Proceedings Paper:

Hotrakool, W. and Abhayaratne, C. (2014) Running Gaussian reference-based reconstruction for video compressed sensing. In: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). ICASSP 2014 - 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 04-09 May 2014, Florence, Italy. IEEE, pp. 2001-2005. ISBN: 9781479928934. ISSN: 1520-6149. EISSN: 2379-190X.

https://doi.org/10.1109/icassp.2014.6853949

RUNNING GAUSSIAN REFERENCE-BASED RECONSTRUCTION FOR VIDEO COMPRESSED SENSING

Wattanit Hotrakool and Charith Abhayaratne

Deparment of Electronic and Electrical Engineering
The University of Sheffield, United Kingdom

ABSTRACT

Our recent work has shown that quality of compressed sensing reconstruction can be improved immensely by minimising the error between the signal and a correlated reference, as opposed to the conventional l_1-minimisation of the data measurements. This paper introduces a method for online estimating suitable references for video sequences using the running Gaussian average. The proposed method can provide robustness to video content changes as well as reconstruction noise. The experimental results demonstrate the performance of this method to be superior to those of the state-of-the-art l_1-min methods. The results are comparable to the lossless reference reconstruction approach.

1. INTRODUCTION

Compressed sensing allows a complete signal to be reconstructed from its under-sampled measurements. The reconstruction is done by maximising the sparsity of the signal, giving the sparsest solution to the problem which is the closest approximation to the original signal [1]. It is well established that l_1-minimisation (l_1-min) – which minimises the l_1-norm of signal – can be used to achieve such sparsest solutions [2, 3, 4]. However, in many real-world applications, such as, natural video reconstruction or medical imaging, the sparsest solution is not always the best approximation. The reason is the sparsity of the original signal itself is not sufficiently sparse. In this case, the absence of many small coefficients in the sparsest solution actually produces reconstruction noise.

In order to improve reconstruction results, many attempts have used side-information or some prior-knowledge about the signal to the reconstruction process. The motivation of these attempts is that in most applications, some characteristics of the signal can be predetermined or approximated from its neighbours. This observation can be seen clearly in Magnetic Resonance Imaging, sensor networks, multiview imaging , and surveillance camera applications, *etc.* [5, 6]. Many kinds of side-information have been employed for successfully improving the quality of reconstructed signal significantly. These side-information types include sparsity patterns, signal's upper and lower bounds [7], and group reconstruction [8]. There are many variants of the use of sparsity patterns, including the case where only the location of sparse support is known [9, 10], Model-based compressed sensing, where both location and structure (such as wavelet coeeficient distribution) is known [11, 12], and Kalman-filtered compressed sensing, where the sparsity pattern is assumed to change slowly over time [13].

One of the side-information which contains a lot of useful information about the signal is the signal itself from another instance of time. That is, in a sequence of time-series signal, such as video sequence, a frame is highly temporally correlated to its neighbours. Therefore any frame can be used as a side-information to reconstruct its consecutive frames. Our previous work in [14] shows that by using a correlated (either temporally or spatially) reference of the measured data, it improves reconstruction greatly. Many compressed sensing algorithms employing temporally or spatially acquired side-information [15, 16, 17, 8] also demonstrate the same improvement over the conventional l_1-min. However, even though using lossless references can lead to much better results, acquiring such references is impractical in most applications. Apart from few applications that lossless references can be obtained using specially designed sensing schemes, such as MRI or multiview imaging, most applications do not have the comfort of acquiring some frames fully to used as references. Thus a reference estimation method is needed to create references directly from the reconstruction results.

This paper introduces a method to construct references without the need of uncompressed acquisition by using the running Gaussian average (RGA) of previously reconstructed frames as the reference estimator. This paper shows that by using RGA to estimate references, the performance of a lossless reference reconstruction is maintained. The proposed method also increases the stability of the system, making it less susceptible to reconstruction noise and content's motions. The rest of the paper is as follow: Section 2 presents the reconstruction method based on correlated references, Section 3 discusses the use of running Gaussian average as the reference estimator. The experimental results are shown in Section 4, followed by the conclusions in Section 5.

2. RECONSTRUCTION USING CORRELATED REFERENCES

Our previous work [14] shows that by minimising the error between the signal and its correlated reference during its reconstruction instead of the sparsity, the result can be greatly improved over the conventional l_1-min. Such references can be correlated temporally or spatially. It also shows that by incorporating such references in reconstruction, the least squares method – *i.e.* the l_2-norm minimisation – can be used in place of the l_1-min. Doing so provides the results comparable to those of the l_1-min but with a much lower complexity.

