

Video Super-resolution Reconstruction Algorithm Based on Total Variation Regularization

Ling Tang

Sichuan University of Science & Engineering, Zigong 643000, China.
fashion-kitty@163.com

A video super-resolution (SR) reconstruction algorithm based on total variation (TV) regularization is proposed in this paper. With the analysis of image degradation, the image formation model is introduced. Total variation method is chosen to regularize the ill-posed problem and reconstruct the SR image. Instead of computing the complicated nonlinear partial differential equation (PDE) of the TV regularization in SR reconstruction, the quadratic upper bound function minimization method is used to efficiently solve the optimization. Results of the experiments show that the proposed algorithm has better subjective visual effect, and better peak signal to noise ratio, structural similarity and convergence, which confirms the effectiveness of the method.

Keywords : Super-resolution; Video reconstruction; Total variation regularization; Quadratic upper bound function

1. Introduction

In recent years, super-resolution reconstruction technology has become a hot issue in the field of computer vision and image processing, which has aroused wide attention of the researchers. The idea of super-resolution image reconstruction is first proposed by Huang and Tsai in 1984. Over the next thirty years, the super-resolution reconstruction technique has been developed rapidly, and many effective algorithms have been proposed (Chung (2013), LINGALA et al. (2011), Watanabe et al. (2012)). However, with the popularity of video capture devices, people's research on super-resolution reconstruction is no longer limited to the static images, but to the videos. Video super-resolution reconstruction is combined with the sliding window model (Narayanan et al. (2007)) and sub pixel level complementary information between the consecutive frames. A high resolution (HR) (Tu (2014)) video is reconstructed from a video or multi video sequences with low quality and low resolution (LR), so as to get more details and information. Many researchers study the video super-resolution reconstruction, and have achieved many excellent results (Qin et al. (2009), Agrawal et al. (2010)).

Taking into account that the reconstruction is actually an ill-posed inverse problem, regularization method is an effective method for solving the ill-posed problem. Tikhonov regularization has smooth effects on the solutions, while the actual image has a lot of details constituted by edges and points, losing these details means the loss of information. Thus TV regularization (NG et al. (2007), JIANG et al. (2014)) can suppress the noise in image reconstruction, but not make smooth effects on the solutions, so that the edges of the solutions can be maintained. In this paper we present a video super-resolution reconstruction algorithm based on TV regularization. The problem of the reconstruction is defined as the minimization problem of the objective function with the regularization parameter. And through the quadratic upper bound function minimization, the optimization solution of the TV regularization can be achieved.

2. Video super-resolution reconstruction algorithm

2.1 Degradation model

In the process of video acquisition, many factors can cause the degradation of the video, such as the aberration of optical system, atmosphere disturbance, the motion of scene or object, and the system noise and so on. In

order to study the super-resolution reconstruction, we must establish a realistic imaging model. The relationship between the LR video frame and the original HR image can be expressed as:

$$g_k = \mathbf{DBM}_k f + n_k, \quad 1 \leq k \leq N \quad (1)$$

Where, N is the number of the LR frames involved to reconstruction, f , g_k and n_k respectively is the HR reconstruction image, the k th LR frame and noise, \mathbf{D} , \mathbf{B} and \mathbf{M}_k is the downsampling matrix, the fuzzy matrix and the motion matrix composed of the motion vector between the reference frame and the k th frame respectively.

2.2 Video reconstruction model

In the reconstruction of a frame in the LR video, the frame and the adjacent LR frames are required. When reconstruct a video, it is need to continuously output each super-resolution reconstruction frame. Therefore, we need to establish an effective video reconstruction model. The sliding window model is the most popular video reconstruction model, Fig. 1 shows the schematic diagram of the sliding window with length $L=5$, which determines the number of LR frames for reconstruction. As shown in the figure, with sliding window from the video of the first mobile terminal to the end, each frame after super-resolution reconstruction can continuously output, so as to reconstruct the entire video.

