The Wiener Filter-based Adaptive Denoising for Pseudo Analogy Video Transmission

WANNING HE, XIN-LIN HUANG*, (Senior, IEEE), PENGFEI LI

1Department of Electronic and Information Engineering, Tongji University, Shanghai 201804, China
Corresponding author: Xin-Lin Huang (e-mail: xlhuang@tongji.edu.cn).

This work was supported in part by the National Natural Science Foundation of China under Grants 62071332 and in part by the Fundamental Research Funds for the Central Universities.

ABSTRACT With the popularity of video conferences, video calls and other activities, video transmission has been widely used. To meet a huge number of subscribers’ requirements, the mobile video transmission scheme needs to overcome some disadvantages, such as resources limitation and noise interference. The knowledge-enhanced mobile video broadcasting (KMV-Cast) is a scheme utilizing joint source-channel coding and correlated information in clouds. However, there is an item of noise that cannot be removed in the original KMV-Cast scheme. In this paper, an adaptive Wiener filtering denoising algorithm is proposed to reduce such noise at the receiver in order to maximize the signal-to-noise ratio (SNR) of the reconstructed video frame. The simulation results show that the proposed Wiener filter algorithm is superior to other schemes without the Wiener filter under different sources and channel qualities. At lower-SNR channels (i.e., -5dB), the proposed algorithm achieves 2dB gains in terms of peak signal-to-noise ratio (PSNR), while at higher-SNR channels (i.e., 10dB), the proposed algorithm achieves 3dB gains in terms of PSNR.

INDEX TERMS Wireless Video Transmission, Denoising Scheme, Wiener Filter

I. INTRODUCTION

With the prediction of Cisco Annual Internet Report 2020, nearly two-thirds of the global population will have Internet access by 2023. There will be 5.3 billion total Internet users (66 percent of global population) by 2023, up from 3.9 billion (51 percent of global population) in 2018[1]. With the increasing total number of the Internet users, the video transmission service and its data-stream traffic also increase accordingly, which puts pressure on the current wireless video transmission technology. In a wireless communication system, when the wireless resources are limited, it is necessary to reduce the transmission rate of redundant data. Meanwhile, channel noise affects the quality of reconstructed video, and thus it is also necessary to remove noise in the reconstructed video at the receiver.

As we all know, the traditional wireless video transmission scheme is based on the Shannon source-channel separation theory, while in 2003, M. Gastpar, B. Rimoldi and M. Vetterli proposed that the joint source-channel coding could be optimal in some conditions [2]. When the channel quality is below a certain threshold in mobile communications, the quality of the received video declines sharply, which is called cliff effect. To overcome such cliff effect, a typical scheme called Softcast was proposed with the joint source-channel coding scheme [3]. During the process of video coding and decoding, only linear transformation is used, and thus the quality of the reconstructed video is linearly correlated with the channel quality. Then, based on the Softcast scheme, some improved pseudo-analog transmission schemes were proposed in [4]-[7]. Besides, many companies are using cloud services, and move their applications and services from local to the cloud [8]. As a result, the video download from cloud services becomes a hotspot. The knowledge-enhanced mobile video broadcasting (KMV-Cast) is a brand-new video transmission framework [9]. Compared with other pseudo-analog video transmission schemes in [3]-[7], it leverages cloud related information in video coding and decoding. During the reconstruction at the receiver, KMV-Cast eliminates the noise of mutual interference. In this paper, we focus on reducing the residual noise in KMV-Cast scheme further.

Generally, noise is classified into additive noise and multiplicative noise [10]. The distribution of noise is very impor-
tant and widely used in conventional denoising algorithms. Recently, the convolutional neural network (CNN) has been widely used for denoising in image processing [11]-[13]. However, such algorithms do not exploit the relevant information in clouds [11]-[13].

Based on the statistical characteristics of Gaussian noise, the least mean square (LMS) error has been applied as the optimization goal [14]-[16]. For a video frame, the values of adjacent pixels are usually near with less difference. As a result, the Wiener filter has been widely used for the image denoising [17]-[21]. The goal of the Wiener filtering is minimizing mean square error (MMESE). It should be noted that the prior knowledge on the power spectral density of the noise should be previous given in the Wiener filter. Hence, Wiener filtering maybe a promising method to remove the second noise item in KMV-Cast scheme and this paper will propose an adaptive denoising algorithm with Wiener filter.

The rest of the paper is organized as follows. We take a brief review on KMV-Cast scheme in Section 2. We propose a denoising algorithm in Section 3 and each pixel block is considered as a unit which is selectively transmitted and adaptively passing a Wiener filter. In Section 4, we present the detailed frameworks of the proposed algorithm and the simulation results are shown in Section 5. Finally, Section 6 concludes this paper.

II. KMV-CAST SCHEME: A BRIEF REVIEW

The KMV-Cast transmission scheme was proposed in [9] as shown in Fig. 1. As one can see from Fig. 1, there are two cloud servers at the transmitter and receiver separately, and the correlated information of the transmitted video is available at both sides. At the transmitter, the correlated information in the cloud is used to remove the redundant information in the broadcasting video. At the receiver, with the help of correlated information in the local cloud, the video is reconstructed by maximizing SNR.

