A NOVEL COMPRESSION AND RECONSTRUCTION ALGORITHM FOR AERIAL IMAGE SEQUENCE

Xu Cheng[1]*, Member, IEEE, Huang Daqing*[2]*, Han Wei*[1]
College of Electronic Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China
Research Institute of UAV, Nanjing University of Aeronautics and Astronautics, Nanjing, China
xc88@vip.qq.com, nuaauav@126.com, 62458397@qq.com

ABSTRACT

According to the characteristics of aerial image sequence shot by UAV, a novel compression and reconstruction algorithm based on compressed sensing is proposed. The random measurements of each image and the flight parameters are produced at the encoder, and then transmitted to the decoder. At the decoder, a motion estimation model is built based on analyzing the geometric relationship between camera and object, which can effectively reduce the redundancies between images. The de-correlation image is sparser, which can be reconstructed easily and the recovered images have a higher quality. Experimental results show that the proposed algorithm has improved reconstruction performance and consumes less operating time at the encoder, which is more suitable for hardware implementation on UAV.

Index Terms**—**Aerial image sequence, UAV, compressed sensing, motion estimation, image reconstruction

1. INTRODUCTION

The Unmanned Aerial Vehicle (UAV) is a kind of aircraft, which has the advantages of simple structure, small size, light weight, good flexibility, long flight time, low cost etc.. UAV has been used in the field of military, civilian and scientific research widely. In the military , UAV can be used for reconnaissance, surveillance, communications relay, electronic warfare, fire guidance, victories assessment, target localization , early warning etc.. In the civil, UAV can be used in geodesy, meteorology, city environmental monitoring, earth exploration, forest fire prevention etc.. In scientific research, UAV can be used for atmospheric research, validation of new technology and new equipment etc..

In all applications, the UAV should capture images and transmit them to the ground station. With the development of society, the requirements for image clarity and resolution have increased. However, a significant problem is followed, namely, the amount of images increases dramatically. Such enormous amount of information brings serious challenges, particularly, to the embedded systems, such as aircrafts, where the power consumption, memory storage, computational complexity and bandwidth are posing tight constraints on system implementation. Therefore, effective image compression technology for UAV image sequences is an urgent need to resolve this problem.

Many compression methods have been developed over the last twenty years. Among all those works, JPEG and H.264 compression methods are widely used on UAV. However, JPEG does not consider the correlation between image frames, so it is not suitable for image sequence compression. H.264 is good at image sequence compression, but the encoding complexity is high, especially the motion estimation occupies about 80% computation of the entire encoder, which will lead to larger hardware cost and greater power consumption.

In recent years, compressed sensing(CS) has received much attention in the signal processing field. The CS framework was introduced by Candes et al. in [1] - [3] and Donoho in [4]. CS theory enables sparse or compressible signals to be captured and stored at a rate much below the Nyquist rate. As long as the measurements satisfy certain conditions (such as incoherence, RIP), the recovery of the original signal from its randomized projections can be made possible by means of an optimization process. The CS theory can be applied to the field of image compression and reconstruction as well, because a image is sparse in some basis. The image can be reconstructed with high quality via simple random sampling at the encoder and -norm optimization at the decoder. CS has the characteristics of coding easily, decoding complicatedly, and it can reconstruct the image either exactly or with provably small probability of error.

In this paper, we propose a novel image compression and reconstruction framework for UAV image sequences. In this framework, a motion estimation model is built based on analyzing the geometric relationship between camera and object, which can effectively reduce the redundancies between image frames. Furthermore, image reconstruction algorithm base on CS is designed. This framework has slight coding complexity and can be easily implemented in embedded hardware, providing the possibility for the real-time coding and transmission for aerial image sequences. This paper is organized as follows. Section II gives an overview on the proposed scheme firstly, and then introduces the motion estimation model, the basic compressed sensing theory, and compressed sensing implement for aerial image compression and reconstruction in detail. Numerical simulation results are given Section III. Section IV concludes the work.

2. PROPOSED SCHEME

The characteristics of UVA image sequences are different from the general images. The same ground objects often have the corresponding images in multi-frames, which results in tight correlation between image sequences. Furthermore, the flight parameters, such as UAV attitude angle (yaw, pitch, roll), flight height and flight speed can be measured by airborne sensors. Obviously, the common compression algorithms for general image cannot take full advantage of these characteristics of UVA image sequences. In this case, we design a compressive sensing framework for UVA image sequences compression and reconstruction based on motion estimation, which can reduce the correlation among adjacent images, the overall block diagram is shown in Fig. 1.

