



Deep Feature Interpolation for Image Content Changes¹

Steven Lang, Manfred Faldum

February 9, 2017

¹Upchurch et al., “Deep Feature Interpolation for Image Content Changes”

Motivation

Basic Idea

Technical implementation

Results

Motivation

What we want:



Motivation

What we want:

- ▶ Add attribute in pixel space



Motivation

What we want:

- ▶ Add attribute in pixel space
- ▶ New idea: add attribute in deep feature space



Basic Idea

How to get the attribute in deep feature space?

Basic Idea

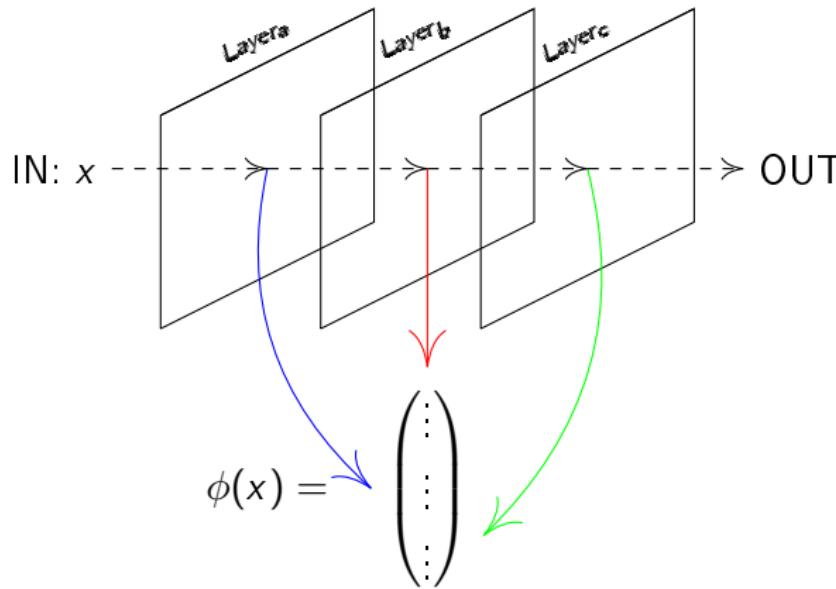
How to get the attribute in deep feature space?

- ▶ Let $\phi(x)$ be the mapping from pixel space into deep feature space by concatenating an arbitrary number of layers

Basic Idea

How to get the attribute in deep feature space?

- Let $\phi(x)$ be the mapping from pixel space into deep feature space by concatenating an arbitrary number of layers



Basic Idea

- ▶ Take k nearest neighbor images with existing attribute: S^+
- ▶ Take k nearest neighbor images with missing attribute: S^-

Basic Idea

- ▶ Take k nearest neighbor images with existing attribute: S^+
- ▶ Take k nearest neighbor images with missing attribute: S^-
- ▶ $\phi^+ = \phi(S^+)$ and $\phi^- = \phi(S^-)$
- ▶ Build the mean $\overline{\phi^+}$ and $\overline{\phi^-}$

Basic Idea

- ▶ Take k nearest neighbor images with existing attribute: S^+
- ▶ Take k nearest neighbor images with missing attribute: S^-
- ▶ $\phi^+ = \phi(S^+)$ and $\phi^- = \phi(S^-)$
- ▶ Build the mean $\overline{\phi^+}$ and $\overline{\phi^-}$
- ▶ Representation of attribute: $w = \overline{\phi^+} - \overline{\phi^-}$

Basic Idea

How to get the output picture?

Basic Idea

How to get the output picture?

- ▶ $\phi(z) = \phi(x) + \alpha w$
- ▶ Reverse mapping of $\phi(z)$ into pixel space:

Basic Idea

How to get the output picture?

