Proposed Method

- We design the learning problem as finding a generative function which is conditioned on the three representations (i.e., identity, viewpoint, and residues) that wield independent effects on the output.
- In order to handle this challenge, we propose two learning schemes.
  1. Viewpoint substitution
  2. Identity substitution
- Also, we suggest a disentangling loss function using distance covariance.

Problem Definition

- Face is generally captured along with diverse factors of variations such as identity, viewpoint, and illumination. These variations pose challenges in face recognition methods in having robust performance in a wild environment.
- In order to handle this challenge, several works have proposed disentangling methods that achieve robust performance in a wild environment by disentangling identity and non-identity variations (i.e., viewpoint and illumination).
- However, they need annotations of non-identity variations such as viewpoint and illumination. It is not easy to collect such pose or illumination information for all subjects in facial databases.
- In this paper, we propose a learning method of disentangling identity and viewpoint representations without any auxiliary supervision of the variations.
- Furthermore, we disentangle not only the identity and viewpoint but also residues (e.g., illumination and color variations) that inevitably exist in a face.
- By disentangling the non-identity variations from a face, we set a new state-of-the-art face recognition method on CFP and Multi-PIE datasets that have large pose variations.

1. Learning the viewpoint representation

- We propose to use a simple transformation that changes the viewpoint while maintaining the identity of a face image.
- The transformation can be affine, perspective, or thin plate spline transformation.
- By using the transformation, we can access the pair of images that contain different viewpoints.

2. Learning the Identity representation

- The generated image should contain the identity of the source identity image while other representations keep remaining.

3. Disentangling loss

- The disentangled representations should contain different information from one another. To guarantee the independence between learned representations, we use distance covariance as a disentangling loss function.
- Distance covariance is a metric that measures dependency between random vectors and becomes zero when the two random vectors are independent from each other.

Experimental Results