In technical, Fig. 1 mainly contains 1) hierarchical Bayesian model, 2) related information and prior knowledge extraction, and 3) reconstruct received signal.

A. HIERARCHICAL BAYESIAN MODEL

In order to make full use of the relevant information, KMVCast uses the hierarchical Bayesian model to describe the relationship between transmitted video and its correlated information in the cloud.

At the transmitter, video frames are evenly divided into small pixel blocks (i.e., 8 by 8). For each block, the 2D-DCT transform is used to reduce the redundant information. The high-frequency DCT coefficients, close to zero, can be discarded to save the transmission bandwidth. Then, we scale the DCT coefficients and reshape them into an \( m \times 1 \) vector (i.e., \( \alpha \theta \)), where \( \alpha \) is the power scaling factor and \( \theta \) is a normalized vector. The original DCT coefficients can be represented as \( \lambda \theta \), where \( \lambda \) is the amplitude value of the block. Besides, an \( m \times m \) unitary matrix (\( \Phi \)) is multiplied to reduce the peak-to-average power ratio. As a result, the received signal using the pseudo analog modulation can be represented as

\[
y = \alpha \Phi \theta + v
\]

where \( v \) is an \( m \times 1 \) vector and represents independently and identically distributed Gaussian noise with zero-mean and a known variance \( \sigma_v^2 \).

At the receiver, the work in [9] proceed the video reconstruction based on Eqn. (1) with the hierarchical Bayesian model, and the likelihood function of the received signal can be represent as

\[
p \{ y | \theta, \sigma_\theta^2 \} = \frac{1}{\sqrt{2\pi \sigma_\theta^2}} \exp \left\{ -\frac{1}{2} \left( y - \alpha \Phi \theta \right)^2 / \sigma_\theta^2 \right\}
\]

To simplify, the work in [9] also use a Gaussian distribution to approximate the probability density function of the DCT coefficients \( \theta \), i.e., \( \theta \sim (0, \Omega) \). Given the received signal \( y \) and hyperparameter \( \Omega \), the posterior probability density function of \( \theta \) can be written as a multivariate Gaussian distribution, that is,

\[
p \{ \theta | y, \sigma_\theta^2 \} = \frac{p \{ y | \theta, \sigma_\theta^2 \} p \{ \theta | \Omega \}}{\int p \{ y | \theta, \sigma_\theta^2 \} p \{ \theta | \Omega \} d\theta}
\]

Since the posterior probability in Eqn. (3) is a multivariate Gaussian distribution, it can be denoted as

\[
\theta | y, \sigma_\theta^2 \sim N ( u, \Sigma )
\]

In Eqn. (4), \( u \) and \( \Sigma \) are the mean value and the variance of the Gaussian distribution respectively, which are given as [9]

\[
u = \Sigma \alpha \Phi^T \sigma_\theta^2 y
\]

\[
\Sigma = ( \alpha^2 \sigma_\theta^2 \Omega - I )^{-1}
\]

If we take the mean value of the posterior probability in Eqn. (4) as the reconstructed video block, there is a mutual interference at the reconstructed signal (see Eqn. (5)). The KMV-Cast scheme proposed a method to cancel such mutual interference at the transmitter by rotating original DCT coefficient as [9]

\[
\tilde{\theta} = \left[ I - \left( \alpha^2 \sigma_\theta^2 \Omega + I \right)^{-1} \right]^{-1} p \theta
\]

where \( p \) is a constant and will be defined in Eqn. (14), and thus the new received signal like Eqn. (1) can be rewritten as

\[
y = \alpha \Phi \tilde{\theta} + v.
\]
Thus, the Bayesian estimation of the reconstructed video can be rewritten as

$$\hat{\theta} = p\theta + (\alpha^2\sigma_0^{-2}I + \Omega^{-1})^{-1} \alpha\Phi^T \sigma_0^{-2}v. \quad (9)$$

**B. RELATED INFORMATION AND PRIOR KNOWLEDGE EXTRACTION**

Bayesian estimation is used again for relevant information in order to find the video structure information $\Omega$. Assume that there are $N$ related video pixel blocks available in clouds, that is,

$$\{\tilde{\theta}_i|i = 1, 2, ..., N\}. \quad (10)$$

Here, we choose the most correlated pixel block in the cloud, and the corresponding Bayesian estimation of video structure information with the maximal SNR can be written as [9]

$$\hat{\Omega} = rI + \tilde{\theta}_0\tilde{\theta}_0^T. \quad (11)$$

Redefine $C = \alpha^2\sigma_0^{-2}$ as the power scaling parameter, $r$ is an undetermined parameter and the expression (7) can be rewritten as:

$$\hat{\theta} = \left[I - (C\Omega + I)^{-1}\right]^{-1} p\theta \quad (12)$$

where

$$(C\Omega + I)^{-1} = \left[C(rI + \tilde{\theta}_0\tilde{\theta}_0^T) + I\right]^{-1} = \frac{I}{Cr + 1 - \frac{1}{(Cr + 1)(Cr + C + 1)}}. \quad (13)$$

Since $\theta$ and $\hat{\theta}$ are both normalized vectors, we have $\|\hat{\theta}\|^2 = 1$ and thus

$$p^2 = (\frac{Cr}{Cr + 1})^2 \left\{\frac{K^2}{(Cr + C + 1)^2} - \frac{2K^2}{(Cr + C + 1)(Cr + t + 1)}\right\}^{-1}. \quad (14)$$

In Eqn. (14), $K = (\tilde{\theta}_0^T \theta)$ represents the correlation coefficient.

