

# A Novel Video Compression Method Based on Underdetermined Blind Source Separation

Jing Liu, Fei Qiao, Qi Wei and Huazhong Yang

**Abstract** If a piece of picture could contain a sequence of video frames, it is amazing. This paper develops a new video compression approach based on underdetermined blind source separation. Underdetermined blind source separation, which can be used to efficiently enhance the video compression ratio, is combined with various off-the-shelf codecs in this paper. Combining with MPEG-2, video compression ratio could be improved slightly more than 33 %. As for combining with H.264, twice compression ratio could be achieved with acceptable PSNR, according to different kinds of video sequences.

**Keywords** Underdetermined blind source separation • Sparse component analysis • Video surveillance system • Video compression

## 1 Introduction

Digital video is famous for its abundant information and its rigid demand for the bandwidth and process power as well. It leads to the emergence of multiple video coding standards, such as MPEG-2, H.264. However, new thoughts can be applied to compress video as well.

Blind Source Separation (BSS) provides a solution to recover original signals from mixed signals. It can be used in multiple fields, such as wireless communication to separate mixed radio signals, biomedicine to separate fetal electrocardiogram signals recorded by sensors, and typical “cocktail party” problem to separate mixed speech signals.

---

J. Liu · F. Qiao (✉) · Q. Wei · H. Yang

Department of Electronic Engineering, Tsinghua University, Beijing 100084, China  
e-mail: qiaofei@tsinghua.edu.cn

Independent Component Analysis (ICA) was widely accepted as a powerful solution of BSS since the past 20 years [1]. In 1999, A. Hyvarinen presented an improved ICA algorithm, called FastICA [2]. A detailed overview of many algorithms on BSS is made and their usages on image processing are presented as well [3]. However, few researchers focused on utilizing BSS into video processing.

In this paper, we apply Underdetermined BSS (UBSS, meaning the number of original signals is more than that of mixed signals) to compress video sequences. A new codec defined as Underdetermined Blind Source Separation based Video Compression (UBSSVC) is developed. As we explained later in detail, UBSSVC has good performance on video compression.

This paper is organized as follows. The next section briefly reviews BSS problem. In Sect. 3, detailed structure of UBSSVC is stated. And Sect. 4 shows simulation results. Finally, Sect. 5 summarizes the superior and deficiency of this video compression method. Also, future work is proposed in this section.

## 2 Blind Source Separation and Solution to UBSS

BSS was first established by J. Herault and C. Jutten in 1985 [1]. It can be described as following: multiple signals from separate sources  $s$  are somehow mixed into several other signals, defined as mixed signals  $x$ . Here  $n$  represents the number of source signals and  $m$  represents the number of mixed signals. The objective of BSS is to design an inverse system to get the estimation of source signals. The reason for the “Blind” here is neither the source signals nor the mixed process is known to the observer. The mixing model can be expressed as,

$$x = As; \quad A \in R^{m \times n}, s \in R^{n \times T} \quad (1)$$

where  $A$  is an  $m \times n$  mixing matrix. Both  $A$  and  $s$  are unknown, while  $x$  is known to observer.

Independent Component Analysis (ICA) is the main solution for overdetermined ( $n < m$ ) and standard ( $n = m$ ) case. However, it is not suitable for UBSS ( $n > m$ ). Other methods like Sparse Component Analysis (SCA) [4–7] and over-complete ICA [8, 9], are investigated for UBSS recent years.

In this work, SCA is adopted to solve the UBSS. SCA uses the sparsity of source signals to compensate information loss in the mixing process. So specific assumptions of mixing matrix  $A$  and source matrix  $s$  should be considered as follows [4]. Assumption 1: any  $m \times m$  square sub-matrix of mixing matrix  $A \in R^{m \times n}$  is nonsingular; assumption 2: there are at most  $m - 1$  nonzero elements of any column of matrix  $s$ . If the above assumptions are satisfied, the source matrix  $s$  can be recovered by SCA.