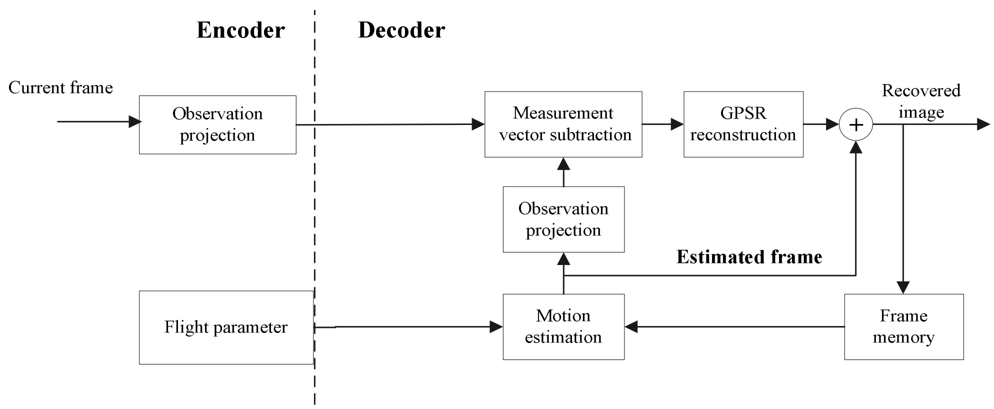


Fig. 1. Scheme structure.

At the encoder, the current frame is observation projected and the flight parameters as auxiliary information is transmitted to the decoder. At the decoder, the motion estimated frame is obtained from the previous recovered frame with motion estimation model. Then the motion estimated frame is observation projected. The residual measurement vector is equal to the current frame measurement vector minus the motion estimated frame measurement vector. After reconstructing the residual image using GPSR (Gradient Projection for Sparse Reconstruction) algorithm [5], the current frame can be recovered combined with the motion estimated frame. Finally, the recovered image is stored in the frame memory, preparing for the next frame image reconstruction.

2.1. Motion estimation model

It is true that there are tight correlations between adjacent video frames. In this section, a motion estimation model is designed for removing the inter-frame correlation.

As can be seen from Fig. 2, the UAV is shooting the ground scenery, set the airborne camera FOV (field of view) as (α, β), the image size as $w_p \times h_p$ and the flight height as H. Obviously, the shooting range of the camera can be calculated by Equation 1.

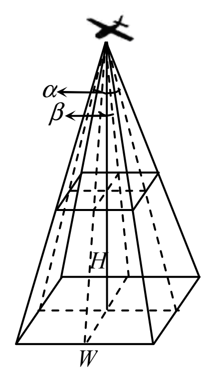


Fig. 2. The geometric relationship of image and scenery.

$$W = 2H\tan(\frac{\alpha}{2}) \tag{1}$$

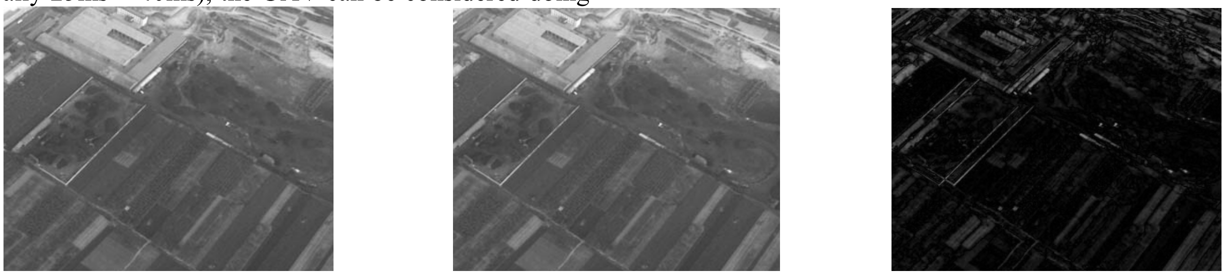
Thus, one pixel represents the length of d

$$d = \frac{W}{w_p} \tag{2}$$

Set the video frame rate as f_{rate}, and then the time interval between two video frames is $t = 1/f_{rate}$. When the flight speed is v, the aircraft movement distance is $s = vt$ between the two frames, and pixels movement distance in the image can be estimated as:

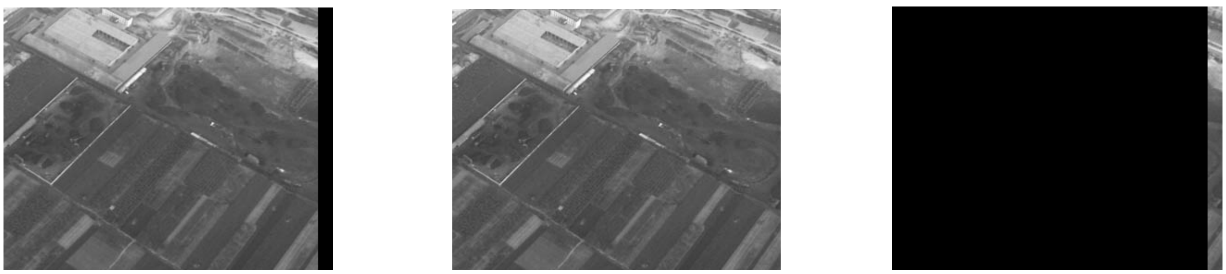
$$p = \frac{s}{d} = \frac{v \cdot w_p}{f_{rate} \cdot 2H \tan(\frac{\alpha}{2})} \tag{3}$$

Due to the time between two frames is rather short (typically 25ms ~ 40ms), the UAV can be considered doing straight flight in a fixed direction and the flight attitude is almost constant. In this case, there are the same ground objects in the area $(w_p - p) \times h_p$ between two adjacent frames, as shown in Fig.3.



(a) image at *t*-1 time (b) image at *t* time (c) residual image

Fig. 3. Correlation between adjacent images.



(a) motion estimated image at *t*-1 time (b) image at *t* time (c) residual image

Fig. 4. Motion estimated image and residual image.

Fig.3(a) is the image taken at t time and Fig.3(b) is the image taken at t-1 time. Fig.3(c) is the residual image obtained by (4).

$$I_r = \left| I_t(x, y) - I_{t-1}(x, y) \right| \tag{4}$$

As can be seen from Fig.3(c), although Fig.3(a) is very similar to Fig.3(b), the residual image still has a relatively clear texture. In order to measure the amount of information of the residual image, discrete image entropy is defined as:

$$H = \sum_{i=0}^{255} p_i \log p_i \tag{5}$$

where $p_i = \dfrac{f(i)}{w_p \times h_p}$, and $f(i)$ is the occurrence rate of gray value i. The entropy value of Fig.3(c) is 5.51, this value is rather large.

In order to reduce the entropy value of the residual image, we construct a motion estimation model as follows:

$$\begin{cases} I_m(x, y) = I_{t-1}(x+p, y) & x \in [1, w_p - p] \\ 0 & x \in (w_p - p, w_p] \end{cases} \tag{6}$$

where $I_m(x, y)$ is the motion estimated image of $I_{t-1}(x, y)$, as shown in Fig.4(a). The residual image can be calculated by (7).

$$I_r = \left| I_t(x, y) - I_m(x, y) \right| \tag{7}$$

Compared with Fig.3(c), Fig.4(c) is sparser, and the entropy value of Fig.4(c) is 0.56, which is far less than that of Fig.3(c).

2.2. CS basic principles

To be able to outline the basic principles of CS, we first introduce some necessary definitions and notation. Signals are considered to be real-valued vectors in an N-dimensional normed Euclidean vector space $\mathbb{R}^N$. For the purposes of CS, mainly the ℓ_0 and ℓ_1 norm are of importance. The sparsity k of a vector x is defined as the number of non-zero components, i.e., $\|x\|_0 = k$. A vector x with sparsity k is said to be k-sparse.

Compressed sensing takes advantage of the sparsity of a signal $s \in \mathbb{R}^N$ in some fixed basis $\mathbf{\Psi} \in \mathbb{C}^{N \times N}$ in order to recover it from a reduced measurement $y \in \mathbb{R}^M$, where $M < N$.

Acquiring a signal by CS consists of two main steps:

1) Measurement: apply a measurement matrix $\mathbf{\Phi}$ to obtain the measurement y,

$$y = \mathbf{\Phi} s = \mathbf{\Phi \Psi} x = \boldsymbol{A} x \tag{8}$$

with $s \in \mathbb{R}^N$ the original signal, $x \in \mathbb{R}^N$ its sparse decomposition, $\boldsymbol{\Psi} \in \mathbb{C}^{N \times N}$ the fixed basis, $\boldsymbol{\Phi} \in \mathbb{R}^{M \times N}$ the measurement matrix, $\boldsymbol{A} = \boldsymbol{\Phi\Psi} \in \mathbb{C}^{M \times N}$ the sensing matrix, and $y \in \mathbb{R}^M$ the measurement.