**C. RECONSTRUCTION FROM THE RECEIVED SIGNALS**

At the receiver, the goal is to maximize the quality of the reconstructed video. The demodulated signal at the receiver can be written as

$$\hat{\theta} = (\alpha^2\sigma_0^{-2}I + \Omega^{-1})^{-1} \alpha\sigma_0^{-2}\Phi^Tv$$

$$= p\theta + \frac{\alpha\sigma_0^{-2}r\Phi^Tv}{Cr + 1} + \frac{\alpha\sigma_0^{-2}\hat{\theta}_0\tilde{\theta}_0^T (\Phi^Tv)}{(Cr + 1)(Cr + C + 1)}. \quad (15)$$

Since the transmitter and the receiver both have the information $\tilde{\theta}_i$ in their cloud, we can multiple $\tilde{\theta}_i^T$ at both sides of Eqn. (15) simultaneously to get

$$\tilde{\theta}_i^T \hat{\theta} = \tilde{\theta}_i^T (\alpha^2\sigma_0^{-2}I + \Omega^{-1})^{-1} \alpha\sigma_0^{-2}\Phi^Tv$$

$$= p\tilde{\theta}_i^T \theta + \frac{\alpha\sigma_0^{-2}\tilde{\theta}_i^T (\Phi^Tv)}{Cr + 1} + \frac{\alpha\sigma_0^{-2} (\tilde{\theta}_i^T \hat{\theta}) \tilde{\theta}_i^T (\Phi^Tv)}{(Cr + 1)(Cr + C + 1)}. \quad (16)$$

Thus, we can calculate

$$\tilde{\theta}_i^T (\Phi^Tv) = \frac{\tilde{\theta}_i^T \hat{\theta} - pK}{\frac{\alpha\sigma_0^{-2}r}{Cr + 1} + \frac{\alpha\sigma_0^{-2}}{(Cr + 1)(Cr + C + 1)}}. \quad (17)$$

Then, the third item in Eqn. (15) can be removed given the result in Eqn. (17) and Eqn. (15) can be rewritten as

$$\hat{\theta} = p\theta + \frac{\alpha\sigma_0^{-2}r\Phi^Tv}{Cr + 1}. \quad (18)$$

From Eqn. (18), we obtain the noise power as:

$$P_N = E\left\{tr\left(\left(\frac{\alpha\sigma_0^{-2}r\Phi^Tv}{Cr + 1}\right)^2\right)\right\}$$

$$= E\left\{tr\left(\frac{\alpha^2\sigma_0^{-4}r^2\Phi^Tv\Phi^Tv}{(Cr + 1)^2}\right)\right\} = mCr^2. \quad (19)$$

The corresponding signal-to-noise ratio $SNR_1$ of the reconstructed signal in Eqn. (18) is

$$SNR_1 = \frac{P_S}{P_N} = \frac{p^2}{mCr^2} = \frac{(Cr + 1)^2 p^2}{mCr^2}. \quad (20)$$

Two new variables are denoted in [9] for simplicity,

$$t = r \left(Cr + C + 1\right) \quad (21)$$

$$\sqrt{C} = \frac{\sqrt{K^2 - 2K^2 (t + 1) + (t + 1)^2}}{t} \quad (22)$$

Substitute the Eqn. (22) into Eqn. (14) to get the power expression,

$$p^2 = \frac{Cr^2A^2}{(Cr + 1)^2} \quad (23)$$

and the expression of $SNR_1$ can be transformed into

$$SNR_1 = \frac{A^2}{m}. \quad (24)$$

**III. ADAPTIVE DENOISING WITH WIENER FILTER FOR THE KMV-CAST SCHEME**

This section will propose the adaptive denoising algorithm with Wiener filter, which will further reduce the noise existing in KMV-Cast scheme. The diagram of this algorithm is highlighten in Fig. 2.
In Fig. 2, the dashed lines represent the original transmission framework of KMV-Cast scheme, and the solid lines at the receiver are the framework of the proposed adaptive denoising algorithm. The proposed algorithm mainly contains three parts: 1) comparing the SNR values of the reconstructed pixel block with or without Wiener filtering; 2) selectively adding a Wiener filter to maximize the SNR of the pixel block; 3) optimizing parameters to maximize the whole reconstructed frame’s SNR.

A. WIENER FILTERING

The essence of Wiener filtering is to minimize the mean square error (MMSE) of the estimated signal. The process of Wiener filtering can be represented as Fig. 3.

