Task-Oriented and Semantics-Aware Communications for 6G

Dr Yansha Deng

King's College London

IEEE ICC
2023

May 28th

XR-aided Teleoperation

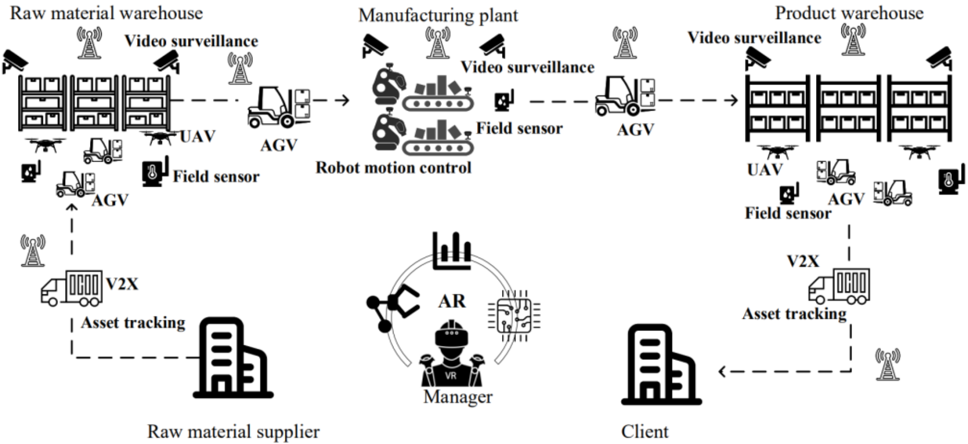
Cellular-Connected XR Networks

Challenges: High transmission bit rate: 35 Mbps – 4.42 Gbps; Low latency: 5 – 10 ms

[1] F. Hu, **Y. Deng***, W. Saad, M. Bennis, A. H. Hamid, "Cellular-Connected Wireless Virtual Reality: Requirements, Challenges, and Solutions", in IEEE Communications Magazine, 2020.

[2] Qualcomm, "VR and AR pushing connectivity limits," Qualcomm Technologies. Inc., Tech. Rep., 2018 (Accessed on 2019-12-19). [Online]. Available: https://www.qualcomm.com/invention/extended-reality/virtual-reality

Cellular-Connected Robotics Networks

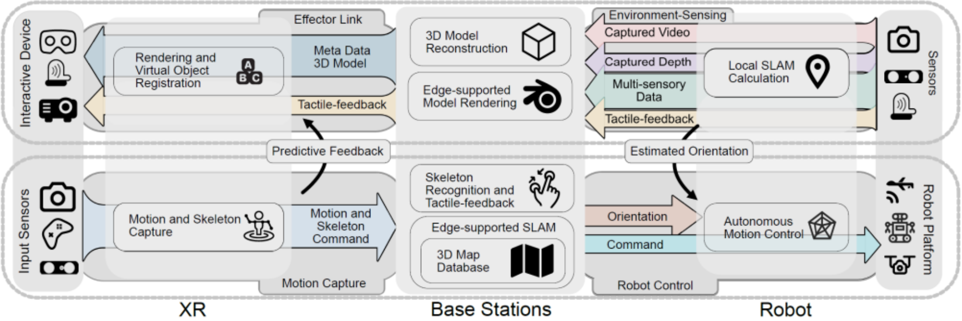


[3] H. Zhou, S. Yang, **Y. Deng***, M. Dohler, A. Nallanathan. "Machine Learning for Massive Industrial Internet of Things" in IEEE Wireless Communications, 2021.

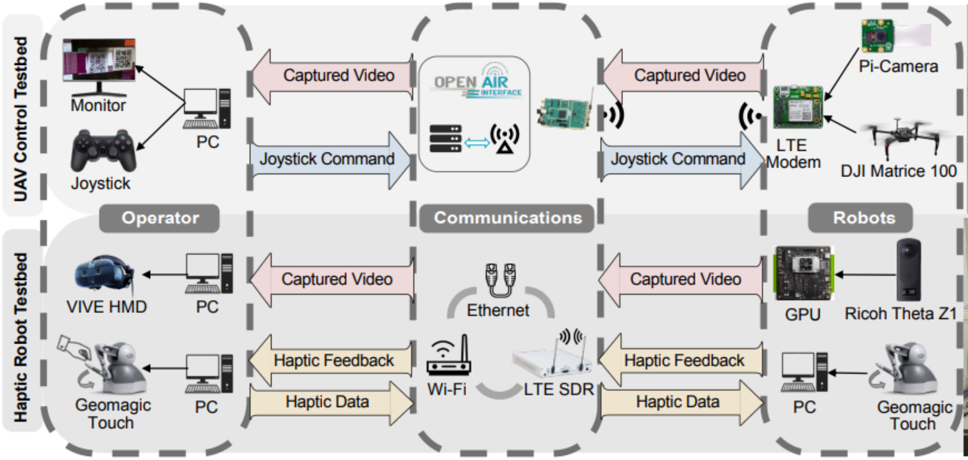
I: Testbeds and Trials

[3] H. Zhou, F. Hu, M. Juras, A. B. Mehta and **Y. Deng***, "Real-time Video Streaming and Control of Cellular-Connected UAV System: Prototype and Performance Evaluation," in IEEE Wireless Communications Letters, 2021.
[4] F. Hu, **Y. Deng***, H. Zhou, T. H. Jung, C. B. Chae, A. H. Hamid, "A Vision of XR-aided Teleoperation System Towards 5G/B5G", in IEEE Communications Magazine, 2021.

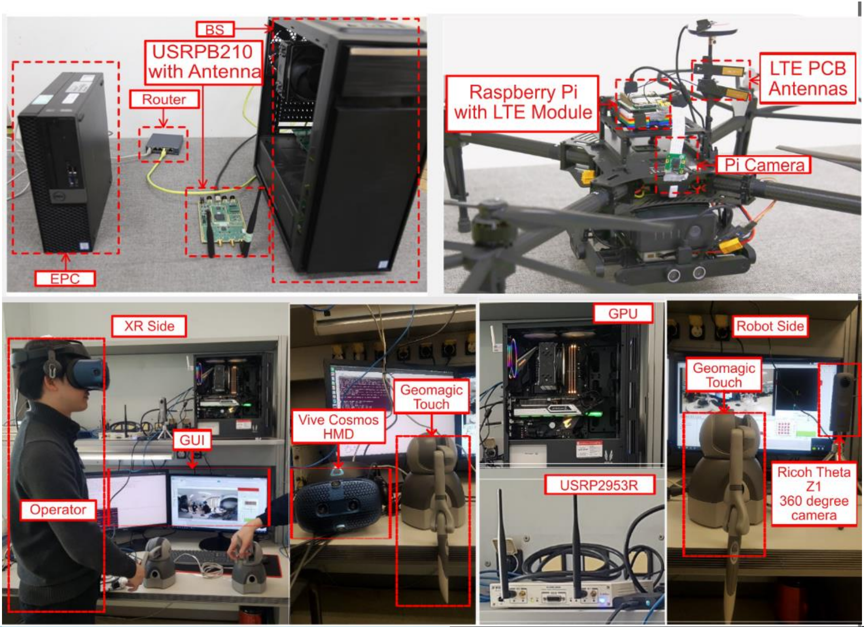
XR-aided Teleoperation



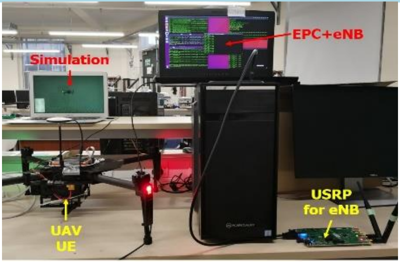
Testbeds and Trials: XR-aided Teleoperation

