VIDEO ENHANCEMENT USING AN ITERATIVE MULTIFRAME SRR BASED ON A ROBUST STOCHASTIC ESTIMATION WITH AN IMPROVED OBSERVATION MODEL

V. Patanavijit[†] *and S. Jitapunkul*[††]

ABSTRACT

This paper proposes a video enhancement method using a novel Super-Resolution Reconstruction (SRR) framework for real standard sequences that are corrupted by any noise models. The traditional SRR algorithms are very sensitive to their assumed model of data and noise, which limits their utility. The real noise models that corrupt the measure sequence are unknown; consequently, SRR algorithm using L1 or L2 norm may degrade the image sequence rather than enhance it. The robust norm applicable to several noise and data models is desired in SRR algorithms. First, this paper proposes a robust SRR algorithm based on the stochastic regularization technique of Bayesian MAP estimation by minimizing a cost function. The Huber norm with Tikhonov regularization is used for measuring the difference between the projected estimate of the high-resolution image and each low resolution image, removing outliers in the data. Second, in order to cope with real sequences and complex motion sequences, this paper proposes an improved SRR observation model, affine block-based transform, devoted to the case of nonisometric inter-frame motion. The experimental results show that the proposed reconstruction can enhance real complex motion sequences, such as Suzie and Foreman sequence, and confirm the effectiveness of our algorithm and demonstrate its superiority to other SRR algorithms based on L1 and L2 norm for several noise models such as AWGN, Poisson, Salt&Pepper and Speckle noise.

Index Terms— Video signal processing, Image enhancement, Image reconstruction

1. INTRODUCTION

In the last two decades, the enlargement in the extensive use of digital imaging technologies in consumer (e.g., digital video) and other markets (e.g., security and military) has brought with it a simultaneous demand for higher-resolution (HR) images. The demand for such images can be partially met by algorithmic advances in SRR technology in addition to hardware development. Such HR images not only give the viewer a more pleasing picture but also offer additional details that are important for subsequent analysis in many applications. SRR [2, 10, 12, 15, 19] is considered to be one of the most promising techniques that can help overcome the limitations due to optics and sensor resolution. In general, the problem of super-resolution can be expressed as that of combining a set of aliased, noisy, low-resolution, blurry images to produce a higher resolution image or image sequence. The idea is to increase the information content in the final image by exploiting the additional spatio-temporal information that is available in each of the LR images.

This section reviews some literature from the estimation point of view because the SRR estimation is one of the most importance parts of the SRR research areas and directly impact to the SRR performance. Schultz and Stevenson [13-14] proposed the SRR algorithm using ML estimator (L2 Norm) with HMRF Regularization in 1996. In 1997, Elad and Feuer [5] proposed the SRR algorithm using the ML estimator (L2 Norm) with nonellipsoid constraints. Next, they [6] proposed the SRR algorithm using R-SD and R-LMS (L2 Norm) in 1999. They [7-8] proposed the fast SRR algorithm ML estimator (L2 Norm) for restoration the warps are pure translations, the blur is space invariant and the same for all the images, and the noise is i.i.d. Gaussian in 2001. Patti and Altunbasak proposed [1] a SRR algorithm using ML (L2 Norm) estimator with POCS-based regularization in 2001 and Altunbasak et al. [22] proposed a SRR algorithm using ML (L2 Norm) estimator for the MPEG sequences in 2002. Rajan and Chaudhuri [2-3] proposed SRR using ML (L2 Norm) with MRF regularization to simultaneously estimate the depth map and the focused image of a scene in 2003. Farsiu and Robinson [16-17] proposed SRR algorithm ML estimator (L1 Norm) with BTV Regularization in 2004. Next, they propose a fast SRR of color images [18] using ML estimator (L1 Norm) with BTV and Tikhonov Regularization in 2006.