For a compressed sensing system $\mathbf{y} = \mathbf{A}\mathbf{x}$, where $\mathbf{y} \in \mathbb{R}^m$ is a compressed measurement of $\mathbf{x} \in \mathbf{X}(R)$, where

$$\mathbf{X}(R) = \{\mathbf{x} | \|\mathbf{x} - \mathbf{r}\|_1 \leq R, \mathbf{x} \in \mathbb{R}^n\},$$

and $m \ll n$, if a ful-length lossless reference $\mathbf{r} \in \mathbb{R}^n$ is known, the reconstruction result $\hat{\mathbf{x}}$ that is obtained from

$$\min \|\hat{\mathbf{x}} - \mathbf{r}\|_1 \text{ subject to } \mathbf{A}\hat{\mathbf{x}} = \mathbf{y}, \tag{1}$$

must satisfy

$$\sup \|\mathbf{x} - \hat{\mathbf{x}}\|_2 \leq \|\mathbf{x} - \mathbf{r}\|_2. \tag{2}$$

The proof can be found in [14]. This means the closer the reference $\mathbf{r}$ is to the original $\mathbf{x}$, the better reconstruction result can be obtained.

We have also shown that the reconstruction based on the reference can be incorporated with the least squares method. That means a compressed sensing $\mathbf{y} = \mathbf{A}\mathbf{x}$, given a full-length lossless reference $\mathbf{r}$, its reconstruction result can be obtained from

$$\hat{\mathbf{x}} = \mathbf{r} + \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}(\mathbf{y} - \mathbf{A}\mathbf{r}). \tag{3}$$

The proof of Eq. (3) can also be found in [14]. The use of the least squares method can provide a much lower complexity compares to the l_1-min. This combination of methods is demonstrated to provide results comparable to the conventional l_1-min's results. Again, the reference $\mathbf{r}$ is preferred to be as close to the original $\mathbf{x}$. However, employing a full-length lossless reference in practical sensing applications is difficult. Thus we present the RGA-based reference estimator in Section 3.

3. REFERENCE ESTIMATION USING RUNNING GAUSSIAN AVERAGE

The most straightforward reference is to use the current reconstruction result as a reference for the next frame. That is, except the first frame where the reconstruction is done using l_1-min, the reconstruction result $\hat{\mathbf{x}}_t$ at time t is used as a reference for reconstructing $\hat{\mathbf{x}}_{t+1}$. This scheme is easy to implement and can provide references that is very close to the current frame. Its drawback is, however, it can reduce the stability of the system with presences of reconstruction noise and motion.

In order to increase the stability of the system by creating a more robust reference, a probability-based approach can be used. The running Gaussian average (RGA) is a method popular among many computer vision techniques and is mostly used to model a background of image sequence [18, 19]. The RGA models each pixel as a Gaussian distribution of pixel values. In another words, at a given time t and pixel location index i, a reference pixel $r_{i,t} \in \mathbf{r}_t | r_{i,t} \in \mathbb{R}$ is model as

$$r_{i,t} = \mathcal{N}(\mu_{i,t}, \sigma^2_{i,t}),$$

where $\mu_{i,t}$ and $\sigma^2_{i,t}$ are the mean and variance of r_i at time t. The model is updated at every frame using the following equations:

$$\mu_{i,t} = \alpha\hat{x}_{i,t} + (1 - \alpha)\mu_{i,t-1}, \tag{4}$$

$$\sigma^2_{i,t} = \alpha(\hat{x}_{i,t} - \mu_{i,t})^2 + (1 - \alpha)\sigma^2_{i,t-1}, \tag{5}$$

where $\hat{x}_{i,t} \in \mathbb{R}$ is a pixel from the reconstructed result $\hat{\mathbf{x}}_t$ at time t.

This paper chooses the RGA as the reference estimator because of its simplicity. A single parameter $0 \leq \alpha \leq 1$ is the refresh rate of the reference which plays an important role to determine the trade off between the reference's rate of response and the system stability. As $\alpha \to 1$, the model responses faster to changes of image contents, ultimately $\alpha = 1$ is equivalent to the simple reference reconstruction discussed earlier. As α decreases, the response of the reference decreases together, which increases the distance between the reference and the original. However, the stability of the system is increased and is more robust to the reconstruction noise and content movements. The effect between the refresh rate α and the stability of the reconstruction system is demonstrated in Section 4. Fig. 1 shows the comparison between the simple references and the references from the RGA-based reference estimator for two different sequences (one with low motion and the other with high motion).

