

SINGLE IMAGE SUPER-RESOLUTION VIA PHASE CONGRUENCY ANALYSIS

Licheng Yu, Yi Xu, Bo Zhang

Department of Electronic Engineering, Shanghai Jiaotong University, Shanghai 200240, China
Shanghai Key Lab of Digital Media Processing and Communication
xuyi@sjtu.edu.cn

ABSTRACT

Single image super-resolution (SR) is a severely unconstrained task. While the self-example-based methods are able to reproduce sharp edges, they perform poorly for textures. For recovering the fine details, higher-level image segmentation and corresponding external texture database are employed in the example-based SR methods, but they involve too much human interaction. In this paper, we discuss the existing problems of example-based technique using scale space analysis. Accordingly, a robust pixel classification method is designed based on the phase congruency model in scale space, which can effectively divide images into edges, textures and flat regions. Then a super-resolution framework is proposed, which can adaptively emphasize the importance of high-frequency residuals in structural examples and scale invariant fractal property in textural regions. Experimental results show that our SR approach is able to present both sharp edges and vivid textures with few artifacts.

Index Terms— Phase congruency, quaternion Gabor, super-resolution, example-based synthesis, fractal-based enhancement

1. INTRODUCTION AND RELATED WORK

Super-resolution refers to the process of obtaining higher-resolution (HR) images from one lower-resolution (LR) image or multiple LR ones with sub-pixel displacement. In this paper, we focus on single-image super-resolution method without consideration of motion estimation problem. Current single-image super-resolution methods can be divided into three categories: interpolation-based method, reconstruction-based method and example-based method.

Among the interpolation-based approaches, bi-linear and bi-cubic are most commonly used in real-time SR tasks for high speed and simplicity. However, they are inclined to produce blurry and jaggy edges. Thus the edge-directed reconstruction-based methods are proposed for reproducing

This work is supported in part by NSFC (60932006, 60902073), the 111 project (B07022), STCSM (12DZ2272600) and the High Technology Research and Development Program of China (2011AA01A107, 2012AA011702).



Fig. 1: 4X results of Bi-cubic, gradient profile transformation[2] and ours. The reproduction of fine edges and vivid textures varies for each method, while our SR work that combines phase-congruency-based segmentation and adaptive up-sampling method is able to highlight both.

smoother edges. By applying different kinds of gradient priors[1, 2, 3], they are effective in reversing the image degradation process. Yet, these prior models focusing on edge recovery tend to leave texture regions untouched. Compared with them, fractal analysis [4] is applied on image gradient and reproduced vivid textures via fractal-based enhancement.

Another hot issue is the example-based method, which takes efforts to fill in the missing high-frequency band using the provided example database. Generally, the example patches are decomposed into the smoothed version and the residual high-frequency version[5, 6], where the former bands are used for sample matching and the latter ones are for missing-information filling. However, a universal external database tends to produce fairly noisy effect, since it is incapable of maintaining high relevance for every image patch. To release the strong reliance, self-similar scale-invariance property was proposed in [7, 8], which constructs the example database from the input LR image itself across different scales. However, some wrongly hallucinated textures were meanwhile introduced and the analysis of such problem was lacked.

More recently, [9, 10] tried to combine the texture-based segmentation and example-based hallucination. Although this kind of methods provides some encouraging visual results in textures, it requires too much human interaction in both content-based segmentation and user-supplied example database, which limits the general application.

In this paper, we investigate the performance of example-based method on edges and textures. It is found that these two kinds of structures present different properties in scale space. Therefore, we propose a robust pixel classification method using quaternion phase congruency map, decomposing images into edge regions, textural regions and flat regions. A unified super-resolution framework is then proposed, which can adaptively emphasize the importance on high-frequency residuals in structural examples and scale invariant fractal property[4] in textural regions. The advantage of our SR method can be seen in Figure 1, which is able to highlight both fine edges and vivid textures.

The rest of our paper is organized as follows: Section 2 discusses the open problems of existing example-based SR methods. Section 3 introduces the pixel classification method using quaternion phase congruency map. Section 4 provides the implementation details of our SR framework. Section 5 compares our method with the state-of-the-art SR methods. And finally, the conclusion of this work is drawn in Section 6.

