PREIM3D: 3D Consistent Precise Image Attribute Editing from a Single Image.

What is the problem?

- **Photorealistic 3D Modeling:** Generating believable 3D models from a single image.
- **Realistic View Synthesis:** Simulating how a scene would look from multiple angles, which is essential for applications like virtual reality and 3D video.
- **Limited Encoder Sharing:** Many methods train a separate encoder for each image, leading to a further increase in computation.

What has been done earlier?

Previous Work:

- **Shared Encoder for All Images:** These methods train a shared encoder that maps images into a 3D generator's latent space.
- **Per-Image Latent Code Optimization:** Each image is encoded separately, which is resource-intensive. After encoding, editing is performed in the latent space.
- **Editing in the Original Latent Space:** Previous editing methods often operate directly within the original latent space of a GAN.

Hemant Sah, B421025

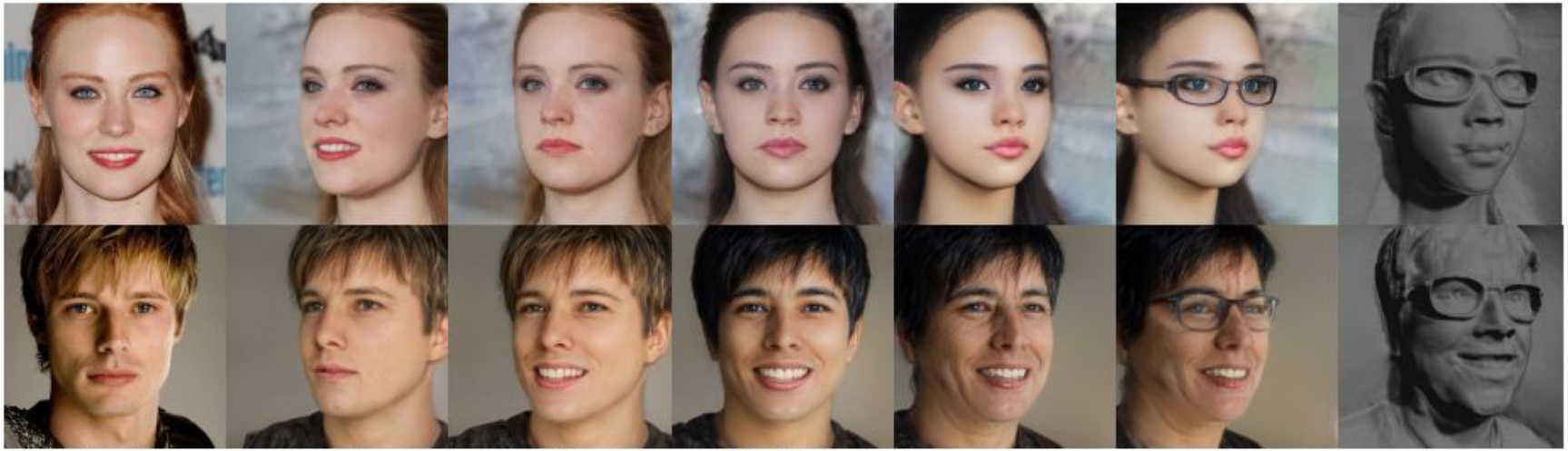

Figure 1. 3D consistent precise inversion and editing. Our method enables reconstructing texture and geometry from a single real image and allows one to perform a list of attributes editing sequentially. The yaw angles of the second to sixth columns are $[-30°, -20°, 0°, 20°, 30°]$. The last column is the shape of the sixth column.

PREIM3D: 3D Consistent Precise Image Attribute Editing from a Single Image

What are the remaining challenges? What novel solution proposed by the authors to solve the problem?

Remaining Challenges :

- **3D Inconsistency:** Images edited with these methods often appear unrealistic or distorted when viewed from different camera angles.
- **Imprecise Editing:** Editing in the original latent space can unintentionally affect attributes that weren't targeted, leading to undesirable artifacts.

Novel Solution Proposed :

- **Shared Encoder for Efficiency:** Similar to previous works, a shared encoder is trained for all images. This significantly improves the efficiency of the inversion process.

- **Alternating Training Scheme and Multi-View Identity Loss:** The encoder is trained using both real images (out-domain) and synthetic images generated by the GAN (in-domain). A multi-view identity loss is introduced to ensure that the edited images remain consistent even when viewed from different camera angles.

- **Editing in the Inversion Manifold:** The authors demonstrate that directly editing in the latent space of real images (referred to as the inversion manifold) leads to more precise results compared to editing in the original latent space of the GAN.

Hemant Sah, B421025