# 会议纪要

**会议主题：基于掩码自编码器的图像传输安全语义通信**

**会议时间：** 2025年2月10日19：00至21：00

**会议地点：** 西安电子科技大学北校区会议中心303-1会议室

**线下人员：** 承楠 孙瑞锦 尹志胜 王葳 王秀程 马龙飞 周新阳 权赟昊 张玉洁 胡陆莹 齐阁 邓川 孙兴栋 方忠盛 贾宏刚 解思舀 郑雯馨 祝馨平 贾昊燏 张智杰 杨杰 邱子仪 刘永红 万佳林 朱煜朋 韩松明 刘苏 郑佩林 孙路路 王天宇 于凡迪 尉家豪 杨双宇 傅连浩 郭译凡 侯毓真 李青壮 许嘉洁 冉艺泉 闵昕阳

**线上人员：** 尚佳瑶 黄蕾 李成成 张岳 赵璇 陈梦豪

**记 录 人：** 张智杰

**会议内容：**刘苏介绍了基于掩码自编码器的图像传输安全语义通信系统

* **承楠：**我认为视频语义通信是语义通信中最重要的应用。通常传输视频的信道条件较差，例如海洋或深空环境，发送的视频传输速度较慢，利用语义通信传输的是就是文字信息吗？
* **刘苏：**对视频领域没有深入研究。目前我的研究主要集中在文本和图像。
* **承楠：**假设语义通信是一张图片，提取出它的语义信息并发送。传输结束后，是否还需要将其恢复成图片？
* **王秀程：**目前需要恢复。华为提出了一个概念，即无译码的语义通信系统。他们认为只需传输语义信息，无需进行译码。
* **尹志胜：**实际上，无论是无译码还是无编码，本质上并不重要。虽然称为“无译码”，但其核心思想是不进行解码，而直接利用语义信息。不过，通信系统仍需有一个评价指标，即恢复后的效果必须进行评估，否则通信系统将失效。如果只发送而不关注传输质量，这是不可行的。
* **承楠：**首先，你是否对图像进行了两种遮罩处理？然后将其合并，经过信道编码和传输。这意味着接收方收到的应该是相同的内容，只是信道条件不同，噪声也不同。那么，如何实现这种效果？两者的解码器是否不同？
* **刘苏：**解码器是相同的。
* **承楠：**如果解码器相同，为什么会出现不同的效果？如果H变成G，G变成H，传输效果会相同吗？
* **刘苏：**这里使用了神经网络，并根据窃听者和合法接收者的信道差异进行对抗训练，从而训练编码器和解码器。如果H变成G，G变成H，那样传输的内容可能会反转。在这个系统中，假设窃听者的信道条件比合法接收者差，因此不会出现这种情况
* **承楠：**这是可行的。所以，在训练过程中，需要确保信道条件较好的解码器恢复效果优于信道条件较差的解码器。这需要大量数据进行训练。
* **尹志胜：**我认为他的表述有些不准确。实际上，原理并不在于H变成G或G变成1/2G，也不一定要求窃听者的信道条件更弱。只要两者存在差异，就可以利用这种差异性进行设计。仿真的目的是为了增大差异性，系统通过学习这种差异性来设计遮罩比例和随机化处理，从而将差异性融入语义编码中。
* **承楠：**在通信开始前，如何确定合法用户和非法用户之间的信道差异来进行遮罩处理。如何根据信道差异来确定遮罩的位置？
* **尹志胜：**可以这样理解，G一定在H的零空间中。H和H的零空间之间存在差异，然后学习它的差异性。
* **承楠：**只要保证H的条件良好，而不是H的条件较差，这是可行的。
* **孙瑞锦：**测试的时候信道H发生变化会影响结果吗？
* **承楠：**是不是有神经网络将合法信道H作为输入，来指导遮罩的pattern,这里的H是怎么使用的，在架构图中没有体现出来。
* **刘苏：**这里的mask不涉及神经网络，使用人工遮罩算法提前决定遮罩图片的方式，没有输入。