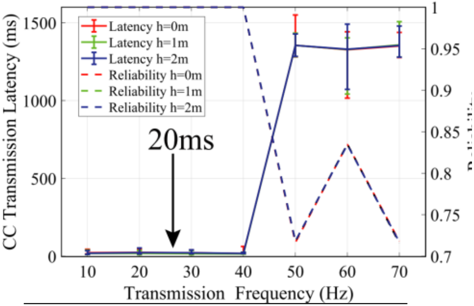


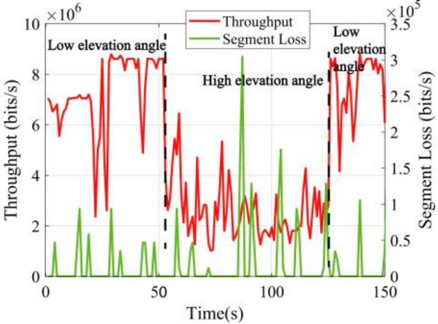
Testbeds and Trials: XR-aided Teleoperation



Testbeds and Trials: SDR-based UAV Network





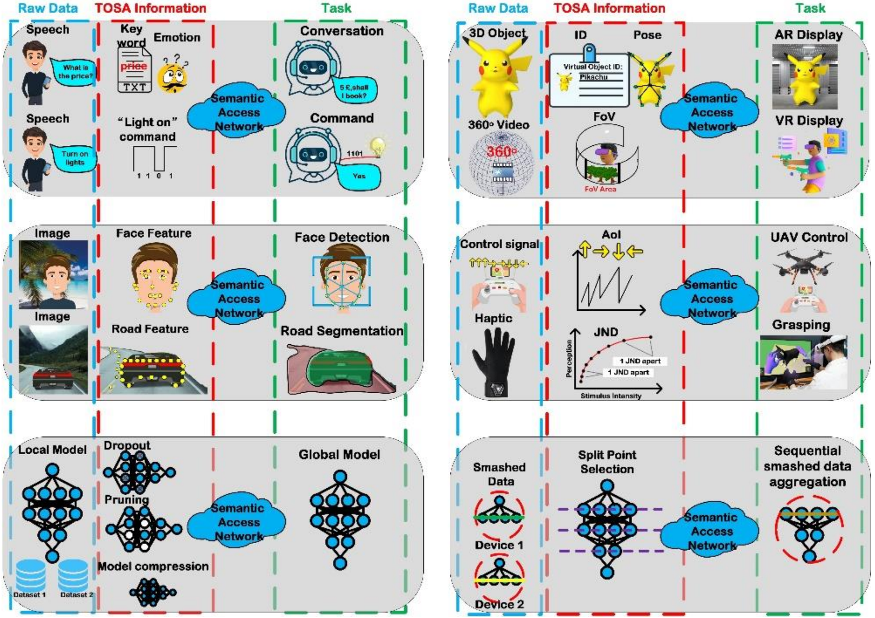


[3] H. Zhou, F. Hu, M. Juras, A. B. Mehta and **Y. Deng***, "Real-time Video Streaming and Control of Cellular-Connected UAV System: Prototype and Performance Evaluation," in IEEE Wireless Communications Letters, 2021.

Task-Oriented Semantics-Aware Communication

[5] H. Zhou, X. Liu, **Y. Deng***, N. Pappas, and A. Nallanathan, “Task-Oriented and Semantics-Aware 6G Networks,” arXiv preprint arXiv:2210.09372, 2022 .

Task-oriented Semantics-aware Coms Architecture



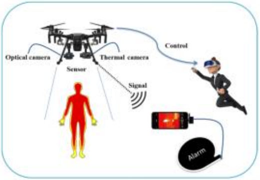
II: Task-Oriented Semantics-Aware Communication for Wireless UAV Control and Command Transmission

[6] Y. Xu, H. Zhou, **Y. Deng*** "Task-Oriented and Semantics-Aware Communication for Wireless UAV Control and Command Transmission," IEEE Communications Letter, 2023

Introduction

UAV applications

Food delivery



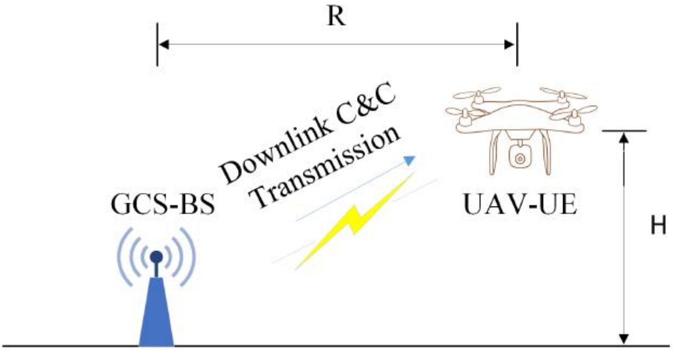
Epidemic monitoring and control

Environmental exploration

Agriculture

Challenge: **Dramatically increasing C&C data** under high transmission frequency **brings a heavy burden** on the existing bit-oriented cellular network design.

System model



- BS serves as the ground control station(GCS) and **periodically generates the C&C signal.**
- GCS will **receive one acknowledge character(ACK)** if **the C&C signal is received successfully**.
- UAV user equipment (UAV-UE) **flies in a circular horizontal disk** with radius R and height H.

System model

- The path loss from BS to UAV

$$h = (P_{\mathrm{LoS}}\eta_{\mathrm{LoS}} + P_{\mathrm{NLoS}}\eta_{\mathrm{NLoS}})\left(\frac{4\pi d f_c}{c}\right)^{\alpha}\beta \qquad \textbf{(1)}$$

where P_{LoS} and P_{NLoS} are the LoS and NLoS probility, η is the path loss coefficient, α is the pass loss exponent, β is the small $-$ scale Rayleigh fading.

- The SNR of the received signal at UAV

$$\mathrm{SNR} = \frac{\mathrm{P}h}{\sigma^2} \qquad \textbf{(2)}$$

where P represents the transmit power of BS, and σ^2 is Additive White Gaussain Noise (AWGN) power.

Problem Formulation

- Successfully decode the C&C message or not

$$\delta = \begin{cases} 1, & \text{SINR} > \gamma \\ 0, & \text{else} \end{cases} \quad \textbf{(3)}$$

We aim to optimize the TOSA information of the downlink C&C transmission based on the similarity and AoI of C&C data

$$\text{(P1):} \quad \max_{\{\pi(A^t|O^t)\}} \sum_{k=t}^{\infty} \gamma^{k-t}\, \mathbb{E}_\pi[R^t], \quad \textbf{(4)}$$

In (4), A^t decides whether to transmit or drop the C&C signal at GCS at the t^{th} TTI, U^t is all prior historical observations before the beginning of every transmission time interval (TTI) t, O^t is observed history and can be defined as $O^t = \{A^{t-1}, U^{t-1}, \dots, A^1, U^1\}$, R^t is the reward, and $\gamma \in [0,1)$ is the discount rate for the performance in future TTIs.

Reward Design

TOSA Information: we define the reward R_{t+1} as a function related to the observed similarity between adjacent C&C signals L_t and L_{t-1} and the AoI I_t and I_{t-1} , which is defined as

$$R^{t+1} = \begin{cases} f(L^t)g(I^t) & \text{Successful transmission} \\ 0 & \text{Faliure transmission,} \end{cases}$$

where the similarity quantifies the importance of C&C related to the task, and AoI quantifies the freshness of the C&C signal.