Let  $x_i$ ,  $i = 1, 2, \dots, m$  and  $s_i$ ,  $i = 1, 2, \dots, n$  represent mixed signals and source signals respectively; and  $a_j$ ,  $j = 1, \dots, n$  is the  $j$ th column of mixing matrix  $A$ . Therefore, the mixing process can also be described as following.

$$x(t) = (x_1(t) \ x_2(t) \ \cdots \ x_m(t))^T = a_1 s_1(t) + a_2 s_2(t) + \cdots + a_n s_n(t) \quad (2)$$

Given the mixing matrix  $A$  satisfies the assumption 1, any  $m - 1$  columns of  $A$  span a  $m$ -dimensional linear hyperplane  $\mathcal{H}_q$ , which can be denoted as  $\mathcal{H}_q = \{h | h \in R^m, \lambda_{ik} \in R, h = \lambda_{i_1} a_{i_1} + \cdots + \lambda_{i_{m-1}} a_{i_{m-1}}\}$ , where  $q = 1, \dots, C_n^{m-1}$ . If source matrix  $s$  satisfies assumption 2, it is reasonable to suppose that at the  $t$  moment, all source signals except for  $s_{i_1}, s_{i_2}, \dots, s_{i_{m-1}}$  are zero, where  $\{i_1, i_2, \dots, i_{m-1}\} \subset \{1, 2, \dots, n\}$ . Consequently, at  $t$  moment, Eq. (2) can be rewritten as

$$(x_1(t) \ x_2(t) \ \cdots \ x_m(t))^T = a_{i_1} s_{i_1}(t) + a_{i_2} s_{i_2}(t) + \cdots + a_{i_{m-1}} s_{i_{m-1}}(t) \quad (3)$$

From (3), it can be concluded that the  $t$ th column vector of observed signals matrix  $x$  is in one of  $C_n^{m-1}$  hyperplanes  $\mathcal{H}$ . Therefore, mixed frames can be recovered by the following algorithm.

- (a) Get the set  $\mathcal{H}$  of  $C_n^{m-1}$   $m$ -dimensional hyperplanes which are spanned by any  $m - 1$  columns of  $A$ ;
- (b)  $j$  repeat from 1 to  $m$ ,
  - (i) If  $x_j$ , which stands for the  $j$ th column of mixed signals matrix  $x$ , is in a hyperplane  $\mathcal{H}_q$ , then the following equation can be gotten

$$x_j = \sum_{v=1}^{m-1} \lambda_{i_v j} a_{i_v} \quad (4)$$

- (ii) Comparing Eqs. (3) and (4),  $s_i$ , the  $i$ th column of source signals matrix  $s$ , can be recovered: its components are  $\lambda_{i_v j}$  in the place  $i_v, v = 1, \dots, m - 1$ , and other components equal to zero.

### 3 Proposed UBSSVC Method

As explained above, for UBSS the number of mixed signals is less than that of source signals. Therefore, the mixing process of UBSS could be used to compress video sequences, and the separating process is used to decode the compressed video sequences.

#### 3.1 Mapping UBSS to Video Compression

Consider a video sequence with  $L$  frames,  $s_1, s_2, \dots, s_L$ , where  $s_i \in R^T$  is a T-pixel frame; we firstly divide the  $L$  video frames into  $b$  groups and in each group there are  $n$  frames. The encoder first chooses a matrix  $A \in R^{m \times n} (m < n)$  to mix  $n$  frames in each group. Thus, the compression ratio is  $n/m$ .

At the encoder side, unlike the traditional scenario of the UBSS issue, the mixing process is factitious in this proposed method. Thus, a specific mixing matrix  $A$ , known by both encoder and decoder, is chosen to mix raw video frames.

