathbf{B} \in \mathbb{R}^{3n\times 3n}$ describes camera optical blur, downsampling matrix $\mathbf{D} \in \mathbb{R}^{m\times 3n}$ models the aliasing, and $\mathbf{n}_k \in \mathbb{R}^m$ is the noise vector that corrupts $\mathbf{y}_k$.

Single-channel super-resolution

The standard SR reconstruction problem recovers an estimate of $\mathbf{x}$ by minimizing the error between the warped, blurred and downsampled versions of $\mathbf{x}$ as predicted by the imaging

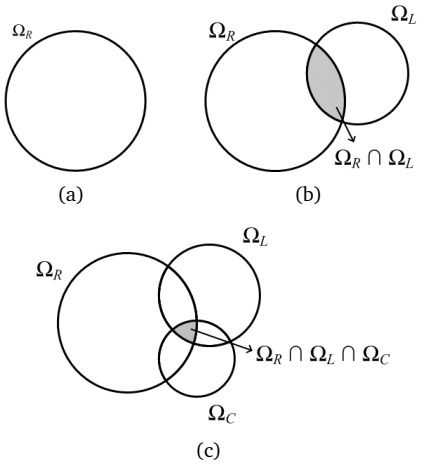


Figure 1. *Venn diagram interpretation of the role of regularization. (a) The set Ω_R of all images* $\mathbf{x}$ *that minimize* reconstruction error *in* (3). *(b) The set Ω_L consists of images* $\mathbf{x}$ *with significant luminance edge information, which is distributed among the RGB channels. Images in the intersection $\Omega_R \cap \Omega_L$ are obtained by minimizing the cost function in* (5). *(c) The set Ω_C of all images* $\mathbf{x}$ *that possess chrominance edges. Images $\mathbf{x} \in \Omega_R \cap \Omega_L \cap \Omega_C$ preserve luminance* and *chrominance edges.*

model in Eq. (1) and the observed lo-res images $\{\mathbf{y}_k\}_{k=1}^K$:

$$\mathscr{C} = \sum_{k=1}^{K} \|\mathbf{y}_k - \mathbf{DBT}_k\mathbf{x}\|_p + \gamma\rho(\mathbf{x}), p \geq 1, \tag{2}$$

where $\rho(\mathbf{x})$ is a regularization term (or equivalently, a prior in a MAP estimation set-up). We drop the generic $\rho(\mathbf{x})$ term in the subsequent algorithmic development for notational convenience. Traditionally, the warping matrix $\mathbf{T}_k$ is parameterized by a registration vector $\boldsymbol{\theta}_k$, obtained from the homography matrix. Although it has been pointed out in literature [5] that the joint minimization of $\mathscr{C}$ under $\mathbf{x}$ and $\{\boldsymbol{\theta}_k\}$ leads to better results than sequential optimization, a key challenge that still remained was the tractability of the resulting optimization problem, since $\mathscr{C}$ is not jointly convex in $\mathbf{x}$ and $\{\boldsymbol{\theta}_k\}$.

The framework in [6] employs a transformation-of-variables trick to arrive at a separably convex optimization problem, together with the enforcement of constraints motivated by real-world imaging physics. The optimization problem is stated as follows:

$$\begin{array}{ll} \text{minimize} & \sum_{k=1}^{K} \|\mathbf{y}_k - \mathbf{DBT}_k\mathbf{x}\|_p \\ \text{subject to} & \mathbf{0} \leq \mathbf{x} \leq \mathbf{1} \\ & \mathbf{0} \leq \mathbf{DBT}_k\mathbf{x} \leq \mathbf{1}, \quad 1 \leq k \leq K \\ & \mathbf{T}_k.\mathbf{1} = \mathbf{1}, \quad 1 \leq k \leq K \\ & \mathbf{B}.\mathbf{1} = \mathbf{1} \\ & \mathbf{t}_{k,i}^T \mathbf{m}_{k,i} = 0, \quad 1 \leq i \leq 3n, \quad 1 \leq k \leq K \\ & \mathbf{b}_i^T \mathbf{e}_i = 0, \quad 1 \leq i \leq 3n \end{array} \tag{3}$$

The vectors $\mathbf{0}$ and $\mathbf{1}$ have all their entries equal to 0 and 1 respectively, and their dimension is derived from context. The first two constraints restrict pixel intensities to the range $[0,1]$, for processing convenience. The fact that $\mathbf{T}_k$ is an interpolation matrix and $\mathbf{B}$ is a spatially local kernel operator is enforced by the next two constraints. Finally, prior knowledge of the location of the non-zero elements in the matrices $\mathbf{T}_k$ and $\mathbf{B}$ is enforced via the last two membership constraints. The problem is solved in an iterative manner, alternating between the optimization variables $\mathbf{x}, \{\mathbf{T}_k\}, \mathbf{B}$. The separably convex nature of the problem guarantees convergence to local minima and robustness to the choice of initial values.

