1. Problem

- Attribute to Facial Image
  - blond hair
  - female
  - mouth slightly open
  - arched eyebrows
  - heavy makeup

2. Related Work and Motivation

- Text to image
  - StackGAN++, AttnGAN, MirrorGAN, etc.
    - the input is different: text vs attribute
    - can not well embed the attribute label

- Attribute to image
  - Attribute2sketch2face, Lu et al., Wang et al., etc.
    - the generated images are always low resolution
    - do not consider the relationship between different attributes

3. Overall Structure of AFGAN

4. Details

- AEM: Attribute Embedding Module
  - convert the input face attributes into global and local features respectively
  - two path embedding
    - well reflect their meanings of the input attribute

  \[ S_{local} = embed_1 + embed_2 + embed_3 + (1 - S_{global}) \]
  - self attention layer
    - model the relationships between different attributes

  \[ f(x) = W_1 + S_{local} \quad g(x) = W_2 + S_{global} \quad s_1 = f(x)^T g(x) \quad s_2 = \sum_{i=1}^{N} \exp(s_i) \]
  \[ h(x) = W_3 + S_{local} \quad S_{local} = \sum_{i=1}^{N} s_i h_i(x) \]
  \[ S_{local} = (S_{local_1}, S_{local_2}, \ldots, S_{local_n}) \in \mathbb{R}^C \times \mathbb{R}^N \]

- SIGM: Stacked Image Generation Module
  - gradually generate faces with more details through a three-stage generator
  - can generate images with high resolution

- SCM: Similarity Constrain Module
  - encode the generated images with a pretrained model: \( l_{local} \) and \( l_{global} \)
  - calculate the matching degree between attribute features and image features

  \[ j = \exp(s_1) / \sum_{i=1}^{N} \exp(s_i) \]

5. Objective Function

- Generator
  - Overall
  - In SIGM
    \[ L_{G} = \sum_{i=1}^{C} l_{local} \cdot \log(h_i) + \frac{1}{2} \sum_{i=1}^{C} l_{global} \cdot (h_i - 1) + \lambda \cdot \text{MS-SSIM} \]

- In SCM
  \[ L_{SCM} = -\sum_{i=1}^{C} \log(h_i) - \lambda \cdot \text{MS-SSIM} \]

- Discriminator
  \[ L_{D} = \sum_{i=1}^{C} l_{local} \cdot \log(h_i) + \frac{1}{2} \sum_{i=1}^{C} l_{global} \cdot (h_i - 1) + \lambda \cdot \text{MS-SSIM} \]

6. Experimental Results

- The generated face images of three stages in SIGM
- The comparison of generated images with other methods

- Qualitative results
- Quantitative results

ICPR2020@ Milan, Italy