PARAMETER ESTIMATIONS FOR SUPER RESOLUTION BASED ON THE LAPLACIAN PYRAMID REPRESENTATION

Shuai Yuan, Yasumasa Takahashi, Akira Taguchi***

Department of Electrical and Electronic Engineering, Musashi Institute of Technology
1-28-1 Tamazutsumi, Setagaya-ku, Tokyo 158-8557, Japan
yuan@sp.ee.musashi-tech.ac.jp, *takahashi@sp.ee.musashi-tech.ac.jp, **ataguchi@eng.musashi-tech.ac.jp

ABSTRACT

Image enlargement with higher frequency components prediction can be called as super resolution. In some print or copy applications, arbitrary scale enlargement is necessary. For obtaining better enlargement results, we have proposed one arbitrary scale super resolution method based on the Laplacian Pyramid representation [5]. In this method there are two parameters in the higher frequency prediction processing. In fact, the exact relationship of them is very complex. To expound the meaning of these two parameters for arbitrary scale super resolution, in this paper we will do a mathematic analysis via the unit step signal. But we also point out that theoretical analysis maybe can't get the best effect for the natural image enlargement and in actual applications the experiential way should not be disregarded, too.

Keywords: Laplacian Pyramid, image enlargement, super resolution, step signal, natural image.

Ⅰ. INTRODUCTION

In order to watch the details better or get the bigger size for print or copy processing about the digital images, the enlargement is necessary. Some linear methods are used in many applications, such as the Cubic and the Bilinear interpolations [1]. But the problem of them is the enlarged images will appear blurry because there is no power in the higher spatial frequency band. For solving this problem, several enlargement methods with higher frequency components prediction (super resolution methods) were proposed [2]-[4]. We also proposed one [3] which is based on the Laplacian Pyramid representation (LP based). However, these LP based methods [3],[4] can only be used for expanding an image up by a factor of two in size (so called "zoom in"). In fact we also need the enlarged scales as 1.15 times, 1.41 times and other enlargement scales for some print or copy applications. For obtaining this ability, based on the LP representation we proposed an arbitrary scale super resolution method [5], which utilizes the calculation of non-integer LP stage.

In the [5], there are two parameters α_r and c_r in the higher frequency prediction processing. The exact relationship of them is complex. And the [5] didn't give out it on mathematic level. To expound the mathematic meaning of these two parameters, in this paper we will do a theoretical analysis via the unit step signal. We know that theoretical analysis maybe can't get the best effect for the natural image enlargement and in actual applications the experiential way should not be disregarded. However, show the mathematic base of this arbitrary scale super resolution algorithm is still important.

The arbitrary scale LP super resolution algorithm is described in SectionⅡ. In SectionⅢ-A, we present the parameter evaluation from the analysis of unit step signal. SectionⅢ-B reviews the experiential way via natural image. The simulation results and conclusion are included in Section Ⅳ and Section Ⅴ, respectively.

Ⅱ. THE ARBITRARY SCALE ENLARGEMENT BASED ON LAPLACIAN PYRAMID REPRESENTATION

2.1 Gaussian and Laplacian Pyramid Representations

The image pyramid includes Gaussian- and Laplacian-pyramid representations, which are described in the [6]. Here, we refer to the Gaussian pyramid components as {G_0, G_1, G_2,..., G_n}. The "G_0" shows original input image and the "G_{n+1}" is given by

$$\overline{G}_{n+1} = W * G_n \tag{1}$$

$$G_{n+1} = DownSample(\overline{G}_{n+1}) \tag{2}$$

Where, the * is the convolution operator. "W" is the half band low-pass Gaussian filter.

The Laplacian pyramid is composed by high-pass filtered versions (high frequency components) of the corresponding Gaussian pyramid components. And each stage {L_0,L_1,…,L_n} of the pyramid is constructed by

$$L_n = G_n - Expand\ (G_{n+1}) \tag{3}$$

$$Expand\ (G_{n+1}) = 4 \times (W * G^0_{n+1}) \tag{4}$$

$$G^0_{n+1} = UpSize\ (G_{n+1}) \tag{5}$$

G^0_{n+1} is the image which is obtained from interpolating zeros into G_{n+1}. The size of G_n is 2^{-n} times of the original stage G_0.