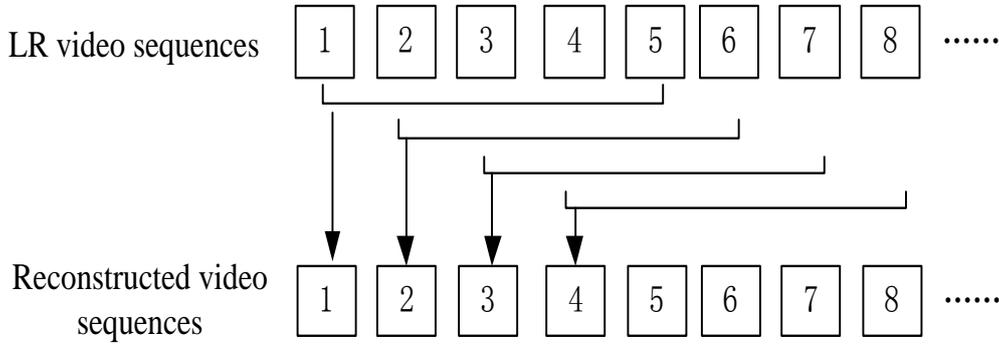


Figure 1. Video super-resolution reconstruction model ($L=5$)

2.3 Algorithm description

SR image reconstruction is an inverse problem. Because of the existence of noise, the process is ill-posed, which means no solution that satisfies the existence, uniqueness and continuous. The reconstruction results are also affected by the noise. So regularization method is usually used to make certain restrictions to the reconstructed image and overcome the ill-posed problem. The cost function is expressed as:

$$J(f) = \sum_{k=1}^N \rho(g_k, \mathbf{DBM}_k f) + \lambda \Gamma(f) \quad (2)$$

Where, λ is the regularization parameter.

The minimum solution of the first item in the formula (2) is the solution of the maximum likelihood (ML) estimation, i.e.

$$\hat{f}_{ML} = \arg \min \left[\sum_{k=1}^N \rho(g_k, \mathbf{DBM}_k f) \right] \quad (3)$$

Where $\rho(\bullet, \bullet)$ is the measure of the representation of distance in space, which is often used as a form of L_p norm.

$$\rho(g_k, \mathbf{DBM}_k f) = \| \mathbf{DBM}_k f - g_k \|_p^p \quad (4)$$

When $p=2$, the formula (3) obtains the optimal solution by the least square. The specific process is equivalent to the average value estimation of the pixels (Elad and Hel-Or (2001)). The advantage of this model is that it is easy to compute and the optimal solution can be obtained when the noise is Gaussian. But the estimation is not robust. Any abnormal points will affect the reconstruction results. When $p=1$, the equation (3) is equivalent to the

median value of the pixels (Farsiu et al. (2004)), which is not sensitive to the abnormal points. Therefore, it has strong robustness. In this paper we take $p=1$.

The second item in the formula (2) is to overcome the ill-posed problem of the regularization penalty function. The widely used penalty function is the Tikhonov, which the purpose of is to limit the total energy or spatial smoothness of the image. Because the noise and the edges of the image contain high frequency components, so Tikhonov regularization in the suppression of the noise also makes the edge of the image become fuzzy. Rudin et al. (1992) proposed a nonlinear regularization method based on total variation minimization which has good edge preserving. The penalty function of TV regularization is:

$$\Gamma_{TV}(f) = \int_{Z_f} \sqrt{|\nabla f|^2 + \beta^2} dx dy = \int_{Z_f} \sqrt{f_x^2 + f_y^2 + \beta^2} dx dy \quad (5)$$

Where, Z_f represents the image space, f_x , f_y are the first order partial derivatives of the image f in the point (x, y) in the horizontal and vertical direction. β is a constant that making $\Gamma_{TV}(f)$ differentiable when $f_x = f_y = 0$.

In summary, the TV regularization cost function is:

$$J(f) = \sum_{k=1}^N \| \mathbf{DBM}_k f - g_k \|_1 + \lambda \Gamma_{TV}(f) \quad (6)$$

At this time, the SR reconstruction problem can be transformed into the minimization of the cost function. That is:

$$\hat{f} = \arg \min_f J(f) \quad (7)$$

2.4 Algorithm solution

In this paper, the optimal solution to the TV regularization is achieved by minimizing the quadratic upper bound function of the TV. Do not need to solve the exact minimum in each step, only make the upper bound function gradually reduced. Compared with the method of traditional fixed point iteration to solve the nonlinear partial differential operator, the proposed algorithm has smaller computations.