2. IMAGE ANALYSIS ON EXAMPLE-BASED SR METHOD

In example-based SR methods, the alternative examples are divided into their smoothed version and high-pass filtered version, where the former band information are used for matching and the latter one for filling the extra details. However, the accuracy of texture matching is usually lower than that of edge. This can be explained from two points of view. Firstly, the highest spatial-frequency components of the smoothed version are most important to the matching phase[5, 6]. For salient edges, they generally contain wider frequency bands than textures, thus are able to provide more features for higher

matching accuracy. On the contrary, the band of textures are relatively short. This might introduce examples with low relevance for synthesis, making the hallucination performance become poor.

Secondly, the edge and line structures have higher redundancies in scale space. As phase encodes the image structure information, the self-similarity of structure across scales can be verified using the scalogram of phase. In Figure 2, we investigate the performance of some structural edge and texture, which are marked with red and yellow circles. As shown, a rich redundancy of edge structures can be found across large scale space, which is consistent with local self-similarity assumption. However, it is not stable for textural regions since their phases vary a lot beyond a small scale range, which implies textures exhibit low self-similarity across scales. This phenomenon can be seen in Figure 3, where the up-scaled line structures are fine and smooth but the textures are wrongly hallucinated with false line-like artifacts. One possible solution to the example deficiency problem is to enlarge local searching window size for including more alternatives, but it is a time-consuming way with limited improvement.

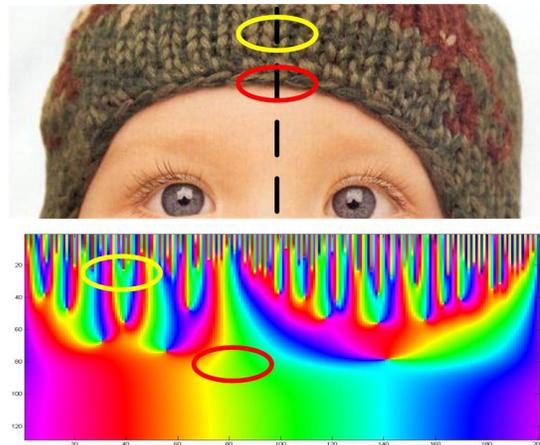


Fig. 2: A one dimensional signal that contains both step structure and texture, and its phase scalogram. Each row of the scalogram is the result of convolving the signal with a quadrature pair of wavelets at a certain scale and the vertical axis corresponds to a logarithmic frequency scale.

3. PIXEL CLASSIFICATION VIA PHASE CONGRUENCY

According to the former analysis, it is difficult to derive a universal assumption for effectively selecting good examples for every types of structures. The works in [10, 9] used a high-level texture discrimination and performed the synthesis via user-supplied example database, but they involved too much human interaction. Different from performing the content-based segmentation, we employ the phase information of pixels to classify them into three structural types: edges, textures

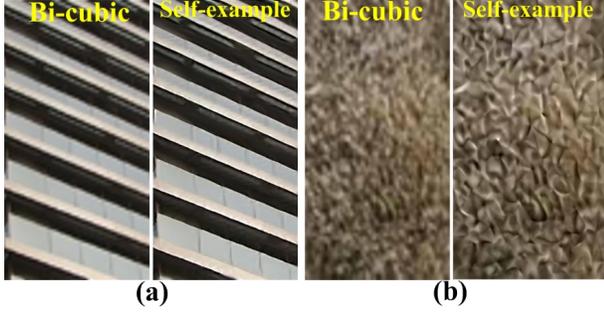


Fig. 3: 4X results of Bi-cubic versus Local self-example on (a) Line structure and (b) Textures.