2.2 The arbitrary scale image enlargement method

If we are able to achieve the higher frequency LP stage L_{-1} from the L_0, ideal 2 times super resolution image G_{-1} can be calculated as the following equation:

$$G_{-1} = L_{-1} + Expand(G_0) \quad (6)$$

We found that the relationship of the LP transmission exists not only among the integer stages but also in the un-integer stages. The "r" in enlargement scale $S = 2^{-r}$ can be looked as the image pyramid scale. Then, there is the relationship equation, between the arbitrary larger size stages (integer or un-integer) and the original stage (input image) G_0, as:

$$G_r = L_r + Expand_r(G_0) \qquad (r < 0) \quad (7)$$

The size of G_r is $S = 2^{-r}$ times of the G_0. Next, we will show how to calculate $Expand_r(G_0)$ and predict L_r. We'd like to use one-dimensional signal to explain this calculation. From the Eq.(7), we know the size of $Expand_r(G_0)$ ($= \tilde{G}_r$) is 2^{-r} times of the G_0. And the image $\tilde{G}_r$ should keep the same frequency components as the G_0 image. Utilizing the Gaussian all pass filter for the up-sample processing, we can approximately retain all of the G_0 image frequency components. So that, if we can expand the G_0 image to the expected size by the Gaussian all pass filter processing, the $\tilde{G}_r$ can be obtained. We refer to the Gaussian function as $W(x)$. Thus, there is

$$W(x) = \frac{1}{2\sqrt{\rho\pi}} e^{-\frac{x^2}{4\rho}} \quad (8)$$

$$\rho = \sigma^2 / 2 \quad (9)$$

There is the $\rho = 4/9\pi$ for Gaussian all pass filter. In the following, we refer to the $g_0(i)$ $(i = 1,2,3,\ldots,M)$ as the pixel value of original image G_0 and the $\tilde{g}_r(j)$ $(j = 1,2,3,\ldots,(M \times 2^{-r}))$ as the pixel value of $\tilde{G}_r$. From the following equations, we can get the image $\tilde{G}_r$.

$$\tilde{g}_r(j) = \sum_{n=-2}^{2} W_{all}(n) \cdot g_0(\mathrm{int}(\frac{j}{2^{-r}}) + n) \quad (10)$$

$$W_{all}(n) = \frac{3}{4} e^{-\frac{9\pi}{16}(\Delta k + n)^2} \quad (11)$$

$$\Delta k = \mathrm{int}(\frac{j}{2^{-r}}) - (\frac{j}{2^{-r}}) \quad (12)$$

The int(•) is for taking the integer part. The L_0 enlarged version $Expand_r(L_0)$ ($= \tilde{L}_r$) can be calculated from the similar way:

$$\tilde{l}_r(j) = \sum_{n=-2}^{2} W_{all}(n) \cdot l_0(\mathrm{int}(\frac{j}{2^{-r}}) + n) \quad (13)$$

Where, $\tilde{l}_r(j)$ is the value of $\tilde{L}_r$.

From the [3][4], it is known that sharpening the $Expand_r(L_0)$ near the zero crossing can get the super resolution LP stage (higher frequency components). After predicting L_r from $\tilde{L}_r = Expand_r(L_0)$, the arbitrary scale super resolution enlarged image will be achieved. The L_r prediction processing is as the following:

$$\bar{l}_r(j) = \alpha_r \times \begin{cases} T & \text{if } T \le \tilde{l}_r(j) \\ \tilde{l}_r(j) & \text{if } -T < \tilde{l}_r(j) < T \\ -T & \text{if } -T \ge \tilde{l}_r(j) \end{cases} \quad (14)$$

$$T = (1 - c_r) \times \tilde{L}_{r\max} \quad (15)$$

$$L_r = \bar{L}_r - W_r * \bar{L}_r \quad (16)$$

There are two parameters, α_r and T, in the Eq.(14). The T is threshold, which is decided by the clipping parameter c_r. The "α_r" is a constant which can make the slope of $\tilde{L}_r$ sharper. Both of them should be pre-determined. The $\bar{l}_r(j)$ is the value of $\bar{L}_r$. And the $\bar{L}_r$ is high-passed (Eq.(16)) in order to leave only the higher frequency components in the L_r, utilizing the variable band low pass Gaussian filter W_r, which depends on the enlargement scale. Next section we will show how to evaluate the parameter α_r and c_r.