Assume the quadratic upper bound function $Q(f | f^j)$ satisfies the following conditions:

- (1) $L(f^j) = Q(f^j | f^j)$
- (2) $L(f) \leq Q(f | f^j), f \neq f^j$
- (3) $f^{(j+1)} = \arg \min Q(f | f^j)$

The sequence $\{L(f^j), j=1,2,3,\dots\}$ is convergent. In addition, if $L(f)$ is strictly convex, $\{f^j, j=1,2,3,\dots\}$ can converge to the global minimum of $L(f)$ (WU (1983)). Make sure $Q(f^{(j+1)} | f^j) \leq Q(f^j | f^j)$, then $L(f^{(j+1)}) \leq L(f^j)$. It is very favorable when the global minimum point of the Q function is difficult to accurately be calculated.

Assume that the matrix $\mathbf{G}_x f$ and $\mathbf{G}_y f$ are represented the first order difference matrices in the horizontal and vertical directions of the image. Then the TV minimization problem of the quadratic upper bound function can be expressed as follow (Xu et al. (2012)):

$$Q(f | f^j) = \| \mathbf{DBM}_k f - g_k \|_1 + Q_{TV}(f | f^j) \quad (8)$$

Where $Q_{TV}(f | f^j) = f^T \mathbf{G}_x^T \mathbf{W}^j \mathbf{G}_x f + f^T \mathbf{G}_y^T \mathbf{W}^j \mathbf{G}_y f + c_k$, c_k is a constant.

$$\mathbf{W}^j = \text{diag} \left(\frac{\lambda}{2\sqrt{(f_x^j)^2 + (f_y^j)^2 + \beta^2}} \right) \quad (9)$$

In theory, the exact value of $f^{(j+1)}$ can be obtained by minimizing the equation (8). But this solution refers to a large linear system which makes the computation complex. So we can use the steepest descent method to

solve the problem. Only need to make $f^{(j+1)}$ satisfies $Q(f^{(j+1)} | f^j) \leq Q(f^j | f^j)$, which ensures that the quadratic upper bound function is gradually reduced during the iterations. In this way we don't need the exact solution of the equation.

The process is expressed as follows:

- (1) Use a LR image in the video sequences as a reference frame for the bilinear interpolation, in order to determine the initial value f^1 . And set the threshold ε ;
- (2) According to equation (9) calculate \mathbf{W}^j which is insert into the formula (8) to get the quadratic upper bound function $Q(f^j | f^j)$;
- (3) Update the solution of f^j by using the steepest descent method;
- (4) If $\|f^{(j+1)} - f^j\|^2 / \|f^j\|^2 \leq \varepsilon$, the iteration is over, and $f^{(j+1)}$ is the solution. Otherwise return to (2).

3. Simulations and analysis

Take three groups of video sequences of size 256×256 pixels with 10 frames per group, which are Flower, Foreman and Calendar. And different LR sequences of size 128×128 are created from them by sub-pixel shift, Gauss blur (3 × 3, variance is 0.5) and 2 times downsampling. Add Gauss white noise in the experiment. Fig.2 (a) is a HR image of the original video sequences. Fig.2 (b) is the result of a reference frame with the bilinear interpolation. Fig.2 (c) is a reconstructed image obtained by Tikhonov regularization (NGUYEN and MILANFAR (2001)). Fig.2 (d) is the result of the proposed algorithm. As seen from the results, the effect of bilinear interpolation is the worst. The reconstruction effect of the proposed algorithm is better than the Tikhonov regularization, and with better restoration to the edge details.

For the objective evaluation to the experiment results, we use the peak signal to noise ratio (PSNR) and the structural similarity (SSIM) as the evaluation index. Tables 1 and 2 give the comparisons of SSIM and PSNR with different methods and different video sequences. It can be seen that the algorithm in this paper can obtain higher PSNR and SSIM, which shows good performance.