Color super-resolution: state of the art

The problem of color demosaicking (recovering full-color images from sub-sampled single-sensor images) has been addressed with the use of regularizers that explicitly maximize inter-channel correlation. Such terms have also been found useful for color SR [4, 7, 8]. Formally, let the matrices $\mathbf{S}_r, \mathbf{S}_g, \mathbf{S}_b \in \mathbb{R}^{3n \times 3n}$ represent gradient operators on the red, green and blue color channels respectively. Then, $\mathbf{S}_r\mathbf{x}$ represents the red channel edge map; similar descriptions hold for $\mathbf{S}_g\mathbf{x}$ and $\mathbf{S}_b\mathbf{x}$. For images with dominant luminance edges, we then have a luminance regularization term of the form:

$$\rho_L(\mathbf{x}) = \|(\mathbf{S}_r - \mathbf{S}_g)\mathbf{x}\|_1 + \|(\mathbf{S}_g - \mathbf{S}_b)\mathbf{x}\|_1 + \|(\mathbf{S}_r - \mathbf{S}_b)\mathbf{x}\|_1, \tag{4}$$

such that differences between pairs of color channel edge maps are minimized. The ℓ_1-norm offers robustness to outliers. In [7], the ℓ_2-norm version of this regularizer is used. The modified optimization cost function incorporates this regularization term with corresponding parameter α_L as follows:

$$\mathscr{C}_1 = \sum_{k=1}^{K} \|\mathbf{y}_k - \mathbf{DBT}_k\mathbf{x}\|_p + \alpha_L\rho_L(\mathbf{x}). \tag{5}$$

A qualitative explanation of the role of the regularization term is provided in Fig. 1. The ill-posed nature of the SR problem results in a (convex) set Ω_R of hi-res images $\mathbf{x}$ that minimize the reconstruction error term in (3). For the color SR problem, $\mathbf{x} \in \Omega_R$ may be interpreted as the optimal solutions by the independent-channel SR approach. In Fig. 1(b), the set Ω_L corresponds to the set of all images with significant luminance geometry that is distributed among the red, green and blue channels. The intersection set $\Omega_R \cap \Omega_L$ contains precisely those images which minimize the constrained optimization problem with the modified cost function in (5). The regularization parameter α_L controls the "size" of Ω_L, and by extension, the intersection set.

Proposed approach: chrominance regularization

The $\rho_L(\mathbf{x})$ term guides the reconstruction towards images that preserve luminance edge information. In fact, the success of color SR formulations incorporating such luminance priors can be attributed to the fact that images with dominant luminance geometry are encountered most frequently in the universe of all images. However, luminance-based priors fail to capture the useful chrominance information present in images with significant chrominance geometry. We propose the inclusion of a novel chrominance-based regularization term for such images.

The YCbCr representation of a color image decorrelates it into luminance and chrominance components. While a luminance edge appears in each channel of the equivalent RGB representation, a chrominance edge, in Cb for example, appears

prominently in the blue channel and not so in the other two channels. Thus, for images with reasonable degree of *chrominance edge information*, we in fact expect *low* correlation between the RGB color channels. For such images, we seek to minimize edge correlation between the different color channels in the desired hi-res image as follows:

$$(\mathbf{S}_r\mathbf{x})^T(\mathbf{S}_g\mathbf{x}) < \epsilon_{rg}, (\mathbf{S}_g\mathbf{x})^T(\mathbf{S}_b\mathbf{x}) < \epsilon_{gb}, (\mathbf{S}_b\mathbf{x})^T(\mathbf{S}_r\mathbf{x}) < \epsilon_{br}, \quad (6)$$

where $\epsilon_{rg}, \epsilon_{gb}$ and ϵ_{rb} are suitably chosen constants. From optimization theory, the Lagrangian formulation incorporates this constraint into the cost function as another regularization term:

$$\rho_C(\mathbf{x}) = (\mathbf{S}_r\mathbf{x})^T(\mathbf{S}_g\mathbf{x}) + (\mathbf{S}_g\mathbf{x})^T(\mathbf{S}_b\mathbf{x}) + (\mathbf{S}_b\mathbf{x})^T(\mathbf{S}_r\mathbf{x}), \quad (7)$$

with an accompanying regularization parameter α_C.