Ⅲ. PARAMETER ESTIMATION

A. Theoretical Evaluation – A Special Case

In the following, we will expand the analysis of [4] to arbitrary scale super resolution. In our theoretical analysis, the last HPF processing "Eq.(16)" is ignored.

It is known that the step signal can show the relationship of super resolution very well, which is educed by the variation of sampling distance. Therefore, the unit step signal is used as an ideal step edge input for our analysis in special domain. It is known that the optical system's unit impulse signal response is PSF (point spread function), which is a Gaussian low-pass function with the standard deviation σ_0 (normally $\sigma_0 = 0.9$). Thus, the optical system response of ideal unit step edge will be the convolution of PSF and ideal unit step signal, which is equal to the integral probability distribution of the PSF from $-\infty$ to x. We look this response as Gaussian pyramid stage G_0 and it will be sampled as the distance of 1 for the A/D processing. Then, the ideal 2^{-r} times super resolution Gaussian pyramid stage G_r should be sampled at a distance of $1/2^{-r}$ and the value of it is equal to the integral probability distribution of Gaussian low-pass function from $-\infty$ to x, with the standard deviation $\sigma_r = (\sigma_0)/(2^{-r})$.

In the Image Pyramid representation algorithm, the half band Gaussian low-pass filter that we used is approximation of a normalized Gaussian function with a standard deviation of $\sigma_w = 1.0$. Then, the standard deviation of the Gaussian pyramid stage G_1 is $\sigma_1 = 1.345$ ($\sigma_1{}^2 = \sigma_0{}^2 + \sigma_w{}^2$). We define the unit step edge as $U(x)$, and its basic Laplacian pyramid stage as $L_0\{U(x)\}$. There is

$$L_0\{U(x)\} = G_0\{U(x)\} - G_1\{U(x)\} \cdot \tag{17}$$

The $G_i\{U(x)\}$ is Gaussian pyramid i th stage of the ideal unit step edge. Where,

$$G_i\{U(x)\} = F(x/\sigma_i) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{x^2}{2\sigma_i^2}} dx \cdot \tag{18}$$

From the Eq.(17) and the Eq.(18) we can get

$$L_0\{U(x)\} = F(x/\sigma_0) - F(x/\sigma_1) \cdot$$

The x position, which lets the $L_0\{U(x)\}$ with the max value, is $x = x_{0\max}$. At this position there is

$$\left(\frac{dL_0\{U(x)\}}{dx}\right)_{x=x_0 \max} = L_0'\{U(x)\}_{x=x_0 \max} = 0 \cdot$$

Then, $\quad F'(x_0 \max/\sigma_0) - F'(x_0 \max/\sigma_1) = 0$

$$\frac{1}{\sqrt{2\pi}\sigma_0} e^{-\frac{x_{0\max}^2}{2\sigma_0^2}} - \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{x_{0\max}^2}{2\sigma_1^2}} = 0 \tag{19}$$

The $F'(x/\sigma_i) = \dfrac{dF(x/\sigma_i)}{dx}$ is Gaussian function with the standard deviation σ_i. The solution of Eq.(19) is

$$x_0 \max = sqrt(2\log(\sigma_1/\sigma_0)/(1/\sigma_0{}^2 - 1/\sigma_1{}^2))$$

From $\sigma_0 = 0.9$ and $\sigma_1 = 1.345$, we can get

$$x_0 \max = sqrt(2\log(1.345/0.9)/(1/0.81 - 1/1.8)) = 1.085$$

and

$$L_0\{U(x)\}_{(x=x_0 \max)} = F(1.085/0.9) - F(1.085/1.345) = 0.097$$

The max value searching-calculation of ideal $L_r\{U(x)\}$ is done in a similar way using the σ_0 and the σ_r . There are

$$x_r \max = sqrt(2\log(\sigma_0/\sigma_r)/(1/\sigma_r{}^2 - 1/\sigma_0{}^2))$$

and

$$L_r\{U(x)\}_{(x=x_r \max)} = F(x_r \max/\sigma_r) - F(x_r \max/\sigma_0) \tag{20}$$

Because the predicted Laplacian component should have the same max value with the ideal $L_r\{U(x)\}$. Thus, we can get the following equation:

$$\alpha_r \times (1 - c_r) \times L_0\{U(x)\}_{(x=x_0 \max)} = L_r\{U(x)\}_{(x=x_r \max)} \tag{21}$$

The slope value of $L_0\{U(x)\}$ is its differential calculus $L_0'\{U(x)\}$. And at the zero-crossing position ($x = 0$), we would like to fit the predicted LP stage to the ideal $L_r\{U(x)\}$.