Value of Information: to quantify the difference between consecutive C&C signals, we define the **similarity** as

$$L^t = \sum_i \frac{\mu_i(M_i^t - M_i^{t-1})}{R_i}$$

TABLE I

INFORMATION OF M_i^t AND R_i

i	Length	M_i^t Range	R_i
ROW	4 bytes	$-35°$ to $35°$	70
PITCH	4 bytes	$-35°$ to $35°$	70
YAW	4 bytes	$-150°/s$ to $150°/s$	300
THRUST	4 bytes	$-5m/s$ to $5m/s$	10

Reward Design

To normalize the similarity value, we utilize the sigmoid function as

$$f(L^t) = \frac{2}{1 + e^{-\theta(L^t - \zeta)}} - 1$$

where κ controls the gain, and ζ controls the cutoff.

Age of Information (AoI) is the **metric** to characterize **the freshness of the information**, it can be derived as:

$$I^t = t_{rcv} - t_{gen} = \frac{\mathrm{N_{cc}}}{B\log(\mathrm{SNR} + 1)},$$

where $\mathrm{N_{cc}}$ is the size of C&C signal, and B is the bandwidth. In I^t, t_{rcv} is the time that the packet is received by the UAV-UE and confirmed by ACK, and t_{gen} is the generated time of C&C signal at the GCS.
We then normalize I^t as

$$g(I^t) = 1 - \frac{I^t}{\Delta T}$$

where ΔT is the duration of each TTI.

DRL-based TOSA Com for C&C signal

We adopt DQN algorithm.

- State $$S^t = [L^t, I^t, A^t]$$
- Action $$A^t = 0 \text{ or } 1$$

Algorithm 1 DRL-based TOSA Communication Framework for C&C transmission

Input: The set of available action $\mathcal{F}$.

Algorithm hyperparameters: learning rate $\lambda_{\text{RMS}} \in (0,1]$, discount rate $\gamma \in (0,1]$, ϵ-greedy rate $\epsilon \in (0,1]$, target network update frequency K

Initialization of replay memory M to capacity C, the primary Q-network $\boldsymbol{\theta}$, and the target Q-network $\overline{\boldsymbol{\theta}}$

for $t = 1, ..., T$ **do**

    Update the traffic

    if $p_\epsilon < \epsilon$ **then**

        select a random action A^t from $\mathcal{A}$

    else

        select $A^t = \text{argmax} Q(S^t, a, \boldsymbol{\theta})$

    end if

    The BS executes the decision, transmits tth C&C signal or drops tth C&C signal

    The central server observes S^{t+1}, and calculate the related R^{t+1} using Eq. 6

    Store transition $(S^t, A^t, R^{t+1}, S^{t+1})$ in replay memory M

    Sample random minibatch of transitions $(S^t, A^t, R^{t+1}, S^{t+1})$ from replay memory M

    Perform a gradient descent for $Q(s, a, \boldsymbol{\theta})$ using (13)

    Every K steps update target Q-network $\overline{\boldsymbol{\theta}} = \boldsymbol{\theta}$.

end for

Simulation Environment and Results

parameters	values
f_c	5GHz
γ_{th}	5.5dB
σ^2	-104dBm
P	18dBm
ϵ	1
batch size	32
λ_{RMS}	10^{-5}
γ	0.1

Table : Simulation parameters

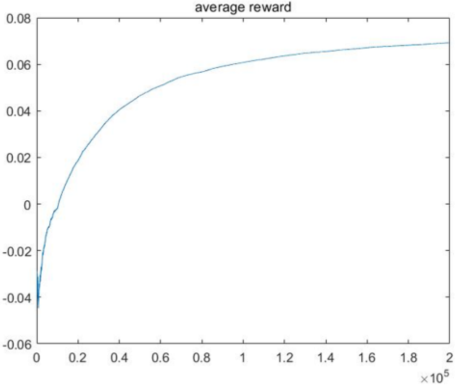


Figure: Reward for convergence

Simulation Results

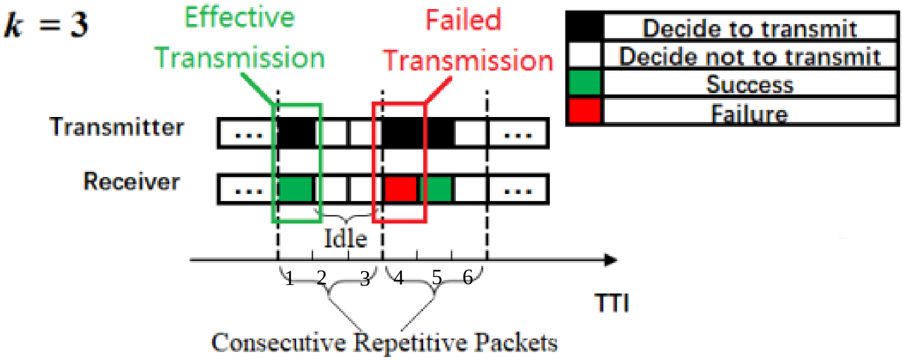


Figure: Agent decisions with repetitive times k = 3.

Simulation results

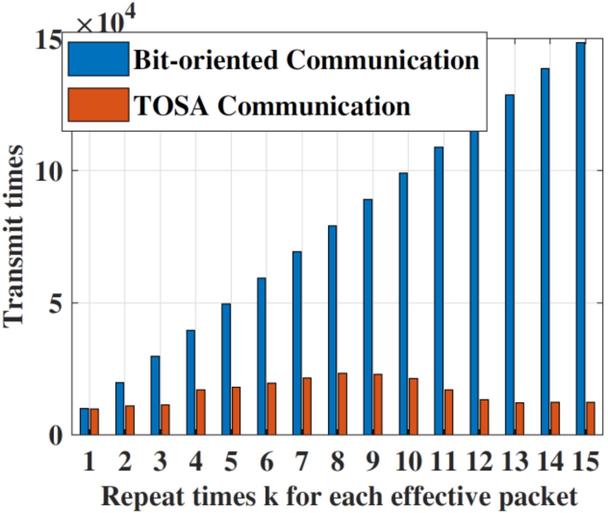


Figure: Number of C&C Transmission Times

Simulation results

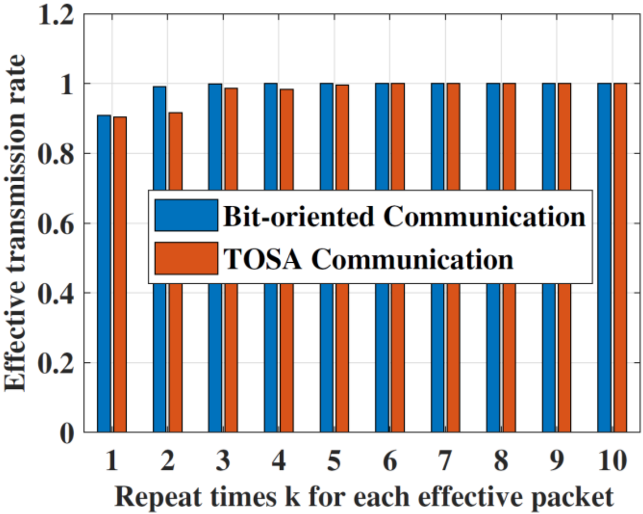


Figure: Effective Transmission Rate.

Conclusions

- We developed a general TOSA communication framework for UAV C&C transmission

- We defined age of information (AoI) and similarity to quantify the semantic-level and effectiveness-level performance, respectively.

- We proposed a general DRL algorithm to optimize the TOSA information of the downlink C&C transmission based on the similarity and AoI of C&C data.

- Our numerical results shed light on that our proposed TOSA framework can guarantee the C&C task execution with much fewer communication resources.