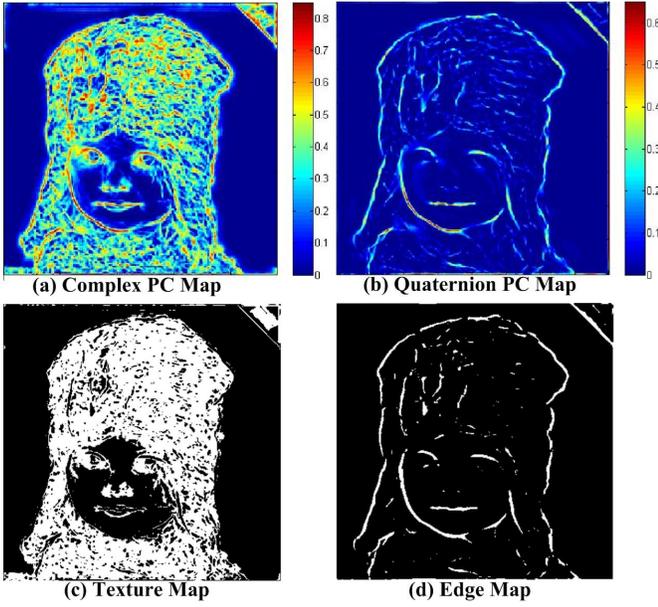


Fig. 4: Phase congruency map and computed texture/edge map.

and flat regions.

The key idea of our classification lies in that image features can be investigated in a viewpoint of phase congruency. Given a square wave which consists of infinite number of sine Fourier components, we can observe that they are exactly in phase at the point of the upward (downward) step at the angle of 0 (180) degree, while all other points in the square wave exhibits low phase consistency. Besides step edges, a wide range of feature types give rise to points of high phase congruency (PC), like lines, singularities and corners.

The measure of phase congruency[11] at point x is calculated as,

$$PC(x) = \frac{|E(x)|}{\sum_n A_n(x)}, \quad (1)$$

which is the ratio of Local Energy $|E(x)|$ over the sum of the amplitude of every Fourier component $A_n(x)$.

In [12, 13], the local frequency information of (1) is obtained using banks of complex Gabor wavelets tuned to dif-

ferent spatial frequencies. However, the complex phase congruency map tends to assign image details with a rather high importance, as shown in Figure 6(a), which makes it hard to differentiate edges from them. Thus, some efforts need to be made to improve the phase congruency model for presenting a hierarchical saliency of different image structures.

The previous works in [14] employ quaternion wavelet transform (QWT) to compute phase congruency instead of complex wavelet transform (CWT). The quaternion Gabor wavelet function is formulated as,

$$G^q(x, \mathbf{u}, m) = \frac{uv}{2\pi m^2 \sigma_f^2} e^{-0.5(\frac{xu}{m\sigma_f})^2 - 0.5(\frac{yv}{m\sigma_f})^2 - i2\pi ux - j2\pi vy},$$

where $\mathbf{u} = (u, v)^T$ is the radial center frequency of the filter, and the Gaussian envelope is truncated by the window $x \in [-\frac{m}{2u}, \frac{m}{2u}]$, $y \in [-\frac{m}{2v}, \frac{m}{2v}]$. Here, the parameter m stands for the number of wavelength included in the window, and σ_f is the fraction of the window size corresponding to one standard deviation of the Gaussian envelop. Further, $G^q(x, \mathbf{u}, m)$ can be reformed as $G^q(x, \mathbf{u}, m) = q_{ee} - iq_{oe} - jq_{eo} + kq_{oo}$, where i, j, k are the imaginary parts of quaternion systems which obey the rule of $i^2 = j^2 = k^2 = ijk = -1$, and the filters $q_{ee}, q_{oo}, q_{oe}, q_{eo}$ are formulated as,

$$q_{ee} = A \cos 2\pi u x \cos 2\pi v y, q_{oe} = -A \sin 2\pi u x \cos 2\pi v y, \\ q_{eo} = -A \cos 2\pi u x \sin 2\pi v y, q_{oo} = A \sin 2\pi u x \sin 2\pi v y,$$

where the amplitude A of these frequency components is calculated as follows,

$$A = \frac{uv}{2\pi m^2 \sigma_f^2} e^{-0.5(\frac{xu}{m\sigma_f})^2} e^{-0.5(\frac{yv}{m\sigma_f})^2}.$$