![FIGURE 3. The Wiener filter model.](image)

As we can see from Fig. 3, the input of the filter is the sum of the original signal and the noise, and the ideal output is the original signal \( \hat{p} \). But the actual output \( \hat{H} \) cannot be the optimal. Based on MMSE criterion, the transfer function can be formulated as:

\[
H = R_{ss} (R_{ss} + R_{vv})^{-1}.
\]

(25)

In Eqn. (25), \( H \) is the matrix of the optimal transfer function, \( R_{ss} \) is the autocorrelation matrix of the original signal, and \( R_{vv} \) is the autocorrelation matrix of the noise. Then, the two autocorrelation matrices in Eqn. (25) can be calculated as:

\[
R_{ss} = E \left\{ p^2 \theta \theta^T \right\} = p^2 E \left\{ \theta \theta^T \right\} = p^2 \Omega
\]

(26)

\[
R_{vv} = E \left\{ \frac{\alpha_0^2 \sigma_v^4 r^2 (\Phi_v \Phi_v^T)^T}{(Cr + 1)^2} \right\} = \frac{Cr^2}{(Cr + 1)^2} I.
\]

(27)

Based on the expression of \( \Omega \) in Eqn. (11), the autocorrelation matrix can be represented as

\[
R_{ss} = p^2 r I + p^2 \hat{\theta} \hat{\theta}^T.
\]

(28)

So, the transfer function can be represented as:

\[
H = \frac{(Cr + 1)^2 p^2 I}{(Cr + 1)^2 p^2 + Cr} + \frac{Cr (Cr + 1)^2 p^2 \hat{\theta} \hat{\theta}^T}{[Cr + 1]^2 p^2 + Cr + Cr^2}.
\]

(29)

Considering the Eqn. (29) is independent of time, the output signal through the Wiener filter can be written as

\[
\hat{H} = H \left( p \theta + \frac{\alpha_0^2 r \Phi_v \Phi_v^T}{Cr + 1} \right).
\]

(30)

As a result, the noise of the signal processed again is changed into

\[
\text{noise} = \hat{H} \theta - p \theta = H \frac{\alpha_0^2 r \Phi_v \Phi_v^T}{Cr + 1} - (I - H) \theta.
\]

(31)

Similar to the KMV-Cast, we can calculate the noise power through

\[
P_N = E \left\{ \text{noise} \cdot \text{noise}^T \right\} = \frac{(Cr + 1)^2}{Cr^2} \frac{\alpha_0^2}{\sigma_v^2} r^2 E \left\{ \Phi_v \Phi_v^T \right\}.
\]

(32)

and the noise is the sum of the following three parts:

\[
P_{N1} = E \left\{ \theta^T (I - H) \theta \theta^T (I - H)^T \right\},
\]

(33)

\[
P_{N2} = E \left\{ \frac{\alpha_0^2}{\sigma_v^4} r^2 \Phi_v \Phi_v^T \theta \theta^T (I - H)^T \right\},
\]

(34)

\[
P_{N3} = E \left\{ \frac{\alpha_0^2}{\sigma_v^4} r^2 \Phi_v \Phi_v^T \theta^T (I - H) \right\}.
\]

(35)

We can use the notation in Eqn. (22) to rewrite the \( SNR_2 \) of the KMV-Cast scheme after Wiener filtering as:

\[
SNR_2 = \frac{PS}{P_N} = \left( \frac{P_{N1} + P_{N2} + P_{N3}}{p^2} \right)^{-1}
\]

(36)

From Eqn. (24) and Eqn. (30), two SNRs are determined by variances \( t, r, A \). From Eqn. (21) and Eqn. (22), one can see that the variance \( r \) and \( A \) are both the function of \( t \), and one solution of variance \( r \) can be calculated as

\[
r = \frac{- (C + 1) + \sqrt{(C + 1)^2 + 4Cr}}{2C}.
\]

(37)

Then the two SNRs in Eqn. (24) and Eqn. (36) can be represented by \( t \). Therefore, we can plot the two SNRs’ curve with respect to \( t \). Under different values of power scaling parameter \( C \) and the same value \( K \), we can get the results in Fig. 4.

![FIGURE 4. The signal-to-noise ratio curves of the two methods under the same \( K^2 = 0.9 \) and different values of power scaling parameter \( C \), where \( SNR_1 \) is the signal-to-noise ratio of the KMV-Cast scheme (blue) and \( SNR_2 \) is the signal-to-noise ratio of the scheme with a Wiener filter (red). (a) \( C = 100 \), (b) \( C = 300 \), (c) \( C = 500 \).](image)
B. ADAPTIVE USING WIENER FILTER

Sometimes, there are some correlated pixel blocks which have high similarity as the transmission pixel block. In this case, it is better to reconstruct the pixel block directly with the index of the corresponding pixel block received instead of DCT coefficients. Such case will not only reduce bandwidth consumption, but also improve the quality of the reconstructed video.

If we only use the relevant information and transmit the index of the similar pixel block \( \theta_i \), the SNR of the reconstructed video block is

$$ SNR_0 = \frac{1}{|\Delta|^2} = \frac{1}{2(1 - |K|)} . $$

(38)

In Eqn. (38), one can see that the SNR is only determined by the similarity coefficient \( K \).