III: Task-oriented and Semantics-aware Communication for Augmented Reality

[7] Z. Wang, **Y. Deng*,** A. H. Hamid**,** "Task-Oriented and Semantics-Aware Communication for Augmented Reality," 2023.

AR application

Interactive Gaming

Virtual Meeting

Virtual Gym

Remote Control

Fig. 1: AR applications

Traditional Communication Framework for Augmented Reality:

- ❑ Use the point cloud as the basic recovery unit.
- ❑ Low reliability and high latency.

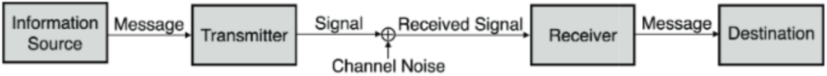


Fig. 2: Traditional Communication Framework

Semantic Communication Framework for Augmented Reality:

- ❑ Support real-time changes for the vast amount of associated information.
- ❑ Focused on identifying the semantic content of traditional data, such as text, speech, and images.

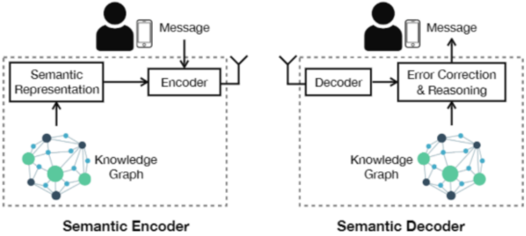


Fig. 3: Semantic Communication Framework

Challenges:

- Task-oriented semantics-aware/ semantic communication framework in AR applications has been proposed.

- Current research in wireless AR applications has not fully addressed the representation data concerning the effectiveness of avatar transmission.

- The definition and extraction methods for task-oriented semantics-aware communication in AR applications have not been explored.

System model

Traditional Point Cloud Communication Framework

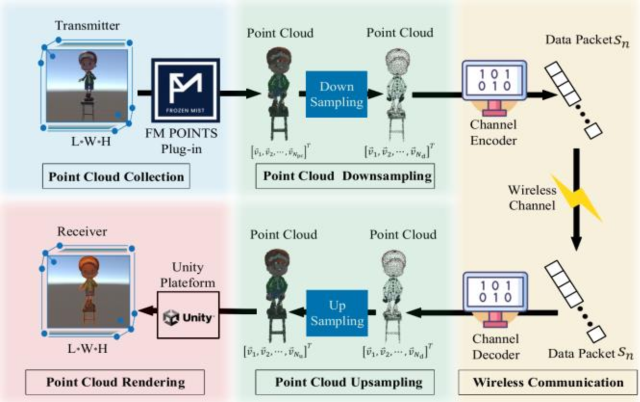


Fig. 4: Traditional point cloud wireless AR framework.

- ❑ Each model requires over 15,000 point cloud representations[8].
- ❑ Needs less than 10ms latency to achieve a satisfied QoE[9].

[8] Z.-L. Zhang, U. K. Dayalan, E. Ramadan, and T. J. Salo, "Towards a Software-Defined, Fine-Grained QoS Framework for 5G and Beyond Networks," in Proceedings of the ACM SIGCOMM Workshop on Network-Application Integration (NAI), August 2021, pp. 7-13.
[9] S. Van Damme, M. T. Vega, and F. De Turck, "Human-Centric Quality Management of Immersive Multimedia Applications," in Proceedings of the IEEE Conference on Network Softwarization (NetSoft), June 2020, pp. 57-64.

Point Cloud Collection

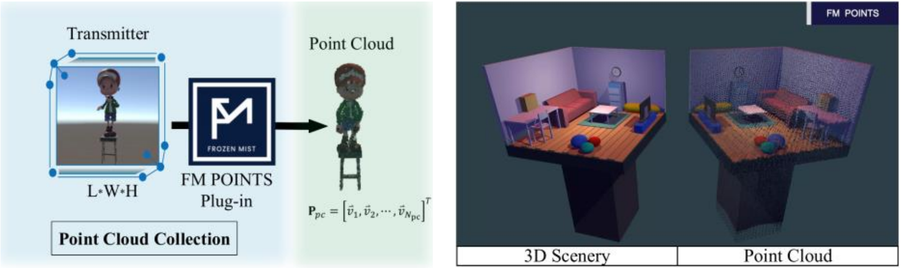


Fig. 5: Point Cloud Collection with FM Points [10].

- The generated point clouds P_{pc} of AR application with FM POINTS plug-in consist of thousands of points v_i, which is denoted as

$$\mathbf{P}_{\mathrm{pc}} = \left[\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_{N_{\mathrm{pc}}}\right]^T \text{ where } \quad \vec{v}_i = \left(l_{\mathrm{x}}, l_{\mathrm{y}}, l_{\mathrm{z}}, c_{\mathrm{r}}, c_{\mathrm{g}}, c_{\mathrm{b}}\right) \qquad \textbf{(1)}$$

N_{pc}: the number of generated point cloud.
$\left(l_{\mathrm{x}}, l_{\mathrm{y}}, l_{\mathrm{z}}\right)$: The location parameters.
$(c_{\mathrm{r}}, c_{\mathrm{g}}, c_{\mathrm{b}})$: The color parameters.

[10] FM POINTS plug-in in Unity Store: https://assetstore.unity.com/packages/tools/modeling/fm-points-163045

Point Cloud Downsampling

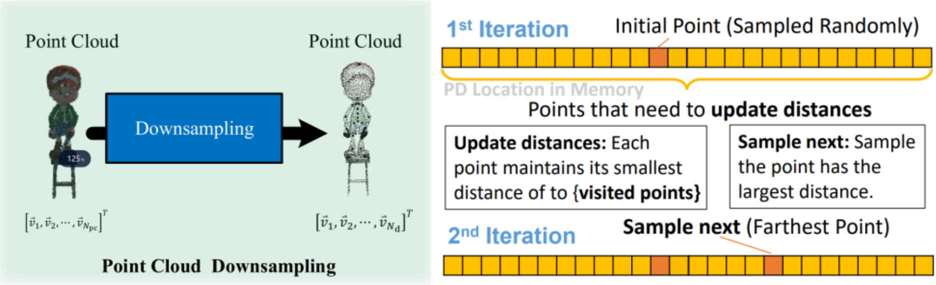


Fig. 6: Point Cloud Downsampling and Farthest Point Sampling[11].

- The process of point cloud downsampling with the farthest Point Sampling (FPS), denoted as $\mathcal{D}(\cdot)$, can be expressed as:

$$\mathbf{P}_{\mathrm{dpc}} = [\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_{N_\mathrm{d}}]^T = \mathcal{D}(\mathbf{P}_{\mathrm{pc}}) \qquad \mathbf{(2)}$$

$\mathbf{P}_{\mathrm{dpc}}$: The point cloud data awaiting transmission.
N_d : The number of downsampled points.

[11] Li, Jingtao, et al. "An Adjustable Farthest Point Sampling Method for Approximately-sorted Point Cloud Data." *2022 IEEE Workshop on Signal Processing Systems (SiPS)*. IEEE, 2022.

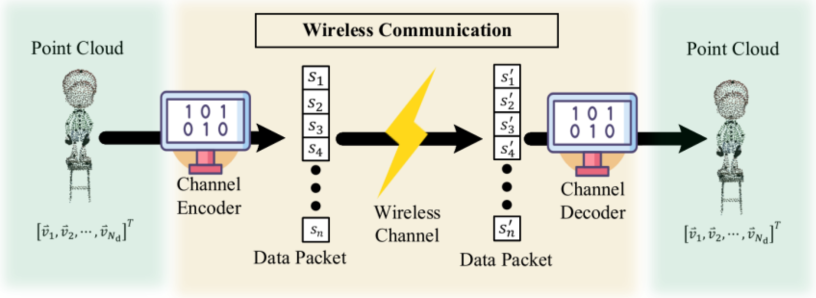
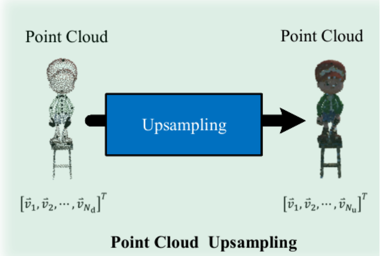


Fig. 7: Wireless Communication.