- ▶ $\phi(z) = \phi(x) + \alpha w$
- ▶ Reverse mapping of $\phi(z)$ into pixel space:
$$\tilde{z} = \operatorname{argmin}_{\tilde{z}} \frac{1}{2} \|\phi(z) - \phi(\tilde{z})\|_2^2 + \lambda R_\beta(\tilde{z})$$

with $R_\beta(\tilde{z}) = \sum_{i,j} ((\tilde{z}_{i,j+1} - \tilde{z}_{i,j})^2 + (\tilde{z}_{i+1,j} - \tilde{z}_{i,j})^2)^{\frac{\beta}{2}}$

Technical

- ▶ Model: VGG19 pretrained on IMAGENET dataset
- ▶ $\phi(x)$: third, fourth and fifth Relu Layer
- ▶ Regularization parameters: $\beta = 2$, $\lambda = 0.001$
- ▶ Test set: labeled faces in the wild (LFW dataset)²
- ▶ Optimizer: Adam
- ▶ KNN: on discrete attribute features (mustache, smiling, ...)
- ▶ Normalization: $\phi(x) = \frac{\phi(x)}{||\phi(x)||}$, $w = \frac{w}{||w||}$
- ▶ Initialize z for optimization with the start image

²Huang, Ramesh, et al., *Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments*, Huang, Mattar, et al., "Learning to Align from Scratch"