Considering the use of the related information in cloud, we need to choose the best way among three possible cases to reconstruct the image, corresponding to the peak values of \( SNR_1 \) and \( SNR_2 \) as \( SNR_{1_{\text{max}}} \) and \( SNR_{2_{\text{max}}} \) respectively. Through comparison, we decide whether or not to transmit the coded pixel block, and whether or not to adopt a Wiener filter at the receiver. The detailed algorithm is in TABLE I.

| TABLE 1. Framework of transmission and filter |
|---------------------------------------------|
| Compared the three SNR values in Eqns. (24), Eqn. (36) and Eqn. (38), choose one with the highest SNR to reconstruct the pixel block. |
| if \( \text{max}_{i=0}^2 \{ SNR_{i_{\text{max}}} \} = SNR_{2_{\text{max}}} \) |
| transmit the DCT coefficient \( \theta \) and add a Wiener filter at the receiver. |
| else |
| if \( \text{max}_{i=0}^2 \{ SNR_{i_{\text{max}}} \} = SNR_{1_{\text{max}}} \) |
| transmit the DCT coefficient without adding a Wiener filter. |
| else |
| only transmit the index of the similar pixel block \( \theta_i \) |
| end |
| end |

C. OPTIMIZED PARAMETERS

| TABLE 2. Framework of maximizing SNR of pixel block |
|---------------------------------------------------|
| Optimize parameters \( A, r, t \) to achieve the maximal SNR of the pixel block. |
| if only transmit the index of the similar pixel block \( \theta_i \) |
| set \( A = r = t = 0 \). |
| else |
| if add a Wiener filter at the receiving side |
| based on Eqn. (36), solve the parameters \( A, r, t \), corresponding to the peak value. |
| else |
| based on Eqn. (24), solve the parameters \( A, r, t \), corresponding to the peak value. |
| end |
| end |

From the expressions of SNRs, one can see that there are four parameters should be calculated out, i.e., \( A, C, r, t \). Among them, \( A \) and \( r \) can be represented by \( t \). Based on maximizing the SNR of each pixel block, we can calculate parameters \( A, r, t \), while based on maximizing the SNR of the whole reconstructed video frame, we can calculate the parameter \( C \). The detailed algorithm is presented in TABLE II.

In order to maximize the SNR of the reconstructed video frame, we should minimize the total noise power of all transmitted pixel blocks with the given constrain of signal power \( P \) and the noise power is calculated in TABLE III.

| TABLE 3. Framework of calculate noise power |
|--------------------------------------------|
| Assume \( l_j \) is the noise power of the \( j \)th reconstructed block and it can be respectively represented with three different conditions. |
| if \( \text{max}_{i=0}^2 \{ SNR_{i_{\text{max}}} \} = SNR_{2_{\text{max}}} \) |
| \( l_j = p^2 / SNR_{2_{\text{max}}} \). |
| else |
| if \( \text{max}_{i=0}^2 \{ SNR_{i_{\text{max}}} \} = SNR_{1_{\text{max}}} \) |
| \( l_j = p^2 / SNR_{1_{\text{max}}} \). |
| else |
| \( l_j = 0, C_j = 0 \). |
| end |
| end |

Then, the minimum total noise power can be written as

$$ \min \sum_{j=1}^M \frac{\lambda^2_j l_j}{C_j} $$

(39)

with the power constraint condition:

$$ \sum_{j=1}^M C_j \leq \frac{P}{\sigma_0^2} $$

(40)

where \( \lambda^2_j l_j / C_j \) is the noise power of the \( j \)th reconstructed block. In order to minimalize the total noise of the reconstruct video, we need to allocate the power scaling \( C_j \) parameter by Lagrange multipliers as

$$ C_j = \frac{\sqrt{\lambda^2_j l_j}}{\sum_{k=1}^M \sqrt{\lambda^2_k l_k}}, j = 1, 2, ..., M. $$

(41)

IV. THE FRAMEWORK OF THE PROPOSED ALGORITHM

The framework of the proposed algorithm mainly includes the frameworks at the transmitter and the receiver, respectively. The detailed frameworks at both sides are introduced in TABLE IV and TABLE V.

V. EXPERIMENTAL RESULTS

The factors affecting the quality of the reconstructed video include: 1) the similarity of the correlated information; 2) the quality of the channel; 3) the characteristics of the video source. As a result, we analyze the performance of the proposed algorithm considering such three factors.

A. THE EFFECT OF CORRELATED INFORMATION

In this section, we evaluate the performance of the proposed adaptive denoising algorithm in terms of PSNR. Assume that the transmission channel is slow fading and its distortion can be canceled by the equalizer. With the standard video test sequences as the sources, we simulate the video transmission
Assume the correlated pixel blocks \{\theta_k | k = 1, 2, ..., N\}, the total transmission power \(P\), and Gaussian noise power \(\sigma_n^2\). Also assume that the video frame has been evenly divided into \(M\) blocks, and each block includes \(n\) pixel values.