- Rayleigh fading channels, influenced by additive white Gaussian noise and employing an Orthogonal Frequency Division Multiplexing (OFDM) scheme. The overall SNR of the communication process within channel can be expressed as

$$\mathrm{SNR} = \frac{\sum_{n=1}^{N_c} \|h_n \cdot s_n\|^2}{\sum_{n=1}^{N_c} \sigma_n^2} \quad \textbf{(3)}$$

s_n: The transmitted signal. h_n: The channel state gain. σ_n^2: The subchannel noise.



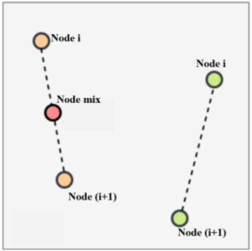


Fig. 8: Point Cloud Upsampling and Linear Interpolation [12].

- The upsampling process, denoted as $\mathcal{U}(\cdot)$, can be expressed as follows:

$$\mathbf{P}_{\mathrm{upc}} = [\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_{N_u}]^T = \mathcal{U}\left(\mathbf{P}'_{\mathrm{dpc}}\right), \quad \textbf{(4)}$$

- The linear interpolation is represented as

$$l_x^{mix} = (1-\lambda)\cdot l_x^i + \lambda \cdot l_x^{i+1} \qquad c_x^{mix} = (1-\lambda)\cdot c_x^i + \lambda \cdot c_x^{i+1}$$
$$l_y^{mix} = (1-\lambda)\cdot l_y^i + \lambda \cdot l_y^{i+1} \qquad c_y^{mix} = (1-\lambda)\cdot c_y^i + \lambda \cdot c_y^{i+1}$$
$$l_z^{mix} = (1-\lambda)\cdot l_z^i + \lambda \cdot l_z^{i+1} \qquad c_z^{mix} = (1-\lambda)\cdot c_z^i + \lambda \cdot c_z^{i+1}$$

(5)

[12] Chen, Yunlu, et al. "Pointmixup: Augmentation for point clouds." Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part III 16. Springer International Publishing, 2020.

Point Cloud Rendering

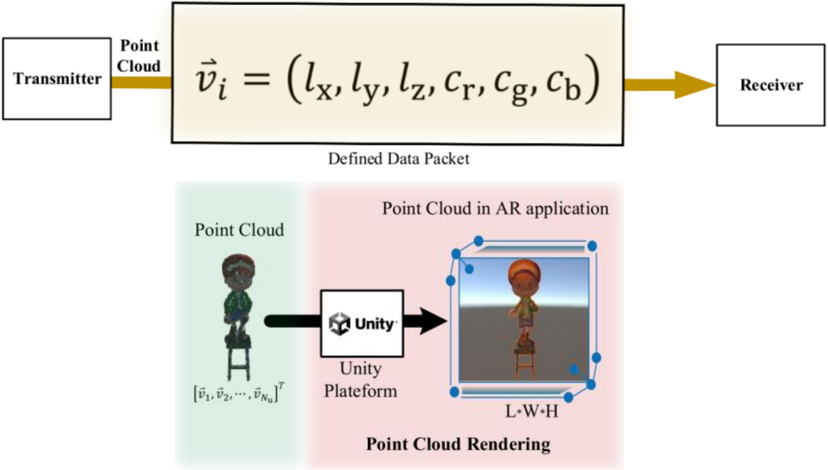


Fig. 9: Point Cloud Rendering and defined data packet.

Level of Detail (LoD): Dynamically adjusts the number of rendered points based on the user viewing distance and save on rendering resources.

Task-oriented and Semantics-aware Communication

- Joint exploitation of information context and its importance to the task.

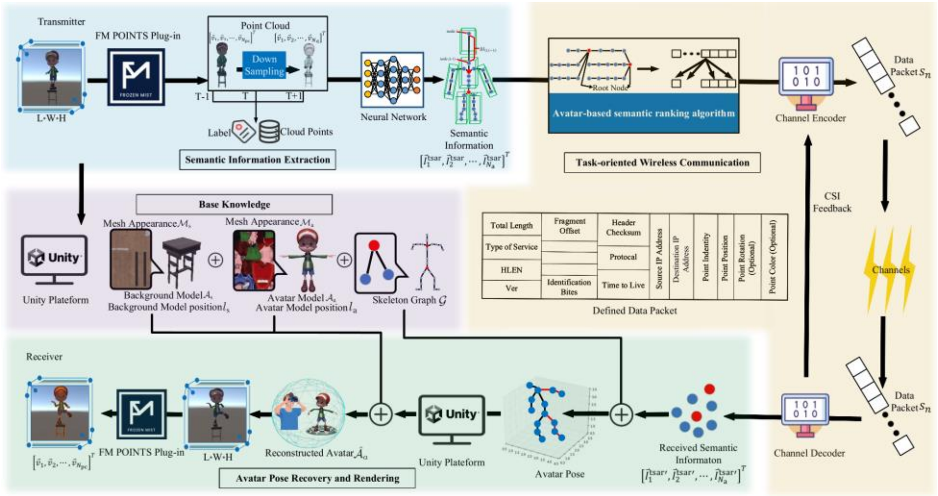


Fig. 10: Task-oriented and semantics-aware communication framework.

Base Knowledge Extraction

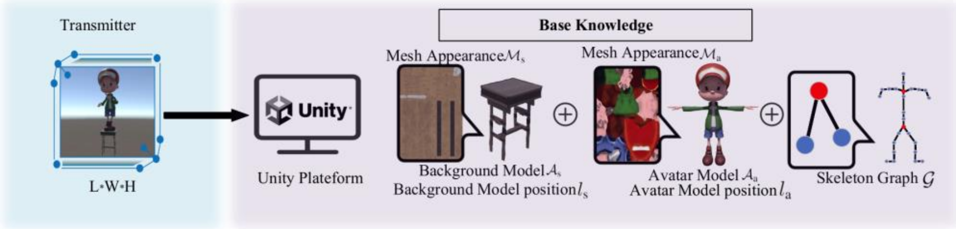


Fig. 11: Base Knowledge.

Base knowledge $\boldsymbol{B}_{tsar}$ is represent as a subset of the following information: avatar skeleton graph $\mathcal{G}$, avatar initial position l_{a}, avatar model $\mathcal{M}_{\mathrm{a}}$, stationary background model $\mathcal{M}_{\mathrm{s}}$, stationary model position l_{s}, and their appearance meshes $\mathcal{A}_{\mathrm{a}}$ and $\mathcal{A}_{\mathrm{s}}$.