A standard 2-D gradient operator like the Sobel matrix seen commonly in edge detection problems can be used to obtain $\mathbf{S}_r, \mathbf{S}_g$ and $\mathbf{S}_b$. However, different images have possibly different edge profiles, motivating the need for *image-dependent* gradient operators which are *optimized* for *resolution enhancement*. Accordingly, we optimize $\mathbf{S}_r, \mathbf{S}_g$ and $\mathbf{S}_b$ together with the other optimization variables in (3) in a similar iterative manner. The modified cost function, with all variables explicitly mentioned, is:

$$\mathscr{C}_2\left(\mathbf{x}, \{\mathbf{T}_k\}, \mathbf{B}, \mathbf{S}_r, \mathbf{S}_g, \mathbf{S}_b\right) = \sum_{k=1}^{K} \|\mathbf{y}_k - \mathbf{DBT}_k\mathbf{x}\|_p + \alpha_L\rho_L(\mathbf{x}) + \alpha_C\rho_C(\mathbf{x}). \quad (8)$$

It is easy to see that the separably convex nature of the cost function in (8) is still preserved after the addition of the color regularization term $\rho_C(\mathbf{x})$. Again, we seek constraints on $\mathbf{S}_r, \mathbf{S}_g$ and $\mathbf{S}_b$ which lead to meaningful gradient operators in practice.

Suppose $h[n_1, n_2]$ represents the spatial domain gradient filter, where n_1 and n_2 represent the 2-D coordinates. Since the gradient is a spatially local operator, $h[n_1, n_2] \neq 0 \ \forall \ (n_1, n_2) \in \mathscr{R}$, a finite region of support. The corresponding 2-D Fourier transform in terms of spatial frequency variables ω_1 and ω_2 is:

$$H(\omega_1, \omega_2) = \sum_{(n_1,n_2)\in} \sum_{\mathscr{R}} h[n_1, n_2] e^{-j\omega_1 n_1} e^{-j\omega_2 n_2}. \quad (9)$$

It follows from the high-pass characteristic of the gradient operator that it has zero magnitude response at DC frequency, so $H(0,0) = 0$, implying that the sum of filter coefficients is 0. In the gradient matrix $\mathbf{S}_r$ (likewise for $\mathbf{S}_r$ and $\mathbf{S}_b$), the set of nonzero elements in every row is exactly the same as the set of nonzero filter coefficients $h[n_1, n_2]$. So, every row of the gradient operator necessarily sums to 0, and this condition is captured by the constraints:

$$\mathbf{S}_r.\mathbf{1} = \mathbf{0}, \quad \mathbf{S}_g.\mathbf{1} = \mathbf{0}, \quad \mathbf{S}_b.\mathbf{1} = \mathbf{0}. \quad (10)$$

The trivial case of convergence to the zero matrix is avoided by enforcing *membership constraints* on the structure of the gradient operators as follows. Based on the filter coefficients of the gradient operator used to initialize the algorithm, the locations of the positive, negative and zero elements are fixed. We generate the corresponding set of membership vectors $\mathbf{f}_{r,i}, \mathbf{f}_{g,i}, \mathbf{f}_{b,i}$ whose entries assume the values +1,-1 or 0 respectively whenever the corresponding element of $\mathbf{S}$ takes positive, negative or zero values. This leads to another set of constraints ($\mathbf{s}$ identifies a row of the corresponding gradient operator):

$$(\mathbf{s}_{r,i})^T\mathbf{f}_{r,i} = 1, (\mathbf{s}_{g,i})^T\mathbf{f}_{g,i} = 1, (\mathbf{s}_{b,i})^T\mathbf{f}_{b,i} = 1; 1 \leq i \leq 3n. \quad (11)$$

Reverting to the set-theoretic interpretation, Fig. 1(c) shows the set Ω_C of images with chrominance edges. If α_L and α_C are chosen suitably, the intersection $\Omega_R \cap \Omega_L \cap \Omega_C$ is non-empty. The images in that intersection set are obtained by solving the ensuing optimization problem in (12), with the upshot that they preserve both luminance *and* chrominance edges in a balanced, image-adaptive manner.