Thus,

$$\alpha_r \times L_0'\{U(x)\}_{(x=0)} = L_r'\{U(x)\}_{(x=0)} \tag{22}$$

Where,

$$L_0'\{U(x)\}_{(x=0)} = F'[(x=0)/\sigma_0] - F'[(x=0)/\sigma_1]$$
$$= (1/(\sigma_0 \times \sqrt{2\pi}) - 1/(\sigma_1 \times \sqrt{2\pi}))$$
$$L_r'\{U(x)\}_{(x=0)} = F'[(x=0)/\sigma_r] - F'[(x=0)/\sigma_0]$$
$$= (1/(\sigma_r \times \sqrt{2\pi}) - 1/(\sigma_0 \times \sqrt{2\pi}))$$

Then, from the Eq.(22) and Eq.(21), the parameters can be obtained as the following:

$$\alpha_r = 3.012 \times (2^{-r} - 1)$$
$$c_r = 1 - (L_r\{U(x)\}_{(x=x_r \max)} / (\alpha_r L_0\{U(x)\}_{(x=x_0 \max)}))$$
$$= 1 - (L_r\{U(x)\}_{(x=x_r \max)} / (0.293 \times (2^{-r} - 1)))$$

Where, the $L_r\{U(x)\}_{(x=x_r \max)}$ is obtained from Eq.(20).

B. Parameter Estimation via Natural Images

In the [5], we have done experiential estimation via natural image. Here we will review it briefly.

This method is mainly used for the natural image processing. Therefore, estimating the parameters via natural image is also necessary. In this scenario, we can account for the arbitrary scale LP algorithm in its entirety, including the last high pass processing "Eq(16)". Seven different kinds of gray level natural images [7] are used, which are containing the "human" pictures, the "script" pictures and the "scenery" pictures, etc. For beginning the experiential analysis, we low-pass filtered the existing 7 images and down-sampled them to the $1/2^{-r}$ times of their original sizes. Next, utilize different parameter sets to enlarge the down-sampled images to their original sizes. And the enlarged results will compare with the original images via minimizing MSE values to estimate the most suitable parameter sets.

We have found that the parameter "c_r" almost doesn't depend on the enlargement scale. And when the value of "c_r" is 0.45, there are always good results in MSE. The parameter α_r is variable with the enlargement scale. The variation curve is shown as the Fig.1. From the Fig.1, we know that the parameter α_r will become small when the enlarged scale becomes large.

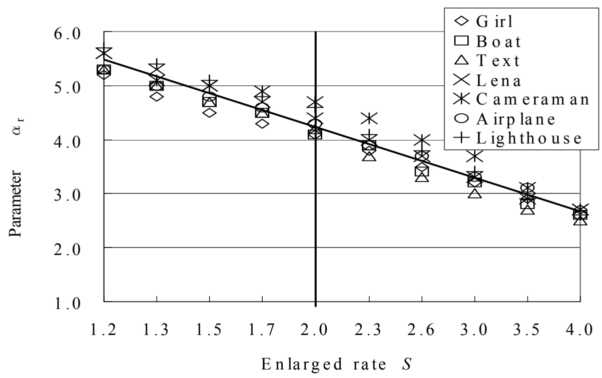


Fig.1 The relationship of parameter α_r and enlarged scale

IV. SIMULATION

In this algorithm, we have presented two kinds of parameter estimation methods. One is via the step signal theoretical analysis and the other is via the natural images.

We show these two kinds of parameters' variation lines in the Fig.2-(a,b).