Whenever a new object appears in the AR scene, the base knowledge at both transmitter and receiver need to be updated synchronously.

$$\boldsymbol{B}_{\mathrm{tsar}} = \{\mathcal{M}_{\mathrm{a}}, \mathcal{M}_{\mathrm{s}}, \mathcal{A}_{\mathrm{a}}, \mathcal{A}_{\mathrm{s}}, l_{\mathrm{s}}\}. \qquad \textbf{(6)}$$

Semantic Information Extraction

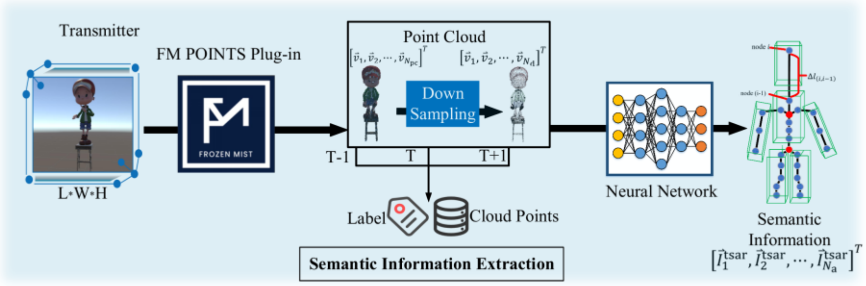


Fig. 12: Semantic Information Extraction.

- The entire process of semantic extraction from cloud point $\mathbf{P}_{\text{upc}}$, denoted as $\mathcal{S}(\cdot)$,, can be expressed as

$$\mathbf{D}_{\text{tsar}} = \left[\vec{I}_1^{\text{tsar}}, \vec{I}_2^{\text{tsar}}, \cdots, \vec{I}_{N_{\text{a}}}^{\text{tsar}}\right]^T = \mathcal{S}\left(\mathbf{P}_{\text{upc}}, \theta_{\text{s}}\right) \quad \textbf{(7)}$$

$\mathbf{D}_{\text{tsar}}$: The semantic information of the AR application.
N_{a} : The total number of skeleton in avatar model.
$\vec{I}_i^{\text{tsar}}$: The semantic information in skeleton of the TSAR.

Semantic Information Extraction

Semantics-aware network(SANet) :

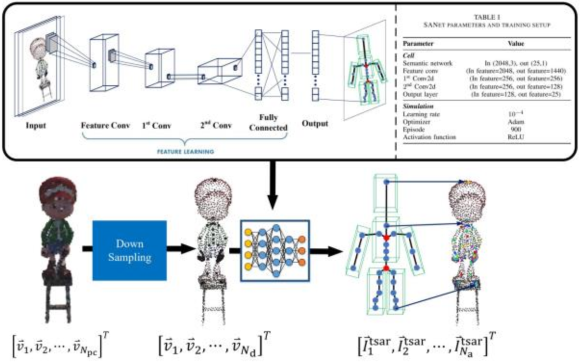


TABLE I
SANET PARAMETERS AND TRAINING SETUP

Parameter	Value
Cell	
Semantic network	In (2048,3), out (25,1)
Feature conv	(In feature=2048, out feature=1440)
1st Conv2d	(In feature=256, out feature=256)
2nd Conv2d	(In feature=256, out feature=128)
Output layer	(In feature=128, out feature=25)
Simulation	
Learning rate	10^{-4}
Optimizer	Adam
Episode	900
Activation function	ReLU

Fig. 13: Semantic extraction process and SANet model parameters.

- ❑ **Input Data**: Downsampled point cloud geometry information $\mathbf{P}_{upc}$.
- ❑ **Output Label**: Semantic information $\mathbf{D}_{tsar}$ of the skeleton point prediction.

$$I_i^{tsar} = (l_x, l_y, l_z, r_w, r_x, r_y, r_z), i \in [0, N_a] \quad \textbf{(8)}$$

Skeleton Graph Formation ($\mathcal{G}$)

We consider three components in each skeleton:

1. Quaternion rotation   2. Euler angle   3.Vector position

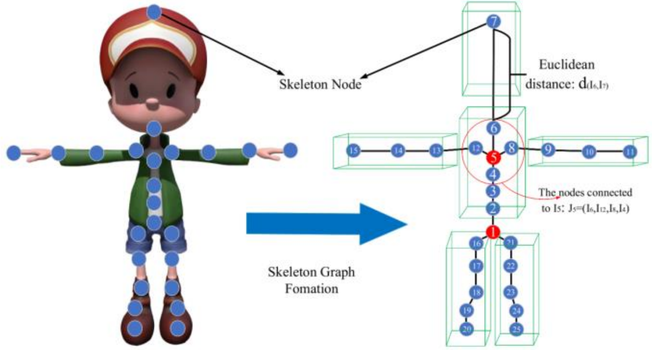
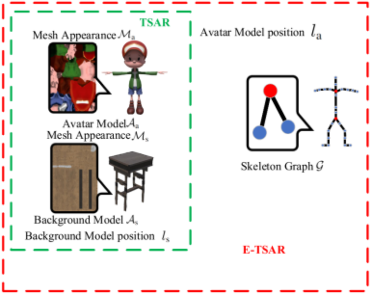


Fig. 14: Avatar Skeleton Graph Formation.

Skeleton Graph notation:

- ❑ **Node**: Each skeleton in avatar is denoted as a node in a graph $\mathcal{G}$.
- ❑ **Neighbor node**: The node connect with each other.
- ❑ **Neighbor nodes distance**: The distance between two skeletons in avatar model.

Semantic Information Definition



$$\begin{bmatrix} e_p \\ e_r \\ e_y \end{bmatrix} = \begin{bmatrix} \arctan \dfrac{2(r_y r_z + r_w r_x)}{1 - 2(r_x^2 + r_y^2)} \\ \arcsin\left(2(r_w r_y - r_x r_z)\right) \\ \arctan \dfrac{2(r_x r_y + r_w r_z)}{1 - 2(r_y^2 + r_z^2)} \end{bmatrix} * \frac{180}{\pi} \quad \textbf{(9)}$$

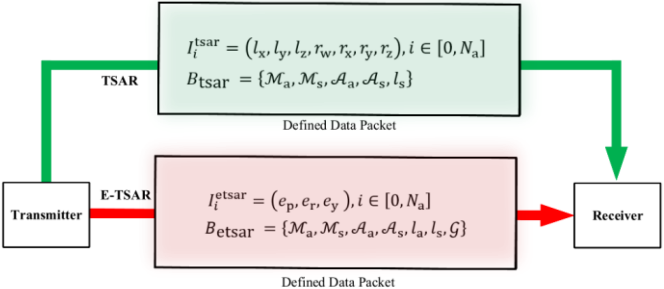


Fig. 15: The semantic information of TSAR and Euler based TSAR (E-TSAR).

Task-oriented Wireless Communication

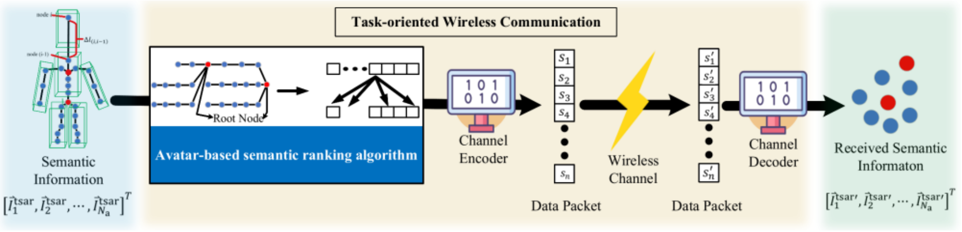


Fig. 16: Task-oriented Wireless Communication.