Optimization framework: key features

In summary, the complete optimization problem is stated as follows:

$$\begin{aligned}
\text{minimize} \quad & \textstyle\sum_{k=1}^{K} \|\mathbf{y}_k - \mathbf{DBT}_k\mathbf{x}\|_p + \alpha_L\rho_L(\mathbf{x}) + \alpha_C\rho_C(\mathbf{x}) \\
\text{subject to} \quad & \mathbf{0} \leq \mathbf{x} \leq \mathbf{1} \\
& \mathbf{0} \leq \mathbf{DBT}_k\mathbf{x} \leq \mathbf{1}, \quad 1 \leq k \leq K \\
& \mathbf{T}_k \cdot \mathbf{1} = \mathbf{1}, \quad 1 \leq k \leq K \\
& \mathbf{B} \cdot \mathbf{1} = \mathbf{1} \\
& \mathbf{t}_{k,i}^T\mathbf{m}_{k,i} = 0, \quad 1 \leq i \leq 3n, \quad 1 \leq k \leq K \\
& \mathbf{b}_i^T\mathbf{e}_i = 0, \quad 1 \leq i \leq 3n \\
& \mathbf{S}_r.\mathbf{1} = \mathbf{0} \\
& \mathbf{S}_g.\mathbf{1} = \mathbf{0} \\
& \mathbf{S}_b.\mathbf{1} = \mathbf{0} \\
& (\mathbf{s}_{r,i})^T\mathbf{f}_{r,i} = 1, \quad 1 \leq i \leq 3n \\
& (\mathbf{s}_{g,i})^T\mathbf{f}_{g,i} = 1, \quad 1 \leq i \leq 3n \\
& (\mathbf{s}_{b,i})^T\mathbf{f}_{b,i} = 1, \quad 1 \leq i \leq 3n
\end{aligned} \quad (12)$$

As discussed earlier, this problem is separably convex in the variables of interest. The variables are optimized in an alternating manner iteratively. A judicious choice of constraints leads to physically meaningful imaging parameters. Under l_2-norm minimization, it can be shown that the individual optimization problems reduce to quadratic programs, providing computational efficiency benefits.

Chrominance information is critical to super-resolution quality

We illustrate the importance of effectively accounting for chrominance geometry in resolution enhancement with the help of an example (Fig. 2). We synthetically generated multiple lo-res (64 × 64) images (Fig. 2(b)) from a hi-res color image of dimension 128× 128 (Fig. 2(a)) using the camera imaging model in Eq. (1). Specifically, the the original hi-res image was subjected to random translation and rotation, blurred using a 5×5 Gaussian blur and sub-sampled by a factor of two along both dimensions to generate the lo-res images.

We reconstructed estimates of the hi-res image (128×128) in four different ways:

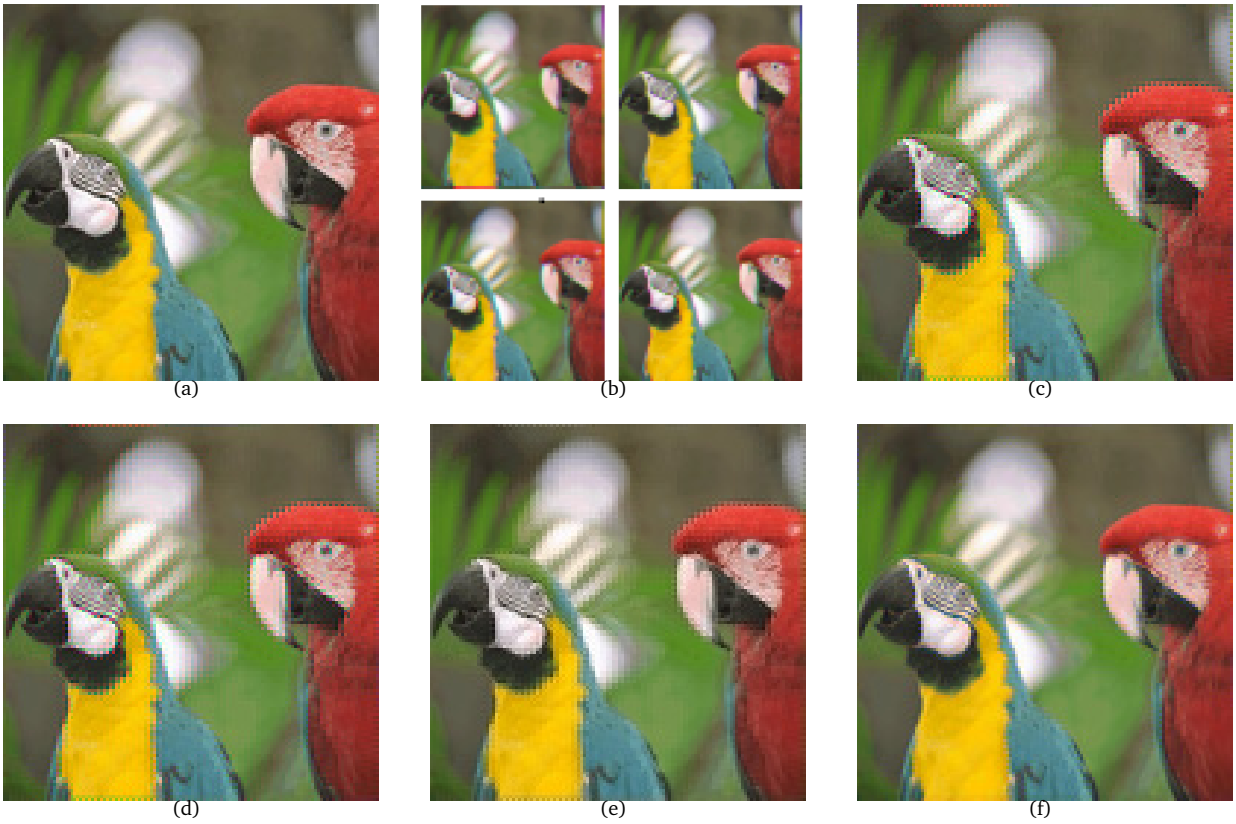


Figure 2. *Illustration of significance of chrominance geometry. (a) Original hi-res image (128× 128). (b) Four sample lo-res images generated synthetically (64×64). (c) SR performed on RGB channels independently. (d) SR performed on luminance channel only (chrominance channels interpolated in a simple manner). (e) SR performed independently on luminance and chrominance channels. (f) SR performed with chrominance geometry-sensitive regularizer - proposed method.*

1. Grayscale SR performed on red, green and blue channels independently (Fig. 2(c)).
2. SR performed only on the luminance channel (Fig. 2(d)).
3. SR performed on luminance and chrominance channels independently [8](Fig. 2(e)).
4. SR by explicitly incorporating chrominance information into the recovery process (Fig. 2(f)).

Color artifacts are visible in Figs. 2(c)-(e). We argue that these artifacts are a result of either ignoring chrominance information completely (Fig. 2(d)), or not adequately accounting for correlation between the RGB channels in reconstruction (Fig. 2(c)). The results of applying the proposed approach (Fig. 2(f)) highlight the importance of effectively using chrominance information for resolution enhancement.

Results and Discussion

A natural question to ask at this juncture is: how can the same optimization framework handle a general class of images possessing reasonable chrominance geometry in addition to the dominant luminance geometry? We now address this issue and show experimental results to validate the approach.

Determining the parameters α_C and α_L

Parameters α_L and α_C control the relative significance given to the reconstruction error and gradient regularization terms in the optimization. Since the corresponding constraints capture different information about the same color image, we assign complementary weights to α_L and α_C, i.e. $\alpha_L = \alpha_{\max} - \alpha_C$, where $\alpha_{\max}$ is the maximum value assumed by α_C for good reconstruction. Our choice of α_C is tied to the degree of chrominance geometry present in an image. We next present an intuitive quantitative estimate of chrominance edge information as a pre-processing step to facilitate a good choice of α_C.

Quantitative estimate of chrominance geometry

In images with high chrominance geometry, it is reasonable to expect a significant amount of edge information in the Cb and Cr channels, in addition to luminance edges present in the Y channel. We define a parameter β as follows:

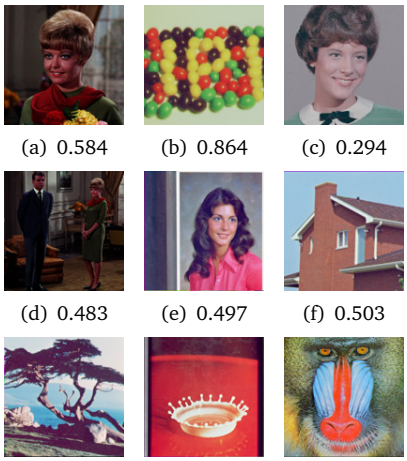
$$\beta = \frac{1}{2}\left(\frac{\|\mathbf{H}_1\mathbf{x}_{Cb}\| + \|\mathbf{H}_1\mathbf{x}_{Cr}\|}{\|\mathbf{H}_1\mathbf{x}_Y\|} + \frac{\|\mathbf{H}_2\mathbf{x}_{Cb}\| + \|\mathbf{H}_2\mathbf{x}_{Cr}\|}{\|\mathbf{H}_2\mathbf{x}_Y\|}\right), \quad (13)$$