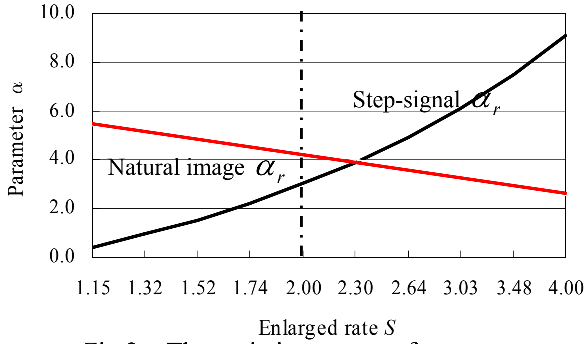


Fig.2-a The variation curves of parameter α_r

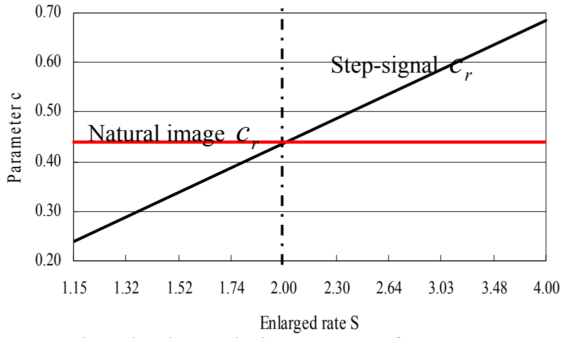


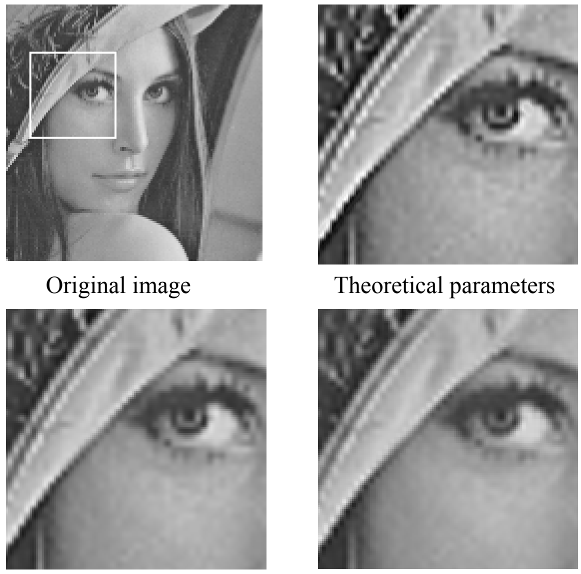
Fig.2-b The variation curves of parameter c_r

In fact only near the 2 times enlargement scale, the parameter sets obtained from the mathematic analysis are close to the parameter sets obtained from the experiential way. We know most of the natural image components are not the step edges. This reason makes that the enlarged results with theoretical parameter sets are worse than the results with experiential parameter sets, except near the 2 times enlargement scale.

In Fig.3 we show the simulation results with these two kinds of parameter sets and Cubic method at 3 times enlargement scale. Since the Cubic method doesn't belong to super resolution method, from the simulations you can see that the Cubic enlarged image was blurred. The enlarged image with the theoretical parameter set is enhanced a little excessively and it makes the edges shown some hard. For getting good balance for natural image super resolution processing, some degree experiential parameter set is better. But the theoretical analysis for parameter evaluation is also very important from algorithm level.

V.CONCLUSION

In this paper we have shown the whole theoretical analysis for arbitrary scale LP super resolution method. Even the simulation results show that for natural image the experiential parameter estimation is important. Explaining the mathematic base of this algorithm is still necessary.



Original image Theoretical parameters

Experiential parameters Cubic method

Fig.3 The simulation results

VI. REFERENCES

[1] M.Takagi and A.Shimoda, "*Image analysis handbook*", pp.441-444, University of Tokyo Press, 1991. (In Japanese)

[2]A.Tanaka, H.Imai, M.Miyakoshi, and A.Date, "Enlargement of digital image using multiresolution analysis", *IEICE Trans. on Information and Systems*, vol.J79-D-II, no.5, pp.819-825, May.1996 (In Japanese)

[3] Y.Takahashi and A.Taguchi, "An Enlargement method of digital image based on Laplacian pyramid representation", *SPIE vol.3961: Nonlinear Image Processing XI*, pp.163-169, 2000.

[4] H.Greenspan, C.H.Anderson and S.Akber, "Image enhancement by nonlinear extrapolation in frequency space", *IEEE Trans. on Image Processing*, Vol.9, No.6, June 2000.

[5] Yasumasa Takahashi and Akira Taguchi, "An Arbitrary Scale Image Enlargement Method with the Prediction of Higher Frequency Components", *IEICE Trans. on Fundermentals*, vol.J82-A, no.9, pp.1192-1201, Sept. 2001. (In Japanese)

[6] P.J.Burt and E.A.Andelson, "The Laplacian pyramid as a compact image code ", *IEEE Trans. on Commun*,Vol.31, pp.532-540,1983.

[7] http://www.sp.ee.musashi-tech.ac.jp/app.html