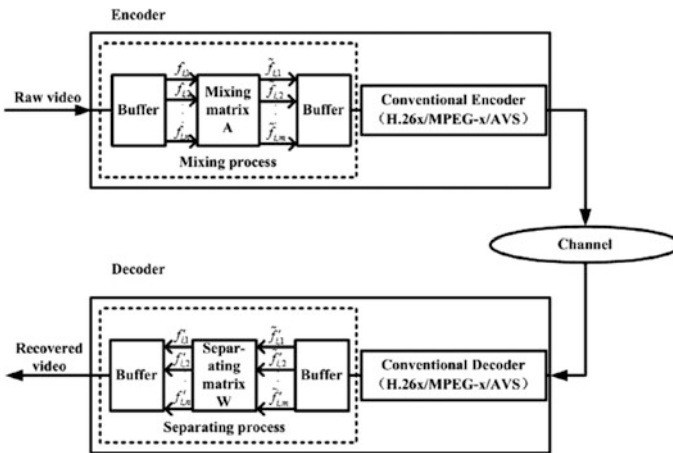
For standard BSS, there is only one restriction of mixing matrix  $A$ , that the columns of  $A$  should be mutually independent. However, in the proposed method, matrix  $A$  not only needs to satisfy the assumption 1, but also has to decrease the information loss in mixing process. Thus, in different mixed frames, the weight of different original frames should be varied. As each component of a row of  $A$  can be treated as the weight of every original frame in a mixed frame, the components of a row of  $A$  should be varied largely from each other. Experiments will be done to show  $A$ 's influence on the separation results in Sect. 4.

At the decoder side, the matrix  $A$  is known exactly, so the frames' order of recovered video sequence is not disturbed by mixing process and separating process, which is different from traditional BSS.

To ensure that the frames could satisfy the assumption 2, mixed frames are first transformed by a 2-D discrete Haar wavelet transform. And then SCA is used to recover the sparse high frequency components, while the recovered low frequency components are equal to multiply generalized inverse of mixing matrix  $A$  by mixed low frequency components.

### 3.2 Proposed UBSSVC Structure

The compression ratio for UBSS is only  $n/m$ . Therefore, to enhance the compression ratio more, we proposed UBSSVC framework that combines UBSS and conventional codec together, shown in Fig. 1.



**Fig. 1** Framework of UBSSVC

At the encoder side,  $n$  frames,  $f_{i,1}, f_{i,2}, \dots, f_{i,n}$ , are mixed into  $m$  frames,  $\tilde{f}_{i,1}, \tilde{f}_{i,2}, \dots, \tilde{f}_{i,m}$ . And then these mixed frames are encoded by traditional encoder such as MPEG-2, H.264. The buffer before mixing is used to buffer enough frames for being mixed. And the buffer after mixing is for storing mixed frames temporarily so that they can be encoded by conventional encoder one by one.

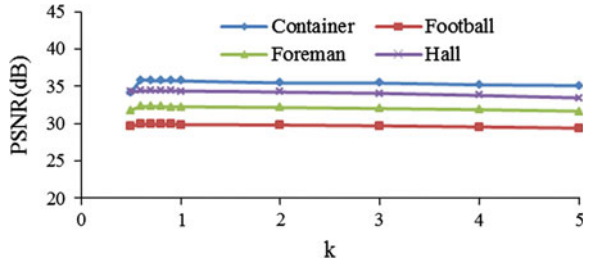
In the proposed decoder structure, received data is firstly decoded by traditional decoder; then the source recovery algorithm of underdetermined BSS is applied to recover original video sequence. In the separating process,  $m$  frames,  $\tilde{f}_{i,1}, \tilde{f}_{i,2}, \dots, \tilde{f}_{i,m}$ , are separated to  $n$  frames,  $f'_{i,1}, f'_{i,2}, \dots, f'_{i,n}$ . The function of two buffers in decoder is similar to that of those two buffers in encoder.

## 4 Experiment Results

In order to validate this approach, multiple simulations are performed on four standard test video sequences: hall, container, foreman and football. The football sequence has the largest temporal variations, followed by foreman, and container ranks the third, while the hall sequence contains the most slowly scene variations. The first 40 frames of each sequence are used for test. Peak-Signal-to-Noise Ratio (PSNR) is used to evaluate the performance of recovery algorithm.