The ranking weights of all semantic information $\mathbf{D}_{\text{tsar}}$ are denoted as $\overrightarrow{W}_{tsar}$ and can be formulated as

$$\overrightarrow{W}_{\text{tsar}} = \mathcal{W}(\mathbf{D}_{\text{tsar}}, \mathcal{G}) = \left(w_1, w_2 \ldots, w_{N_a}\right), \qquad \textbf{(10)}$$

where w_i represents the importance for avatar recovery of I_i^{tsar}. Mapping the skeleton information with higher importance with better channel state:

$$\mathcal{M}\left(\overrightarrow{W}_{\text{tsar}}, \mathcal{G}, \vec{H}_{\text{c}}\right) = \left(I_i \xrightarrow{\mathcal{M}(\cdot)} h_j\right) \; i \in [1, N_{\text{a}}], j \in [1, N_{\text{c}}]. \qquad \textbf{(11)}$$

Avatar-based Semantic Ranking Algorithm

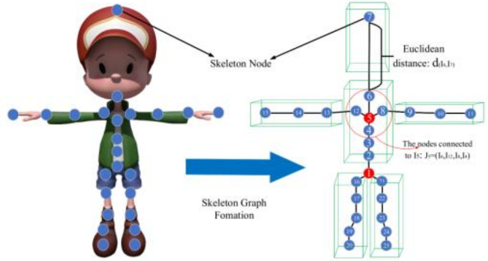


Fig. 17: Avatar Skeleton.

Algorithm 2 Avatar-based Semantic Ranking Algorithm

Initialization: Base Knowledge $\mathcal{B}$
Get $\mathcal{G}, \mathcal{A}_a$ from $\mathcal{B}$,
Get $d_{(I_i,J_j)}$ from $\mathcal{A}_a$
Count skeleton number $N_a = \mathbf{C}_s(\mathcal{G})$
repeat
  $k = k + 1$
  for each i in N_a **do**
    Update $\omega^k_{I_i}$ with $d_{(I_i,J_j)}$ based on Eq. (17)
    $\delta = ||\omega^k_{I_i} - \omega^{k-1}_{I_i}||$
  end for
until $\delta < \varepsilon$
Update $\{I_i, h_j\}$ according to Eq. (18)

Output: Channel Mapping $\{I_i, h_j\}$

Semantic ranking weights calculation:

$$\omega_i = \frac{N_J}{(1-\alpha)} + \sum_{j=0}^{N_J}\left(d_{(I_i,I_j)} \times \omega_j\right) \qquad \textbf{(12)}$$

Semantic information and subchannel mapping:

$$\mathcal{M}\left(\overrightarrow{W_{\text{tsar}}}, \mathcal{G}, \vec{H}_{\text{c}}\right) = \left(I_i \xrightarrow{\mathcal{M}(\cdot)} h_j\right), i \in [1, N_{\text{a}}], j \in [1, N_{\text{c}}]. \qquad \textbf{(13)}$$

Avatar Pose Recovery and Rendering

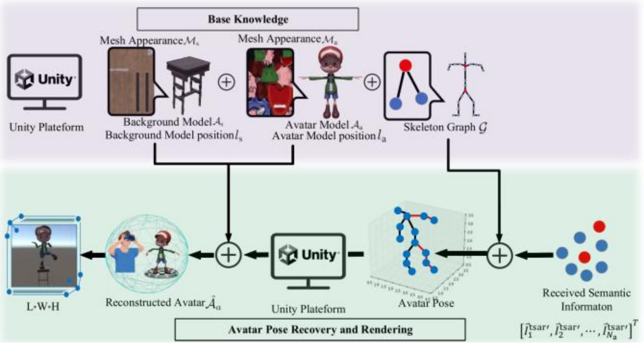


Fig. 18: Avatar Pose Recovery and Rendering.

- The received data D'_{tsar} is denoted as

$$\mathbf{D}'_{\mathrm{tsar}} = \left[\vec{I}_1^{\mathrm{tsar}\prime}, \vec{I}_2^{\mathrm{tsar}\prime}, \cdots, \vec{I}_{N_\mathrm{a}}^{\mathrm{tsar}\prime}\right]^T, \qquad \mathbf{(14)}$$

$\mathbf{D}'_{\mathrm{tsar}}$: Received semantic information

$I_i^{\mathrm{tsar}\prime}$: Single received skeleton information.

- Avatar pose recovery process $\mathcal{R}(\cdot)$ is denoted as

$$\hat{\mathcal{A}}_a = \mathcal{R}(\mathbf{D}'_{\mathrm{tsar}}, \boldsymbol{B}). \qquad \mathbf{(15)}$$

Avatar Pose Recovery

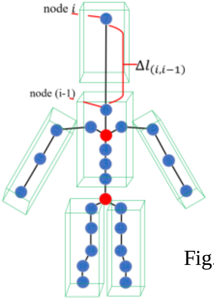


Fig. 19: Skeleton Graph

The relative distance $\Delta l_{(i,i-1)}$ between skeleton node (i-1) and its root node (i) can be represented as follows:

$$\Delta l_{(i,i-1)} = r_i \times l_{i-1} \qquad \textbf{(16)}$$

The actual position of skeleton node i:

$$l_i = l_{i-1} + \Delta l_{(i,i-1)} \qquad \textbf{(17)}$$

Algorithm 1 Avatar Pose Recovery

```
Initialization: Base knowledge B, received data D⃗'
Get skeleton graph 𝒢, avatar initial position l_a avatar
    model ℳ_a, and avatar appearance mesh 𝒜_a from B
Count skeleton number N_a = C_s(𝒢)
Count received data N_r = C_r(D_tsar)
if 𝒢 ∉ B & l_i ∈ D⃗' then
    for each i in N_r do
        Attach I_i^tsar to model ℳ_a
    end for
else
    for each i in N_a do
        update l_i according to Eq. (14) and Eq. (15)
        Attach I_i^etsar to model ℳ_a
    end for
end if
Generate avatar Â_a with appearance mesh 𝒜_a and model
    initial position l_a according to Eq. (16).
Output: Avatar with recovered position Â_a
```

Avatar pose recovery:

$$\hat{\mathcal{A}}_a = \mathcal{R}(\mathbf{D}'_{\text{tsar}}, \boldsymbol{B}) \qquad \textbf{(18)}$$

EC-TSAR

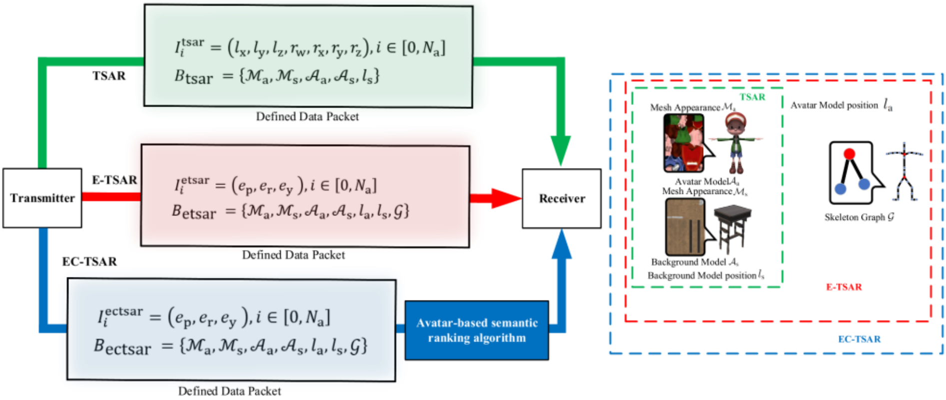


Fig. 20: The semantic information and base knowledge of EC-TSAR

Problem Formulation

The overall TSAR framework aims to optimize task-oriented semantics-aware communication for avatar-centric conferencing and gaming AR applications:

$$\mathcal{P}: \min_{\theta_{S},(I_i,h_j)} \lim_{T\to+\infty} \frac{1}{T}\sum_{t=0}^{T}\sum_{i=0}^{N_A}\left(\vec{I}_{i,t}^{\text{tsar}} - \vec{I}_{i,t}^{\text{tsar}\prime}\right)\cdot\omega_i. \quad \textbf{(19)}$$
$$\text{s.t. } i \in [1, N_A],\ j \in [1, N_c].$$

Where N_{A} and N_{c} represent the number of semantic information and subchannels, respectively. (I_i, h_j) is the channel mapping process. $\vec{I}_{i,t}^{\text{tsar}}$ and $\vec{I}_{i,t}^{\text{tsar}\prime}$ represent the skeleton information at the transmitter and receiver respectively.

