SCIENCE CHINA
Information Sciences

• LETTER •

January 2026, Vol. 69, Iss. 1, 119301:1–119301:2
https://doi.org/10.1007/s11432-025-4622-x

A hybrid bit and semantic communication system

Kaiwen YU[1], Renhe FAN[1], Gang WU[1*] & Zhijin QIN[2,3]

[1]National Key Laboratory of Wireless Communications, University of Electronic Science and Technology of China, Chengdu 611731, China
[2]Department of Electronic Engineering, Tsinghua University, Beijing 100084, China
[3]State Key Laboratory of Space Network and Communications, Beijing National Research Center for Information Science and Technology, Beijing 100084, China

Received 10 June 2025/Revised 24 September 2025/Accepted 29 September 2025/Published online 4 January 2026

Citation Yu K W, Fan R H, Wu G, et al. A hybrid bit and semantic communication system. Sci China Inf Sci, 2026, 69(1): 119301, https://doi.org/10.1007/s11432-025-4622-x

Conventional bit transmission sends all source content as bit sequences, while semantic communication selectively transmits only task-relevant information, reducing data volume and spectrum usage. Therefore, semantic communications are regarded as a potential method to improve transmission efficiency [1,2]. Some research has focused on evolutionary schemes from bit communications to semantic communications. Semantic communications are used to transmit large data content, such as augmented reality (AR), while conventional bit communications are used to transmit other small data content [3]. However, current joint source-channel coding (JSCC) methods map source content from the APP layer directly to symbols in the PHY layer for wireless transmission, which poses significant integration challenges in standardized layered frameworks and hinders their practical deployment.

To enhance transmission efficiency while ensuring compatibility with existing communication infrastructures, this study primarily investigates a hybrid bit and semantic communication system, named HybridBSC, tailored for image transmission. Based on semantic extraction and space transformation, the small-sized semantic features are inserted into another source image for transmission via conventional digital communication systems utilizing the same spectrum resources. Furthermore, we present a Pluto-based software defined radio (SDR) platform as the proof-of-concept, which improves transmission efficiency and system compatibility.

Proposed structure. As shown in Figure 1(a), the transmitter consists of three parts: a semantic encoder, a conventional bit encoder, and a channel encoder. We represent the source images for semantic-level coding as $\mathbf{M}_\mathrm{s} \in \mathbb{R}$ and the source images for bit-level coding as $\mathbf{M}_\mathrm{b} \in \mathbb{R}$. The semantic feature is extracted as $\mathbf{S} = T(\mathbf{M}_\mathrm{s};\boldsymbol{\eta})$, where $\mathbf{S} \in \mathbb{R}^{V_1\times V_2}$ denotes the resized semantic feature, V_1 and V_2 denote the length of the output, and $T(-;\boldsymbol{\eta})$ denotes the semantic encoder with learnable parameters $\boldsymbol{\eta}$. The semantic encoder is trained to convert the source data $\mathbf{M}_\mathrm{s}$ into semantic features, which are then ready for insertion into the source image for bit-level coding, $\mathbf{M}_\mathrm{b}$, for conventional digital communications. The insertion process can be represented as $\mathbf{I} = \mathcal{I}(\mathbf{S},\mathbf{M}_\mathrm{b}) \in \mathbb{R}^{3\times M\times N}$, where $\mathcal{I}$ denotes the insertion processes. To ensure robust and stable semantic insertion, we employ a color space transform to convert the source image $\mathbf{M}_\mathrm{b}$ from RGB to YCbCr space to get the original image luminance $\mathbf{Y}$. The source image for bit-level coding, $\mathbf{M}_\mathrm{b}$, is decomposed into four frequency bands via 2D discrete wavelet transform (DWT): low-frequency (LF), horizontal, vertical, and diagonal high-frequency parts. We then divide the LF band into 4×4 blocks, apply the discrete cosine transform (DCT) to each block, and perform singular value decomposition (SVD) on the transformed output. The whole process can be expressed as $[\mathbf{U}_{i,j},\mathbf{Z}_{i,j},\mathbf{V}_{i,j}] = \mathrm{SVD}\left(\mathrm{FC}_{4\times4}\left(\mathrm{FW}(\mathbf{Y})_\mathrm{LF}\right)\right)$, where FW() denotes the 2D-DWT which embeds information in low-frequency components to reduce noise, FC() denotes the DCT which compresses the image and concentrates energy to enhance insertion quality and transmission robustness, the singular value matrix $\mathbf{Z}_{i,j} \in \mathbb{R}^{4\times4}$ is a diagonal matrix, $\mathbf{U}_{i,j} \in \mathbb{R}^{4\times4}$ and $\mathbf{V}_{i,j} \in \mathbb{R}^{4\times4}$ are orthogonal unitary matrices, and $i,j \in M/4$. Then, we perform following operations on $\mathbf{Z}$ to obtain the new $\tilde{\mathbf{Z}}$:

$$\tilde{\mathbf{Z}}_{i,j} = \begin{cases} \alpha\eta, & \text{if } \mathbf{S}^q_{i,j} = 0, \eta \bmod 2 = 1, \\ \alpha(\eta+1), & \text{if } \mathbf{S}^q_{i,j} = 0, \eta \bmod 2 = 0, \\ \alpha(\eta+1), & \text{if } \mathbf{S}^q_{i,j} = 1, \eta \bmod 2 = 1, \\ \alpha\eta, & \text{if } \mathbf{S}^q_{i,j} = 1, \eta \bmod 2 = 0, \end{cases} \tag{1}$$

where α is the insertion factor, $\eta = \lceil \|\mathbf{Z}_{i,j}\|_{2'}/\alpha \rceil$. $\tilde{\mathbf{Z}}_{i,j}$ denotes the (i,j)-th element of singular value matrix $\tilde{\mathbf{Z}}$. $\|\boldsymbol{\Gamma}\|_{2'} = \sup_{x\in R^n\setminus 0}(|\boldsymbol{\Gamma}x|/|x|)$ denotes the largest singular value of $\boldsymbol{\Gamma}$. $\mathbf{S}^q$ is the q-bit quantization output of semantic feature $\mathbf{S}$. We use $q=4$ bit quantization. Subsequently, we execute inverse SVD, DCT, and DWT to complete the semantic insertion process: $\tilde{\mathbf{Y}} = \mathrm{IFW}\left(\mathrm{IFC}_{4\times4}\left(\mathbf{U}\tilde{\mathbf{Z}}\mathbf{V}^\mathrm{T}\right)\right)$, where IFW() denotes the inverse DWT, $\mathrm{IFC}_{4\times4}()$ denotes the inverse DCT on each 4×4 block. Then, we transform $\tilde{\mathbf{Y}}$ back into RGB space to get the hybrid data $\mathbf{I}$. After conventional source coding, channel coding, and modulation, the information will be transmitted over the wireless channels.

The receiver also consists of three parts: a semantic decoder, a conventional bit decoder, and a channel decoder. The received

* Corresponding author (email: wugang99@uestc.edu.cn)

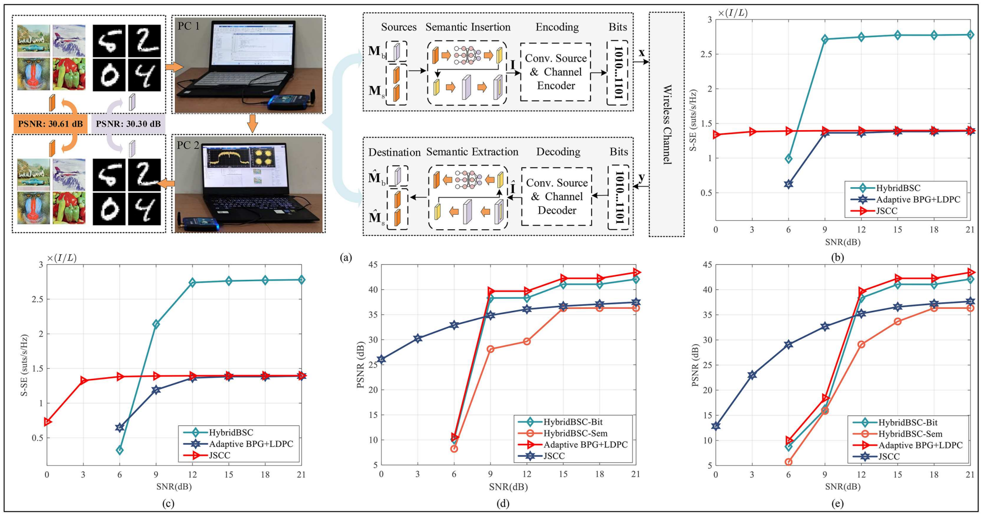

Figure 1 Schematic and measurement results. (a) The structure of the proposed HybridBSC; (b) S-SE versus SNR over AWGN channel; (c) S-SE versus SNR over fading channel; (d) PSNR versus SNR over AWGN channel; (e) PSNR versus SNR over fading channel.

signal can be expressed as $\mathbf{y} = \mathbf{h}\mathbf{x} + \mathbf{n}$, where $\mathbf{x}$ is the complex symbols. $\mathbf{n} \in \mathcal{N}\left(0, \sigma^2\right)$ is the additive circular complex Gaussian noise with zero mean and variance σ^2. Based on the estimated channel, the signal $\hat{\mathbf{x}}$ can be recovered. After demodulation, conventional channel decoding, and source decoding, the hybrid information $\hat{\mathbf{I}}$ is derived from $\hat{\mathbf{x}}$, and can be used to recover $\hat{\mathbf{M}}_{\mathrm{b}}$. Similarly, the recovered hybrid data $\hat{\mathbf{I}}$ is color space transformed to obtain $\hat{\mathbf{Y}}$, which is then subjected to a 2D-DWT. The resulting LF part is partitioned into 4×4 blocks, and a DCT is applied to each block, followed by SVD. The whole process can be expressed as $[\hat{\mathbf{U}}_{i,j}, \hat{\mathbf{Z}}_{i,j}, \hat{\mathbf{V}}_{i,j}] = \mathrm{SVD}\left(\mathrm{FC}_{4\times 4}\left(\mathrm{FW}\left(\hat{\mathbf{Y}}\right)_{\mathrm{LF}}\right)\right)$, where $\hat{\mathbf{Z}}_{i,j}$ is the recovered singular value matrix, $\hat{\mathbf{U}}_{i,j}$ and $\hat{\mathbf{V}}_{i,j}$ are the recovered orthogonal unitary matrices, and $i, j \in M/4$. The feature de-insertion process is conducted according to