In the experiments, we just show an example of mixing 4 video frames into 3 frames, so the compression ratio of UBSS in the experiments is just 4/3. The mixing matrix  $A \in R^{3 \times 4}$ , shown in (5), is chosen to mix raw video sequence, where  $k \in Z, k \neq 0$ . The mixing process is performed as follows: continuous 4 frames are taken as source signals  $s$ , then  $A$  multiplies by  $s$  to calculate the mixed frames  $x$ . 30 mixed frames are generated after the mixing process. And then the above algorithm is applied to separate these mixed frames. Figure 2 shows the recovery PSNR on different video test sequences when  $k = 0.5-5$ . These plots show that the value of  $k$  has little influence on the separation PSNR. Although for some sequence, such as football, PSNR is a little low, it is still enough for monitor applications, which don't have very strict demands on high resolution.

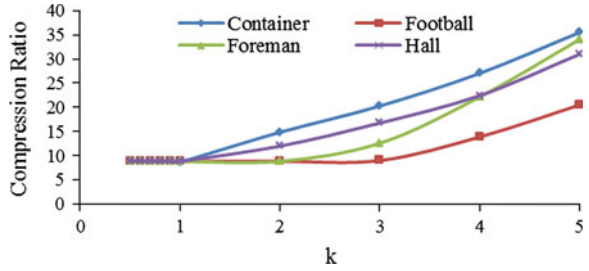
**Fig. 2** Separation PSNR related to different values of  $k$  on four videos



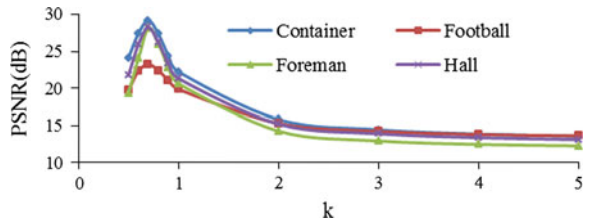
$$A = \frac{1}{k} \begin{pmatrix} 0.30 & 0.45 & 0.15 & 0.10 \\ 0.35 & 0.15 & 0.05 & 0.45 \\ 0.30 & 0.05 & 0.45 & 0.20 \end{pmatrix} \quad (5)$$

Experiments are done as well to show the  $k$ 's effect to the compression ratio and separation PSNR of UBSSVC+MPEG-2 which means that the conventional codec in Fig. 1 is MPEG-2, UBSSVC+H.264 which means that the conventional codec in Fig. 1 is H.264. Results are shown in Figs. 3, 4, 5 and 6. From the results, the  $k$  values indeed affect the UBSSVC+MPEG-2 and UBSSVC+H.264 compression ratio. That's because with the increment of  $k$ , most pixels values of the mixed frames approach to zero. Therefore, the compression ratios of MPEG-2 and H.264 for these mixed frames are much larger than that for the original frames. Meanwhile, it leads to a higher distortion. So the decoding PSNR decreases with the  $k$  increment when  $k > 0.7$ . For these four different test sequences, the largest PSNR and lowest compression ratio are almost gotten at the point  $k = 0.7$ . However, even the lowest compression ratio is larger than the corresponding

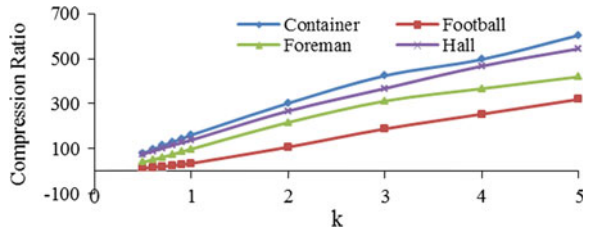
**Fig. 3** Compression ratio of UBSSVC+MPEG-2 related to different values of  $k$  on four video



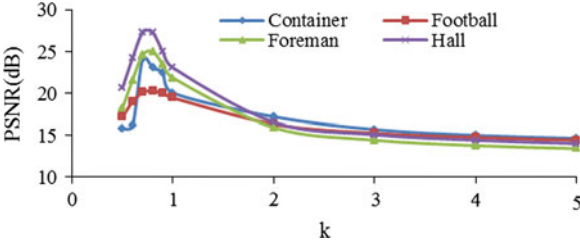
**Fig. 4** Decoding PSNR of UBSSVC+MPEG-2 related to different values of  $k$  on four videos