where the highpass operators $\mathbf{H}_1$ and $\mathbf{H}_2$ are respectively formed using the 2-D Scharr operators:

$$\mathbf{h}_1 = \begin{bmatrix} 3 & 10 & 3 \\ 0 & 0 & 0 \\ -3 & -10 & -3 \end{bmatrix}, \mathbf{h}_2 = \begin{bmatrix} 3 & 0 & -3 \\ 10 & 0 & -10 \\ 3 & 0 & -3 \end{bmatrix}, \quad (14)$$

and $\mathbf{x}_Y, \mathbf{x}_{Cb}, \mathbf{x}_{Cr}$ respectively represent the Y, Cb and Cr channels. While $\mathbf{h}_1$ captures horizontal edges, $\mathbf{h}_2$ captures the vertical edges. We test over a variety of images and determine

that $\beta_0 = 0.75$ is a suitable threshold to classify the chrominance geometry content in images. Fig. 3 shows sample test images and their corresponding β values. Images in Fig. 3(b), (h) and (i) have high chrominance geometry, as confirmed by their large β values in comparison with the other images.



(a) 0.584 (b) 0.864 (c) 0.294

(d) 0.483 (e) 0.497 (f) 0.503

(g) 0.425 (h) 0.828 (i) 0.975

Figure 3. *The computed value of β for color images with varying amounts of chrominance edge information.*

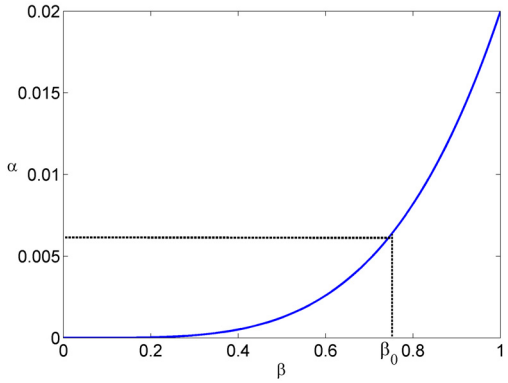


Figure 4. *Curve that maps β values to corresponding α_C.*

Relating β to α_C

From a given set of lo-res images, the mean β is computed as the average of the β values obtained from the individual images. Our aim now is to determine a mapping from β to the α_C eventually used in the optimization problem. Intuitively, when there is minimal chrominance geometry in an image, $\beta = 0$ and the corresponding $\alpha_C = 0$. The complementary α_L term attains a maximum value, respecting the dominant luminance geometry structure of the image. When $\beta = \beta_{\max}$, α_C attains its maximum value and $\alpha_L = 0$. α_C varies monotonically with β. From the same set of images used to compute β_0, we determine suitable α_C values for good reconstruction empirically, shown in Fig. 4. For images with both types of edges, the estimation of β ensures an image-adaptive choice of α_C and α_L which preserves both luminance and chrominance information.

Experimental validation

Fig. 5 shows super-resolution reconstruction performed on a set of lo-res images of dimension 128×128. The β value from Fig. 3(b) indicates that these set of images have significant chrominance geometry. Fig. 5(a) shows a bilinearly interpolated lo-res image for comparison, while Figs. 5(b)-(d) show results from competitive approaches. Fig. 5(e) shows the hi-res image reconstructed using the proposed framework, but with a fixed set of operators $\mathbf{S}_r, \mathbf{S}_g$ and $\mathbf{S}_b$. The improvement in the quality of this image over Fig. 5(b) shows the merits of exploiting the color correlation information. Fig.5(f) shows the result obtained by using the entire framework in (12).