4. EXPERIMENTAL RESULTS

4.1. Implementation details

Algorithm 1 describes the implementation of the proposed method.

The reconstruction of the first frame $\hat{\mathbf{x}}_1$ is reconstructed using the conventional l_1-min method by setting $\mathbf{r}_0 = \mathbf{0}$, *i.e.* a zero vector with the same length as $\mathbf{x}_1$. This $\hat{\mathbf{x}}_1$ is used as the initial reference $\mathbf{r}_1$.

Every other $\mathbf{x}_t$ when $t > 1$ is reconstructed based on the reference $\mathbf{r}_{t-1}$. The reference $\mathbf{r}_t$ is updated using Eq. (4) based on the result $\hat{\mathbf{x}}_t$.

(a) Simple reference (b) RGA reference (c) Simple reference (d) RGA reference

Fig. 1: References of a medium activity sequence (a),(b) and references of a high activity sequence (c),(d)

Algorithm 1 Proposed reconstruction method

Input: a sequence of compressed measurements $\mathbf{Y} = \mathbf{A}\mathbf{X}$, sensing matrix $\mathbf{A}$, refresh rate α
Output: a sequence of reconstruction results $\hat{\mathbf{X}}$
initialise reference $\mathbf{r}_0 := \mathbf{0}$; $t = 1$;
repeat
reconstruct $\hat{\mathbf{x}}_t$ from $\mathbf{y}_t$ and $\mathbf{r}_{t-1}$ using (1) or (3)
if $t = 1$ **then**
$\mathbf{r}_t = \hat{\mathbf{x}}_t$;
else
$\mathbf{r}_t := (1-\alpha)\mathbf{r}_{t-1} + \alpha\hat{\mathbf{x}}_t$;
end if
$t := t + 1$;
until stopped or last frame is reached

The proposed method is applied to a test set of 14 randomly selected video sequences. The reconstructed sequences are compared with several states-of-the-art methods which are the lossless reference reconstruction [14], Kalman-filter CS [13], CS image reconstruction with side-information [20], and the conventional l_1-min using ISAL1 algorithm [21].

4.2. Reconstruction error and visual quality

All measurements are performed with the compression rate of 50%. The proposed method uses the refresh rate $\alpha = 0.5$ through out the experiment. The signal size is 4096 pixels per frame and each sequence contains 300 frames. The reconstruction error (measured by PSNR) and the visual quality (measured by SSIM [22]) are shown in Table 1.

The performance of the proposed method when minimising the l_1-norm of the error with respect to the reference using Eq. (1) (RGA-l_1) is comparable to the quality of the results obtained from lossless reference method. This removes the need of the uncompressed acquisition and therefore allows more practical reconstruction systems than previously possible. The performance of the proposed method is also superior to both state-of-the-art methods. Even though the pixel-wise prediction of Kalman-filter CS provides results with good PSNR, their visual quality is severely compromised. On the contrary, the algorithm of [20] which employ the inter-image correlation provides results with better visual quality despite having lower PSNR. Our method outperforms both algorithms in terms of both metrics.

The another interesting results is that the proposed method when using the least squares to minimise the error with respect to the reference using Eq. (3) (RGA-LS) can provide the results comparable to the l_1-min algorithm. This situation is the same as the one that has been shown in [14], which provide a low complexity alternative to the l_1-min. By using the least squares method with reference, our proposed method can reconstruct the same sequence hundreds times faster than l_1-min method. Again, our method here has an advantage over the previous work since no lossless reference is required.

4.3. Refresh rate and signal stability

As discussed in Section 3, the refresh rate α is related to the stability of the proposed method. This effect of stability is obvious in high-activity sequences, where the correlation between each frame is low.

Fig. 2 demonstrates the effect of the refresh rate to the stability. By decreasing α, the reference's response to the changes in the sequence slows down. In case of the low-activity sequence, reducing α reduces the PSNR accordingly. In the high-activity sequence, however, it can be noticed that the fluctuation in the reconstruction error is reduced when α is lower. As a result, the overall reconstruction quality is better when α is low in the high activity sequence.