### 1. Correlated information extraction

for \(j = 1:M\)
- From the correlated information set \(\{\bar{\theta}_k | k = 1, 2, ..., N\}\), choose two blocks with the maximal and minimal value of \(K^2 = (\bar{\theta}^T \theta)^2\).
- Compare the SNRs of the two blocks to pick that with higher SNR as the similar pixel block.

end

### 2. Whether to transmit the original DCT coefficients

for \(j = 1:M\)
- From Eqn. (41), calculate the power scaling parameter \(C_j\).
- Determine whether to transmit the original DCT coefficients by comparing the value of \(SNR_0\) in Eqn. (38) with the peak values of \(SNR_1\) in Eqn. (24) and \(SNR_2\) in Eqn. (36). If \(SNR_0 \geq SNR_{1max}\) and \(SNR_0 \geq SNR_{2max}\), just transmit the index of similar pixel block \(\bar{\theta}\) to the receiver and donate the noise power \(l_j = 0\).
- Else it is necessary to transmit the original DCT coefficients and denote the noise power \(l_j = p^2/SNR_{imax} (i = 0, 1, 2)\).

end

### 3. Whether to add a Wiener filter

for \(j = 1:M\)
- Determine whether to add a Wiener filter by comparing the peak value of \(SNR_1\) in Eqn. (24) and that of \(SNR_2\) in Eqn. (36). If \(SNR_{2max} > SNR_{1max}\), it is better to add a Wiener filter, and vice versa.
- Each block transmitted needs to add a one-bit-overhead which introduces whether to add a Wiener filter at the receiving side.

end

### 4. Power scaling

(1) The pixel blocks with \(l_j = 0\) will not be assign the power and set \(C_j = 0\).
(2) Process power scaling for all the pixel blocks and calculate the power scaling parameter \(C_j (j = 1, 2, ..., M)\) with Eqn. (41).

end

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/
can be seen that the performance gain increases with the increase of similarities between the transmitted signal and the correlated information in clouds.

B. QUALITY OF CHANNEL

Assume the transmitted signal is highly correlated with the information in clouds. With the standard video test sequence
CARPHONE”, we select 19th frame as the correlated information in clouds and 20th frame as the transmitted signal in Fig. 8. Let us change SNRs of the received signal and analyze the impact of the SNRs on the qualities of reconstructed video frames with different transmission schemes. Set SNRs as -10dB, -5dB, 0dB, 5dB, and the four schemes’ simulation results are shown in Fig. 9–Fig. 12. From Fig. 9 to Fig. 12, we can see that the proposed adaptive denoising algorithm adopted in the KMV-Cast video transmission scheme has the best performance. Compared to the KMV-Cast transmission scheme, the proposed algorithm achieves the PSNR gains of 0.7dB, 1.9dB, 1.4dB and 1.4dB under SNRs as -10dB, -5dB, 0dB and 5dB, respectively, while compared to the SoftCast transmission scheme, the proposed algorithm achieves the PSNR gains of 17dB, 14dB, 14dB and 13dB, respectively. At lower-SNR channel, KMV-Cast transmission scheme utilizes more correlated pixel blocks to reconstruct the video frame instead of transmission, and thus the improvement of the proposed algorithm is limited. As the quality of the channel improves, the number of transmitted blocks increases and the performance of the algorithm gradually appears.

C. SOURCES

Sometimes, different sources can achieve different performances. Assume the transmitted signal is highly correlated with the information in clouds. Choosing the standard video test sequences, “Carphone”, “Container”, “Bridge (close)” and “Hall Monitor” as sources, and changing the channel’s quality SNR as -10dB, -5dB, 0dB, 5dB, 10dB, the simulation results are shown at Fig. 13. One can see that, under different video sources and channel SNRs, the proposed algorithm has the best performance. Specifically, when the reconstructed frame of KMV-Cast scheme contains more Gaussian noise, the proposed algorithm achieves good performance, see Fig. 13(c). At the channel SNR 0dB, the simulation results of reconstructed 20th frame are shown in Fig. 14.

From the simulation results shown in this section, one can see that the proposed algorithm is an enhancement version of the conventional KMV-Cast transmission scheme.

VI. CONCLUSION

In this paper, the Wiener filter-based adaptive denoising algorithm has been proposed for pseudo analogy video transmission, i.e., KMV-Cast video transmission scheme. The residual noise existing in KMV-Cast is reduced to further improve the quality of the reconstructed video at the receiver. Specifically, we set the maximizing SNR as the optimization goal, and adaptively determine whether to add a Wiener filter or not at the receiver. The simulation results have shown that the proposed denoising algorithm performs the best, comparing with the other three typical schemes. Our future work is to consider the application of this algorithm to Ultra HD video transmission.
work will focus on denoising scheme with deep learning on pseudo analogy video transmission.