* **承楠：**信道信息H如何在发端使用？合法用户的H和窃听者的非H是使用在整个架构的哪个位置？H是始终动态变化的，不是一个固定值，但是我们可以假设H是已知的，使用导频等方式获取，在此基础上进行工作。H一定要反映在人工遮罩或者其他的地方。对于不同的H，都要达到相同的效果。
* **孙瑞锦：**训练的时候，H是怎么生成的，只有1个H吗？
* **承楠：**在非H空间找了很多G，H应该是没有变的。后续要想一想怎么把H融入到机制中，哪怕复杂一些。
* **尹志胜：**相当于单次优化，测试等于训练。
* **孙瑞锦：**有使用迫零编码吗？
* **刘苏：**没有。
* **承楠：**现在需要加一个重要的模块，作为你的contribution，以信道H作为输入，输出mask的pattern。
* **尹志胜：**人工遮罩和信道H没有关系吗，人工遮罩是怎么做的？
* **承楠：**遍历查找获取索引，寻找具有更大意义的块，和随机遮罩形成对比。
* **尹志胜：**怎么获得现在的结果图的？
* **承楠：**这是固定H微调语义编解码器得到现在的结果图的，不合理。
* **尹志胜：**关键是要考虑信道变化的情况下，驱动你的训练过程。
* **承楠：**可不可以训练一个mask的神经网络，用来指导人工遮罩。
* **王秀程：**可以。
* **承楠：**为什么要使用random mask？
* **尹志胜：**拉大利用背景和突出特征的差异。
* **承楠：**可以将语义编码器的输入修改为图像和信道。有必要使用MAE模型吗？
* **刘苏：**这样提取的特征比一张完整图片更少，经过两次遮罩提取的特征也只有原始图片的一半，传输的数据量更少。
* **承楠：**也可以直接固定语义编码的编码结果大小。
* **承楠：**MAE的初衷是什么？
* **王秀程：**mask用来预训练网络，训练完后进行图像分类。
* **尹志胜：**他工作中的mask相当于在损失性能不大的前提下，以更少的语义信息进行传输，提升训练速度，减少传输的数据量。
* **承楠：**针对信道有什么好的训练方法，需要强化学习吗？
* **王秀程：**不需要，只要H服从相同的分布。将图片、H、G输入到语义编码器中，预训练语义解码器，只训练编码器，利用损失函数进行端到端的训练，不需要mask。
* **孙瑞锦：**在训练的时候H,G已知，测试的时候只需要H。
* **承楠：**这样的语义编码器是不是需要复杂的结构，用什么架构比较好？
* **刘苏：**现在使用的是VIT。
* **王秀程：**使用U-net容易训练。先训练好U-net,后续进行微调即可，接收端共享一个解码器。
* **承楠：**在做的过程中要学习和理解，把机制搞清楚，将基础抓牢，多向师兄师姐请教。

**会议总结：**

1. 围绕语义通信的相关细节展开了深入讨论。语义通信相比传统视频传输方式，能够提取并传输语义信息，提高通信效率。关键在于利用信道信息与其零空间的差异性以调整遮罩策略，增强语义编码的安全性和鲁棒性。未来需研究如何在通信前有效确定信道差异，并基于此优化遮罩位置。
2. 围绕信道信息H的利用方式与遮罩策略优化进行了深入讨论。当前遮罩为人为设定，与信道信息H无直接关联，这会导致当前系统无法应用到其他信道信息的场景下。因此，会议探讨了利用神经网络生成自适应Mask 的可行性，使其能够根据信道信息H进行动态调整，或者在语义编码器的输入中加入信道信息，并寻找合适的语义编解码网络架构。
3. 科研的过程中要学习和理解，理清机制，打牢基础，多向师兄师姐请教，及时与老师沟通讨论。