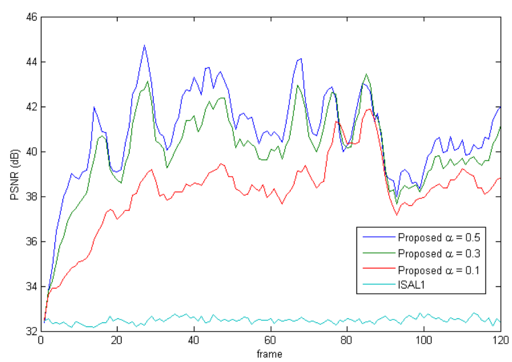
Fig. 3 shows the examples of reconstruction frames from a medium-activity sequence and a high-activity sequence. It can be seen that even when the high motion is present in the sequence, the result of our proposed method still outperforms the result of l_1-min method.

5. CONCLUSIONS

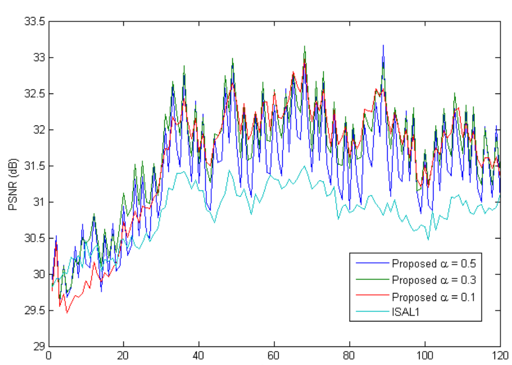
The running Gaussian average reference-based reconstruction method for video compressed sensing has been proposed in this paper. It outperforms other state-of-the-arts algorithms and is comparable to the lossless reference method when using with l_1-min-based optimisation. Moreover, it has the

Table 1: Comparison of reconstruction error and visual quality

Sequence	Proposed RGA-l_1		Proposed RGA-LS		Lossless [14]		KF-CS [13]		CSIR-SI [20]		ISAL1 [21]	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Low-activity sequences												
Akiyo	39.03	0.95	32.02	0.73	40.97	0.94	31.94	0.19	29.44	0.44	32.50	0.75
Claire	36.31	0.95	31.53	0.65	41.12	0.95	31.74	0.18	29.23	0.38	32.36	0.69
Container	38.04	0.94	30.60	0.56	40.66	0.95	32.36	0.16	28.98	0.36	30.86	0.58
highway	39.30	0.93	31.89	0.64	40.19	0.92	31.97	0.23	29.27	0.27	32.43	0.65
Miss USA	41.99	0.96	35.72	0.80	42.79	0.95	32.72	0.19	29.80	0.37	36.33	0.82
Medium-activity sequences												
Carphone	35.42	0.91	30.76	0.63	35.84	0.89	31.86	0.22	28.35	0.30	30.75	0.69
Coastgaurd	33.98	0.82	31.17	0.60	33.96	0.79	32.04	0.14	28.20	0.21	32.07	0.69
hall	36.08	0.92	29.95	0.63	37.17	0.92	32.58	0.19	28.65	0.32	30.58	0.66
Mother	39.70	0.95	32.56	0.73	41.18	0.94	32.32	0.59	29.85	0.44	33.35	0.76
News	35.70	0.92	29.90	0.62	36.94	0.92	31.65	0.27	28.37	0.30	30.26	0.64
Salesman	36.84	0.94	30.60	0.69	37.70	0.93	32.59	0.24	28.73	0.35	31.95	0.76
High-activity sequences												
foreman	32.67	0.84	29.16	0.45	32.99	0.83	32.52	0.18	28.16	0.27	30.69	0.69
Ice skate	30.89	0.66	28.90	0.29	31.17	0.63	32.24	0.18	28.06	0.25	30.51	0.61
Silent	36.38	0.93	30.17	0.64	37.29	0.93	32.11	0.16	28.31	0.27	31.36	0.72



(a) Low activity sequence



(b) High activity sequence

Fig. 2: PSNR comparison between the proposed method and l_1-min method ISAL1 demonstrates the relationship between the refresh rate and stability

(a) Proposed method (PSNR = 31.2 dB, SSIM = 0.92) (b) ISAL1 (PSNR = 30.80 dB, SSIM = 0.62)

(c) Proposed method (PSNR = 32.02 dB, SSIM = 0.7) (d) ISAL1 (PSNR = 30.78 dB, SSIM = 0.57)

Fig. 3: Reconstruction results of the proposed method compares to ISAL1 of medium activity sequence (a),(b) and high activity sequence (c),(d).

performance comparable to the conventional l_1-min method when using the least-squares-based optimisation, despite having much lower complexity. Although high motion in the sequence is still a challenge for good reconstruction, the proposed method is shown to be adjustable to such circumstances by the tuning of its refresh rate.