The corresponding components of local energy in a given LR image $I(x, y)$ can be calculated across scales,

$$ee(x, y) = I(x, y) * q_{ee}, eo(x, y) = I(x, y) * q_{eo}, \\ oe(x, y) = I(x, y) * q_{oe}, oo(x, y) = I(x, y) * q_{oo},$$

$$E(x, y) = \sqrt{(\sum_n ee(x, y))^2 + (\sum_n eo(x, y))^2 + (\sum_n oe(x, y))^2 + (\sum_n oo(x, y))^2}.$$

And the sum of amplitudes of all the frequency components can be expressed by

$$\sum_n A_n = \sum_n \sqrt{(ee(x, y))^2 + eo(x, y))^2 + oe(x, y))^2 + oo(x, y))^2},$$

where n denotes scale n that is selected by varying center frequency (u, v) . Plugging $E(x, y)$ and $\sum_n A_n$ into (1), we can calculate the quaternion phase congruency at $I(x, y)$.

One advantage of quaternion phase congruency is that it can extract the most significant structures as its congruency requires tighter consistency of three phase components across scales, while the ones extracted by complex phase congruency are determined by a single phase. Thus, the most salient

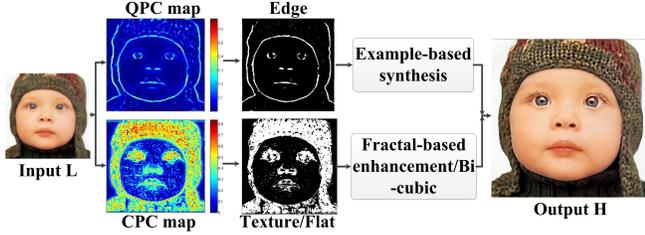


Fig. 5: Our SR framework.

structures like edges, singularities, corners can be more emphasized via quaternion phase congruency, as shown in Figure 4(b). Moreover, phase congruency is a dimensionless measurement that is invariant to illumination and contrast. A threshold of 0.3-0.4 is typically used for determining the feature significance[12]. For higher accuracy, we apply the Otsu’s method[15] to calculate the optimal threshold for picking up the most salient structures. The edge classification result can be seen in Figure 4(d). Another advantage of using quaternion Gabor wavelet is that the complex phase congruency map can be obtained as a co-product during quaternion phase congruency computing[14]. By excluding the extracted edge regions, we further use the complex PC map to efficiently separate the left details and flat regions. The computed binary texture map of “child” is shown in Figure 4(c).

4. OUR SUPER-RESOLUTION FRAMEWORK

Given the classification map that divides given LR image into edges, textures, and flat regions, we can now apply different SR strategies on the proper parts. The adaptive SR framework is presented in Figure 5.

4.1. Edge Synthesis using Local Self-Examples

Compared with most edge-focused SR methods that exploit statistical gradient priors to reproduce sharp edges[1, 2, 3], the self-example-based synthesis is able to present more fine edges, as shown in the lip part of Figure 1. One reason is that the heavily unconstrained interpolation before gradient transformation would make edges diffuse. On the contrary, the redundancy of self-examples across scales can provide sufficient high-frequency band for reproducing fine-scale structures. Thus, the self-example-based method is chosen for up-sampling the edges in our scheme.

Assuming the edge map of the input LR image L is E , the up-sampled image H_E with emphasis on edges can be calculated in an iterative way, as the local structure invariance property holds better between successive small scales[8],

$$Y_0(L * E) \xrightarrow{s_1} Y_1 \xrightarrow{s_2} \dots \xrightarrow{s_i} Y_i \xrightarrow{s_{i+1}} Y_{i+1} \dots \rightarrow Y_{out}(H_E), \quad (2)$$

where Y_i is the i -th synthesized image, s_i is the small magnification factor for the i -th stage.

4.2. Fractal-based Texture Enhancement

Loss of details often leads to unnatural images with large homogeneous texture regions. Those edge-directed SR approaches[1, 2, 3] emphasize little on textures. As for the example-based SR methods, they suffer either deficiency of self-similar examples or strong reliance on the external database. To produce realistic textures, the scale invariance of fractal features in the local gradient field was applied in[4].