Almost SRR algorithms are restricted to globally or locally uniform translational displacement between the measured images or sequences. This implies the measured images or sequences are observed at a high temporal frequency sampling (or high frame rate) but the measured images or sequences are usually observed by the real commercial cameras at low temporal frequency sampling (or low frame rate) such as standard sequences (Foreman, Carphone, Susie, etc.). The measured images or sequences have many complex motions instead of only a simple translational motion therefore the pure translation model can not well represent the real complex motion effectively and image SRR applications can apply only on the sequences that have simple translation motion. In [20], we proposed the SRR using a regularized ML estimator with affine block-based registration for the real image sequence. Later, Rochefort et al. [4] proposed SRR approach based on regularized ML for the extended original observation model [4] devoted to the case of nonisometric inter-frame motion such as affine motion in 2006. This paper proposed the novel SRR observation model to overcome the insufficient temporal sampling frequency and to model the real complex motion sequence that the traditional SSR observation model can not support. To realize the implementation of the proposed SRR observation model, the sub-pixel image registration [20] is designed to calculate the nonisometric inter-frame motion parameter. Moreover, the fast algorithm is proposed to reduce the computational load for the proposed sub-pixel registration.

For the data fidelity cost function, all the above SRR algorithms [1-20,22] are based on the simple estimation techniques such as L1 or L2 Norm therefore these SRR methods are usually very sensitive to their assumed model of data and noise. The success of SRR algorithm is highly dependent on the accuracy of the model of the imaging process. Unfortunately, these models are not supposed to be exactly true, as they are merely mathematically

[†]V. Patanavijit is with Faculty of Engineering, Assumption University, Bangkok, Thailand; E-mail: Patanavijit@yahoo.com.
[††]S. Jitapunkul are with Faculty of Engineering, Chulalongkorn University, Bangkok, Thailand; E-mail: Somchai.j@chula.ac.th.

convenient formulations of some general prior information. When the data or noise model assumptions do not faithfully describe the measure data, the estimator performance degrades. Furthermore, existence of outliers defined as data points with different distributional characteristics than the assumed model will produce erroneous estimates. Most noise models used in SRR algorithms are based on AWGN model at low power therefore SRR algorithms can effectively apply only on the image sequence corrupted by AWGN. With this noise model, L1 and L2 norm error are effectively used in SRR algorithms. For normally distributed data, the L1 norm produces estimates with higher variance than the L2 norm but the L2 norm is very sensitive to outliers and noise because the influence function increases linearly and without bound. The real noise models that corrupt the measure sequence are unknown; consequently, SRR algorithm using L1 norm or L2 norm may degrade the image sequence rather than enhance it. Therefore, the robust norm which is applicable to several noise and data models is desired in SRR algorithms. The robust norm which is applicable to unknown noise models is desired in SRR algorithms. From the robust statistical estimation [11,21], Huber Norm is designed to be more robust than L1 and L2. Huber norm is designed to be robustness and reject outliers, the norm must be more forgiving about outliers; that is, it should increase less rapidly than L2. Hence, this paper proposes a video enhancement method for real standard sequences that are corrupted by arbitrary noise models. The method is based on the iterative robust SRR algorithm [21] using the stochastic regularization technique of Bayesian MAP estimation by minimizing a cost function. The Huber norm with Tikhonov regularization is used for measuring the difference between the projected estimation of the high-resolution image and each low resolution image, removing outliers in the data. In order to cope with real sequences and complex motion sequences, this paper improves the SRR observation model by introducing the affine block-based transform, devoted to the case of nonisometric inter-frame motion.

The organization of this paper is as follows. Section 2 explains the main concepts of SRR improved observation model. Section 3 introduces the proposed video enhancement based on SRR algorithm with improved SRR observation model with Tikhonov Regularization using L1, L2 and Huber norm. Section 4 outlines the proposed SRR algorithm and presents the comparative experimental results between the proposed Huber norm, the L1 norm and L2 norm method. Section 5 provides the conclusion.