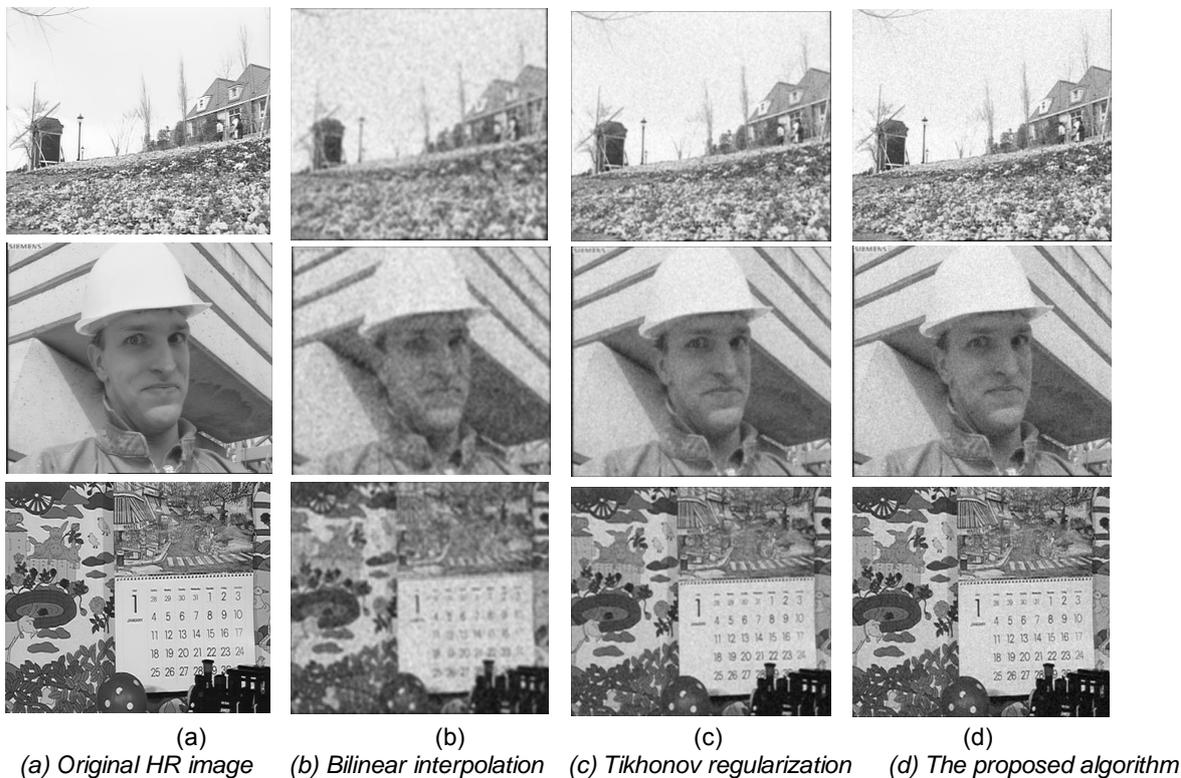


Figure 2: Comparison of reconstruction results

Table 1: Comparison of PSNR (dB)

Video	Bilinear interpolation	Tikhonov regularization	Proposed algorithm
Flower	20.342	25.527	27.876
Foreman	18.201	19.254	20.092
Calendar	21.796	23.860	25.231

Table 2: Comparison of SSIM

Video	Bilinear interpolation	Tikhonov regularization	Proposed algorithm
Flower	0.782	0.902	0.927
Foreman	0.528	0.724	0.828
Calendar	0.711	0.859	0.913

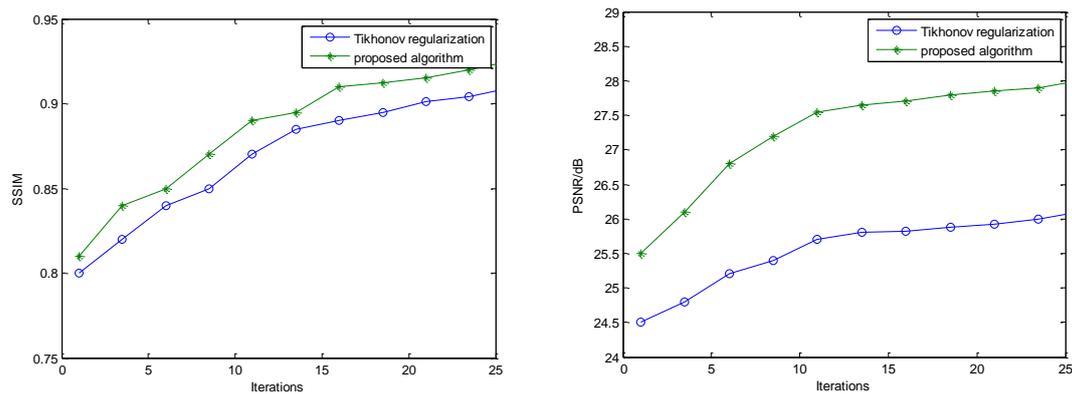


Figure 3: Performance comparison ("Flower" sequences)

4. Conclusions

In this paper, an algorithm for video super-resolution reconstruction based on total variation regularization is proposed. The optimization algorithm is used to optimize the total variation regularization, which improves the robustness of the algorithm, improves the subjective quality of the reconstructed image and the objective index. In the future the research is aim to improve the time performance of the algorithm, so as for the real-time processing of the fuzzy video images.