As a common tool in texture analysis, fractal analysis looks into the texture regions from a geometrical view. Regarding the textures as a fractal set, the fractal dimension is supposed to remain unchanged to bi-Lipschitz transformation, which infers the up-scaling has no effect on its dimension. By choosing gradient as the fractal measures, the invariance property can be turned into gradient recovering,

$$grad(H(x)) = T(x)\beta \frac{\|grad(Y_i(x))\|}{\|grad(Y_i(x))\|^\alpha + \epsilon} grad(H_E(x))^\alpha + (1 - T(x))grad(H_E(x)), \quad (3)$$

where x denotes pixel position, T is the binary texture map, H is the final SR output, and ‘grad’ extracts the gradient. Parameters α and β jointly keep the scale invariance of fractal dimension and fractal length of local gradients[4]. The effect of detail enhancement can be seen in the texture parts of Figure 1(c).

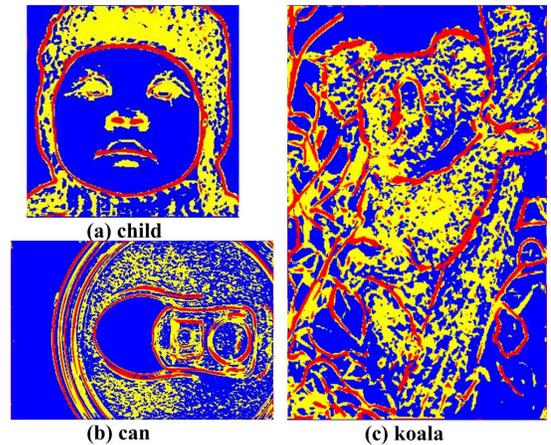


Fig. 6: The pixel classification map, where “red” denotes edges, “yellow” denotes the textures and the left flat regions are marked with “blue”.

5. EXPERIMENTS AND DISCUSSIONS

For evaluating the performance of the proposed pixel classification method and the super-resolution scheme, we compare our approach with the state-of-art methods on some common SR testing images. In Figure 6, we show the image decomposition results using quaternion phase congruency, where the

most salient edges are highlighted with red lines and the texture regions are colored with yellow. We observe that the results are consistent with human perception.

Then we show our SR results as well as those of the state-of-art methods in Figure 7,8,9. It can be observed that the edge-directed methods like [1, 2, 3] present smooth edges but tend to leave large textural areas over-smoothed; The self-example-based methods in [7, 8] can reproduce more fine lines but meanwhile introducing some line-like artifacts into textural regions. On the contrary, our method is able to provide both fine edges and enhanced textures due to performing proper SR methods on corresponding regions. The advantages can be seen more clearly in the parts framed by red rectangles and yellow ones. Besides, we provide the RMSE (Root Mean Square Error) and SSIM (Structural Similarity) values of "child" in Table 1. These objective quality assessments also show the proposed method is comparable to others¹. Moreover, as we only perform example matching on the classified edge regions, the computation load can be reduced compared with other example-based SR methods depending on the amount of edge structures, as listed in the time record of Table 1². (The test platform is MATLAB on Intel Core-i7 CPU with 8GB memory.) Note

Table 1: RMSE/SSIM/TIME of different 4X methods on "child" of size 128×128

Method	RMS /SSIM /TIME(s)	Method	RMS /SSIM /TIME(s)
Bi-cubic	15.85 / 0.71 / 0.3	Freeman02[5]	23.48 / 0.58 / 1228
Glasner09[7]	23.44 / 0.58 / -	Fattal11[8]	23.85 / 0.59 / 124
Fattal07[3]	23.40 / 0.62 / 6	Xu13[4]	25.14 / 0.56 / 7.4
Shan08[1]	17.92 / 0.68 / 12	Ours	23.23 / 0.60 / 106

6. CONCLUSION

In this paper, we analyze the limitation of example-based SR methods using scale analysis. For better performance on textures, image classification becomes necessary for breaking the limits. Different from the methods that semi-automatically segment textures according to the contents, the proposed pixel-classification method applies phase congruency map, dividing images into edge regions, textural regions and flat regions. By performing self-example-based synthesis and fractal-based enhancement on proper regions, our SR method is able to present both fine lines and vivid textures. In the future work, we will take the noise resistance of phase congruency into consideration for upsampling noisy LR images.