2. SRR IMPROVED OBSERVATION MODEL

2.1 Improved Observation Model

In this section, we propose the problem and the model of super-resolution reconstruction. Define a low-resolution (LR) image sequence, $\{\underline{\mathbf{Y}}_k\}$, as our measured data (The size of the LR images is $N_1 \times N_2$ pixels). A HR image $\underline{\mathbf{X}}$ ($qN_1 \times qN_2$ pixels) is estimated from the LR sequences, where q is an integer interpolation factor in both the horizontal and vertical directions. To reduce the computational complexity, each frame is separated into overlapping blocks (the shaded blocks in Fig. 1(a) and Fig. 1(b)). For notation convenience, all overlapping blocked in a frame will be presented as a vector, ordered column-wise lexicographically. Namely, the overlapping blocked in LR frame is $\underline{Y}_k \in \mathbb{R}^{M^2}$ ($M^2 \times 1$) and the overlapping blocked in HR frame is $\underline{X} \in \mathbb{R}^{q^2M^2}$ ($L^2 \times 1$ or $q^2M^2 \times 1$). We assume that the two images are related via the following equation

$$\underline{Y}_k = D_k H_k F_k \underline{X} + \underline{V}_k \quad ; k = 1, 2, \ldots, N \tag{1}$$

where $\underline{Y}_k(t)$ is a blurred, decimated, down sampled and contaminated by additive noise of $\underline{X}$. The matrix F_k ($F \in \mathbb{R}^{q^2M^2 \times q^2M^2}$) stands for the proposed nonisometric inter-frame warp [20] between the images $\underline{X}$ and $\underline{Y}_k$. H_k is the blur matrix which is space and time invariant and $H_k \in \mathbb{R}^{q^2M^2 \times q^2M^2}$. D_k is the decimation matrix assumed constant and $D_k \in \mathbb{R}^{M^2 \times q^2M^2}$. $\underline{V}_k$ is a system noise and $\underline{V}_k \in \mathbb{R}^{M^2}$. The relation between overlapping blocked HR Image and overlapping blocked LR image sequence is shown in Fig 1(c).

2.2 The Proposed Registration for Improved Observation Model of SRR [20]

In this section, we propose a scheme for estimating affine block-based motion vectors for registration step. The estimation can be separated into 2 stages. In the first stage of the estimation algorithm, the current and reference frames are divides into 50% overlapping blocks (16x16). This stage divides the image into small areas in order to detect and estimate the local motions. The advantage of the block processing is the reduction of the computational load and the possibility of parallel processing. In the second stage, the affine motion vector of each block between the current and reference frame is computed by the M3SS (Modified Three Step Search). The M3SS is proposed to reduce a very high computational load in affine motion vector estimation. The M3SS is designed based on the popular 3SS. For practical implementation, the M3SS is proposed to reduce a very high computational load in affine motion vector estimation. The 3SS is one of the popular and fast algorithms used in the translational registration; therefore, this paper develops the M3SS (6 parameters) based on 3SS (2 parameters). For the 7x7 displacement window (translational deformation) and degree (rotation, extraction or expansion deformation), the proposed M3SS algorithm utilizes a search pattern with $3^6 = 729$ check points (the parameters vary in 6 dimensions instead of 2 dimensions) on a search window in the first step. The set of parameters having the minimum error is used as the center of the search area in the subsequent step. The search window is reduced by half in the subsequent step until the search window equals to the pre-determined resolution. The criterion for parameter selection in this paper was based on experiments and the chosen parameters produce the highest PSNR result on 3 standard sequences: Foreman, Carphone and Stefan [20]. From [20], the total number of the M3SS check points is fixed at 3.65E+3. Compared with the classical block-based estimation method (translation block-based estimation method) at 0.25 pixel accuracy and w=9, the total number of the M3SS check points has approximately 3 times more than the FS approach of the classical observation model but the PSNR performance of the M3SS is 5-6 dB higher than that of the classical translational method.