2) Recovery: exactly recover $\hat{s}$ from using y constrained ℓ_p-norm optimization,

$$\hat{x} = \arg\min_z \|z\|_p \quad \text{s.t.} \quad \boldsymbol{A}z = y \tag{9}$$

$$\hat{s} = \boldsymbol{\Psi}\hat{x} \tag{10}$$

with $p \in \{0,1\}$ in typical CS applications.

The actual choice of the ℓ_p norm in the recovery step 2), has considerable implications on the resulting solution $\hat{x}$ as well as on its computation. Firstly, using an ℓ_0 or ℓ_1 norm in the minimization problem will obviously lead to a sparser solution $\hat{x}$ compared to using ℓ_p norms with $p > 1$ .This is expected to provide a more accurate approximation for the sparse signal under consideration. Secondly, optimization problem has distinct properties depending on the ℓ_p norm used in the minimization, and consequently requires distinct optimization methods. Considering ℓ_0-norm minimization, the resulting optimization problem is non-convex, implying that one has to rely on greedy methods such as orthogonal matching pursuit (OMP) [6]. On the other hand, ℓ_1-norm minimization also induces sparsity in the solution and has the advantage of leading to a convex optimization problem, which can be solved by convex optimization methods or dedicated algorithms such as Basis Pursuit (BP) [7]. It is also possible to consider ℓ_p norms with $p \in \{0,1\}$, again leading to non-convex optimization problems.

2.3. Image compression and reconstruction

The entropy value of the residual image after de-correlation is small, as shown in Fig. 3 (c). Compared to the image at t time, the measurement data of de-correlation image is sparser, more easy to reconstruct. The detailed analysis is as follows.

Set the image at t time as x_t , the motion estimated image at t-1 time as x_m , and x_m is calculated by motion estimation model, which is described in the previous section, then the residual image can be expressed as $x_r = x_t - x_m$, taking sparse transformation and random projection of x_r , its measurement vector is:

$$y_r = \boldsymbol{\Phi\Psi}(x_t - x_m) = y_t - \boldsymbol{\Phi\Psi}x_m \tag{11}$$

where $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$ are defined as (8), y_t represents the measurement vector of x_t . In the process of reconstruction, the image x_t is unknown, but y_t is known, it was transformed from the encoder, and the recovered image $\hat{x}_{t-1}$ at t-1 time which has been reconstructed in advance and stored in frame memory is also known, so x_m can be calculated by the motion estimation model. Then it is easy to obtain the measurement vector y_r by (11). The residual image x_r can be reconstructed by GPSR algorithm, and set the recovered image as $\hat{x}_r$, the finally recovered image $\hat{x}_t$ at t time is:

$$\hat{x}_t = x_m + \hat{x}_r \tag{12}$$

The quality of recovered image $\hat{x}_t$ at t time can be evaluated by (13):

$$\begin{aligned} \|x_t - \hat{x}_t\|_2 &= \|x_t - (x_m + \hat{x}_r)\|_2 \\ &= \|x_t - (x_m + x_r + e_r)\|_2 \\ &= \|(x_t - x_m) - x_r - e_r\|_2 = \|e_r\|_2 \end{aligned} \tag{13}$$

where e_r is the reconstruction noise of residual image x_r .The above equation shows that the quality of recovered image is directly related to the reconstruction noise of residual image. The key idea of this approach is reconstructing the residual image instead of reconstructing the image at t time directly, which will lead to higher reconstruction efficiency because the entropy value of the residual image is smaller and its measurement is sparser compared with that of the original one.

3. SIMULATIONS

The experimental image sequences of this paper were shot by a small UAV. We chose two sequences, the first sequence includes 150 frames and the second sequence includes 220 frames. The image size is 1280×720, and the camera Fov is (52.3°,40.1°). The flight height and flight speed for each frame is obtained by sensors.

In order to verify the performance of this algorithm, two numerical experiments are carried out in the same platform (Intel 4-core 1GHZ/2G memory). (1) GPSR algorithm proposed in literature [5]. (2) The proposed algorithm in this paper. Define measurement rate as MR=M/N, M is the length of measurement vector, N is the length of original signal.