Acknowledgments

This work was supported by projects of Sichuan Provincial Department of Education (13ZB0138) and projects of Artificial Intelligence Key Laboratory of Sichuan Province (2013RYY02).

References

- Agrawal A., Gupta M., Veeraraghavan A., et al. 2010. Optimal Coded Sampling for Temporal Super-Resolution. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), San Francisco, CA:[s.n.], 238 (6), pp: 599-606, DOI: 10.1109/CVPR.2010.5540161
- Chu C.H. 2013. Super-resolution image reconstruction for mobile devices. Multimedia Systems. 19 (4), pp: 315-337, DOI:10.1007/s00530-012-0276-y

- Elad M., Hel-Or Y. 2001. A fast super-resolution reconstruction algorithm for pure translational motion and common space invariant blur. *IEEE Transactions on Image Processing*, 10 (8), pp: 1187-1193, DOI: 10.1109/83.935034
- Farsiu S., Robinson M.D., Elad M. et al. 2004. Fast and robust multiframe super resolution. *IEEE Transactions on Image Processing*, 13 (10), pp: 1327-1344, DOI: 10.1109/TIP.2004.834669
- Jiang X.H., Zhao X.J., Li C.J., Zhang X.S. 2014. A Super-Resolution Algorithm Based on Adaptive Weighted Total Variation. *Infrared Technology*, 36 (4), pp: 290-293
- Lingala S.G., Hu Y., Dibella E. et al. 2011. Accelerated dynamic MRI exploiting sparsity and low-rank structure:k-t SLR. *IEEE Trans on Medical Imaging*. 30 (5), pp: 1042-1054, DOI: 10.1109/TMI.2010.2100850
- Narayanan B., Hardie R.C., Barner K.E. et al. 2007. A Computationally Efficient Super-Resolution Algorithm for Video Proceeding Using Partition Filters. *IEEE Transactions on Circuits and Systems for Video Technology*, 17 (5), pp: 621-634, DOI: 10.1109/TCSVT.2007.893833
- Ng M.K., Shen H.F., Lam E.Y. et al. 2007. A total variation regularization based super-resolution reconstruction algorithm for digital video. *EURASIP Journal on Advances in Signal Processing*, 10, pp: 1155-1171, DOI: <http://dx.doi.org/10.1155/2007/74585>
- Nguyen N., Milanfar P. 2001. A computationally efficient superresolution image reconstruction algorithm. *IEEE Transactions on Image Processing*, 10 (4), pp: 573-583, DOI: 10.1109/83.913592
- Qin F.Q., He X.H., Chen W.L. 2009. Video Super-resolution Reconstruction Based on Sub-pixel Registration and Iterative Back Projection. *Journal of Electronic Imaging*, 18 (1), pp: 013007-1-013007-11, DOI: 10.1117/1.3091936
- Rudin L.I., Osher S., Fatimi E. 1992. Nonlinear total variation based noise removal algorithms. *Physcia D Nonlinear Phenomena*, 60 (1-4), pp: 259-268, DOI: 10.1016/0167-2789 (92)90242-F
- Tu J.H. 2014. A novel building boundary extraction method for high-resolution aerial image. *Review of computer engineer studies*. 1 (2), pp: 19-21
- Wu C. 1983. On the convergence properties of the EM algorithm. *The Annals of Statistics*, 11 (1), pp: 95-103, DOI:10.1214/aos/1176346060
- Watanabe M., Sakuta Y., Goto T., Hirano S., Sakurai M. 2012. Super-resolution image processing with total variation regularization and shock filters, *Consumer Electronics (GCCE)*, 2012 IEEE 1st Global Conference on, 2-5 Oct., Tokyo, pp: 570 – 571, DOI:10.1109/GCCE.2012.6379916
- Xu L., Jiang J., Wei Z.Z. 2012. Multiframe super-resolution reconstruction algorithm based on total variation regularization. *Electronic measurement technology*, 35 (1), pp: 76-79