3. THE PROPOSED SRR FRAMEWORK WITH IMPROVED OBSERVATION MODEL

SRR is an ill-posed problem [5–8]. For the under-determined cases, there exist an infinite number of solutions which satisfy (1). The solution for square and over-determined cases is not stable

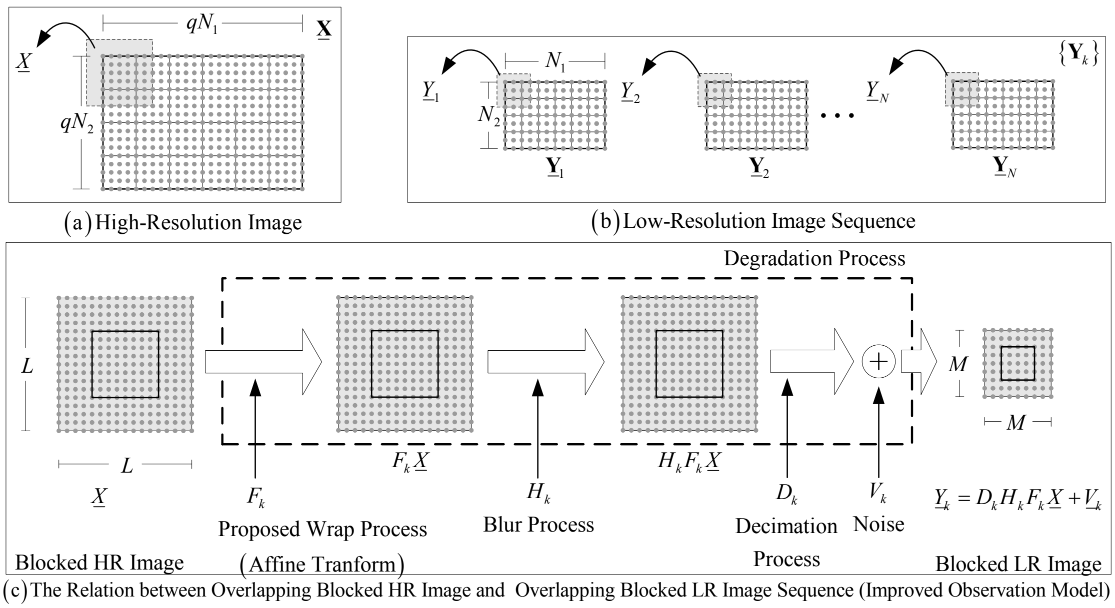

Figure 1. The Improved Observation Model

that means small amounts of noise in measurements will result in large perturbations in the final solution. Therefore, considering regularization in super-resolution algorithm as a means for picking a stable solution is very useful, if not necessary. Also, regularization can help the algorithm to remove artifacts from the final answer and improve the rate of convergence.

3.1 SRR using L1 Norm with Regularized Function

A popular family of estimators is the L1 Norm estimators that are used in SRR problem [4-9]. We rewrite the definition of these estimators in the super resolution context. A regularization term compensates the missing measurement information with some general prior information about the desirable HR solution, and is usually implemented as a penalty factor in the generalized minimization cost function. We rewrite the definition of these estimators in the SRR context as the following minimization problem:

$$\underline{X} = \underset{\underline{X}}{\mathrm{ArgMin}}\left\{\sum_{k=1}^{N}\left\|D_kH_kF_k\underline{X}-\underline{Y}_k\right\| + \lambda\cdot\left(\Gamma\underline{X}\right)^2\right\} \quad (2)$$

The most classical and simplest regularization cost functions is the Laplacian regularization [17] where the Laplacian kernel is

$$\Gamma = \tfrac{1}{8}\left[1 \quad 1 \quad 1 \quad ; \quad 1 \quad -8 \quad 1 \quad ; \quad 1 \quad 1 \quad 1\right] \quad (3)$$

By the steepest descent method, the solution of problem (2) is

$$\hat{\underline{X}}_{n+1} = \hat{\underline{X}}_n + \beta\cdot\left\{\left(\sum_{k=-N}^{N}F_k^TH_k^TD_k^T\,\mathrm{sign}\left(\underline{Y}_k - D_kH_kF_k\hat{\underline{X}}_n\right)\right) - \left(\lambda\cdot\left(\Gamma^T\Gamma\right)\hat{\underline{X}}_n\right)\right\} \quad (4)$$

where β is the step size in the direction of the gradient.