APPENDIX

A. Eqn. (14) Derivation

Set normalized vector \( \| \hat{\theta} \|^2 = 1 \), we can get

\[
\| \hat{\theta} \|^2 = \hat{p} \left( \frac{C_T + 1}{C_T} \right)^2 \frac{I - \frac{\hat{\theta} \hat{\theta}^T}{r(C_T + C + 1)}}{r(C_T + C + 1) + 1} - 1 = 0 \tag{42}
\]

and then, we can get the expression of \( p^2 \) (Eqn. (14)).

B. Eqn. (36) Derivation

The noise power in Eqn. (32) is the sum of the Eqn. (33), (34), (35), and the three parts of the noise can be calculated as:

\[
P_N = E \left\{ \frac{(C_T + 1)^2 p^4}{(C_T + 1)^2 p^2 + C_T} \left[ \frac{[-2(C_T + 1)C_Tp^2 + 2(C_T + 1)p^2 - 2(C_T + 1)p^2(C_T + C + 1)p^2]}{(C_T + 1)^2 p^2 + C_T} \right] \right\}
= \hat{p} \left\{ \frac{1}{r(C_T + 1)^2 p^2 + C_T} \right\}
= \hat{p} \left\{ \frac{1}{r(C_T + 1)^2 p^2 + C_T} \right\}
\]

\[
= \hat{p} \left\{ \frac{1}{r(C_T + 1)^2 p^2 + C_T} \right\}
\]

\[
P_{N2} = E \left\{ \frac{(C_T + 1)^2 p^4}{(C_T + 1)^2 p^2 + C_T} \left( \frac{\Phi_T^2 \Phi_T \phi_T^2}{\sigma^2} \right) \right\}
= E \left\{ \frac{(C_T + 1)^2 p^4}{(C_T + 1)^2 p^2 + C_T} \left( \frac{\Phi_T^2 \Phi_T \phi_T^2}{\sigma^2} \right) \right\}
\]

\[
P_{N3} = -2E \left\{ \frac{(C_T + 1)^2 p^4}{(C_T + 1)^2 p^2 + C_T} \left( \frac{\Phi_T^2 \Phi_T \phi_T^2}{\sigma^2} \right) \right\}
= -2E \left\{ \frac{(C_T + 1)^2 p^4}{(C_T + 1)^2 p^2 + C_T} \left( \frac{\Phi_T^2 \Phi_T \phi_T^2}{\sigma^2} \right) \right\}
\]

Sum Eqn. (33), Eqn. (44), and Eqn. (45), we can get the noise power \( P_N \) in Eqn. (32) to solve \( SNR_2 \).

ACKNOWLEDGMENT

 Portions of this work were presented at the EAI CROWNCOM (16th EAI International Conference on Cognitive Ra-
Diego Oriented Wireless Networks) in 2021, Research on de-noising method in pseudo-analog video transmission.