In addition to the perceptual improvements, we provide a quantitative means of confirming the improvements provided by our proposed approach. For images where we have ground truth for comparison, we compute the mean-squared error (MSE) between the ground truth $\mathbf{x}_{gt}$ and the reconstructed hi-res image $\mathbf{x}$ as:

$$MSE(\mathbf{x}_{gt}, \mathbf{x}) = \frac{\|\mathbf{x}_{gt} - \mathbf{x}\|_2^2}{3n}. \tag{15}$$

MSE values are reported in Table 1 for the parrot image from Fig. 2 as well as for Figs. 3(g)-(h), for which ground truth information is available. Smaller error indicates better reconstruction. For images where we have no such reference image to compare with, we define the following quantity:

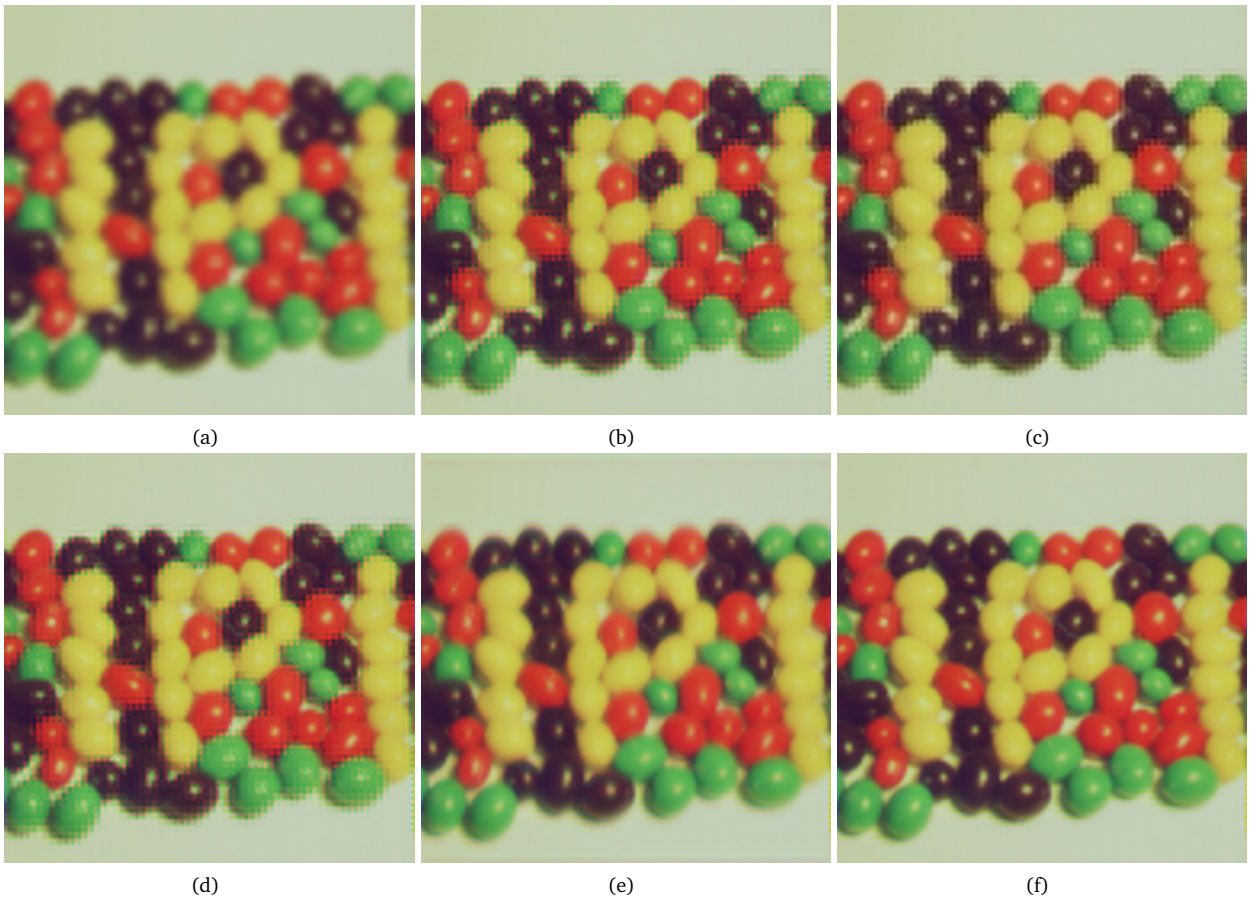
$$J_i = 10\left(\log\left(\frac{\sum_{k=1}^{K}\|\mathbf{y}-\mathbf{DBT}_k\mathbf{x}_i\|_2}{\sum_{k=1}^{K}\|\mathbf{y}-\mathbf{DBT}_k\mathbf{x}\|_2}\right) + (1-\beta)\log\frac{\rho_L(\mathbf{x}_i)}{\rho_L(\mathbf{x})} + \beta\log\frac{\rho_C(\mathbf{x}_i)}{\rho_C(\mathbf{x})}\right), i = 1,2,3 \tag{16}$$