¹Note Bi-cubic and Shan08 however get the highest evaluations but with worse visual results, one reason is that they well maintain the consistency with LR images while other example-based methods might produce some pixel-shift effects as they focus on high-frequency recovery.

²The time of Fattal07 is reported by [3] in C++ and that of Shan08 is calculated using their software under CPU mode; Glasner09 is lack of computation time.

7. REFERENCES

- [1] Q. Shan, Z. Li, J. Jia, and C.K. Tang, "Fast image/video up-sampling," in *ACM Transactions on Graphics (TOG)*. ACM, 2008, vol. 27, p. 153.
- [2] J. Sun, Z. Xu, and H.Y. Shum, "Gradient profile prior and its applications in image super-resolution and enhancement," *Image Processing, IEEE Transactions on*, vol. 20, no. 6, pp. 1529–1542, 2011.
- [3] R. Fattal, "Image upsampling via imposed edge statistics," *ACM Transactions on Graphics (TOG)*, vol. 26, no. 3, pp. 95, 2007.
- [4] H.T. Xu, G.T. Zhai, and X.K. Yang, "Single image super-resolution with detail enhancement based on local fractal analysis of gradient," in *Circuits and Systems for Video Technology, IEEE Transactions on*. IEEE, 2013 (to appear).
- [5] W.T. Freeman, T.R. Jones, and E.C. Pasztor, "Example-based super-resolution," *Computer Graphics and Applications, IEEE*, vol. 22, no. 2, pp. 56–65, 2002.
- [6] J. Yang, J. Wright, T.S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *Image Processing, IEEE Transactions on*, vol. 19, no. 11, pp. 2861–2873, 2010.
- [7] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in *Computer Vision, 2009 IEEE 12th International Conference on*. IEEE, 2009, pp. 349–356.
- [8] G. Freedman and R. Fattal, "Image and video upscaling from local self-examples," *ACM Transactions on Graphics (TOG)*, vol. 30, no. 2, pp. 12, 2011.
- [9] Y.W. Tai, S. Liu, M.S. Brown, and S. Lin, "Super resolution using edge prior and single image detail synthesis," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 2400–2407.
- [10] Y. HaCohen, R. Fattal, and D. Lischinski, "Image upsampling via texture hallucination," in *Computational Photography (ICCP), 2010 IEEE International Conference on*. IEEE, 2010, pp. 1–8.
- [11] M Concetta Morrone, John Ross, David C Burr, and Robyn Owens, "Mach bands are phase dependent," *Nature*, vol. 324, no. 6094, pp. 250–253, 1986.
- [12] Peter Kovési, "Image features from phase congruency," *Videre: Journal of computer vision research*, vol. 1, no. 3, pp. 1–26, 1999.
- [13] Peter Kovési, "Phase congruency detects corners and edges," in *The Australian pattern recognition society conference: DICTA 2003*, 2003.
- [14] Yi Xu, Xiaokang Yang, Li Song, Leonardo Traversoni, and Wei Lu, "Qwt: Retrospective and new applications," in *Geometric Algebra Computing*, pp. 249–273. Springer, 2010.
- [15] Nobuyuki Otsu, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, no. 285–296, pp. 23–27, 1975.