$$\hat{\mathbf{S}}^q_{i,j} = \begin{cases} 1, & \text{if } \hat{\eta} \bmod 2 = 0, \\ 0, & \text{if } \hat{\eta} \bmod 2 = 1, \end{cases} \tag{2}$$

where $\hat{\eta} = \left\lceil \left\| \hat{\mathbf{Z}}_{i,j} \right\|_{2'} / \alpha \right\rceil$, $\hat{\mathbf{S}}^q_{i,j}$ is the (i,j)-th element of quantized semantic feature $\hat{\mathbf{S}}^q$. The semantic feature $\hat{\mathbf{S}}$ is finally obtained by inverse quantization of $\hat{\mathbf{S}}^q$. It is worth noting that the insertion factor α is known to the transceiver, so the receiver can accurately extract semantics from the hybrid data $\hat{\mathbf{I}}$. Ultimately, the semantic decoder reconstructs the content $\hat{\mathbf{M}}_{\mathrm{s}} = R(\hat{\mathbf{S}}; \boldsymbol{\gamma})$, where $R(-; \boldsymbol{\gamma})$ denotes the semantic decoder with learnable parameters $\boldsymbol{\gamma}$.

Experiments. We evaluate the proposed HybridBSC in terms of transmission accuracy and semantic spectral efficiency (S-SE) [4] over AWGN and Rayleigh channels, respectively. For performance comparison, we use JSCC and conventional adaptive coding and modulation as benchmarks. Better portable graphics (BPG) serves as the conventional source coding, and low-density parity-check (LDPC) as the channel coding, referred to as "Adaptive BPG+LDPC". "HybridBSC-Bit" and "HybridBSC-Sem" denote the performance of bit-coded and semantic-coded images, respectively. Figures 1(b) and (c) show the relationship between the semantic spectral efficiency and the SNR over AWGN and Rayleigh channels, respectively. It is clear that all benchmarks exhibit lower semantic spectral efficiency in comparison to HybridBSC. The reason is that HybridBSC is able to transmit two source images accurately and does not require additional time-frequency resources. Figures 1(d) and (e) show the relationship between the peak signal-to-noise ratio (PSNR) score and the SNR over AWGN and Rayleigh channels, respectively. The HybridBSC-Bit is very close to that of the conventional adaptive coding and modulation strategy. At the same time, HybridBSC-Sem is also close to that of JSCC at the high SNR regimes. This proves that the proposed HybridBSC can resist the distortion caused by the coding and noise. A wireless HybridBSC prototype based on IEEE 802.11a OFDM PHY via the Analog Device ADALM-PLUTO SDR module (PlutoSDR) is implemented for experiments over a real wireless channel. Figure 1(a) shows the experience scenario of the PlutoSDR-based HybridBSC prototype. Specifically, the source images are encoded on the PC 1, transmitted wirelessly via PlutoSDR to the PC 2, and then decoded to generate the output images. In the experiment, the average PSNR scores for bit-coded images are 30.61 dB, while for semantic-coded images, the average PSNR is 31.30 dB. These results demonstrate that the proposed HybridBSC effectively mitigates damage from the source coding and real wireless channels.

Conclusion. This work proposed a hybrid bit and semantic communication system, named HybridBSC, which jointly performs the bit and semantic coding for image transmission. With the HybridBSC, the semantic features can be attached to another source image for bit-level coding and transmitted using a conventional digital transceiver. The simulation and experimental results have demonstrated that the HybridBSC can accurately transmit bits and semantics simultaneously in the same time-frequency resources.

Acknowledgements This work was supported by National Key R&D Program of China (Grant No. 2020YFB1806604), and National Natural Science Foundation of China (Grant No. 62293484).

References

1 Qin Z, Liang L, Wang Z, et al. AI empowered wireless communications: from bits to semantics. Proc IEEE, 2024, 112: 621–652
2 3rd Generation Partnership Project (3GPP). Study on 6G use cases and service requirements. Technical Specification (TS) 22.870, 06 2025, version 0.3.1 https://portal.3gpp.org/desktopmodules/Specifications/ SpecificationDetails.aspx?specificationId=4374
3 Wang C, Li Y, Gao F, et al. Adaptive semantic-bit communication for extended reality interactions. IEEE J Sel Top Signal Process, 2023, 17: 1080–1092
4 Yan L, Qin Z, Zhang R, et al. Resource allocation for text semantic communications. IEEE Wireless Commun Lett, 2022, 11: 1394–1398