6. REFERENCES

[1] D.L. Donoho, "Compressed Sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006.

[2] E.J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, February 2006.

[3] V.M. Patel, R. Maleh, A.C. Gilbert, and R. Chellappa, "Gradient-based image recovery methods from incomplete Fourier measurements," *IEEE Transactions on Image Processing*, vol. 21, no. 1, pp. 94–105, January 2012.

[4] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*, Springer, 2010.

[5] A. Majumdar, R. Ward, and T. Aboulnasr, "Compressed sensing based real-time dynamic mri reconstruction," *IEEE Transactions on Medical Imaging*, vol. 31, no. 12, pp. 2253–2266, December 2012.

[6] V. Thirumalai and P. Frossard, "Joint reconstruction of multi-view compressed images," *IEEE Transactions on Image Processing*, vol. 22, no. 5, pp. 1969–1981, 2012.

[7] D.R. Jones, R.O. Schlick, and R.F. Marcia, "Compressive video recovery with upper and lower bound constraints," in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2012, pp. 3325–3328.

[8] G. Yu and G. Sapiro, "Statistical compressed sensing of gaussian mixture models," *IEEE Transactions on Signal Processing*, vol. 59, no. 12, pp. 5842–5858, 2011.

[9] H. Mansour and O. Yilmaz, "Support driven reweighted l1-minimization," *arXiv preprint arXiv:1205.6846*, 2012.

[10] C.J. Miosso, R. von Borries, and J.H. Pierluissi, "Compressive sensing with prior information requirements and probabilities of reconstruction in l1-Minimization," *IEEE Transactions on Signal Processing*, vol. 61, no. 9, pp. 2150–2164, 2013.

[11] R.G. Baraniuk, V. Cevher, M.F. Duarte, and C. Hegde, "Model-based compressive sensing," *IEEE Transactions on Information Theory*, vol. 56, no. 4, pp. 1982–2001, 2010.

[12] X. Wu, W. Dong, X. Zhang, and G. Shi, "Model-assisted adaptive recovery of compressed sensing with imaging applications," *IEEE Transactions on Image Processing*, vol. 21, no. 2, pp. 451–458, February 2012.

[13] N. Vaswani, "Kalman filtered compressed sensing," in *15th IEEE International Conference on Image Processing (ICIP)*. IEEE, 2008, pp. 893–896.

[14] W. Hotrakool and C. Abhayaratne, "Fast compressed sensing reconstruction using the least squares and signal correlation," in *1st IET Intelligent Signal Processing Conference (ISP)*, December 2013, pp. 1.1–1.1(1).

[15] J.Y. Park and M.B. Wakin, "A multiscale framework for compressive sensing of video," in *Picture Coding Symposium, 2009. PCS 2009*. IEEE, 2009, pp. 1–4.

[16] B. Zhao, J.P. Haldar, A.G. Christodoulou, and Z. Liang, "Image Reconstruction from Highly Undersampled (k, t)-Space Data with Joint Partial Separability and Sparsity Constraints," *IEEE Transactions on Medical Imaging*, June 2012.

[17] O. Besson, N. Dobigeon, and J.Y. Tourneret, "Minimum mean square distance estimation of a subspace," *IEEE Transactions on Signal Processing*, vol. 59, no. 12, pp. 5709–5720, 2011.

[18] Z. Tang, Z. Miao, and Y. Wan, *Lecture Notes in Computer Science*, vol. 4740, chapter Background Subtraction Using Running Gaussian Average and Frame Difference, pp. 411–414, Springer Berlin Heidelberg, 2007.

[19] S. Su and Y. Chen, "Moving object segmentation using improved running gaussian average background model," *2008 Digital Image Computing: Techniques and Applications*, pp. 24–31, 2008.

[20] V. Thirumalai and P. Frossard, "Image reconstruction from compressed linear measurements with side information," *18th IEEE International Conference on Image Processing (ICIP)*, pp. 2669–2672, 2011.

[21] D.A. Lorenz, M.E. Pfetsch, and A.M. Tillmann, "Solving basis pursuit: Subgradient algorithm, heuristic optimality check, and solver comparison," *Optimization Online ID 2011-07-3100*, pp. 07–3100, 2012.

[22] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600 –612, April 2004.