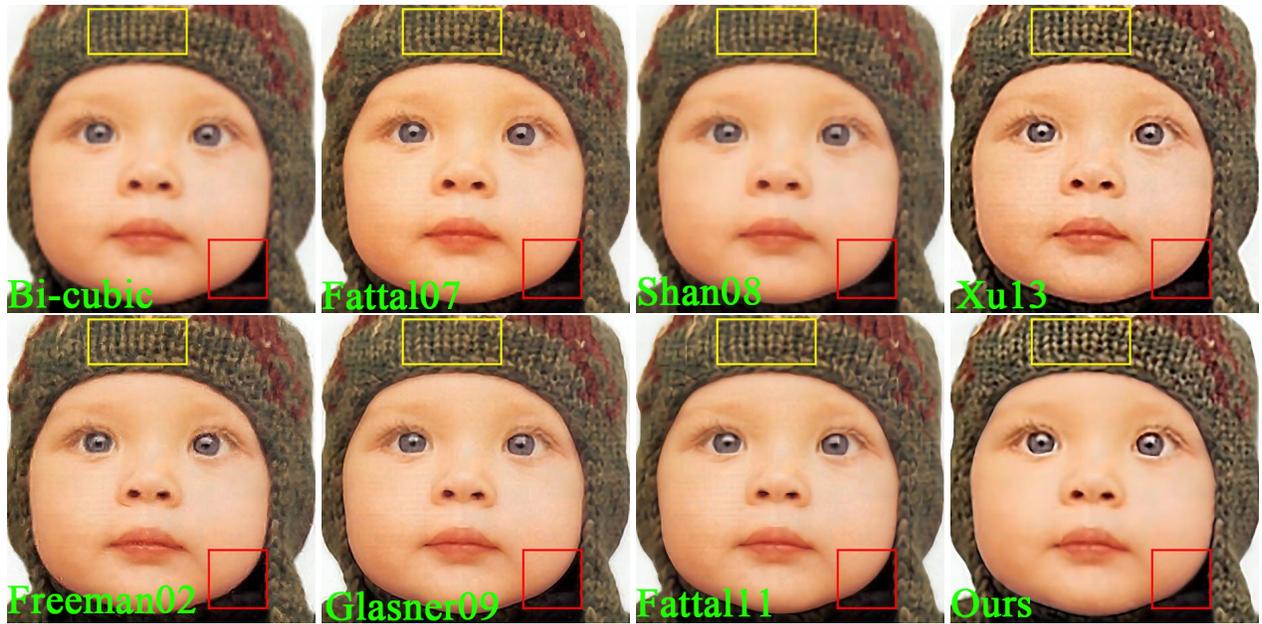


Fig. 7: Comparisons of 4X results on "child".

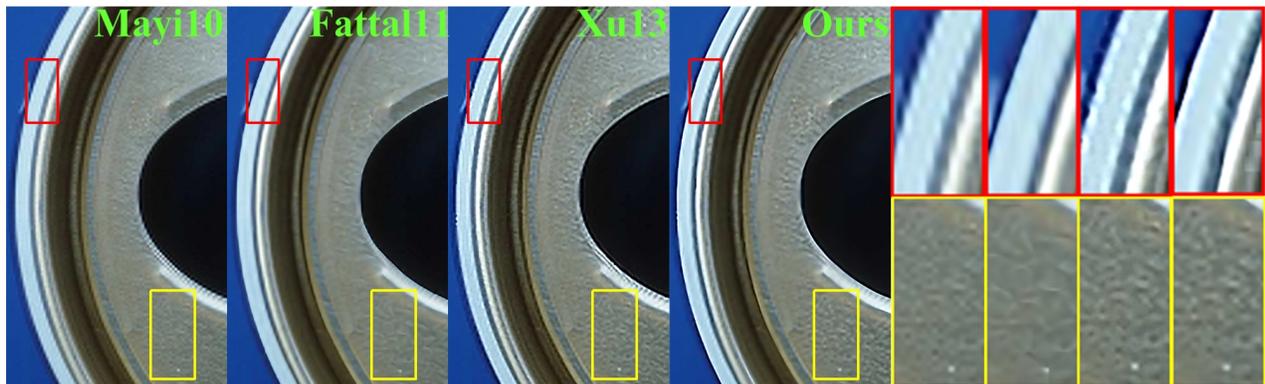


Fig. 8: Comparisons of 4X results on "can" using different methods.



Fig. 9: Comparisons of 4X results on "koala" using different methods.