where $i = 1,2,3$ refer to the methods [4], [7] and [8] respectively and $\mathbf{x}_i$ are the corresponding reconstructed hi-res images, while $\mathbf{x}$ refers to the hi-res image from our method. The first term compares the reconstruction error, with the $\mathbf{B}$ and $\{\mathbf{T}_k\}$ estimated according to the specific approach. The relative significance assigned to the luminance and chrominance regularizer terms is controlled by the image-dependent parameter β. The matrices $\mathbf{S}_r, \mathbf{S}_g$ and $\mathbf{S}_b$ are chosen to be standard gradient operators for the methods being compared, while the optimized matrices are used in the $\rho_L(\mathbf{x})$ and $\rho_C(\mathbf{x})$ terms for our proposed method.

Table 2 lists the values $J_i, i = 1,2,3$ corresponding to [4, 7, 8], for the nine images shown in Fig. 3. Note that J_i is expressed in units of decibels (dB), so that a value of 10 dB indicates a reduction in overall error by a factor of 10. A negative value for J_i indicates that the method under comparison performs better than our approach.

Image	M_1	M_2	M_3	M_4
Parrot	7.859	17.027	14.563	20.533
Fig. 3(g)	11.537	17.081	14.366	18.169
Fig. 3(h)	12.340	30.506	20.429	36.299

Table 1. MSE values for the three images where ground truth is available. M_1 refers to reconstruction using our method, while M_2, M_3 and M_4 refer to reconstruction from [4,7,8] respectively.



(a) (b) (c)

(d) (e) (f)

Figure 5. *(a) Bilinearly interpolated lo-res image. (b) SR using the method in [4]. (c) SR using the method in [7]. (d) SR using the method in [8]. (e) SR using the proposed framework, but with fixed gradient operators. (f) Result using the complete proposed framework.*

Image	J_1	J_2	J_3
(a)	0.182	1.599	16.469
(b)	**7.900**	**6.562**	**26.856**
(c)	-0.493	-1.153	17.711
(d)	0.404	-0.979	12.196
(e)	7.902	4.674	21.647
(f)	7.222	5.260	20.804
(i)	**12.110**	**10.899**	**27.805**

Table 2. Quantitative comparison of our proposed approach with competitive methods in literature. The quantity J_i is defined in (16) (in units of dB), with $i = 1,2,3$ corresponding to [4], [7], [8] respectively. Images (b) and (i) from Fig. 3 have significant chrominance geometry.

Conclusions

Most state-of-the art color image super-resolution approaches treat the problem as an extension of grayscale super-resolution. The few approaches that consider inter-channel color correlation only apply well to images that have spatial features predominantly in the luminance channel. We propose a new approach to color super-resolution that also accounts for chrominance geometry. We formulate the problem as a constrained minimization that provides a convergence guarantee. We show results demonstrating substantial image quality improvement over the state of the art, especially for images with significant chrominance geometry.

References

[1] S. C. Park, M. K. Park, and M. G. Kang. Super-resolution image reconstruction: A technical overview. *IEEE Signal Processing Magazine*, 20: 21-36, 2003.

[2] N. R. Shah and A. Zakhor. Resolution enhancement of color video sequences. *IEEE Trans. Image Processing*, 8: 879-885, 1999.

[3] B. C. Tom and A. Katsaggelos. Resolution enhancement of monochrome and color video using motion compensation. *IEEE Trans. Image Processing*, 10: 278-287, 2001.

[4] S. Farsiu, M. Elad and P. Milanfar. Multiframe demosaicing and super-resolution of color images. *IEEE Trans. Image Processing*, 15:141-159, 2006.

[5] R. Hardie, K. Barnard, and E. Armstrong. Joint MAP registration and high-resolution image estimation using a sequence of undersampled images. *IEEE Trans. Image Processing*, 6: 1621-1633, 1997.

[6] V. Monga and U. Srinivas. A constrained optimization perspective on joint spatial resolution and dynamic range enhancement. *Proc. IEEE Asilomar Conference*, 870-874, 2010.

[7] D. Menon and G. Calvagno. Regularization approaches to demosaicking. *IEEE Trans. Image Processing*, 18: 2209-2220, 2009.

[8] P. Vandewalle, K. Krichane, D. Alleysson and S. Susstrunk. Joint demosaicing and super-resolution imaging from a set of unregistered aliased images. *Proc. IST/SPIE Electronic Imaging: Digital Photography III*, vol. 6502, 2007.