Tab. 1 shows the average PSNR (Peak Signal to Noise Ratio) and the average reconstruction time of two experimental image sequences at MR=0.3 and MR=0.6 using different algorithm . As can be seen from Tab. 1, the proposed algorithm has improved the quality of the reconstruction compared with GPSR algorithm. When MR=0.6, the average PSNR of the proposed algorithm has raised 2.99dB and 3.70dB for Sequence 1 and Sequence 2, respectively. When MR=0.3, the average PSNR of the proposed algorithm has raised 4.31dB and 2.30dB for Sequence 1 and Sequence 2, respectively. Furthermore, the operating time using our algorithm is also shorter than using GPSR algorithm, the operating time of the proposed algorithm is nearly 40% of GPSR algorithm.

Fig. 5 and Fig. 6 show the PSNR of Sequence 1 and Sequence 2 at MR=0.6 using this two algorithms. It is explicit that the PSNR of each frame using our algorithm is higher than or equal to using GPSR algorithm.

Table 1. Performance comparison between proposed method and GPSR method

Image sequence	Total number of frames	MR	Average PSNR (dB)		Average operating time (s)	
			GPSR	***ours***	***GPSR***	***ours***
Sequence 1	150	0.6	66.42	69.41	22.43	8.34
		0.3	62.46	66.77	24.63	9.55
Sequence 2	224	0.6	67.51	71.21	23.21	9.41
		0.3	64.02	66.32	25.87	10.58

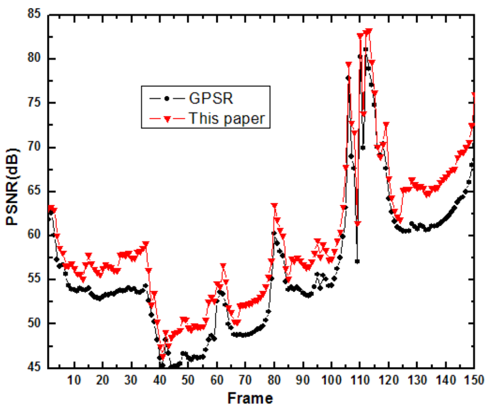


Fig. 5. Comparison of Sequence 1 PSNR.

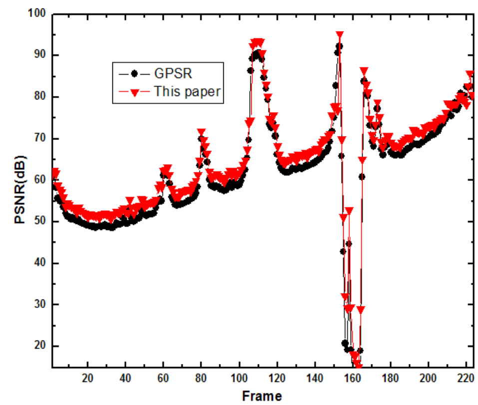


Fig. 6. Comparison of Sequence 2 PSNR.

4. CONCLUSION

In this paper a novel compression and reconstruction algorithm based on compressed sensing is proposed. Our approach has two main innovations: 1) a motion estimation model is built based on analyzing the geometric relationship between camera and object, which can effectively reduce the redundancies between images. 2) the theory of CS has been applied in image compression and reconstruction, which can greatly reduce the computational complexity and data storage at the encoder side. By the simulations it is indicated that the proposed approach can not only raise PSNR of the recovered image, but also reduce operating time. It is true that this approach has great value in application and is very suitable for the field of aerial image sequence.

REFERENCES

[1] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information" , *IEEE Trans. on Information Theory*, vol. 52, pp. 489-509, 2006.

[2] E. Candès, "Compressive sampling", *Proc. International Congress of Mathematics*, vol. 3, pp. 1433-1452, Madrid, Spain, 2006.

[3] E. Candès and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies" , *IEEE Trans. on Information Theory*, vol. 52, pp. 5406-5425, 2006.

[4] D. Donoho, "Compressed sensing" , *IEEE Trans. on Information Theory*, vol. 52, pp. 1289-1306, 2006.

[5] Figueiredo, Mário AT, Robert D. Nowak, and Stephen J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems", *IEEE Journal on Selected Topics in Signal Processing*, vol. 1, pp. 586-597.

[6] Pati, Yagyensh Chandra, Ramin Rezaiifar, and P. S. Krishnaprasad. "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition", *Signals, Systems and Computers*, pp. 40-44,1993.

[7] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit" , *SIAM J. Sci. Comput.*, vol. 20, No. 1, pp. 33–61, 1998.