**Fig. 5** Compression ratio of UBSSVC+H.264 related to different values of  $k$  on four videos



**Fig. 6** Decoding PSNR of UBSSVC+H.264 related to different values of  $k$  on four videos



**Table 1** Compression results of test videos

	Hall (dB)	Container (dB)	Foreman (dB)	Football (dB)
MPEG-2	6.48	6.48	6.48	6.48
UBSSVC+MPEG-2 ( $k = 0.7$ )	8.84	8.84	8.84	8.84
H.264	76.74	92.28	62.04	12.15
UBSSVC+H.264 ( $k = 0.7$ )	102.41	161.06	74.25	25.38

**Table 2** PSNR results of test videos

	Hall (dB)	Container (dB)	Foreman (dB)	Football (dB)
MPEG-2	37.87	30.86	27.04	27.11
UBSSVC+MPEG-2 ( $k = 0.7$ )	28.23	29.18	28.15	23.23
H.264	36.19	35.43	35.13	31.7
UBSSVC+H.264 ( $k = 0.7$ )	27.26	24.16	25.01	20.29

compression ratio of MPEG-2 and H.264. Tables 1 and 2 show the comparison results. The PSNRs of UBSSVC+H.264 ( $k = 0.7$ ), and UBSSVC+MPEG-2 ( $k = 0.7$ ) is lower than those of H.264 and MPEG-2 respectively. Although the PSNR value is a little low, it is enough for some applications which don't have strict demands on high resolution, such as video surveillance system.

### 5 Conclusion

This paper initially develops the novel video compression approach UBSSVC. Furthermore, experiments are conducted to validate the efficiency of recovery algorithm, the influence of  $k$  values on separation PSNR and to measure the video compression ratio improvements of UBSSVC. The proposed method is suitable for video surveillance system perfectly. Firstly, it can achieve higher video compression ratio to decrease the bandwidth resource utilization. Secondly, the computation complexity of mixing process at encoder side is low, when improving the

video compression ratio. What's more, the mixing and separating process of UBSS has great potential in low-complexity video compression.

However, the presented new method still has more issues to be improved in our future work. Like the largest compression ratio the UBSS can achieve, and how to improve the compression ratio gained by the mixing process and enhance the separating results of video quality.

## References

1. Hyvarinen A, Karhunen J, Oja E (2001) Independent component analysis. Wiley, New York
2. Hyvarinen A (1999) Fast and robust fixed-point algorithms for independent component analysis. *IEEE Trans Neural Netw* 10:626–634
3. Cichocki A, Amari SI (2002) Adaptive blind signal and image processing: learning algorithms and applications. Wiley, Chichester
4. Georgiev P, Theis F, Cichocki A (2005) Sparse component analysis and blind source separation of underdetermined mixtures. *IEEE Trans Neural Netw* 16:992–996
5. Li YQ, Cichocki A, Amari SI, Shishkin S, Cao JT, Gu FJ (2004) Sparse representation and its applications in blind source separation. In: Thrun S, Saul K, Scholkopf B (eds) *Advances in neural information processing systems* 16, vol 16. MIT Press, Cambridge, pp 241–248
6. Ren M.-r, Wang P (2009) Underdetermined blind source separation based on sparse component in electronic computer technology, 2009 international conference on. pp 174–177
7. Zhenwei S, Huanwen T, Yiyuan T (2005) Blind source separation of more sources than mixtures using sparse mixture models. *Pattern Recogn Lett* 26:2491–2499
8. Lee TW, Lewicki MS, Girolami M, Sejnowski TJ (1999) Blind source separation of more sources than mixtures using overcomplete representations. *IEEE Signal Process Lett* 6:87–90
9. Waheed K, Salem FM (eds) (2003) Algebraic independent component analysis: An approach for separation of overcomplete speech mixtures. In: *Proceedings of the IEEE, international joint conference on neural networks 2003*, vols 1–4. New York, pp 775–780



Multimedia and Ubiquitous Engineering

MUE 2013

Park, J.H.; Ng, J.K.-Y.; Jeong, H.-Y.; Waluyo, B. (Eds.)

2013, XXXI, 1209 p. 541 illus. In 2 volumes, not available separately., Hardcover

ISBN: 978-94-007-6737-9