3.2 SRR using L2 Norm with Regularized Function

Another popular family of estimators is the L2 Norm estimators that are used in SRR problem [13-14]. We rewrite the definition of these estimators in the SRR context. Combining the Laplacian regularization, we propose the solution of the super-resolution problem as follows:

$$\underline{X} = \underset{\underline{X}}{\mathrm{ArgMin}}\left\{\sum_{k=1}^{N}\left\|D_kH_kF_k\underline{X}-\underline{Y}_k\right\|_2^2 + \lambda\cdot\left(\Gamma\underline{X}\right)^2\right\} \quad (5)$$

By the steepest descent method, the solution of problem (5) is

$$\hat{\underline{X}}_{n+1} = \hat{\underline{X}}_n + \beta\cdot\left\{\sum_{k=1}^{N}F_k^TH_k^TD_k^T\left(\underline{Y}_k - D_kH_kF_k\hat{\underline{X}}_n\right) - \left(\lambda\cdot\left(\Gamma^T\Gamma\right)\hat{\underline{X}}_n\right)\right\} \quad (6)$$

3.3 SRR using Huber Norm with Regularized Function

This paper proposes SRR using Huber norm [10,21] that is more robust than L1 and L2 norm. Huber norm is designed to be robustness and reject outliers, the norm must be more forgiving about outliers; that is, it should increase less rapidly than L2. We rewrite the definition of these estimators in the SRR context as the following minimization problem:

$$\underline{X} = \underset{\underline{X}}{\mathrm{ArgMin}}\left\{\sum_{k=1}^{N}f_{HUBER}\left(D_kH_kF_k\underline{X}-\underline{Y}_k\right) + \lambda\cdot\left(\Gamma\underline{X}\right)^2\right\} \quad (7)$$

$$f_{HUBER}\left(x\right) = x^2; |x| \le T \quad \text{or} \quad T^2 + 2T\left(|x| + T\right); |x| > T \quad (8)$$

where T is Huber constant parameter. Combining the Laplacian regularization, By the steepest descent method, the solution of problem (7) is defined as

$$\hat{\underline{X}}_{n+1} = \hat{\underline{X}}_n + \beta\cdot\left\{\sum_{k=1}^{N}F_k^TH_k^TD_k^T\cdot\psi_{HUBER}\left(\underline{Y}_k - D_kH_kF_k\hat{\underline{X}}_n\right) - \left(\lambda\cdot\left(\Gamma^T\Gamma\right)\hat{\underline{X}}_n\right)\right\} \quad (9)$$

$$\psi_{HUBER}\left(x\right) = f'_{HUBER}\left(x\right) = 2x; |x| \le T \quad \text{or} \quad 2T\cdot\mathrm{sign}\left(x\right); |x| > T \quad (10)$$

4. THE EXPERIMENTAL

This section presents results obtained by video enhancement using the SRR algorithm with the improved observation model. The experiment was implemented in MATLAB and the block size of LR images is fixed at 8x8 (16x16 for overlapping block) and the search window is 7 for affine block-based registration [20] and N=5 (5 Frames) for ML estimation process. We use Susie (40^{th}) and Foreman (110^{th}) sequence, QCIF format and complex-edge characteristic, as our test sequences. Then, to simulate the effect of camera PSF, the images were convolved with a symmetric Gaussian low-pass filter with size of 3x3 and standard deviation of one. The blurred images were subsampled by the factor of 2 in each direction (88x72) and the blurred subsampled images were corrupted by Gaussian noise. The criterion for parameter selection in this paper was to choose parameters which produce both most visually appealing results and highest PSNR. Therefore, to ensure fairness, each experiment was repeated several times with different parameters and the best result of each experiment was chosen. We corrupted images with the following four noise families: AWGN, Poisson, Salt&Pepper and Speckle noise. One level of noise was applied in Poisson noise case. Five Levels were applied in AWGN case. Three levels of noise were applied in the remaining two cases. The noise levels applied are as follow:

1) AWGN:PSNR of the corrupted image=25, 22.5, 20, 17.5, 15 dB
2) Salt&Pepper noise: D=0.005, 0.010, 0.015 (D is noise density)
3) Speckle noise: V=0.01, 0.02, 0.03

The PSNR of Suzie and Foreman result are summarized in Figure 2 and 3 respectively and all comparatively experimental results are concluded as follow: The SRR algorithm using Huber norm with the proposed registration gives the highest PSRN because these robust estimators are designed to be robust and reject outliers (noise and registration error). The norms are more forgiving on outliers; that is, they should increase less rapidly than L1 and L2. Especially, the Huber norm has the higher PSNR than L1 and L2 norm when the noise power or registration error increases. The SRR algorithm using L1 norm gives the higher PSRN than the SRR algorithm using L2 because L2 is more sensitive the outliers such as noise and the registration error than L1 norm. The SRR algorithm using L2 norm with an improved observation model gives the lowest PSRN because L2 norm is more sensitive the outliers such as the noise and registration error. The L2 influence function increases linearly and without bound. When the noise power or registration error increases the PSNR result of the SRR algorithm using L2 norm decreases rapidly.

5. THE CONCLUSION

This paper proposes a novel video enhancement using SRR framework based robust estimation norm with improved observation model. The proposed SRR can be applied on the several noise models and on the real complex sequence such as Susie and Foreman sequence. Experimental results conducted clearly that the proposed algorithm can apply on the general noise models such as AWGN, Poisson and Salt & Pepper Noise and the proposed algorithm can obviously improve the result.

REFERENCES

[1] A. J. Patti and Y. Altunbasak, "Artifact Reduction for Set Theoretic Super Resolution Image Reconstruction with Edge Constraints and Higher-Order Interpolation", IEEE Trans. on IP, Vol. 10, No. 1, Jan. 2001.

[2] D. Rajan, S. Chaudhuri and M. V. Joshi, "Multi-objective super resolution concepts and examples", IEEE SP. Mag., May. 2003.

[3] D. Rajan and S. Chaudhuri, "Simultaneous Estimation of Super-Resolution Scene and Depth Map from Low Resolution Defocuses Observations", IEEE Trans. on PAMI., Sep. 2003.

[4] G. Rochefort, F. Champagnat, G. L. Besnerais and Jean-Francois Giovannelli, "An Improved Observation Model for Super-Resolution Under Affine Motion", IEEE Transactions on IP., Nov. 2006.

[5] M. Elad and A. Feuer, "Restoration of a Single Superresolution Image from Several Blurred, Noisy and Undersampled Measured Images", IEEE Trans. on Image Processing, Vol. 6, Dec. 1997.

[6] M. Elad and A. Feuer, "Superresolution Restoration of an Image Sequence: Adaptive Filtering Approach", IEEE Trans. on IP., Match 1999.

[7] M. Elad and A. Feuer, "Super-Resolution Reconstruction of Image Sequences", IEEE Trans. on PAMI., Sep. 1999.

[8] M. Elad and Y. Hecov Hel-Or, "A Fast Super-Resolution Reconstruction Algorithm for Pure Translational Motion and Common Space-Invariant Blur", IEEE Trans. on IP., 2001.

[9] M. Elad and A. Feuer, "Super-Resolution Restoration of Continuous Image Sequence – Adaptive Filtering Approach", Technical Report, The Technion, Israel Institute of Technology.

[10] M. K. Ng and N. K. Bose, "Mathematical analysis of super-resolution methodology", IEEE SP. Magazine, May. 2003.

[11] M. J. Black, G. Sapiro, D. H. Marimont and D. Herrger, "Robust Anisotropic Diffusion", IEEE Trans. on IP. 1998.