REFERENCES

[1] Cisco, “Cisco Annual Internet Report (2018–2023) White Paper,” 2020.
[2] M. Gastpar, B. Rimoldi and M. Vetterli, “To code, or not to code: lossy source-channel communication revisited,” IEEE Transactions on Information Theory, vol. 49, no. 5, pp. 1147–1158, May 2003, 10.1109/TIT.2003.810631.
[3] Szymon Jakuczuk, Dina Katabi, “A Cross-Layer Design for Scalable Mobile Video,” presented at ACM MobiCom ’11, 289–300, Sep. 2011, 10.1145/2030613.2030646.
[4] Z. Song, R. Xiong, X. Fan, S. Ma and W. Gao, “Transform domain energy modeling of natural images for wireless SoftCast optimization,” 2014 IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1114–1117, July 2014, 10.1109/ISCAS.2014.6865335.
[5] Z. Song, R. Xiong, S. Ma, X. Fan and W. Gao, “Layered image/video softcast with hybrid digital-analog transmission for robust wireless visual communication,” 2014 IEEE International Conference on Multimedia and Expo (ICME), pp. 1–6, July 2014, 10.1109/ICME.2014.6890331.
[6] F. Gao, H. Gao and J. Wu, “A Reconfigurable SoC for SoftCast Wireless Video Transmission,” 2018 IEEE International Conference on Industrial Internet (ICI), pp. 169–170, Oct. 2018, 10.1109/ICII.2018.00029.
[7] L. Luo, T. Yang, C. Zhu, Z. Jin and S. Tang, “Joint Texture/Depth Power Allocation for 3-D Video SoftCast,” IEEE Transactions on Multimedia, vol. 21, no. 12, pp. 2973–2984, Dec. 2019, 10.1109/TMM.2019.2919474.
[8] A. Ksentini, T. Taleb and F. Messaoudi, “A LISP-Based Implementation of Follow Me Cloud,” IEEE Access, vol. 2, pp. 1340–1347, Sep. 2014, 10.1109/ACCESS.2014.2363032.
[9] X.-L. Huang, J. Wu and F. Hu, “Knowledge-Enhanced Mobile Video Broadcasting Framework With Cloud Support,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 27, no. 1, pp. 6–18, Jan. 2017, 10.1109/TCASVT.2016.2555558.
[10] C. Zhao, J. Liu and J. Zhang, “A Dual Model for Restoring Images Corrupted by Mixture of Additive and Multiplicative Noise,” IEEE Access, vol. 9, pp. 168869–168888, Dec. 2021, 10.1109/ACCESS.2021.3137995.
[11] S. F.-W. Wang, W. -K. Yu and Y. -X. Li, “Multi-Wavelet Residual Dense Convolutional Neural Network for Image Denoising,” IEEE Access, vol. 8, pp. 214413–214424, Nov. 2020, 10.1109/ACCESS.2020.3040542.
[12] H. Aetesam, S. K. Maji and H. Yahia, “Bayesian Approach in a Learning-Based Hyperspectral Image Denoising Framework,” IEEE Access, vol. 9, pp. 169335–169347, Dec. 2021, 10.1109/ACCESS.2021.3137666.
[13] R. S. Thakur, S. Chatterjee, R. N. Yadav and L. Gupta, “Image De-Noising With Machine Learning: A Review,” IEEE Access, vol. 9, pp. 93338–93363, June 2021, 10.1109/ACCESS.2021.3092425.
[14] M. Ghogho, M. Ibnkahla and N. J. Bershad, “Analytic behavior of the LMS adaptive line enhancer for sinusoids corrupted by multiplicative and additive noise,” IEEE Transactions on Signal Processing, vol. 46, no. 9, pp. 2386–2393, Sep. 1998, 10.1109/78.709525.
[15] S. S. Bansal, S. U. Bhandari and P. S. Thorbole, “Implementation of LMS algorithm for image noise cancellation,” 2019 5th International Conference On Computing, Communication, Control And Automation (IC-CUBEA), pp. 1–4, Sep. 2019, 10.1109/ICCUBEA47591.2019.9129121.
[16] S. Ahmad, S. G. Khawaja, N. Amjad and M. Usman, “A Novel Multiplier-Less LMS Adaptive Filter Design Based on Offset Binary Coded Distributed Arithmetic,” IEEE Access, vol.9, pp. 78138–78152, May 2021, 10.1109/ACCESS.2021.3083282.
[17] G. A. Tammana and Y. F. Zheng, “Image Denoising using Multi-Resolution Coefficient Support Based Empirical Wiener Filtering,” 2006 International Conference on Image Processing, pp. 2613–2616, Oct. 2006, 10.1109/ICIP.2006.313022.
[18] Norifumi Kamiya, Eisaku Sasaki, “Pilot-Symbol-Assisted Phase Noise Compensation With Forward/Backward Wiener Smoothing Filters,” IEEE Transactions on Signal Processing, vol.65, no.17, pp. 4443–4453, Sep. 2017, 10.1109/TSP.2017.2709272.
[19] Z. Mbarki, H. Saddik and E. Ben Braiek, “Non blind image restoration scheme combining parametric wiener filtering and BM3D denoising technique,” 2018 4th International Conference on Advanced Technologies for Signal and Image Processing (ATSIPI), pp. 1–5, Mar. 2018, 10.1109/ATSIPI.2018.8364524.
[20] Y. Lai, S. Wang, C. Chen and S. Jhang, “Adaptive Wiener Gain to Improve Sound Quality on Nonnegative Matrix Factorization-Based Noise Reduction System,” IEEE Access, vol.7, pp. 43286–43297, Mar. 2019, 10.1109/ACCESS.2019.2907175.

WANNING HE received the double B.E. degrees in 2021, respectively in Electronic and Information Engineering, Tongji University, Shanghai, China and in Computer Engineering, Polytechnic Institute of Turin, Turin, Italy. She is currently working toward Ph.D. degrees in Tongji University. Her research focuses on multimedia transmission and signal processing.

XIN-LIN HANG (Senior Member, IEEE) received the M.E. and Ph.D. degrees in information and communication engineering from the Harbin Institute of Technology (HIT) in 2008 and 2011, respectively. He is currently a Professor and the Vice-Dean of the School of Electronic and Information Engineering, Tongji University, Shanghai, China. His research focuses on cognitive radio networks, multimedia transmission, and machine learning. He published 80 research papers and 19 patents in these fields. He was a recipient of Scholarship Award for Excellent Doctoral Student granted by Ministry of Education of China in 2010, the Best Ph.D. Dissertation Award from HIT in 2013, the Shanghai High-Level Overseas Talent Program in 2013, and the Shanghai Rising-Star Program for Distinguished Young Scientists in 2019. From August 2010 to September 2011, he was supported by China Scholarship Council to do research with the Department of Electrical and Computer Engineering, University of Alabama, USA, as a visiting scholar. He was invited to serve as the Session Chair for the IEEE ICC. He served as the Guest Editor for IEEE Wireless Communications and the Chief Guest Editor for International Journal of Mobile Networks and Applications and International Journal of Wireless Communications and Mobile Computing. He serves as IG co-chair for IEEE ComSoc MMTC, and an Associate Editor for IEEE Access. He is a Fellow of the EAI.

PENGFEI LI received the M.E. degree in electronics and communication engineering in 2020 from Shanghai University, Shanghai, China. He is currently working toward Ph.D. degrees in Tongji University. His research focuses on deep reinforcement learning and wireless resource management.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2022.3175511, IEEE Access