Semantic Extraction Performance

Semantic-aware network(SANet) :

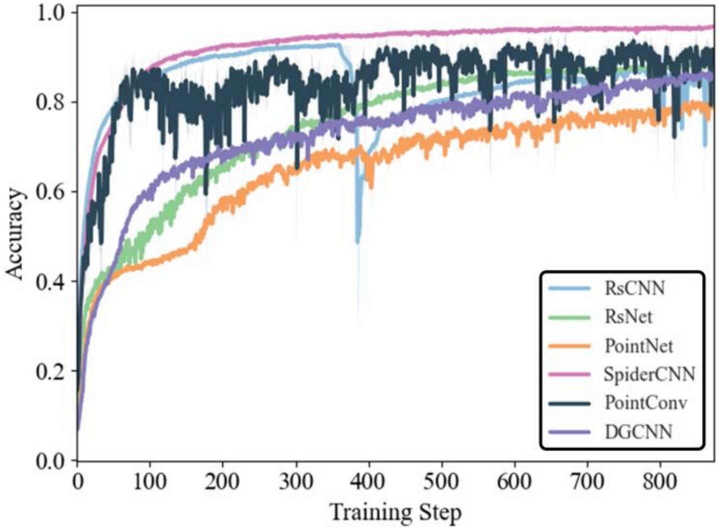


Fig. 21: SANet Accuracy.

Mean Per Joint Position Error (MPJPE）Distance

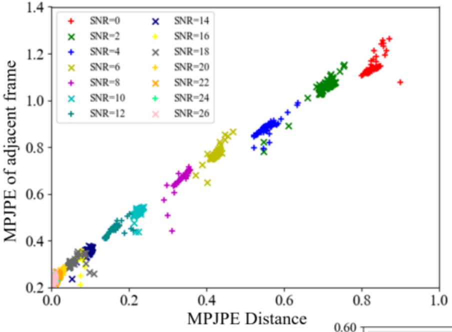


Fig. 22 (a): TSAR

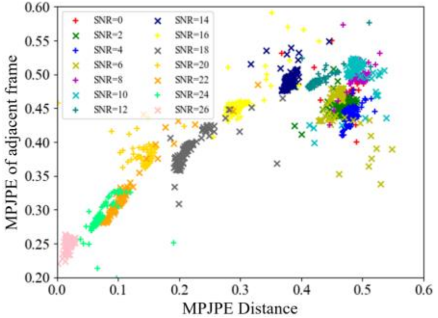


Fig. 22 (b): E-TSAR

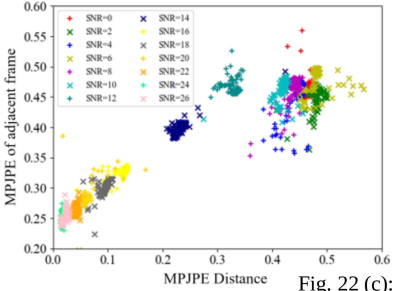


Fig. 22 (c): EC-TSAR

Mean Per Joint Position Error (MPJPE):

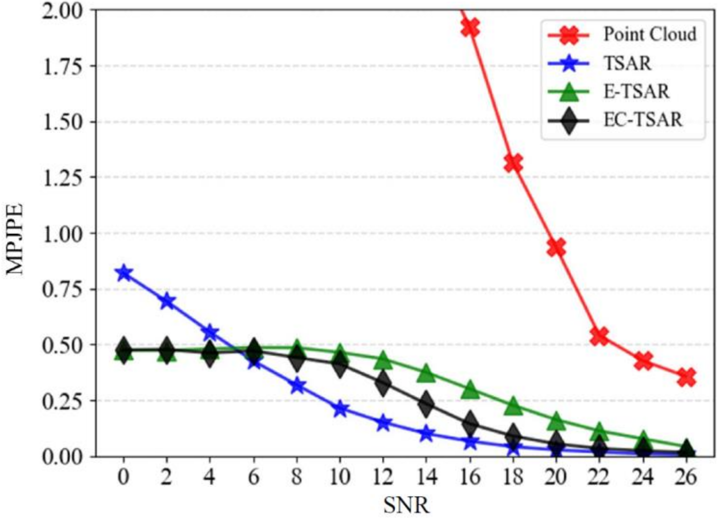


Fig. 23: MPJPE under various SNRs.

Time Delay

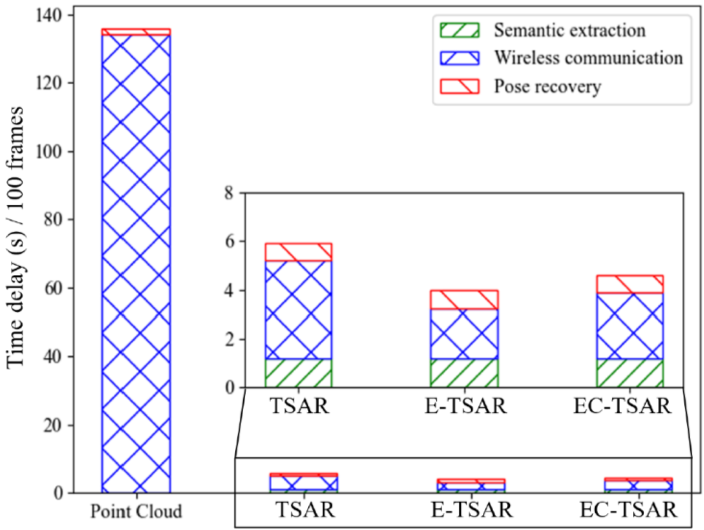


Fig. 24: Time Delay

P2Point and PSNRy:

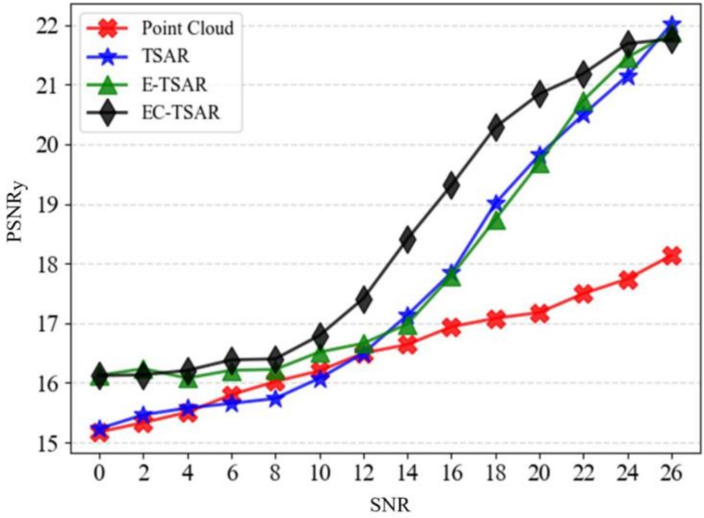


Fig. 25: PSNRy

Demo Video

Conclusion

- We proposed a task-oriented and semantics-aware communication framework in AR (TSAR) for avatar-centric end-to-end communication in AR applications.

- We applied an avatar-based semantic ranking (AbSR) algorithm to abstract features from the avatar transmission and recovery task-level.

- We demonstrated the TSAR superiority via numerous experiments.

Acknowledgement

Yujie Xu

Hui Zhou

Wenchao Wu

Zhe Wang

N. Pappas

A. Nallanathan

Thanks for your attention!
We are recruiting!