[12] M. G. Kang and S. Chaudhuri, "Super-Resolution Image Reconstruction", IEEE Signal Processing Magazine, May. 2003.

[13] R. R. Schultz and R. L. Stevenson, "A Bayesian Approach to Image Expansion for Improved Definition", IEEE Trans. on IP., May 1994.

[14] R. R. Schultz and R. L. Stevenson, "Extraction of High-Resolution Frames from Video Sequences", IEEE Transactions on IP., June 1996.

[15] S. C. Park, M. K. Park and M. G. Kang, "Super-Resolution Image Reconstruction: A Technical Overview", IEEE SP. Magazine, May 2003.

[16] S. Farsiu, D. Robinson, M. Elad, P. Milanfar, "Advances and Challenges in Super-Resolution", Wiley Periodicals, Inc., 2004

[17] S. Farsiu, M. Dirk Robinson, M. Elad and P. Milanfar, "Fast and Robust Multiframe Super Resolution", IEEE Trans. on IP., Oct. 2004.

[18] S. Farsiu, M. Elad and P. Milanfar, "Multiframe Demosaicing and Super-Resolution of Color Images", IEEE Trans. on IP, 2006.

[19] S. C. Park, M. K. Park and M. G. Kang, "Super-Resolution Image Reconstruction: A Technical Overview", IEEE SP. Magazine, May 2003.

[20] V. Patanavijit and S. Jitapunkul, "An Iterative Super-Resolution Reconstruction of Image Sequences using a Bayesian Approach with BTV Prior and Affine Block-Based Registration", IEEE CRV 2006, June 2006.

[21] V. Patanavijit and S. Jitapunkul, "A Robust Iterative Multiframe Super-Resolution Reconstruction using a Huber Statistical Estimation Technique", IEEE CHINACOM 2006, Beijing, China, Oct. 2006.

[22] Y. Altunbasak, A. J. Patti and R. M. Mersereau, "Super-Resolution Still and Video Reconstruction from MPEG-Coded Video", IEEE Trans. on C&S for Video Technology, April 2002.

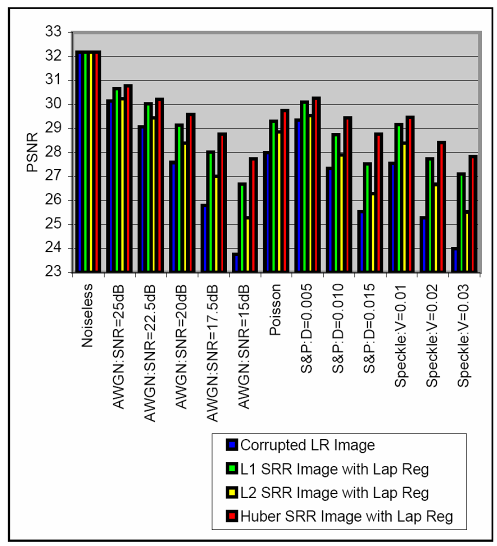

Fig. 2 : The experimental result (Susie Sequence)

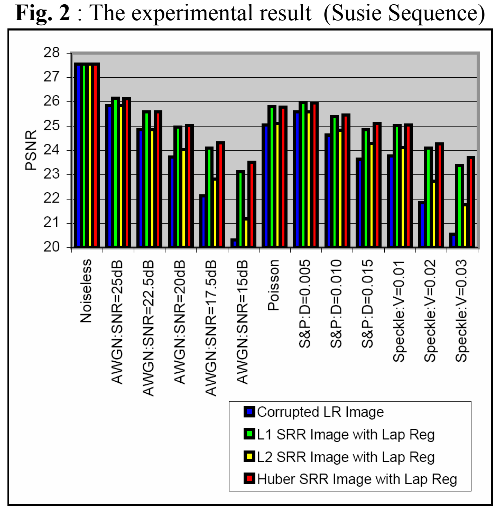

Fig. 3 : The experimental result (Foreman Sequence)