Results: Eyeglasses

$\alpha = 0.4, k = 100$



Results: Sunglasses

$\alpha = 0.4, k = 100$



Results: Mouth open

$\alpha = 0.4, k = 100$



Results: Female

$\alpha = 0.4, k = 100$



Results: Heavy makeup

$\alpha = 0.4, k = 100$



Results: Smiling

$\alpha = 0.4, k = 100$



Results: Mustache

$\alpha = 0.4, k = 100$



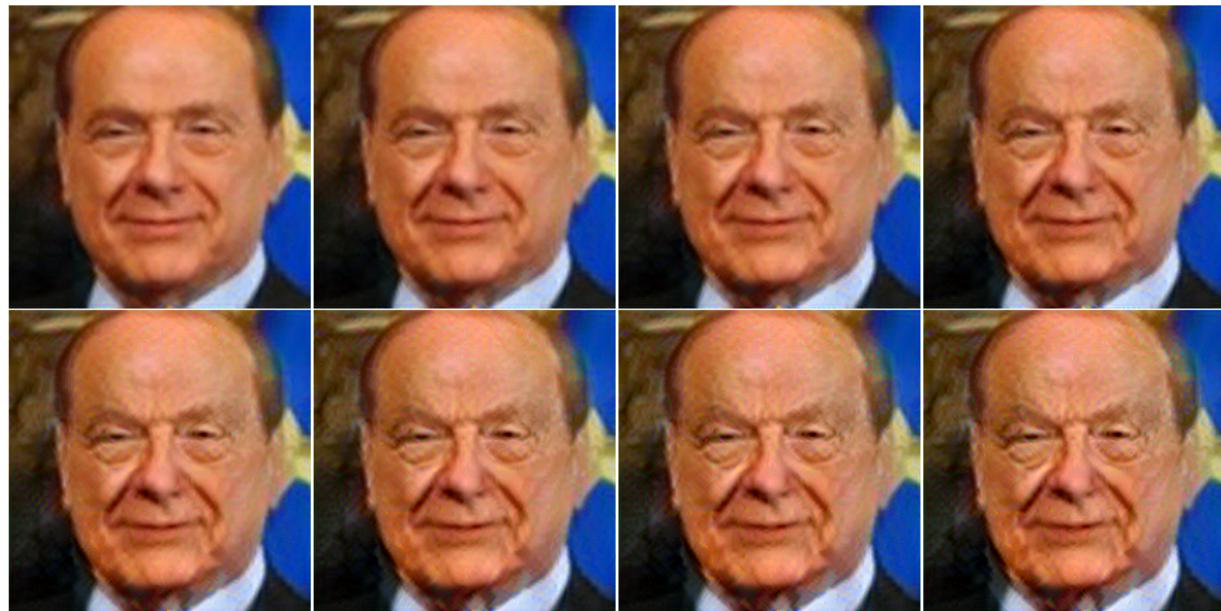
Results: Asian

$\alpha = 0.4, k = 100$



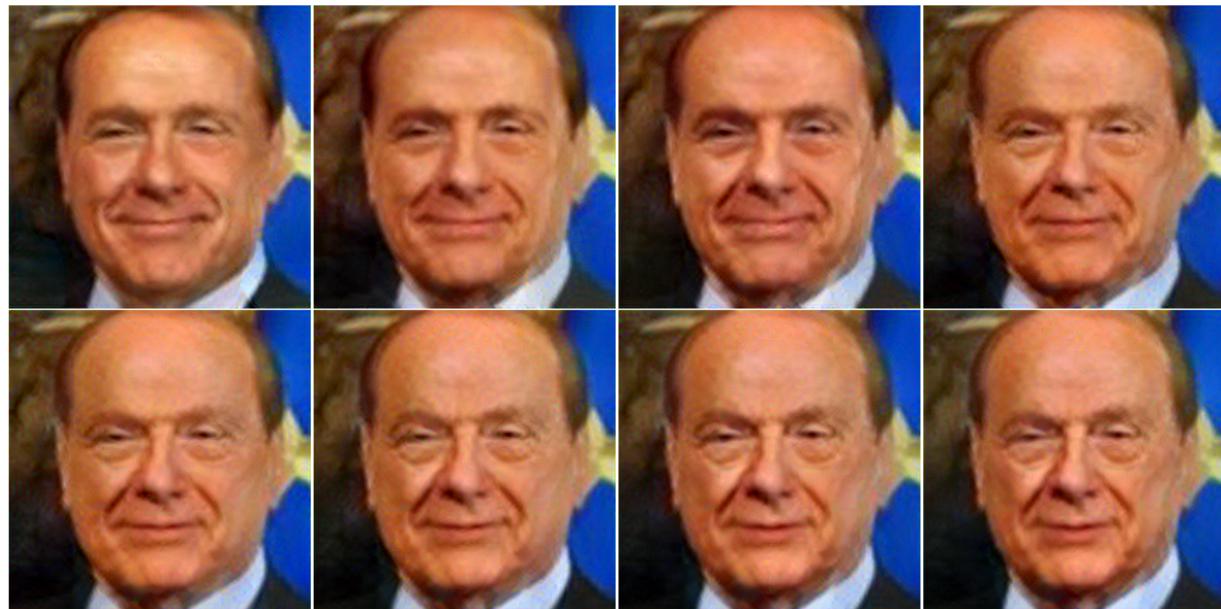
Varying the parameters: Senior

$k = 100, \alpha \in [0.3, 0.4, \dots, 1.0]$



Varying the parameters: Senior

$k \in [1, 10, 20, 50, 75, 100, 150, 200]$, $\alpha = 0.5$



Varying the parameters: Smiling

$k = 100, \alpha \in [0.3, 0.4, \dots, 1.0]$



Varying the parameters: Smiling

$k \in [1, 10, 20, 50, 75, 100, 150, 200]$, $\alpha = 0.4$



Varying the parameters: ϕ , $k = 100$, $\alpha = 0.4$

$conv3_1, conv4_1,$
 $conv5_1$



$conv4_1, conv4_2,$
 $conv4_3, conv4_4$



$conv3_1, conv3_2, conv3_3, conv3_4,$
 $conv4_1, conv4_2, conv4_3, conv4_4,$
 $conv5_1, conv5_2, conv5_3, conv5_4,$



$conv3_1, conv3_2,$
 $conv3_3, conv3_4$



$conv5_1, conv5_2,$
 $conv5_3, conv5_4$



Varying the parameters

Optimization initialization (start image)

Original image



Random image



Outlook

- ▶ Multiple features at once: $\phi(z) = \phi(x) + \alpha \sum_i w_i$
- ▶ Nearest Neighbor in deep feature space
- ▶ Change normalization: $w = w \frac{\|\phi(x)\|}{\|w\|}$

Motivation

What we want:

- ▶ Add attribute in pixel space



Motivation

What we want:

- ▶ Add attribute in pixel space
- ▶ New idea: add attribute in deep feature space



References

-  Gary B. Huang, Marwan Mattar, et al. "Learning to Align from Scratch". In: *NIPS*. 2012.
-  Gary B. Huang, Manu Ramesh, et al. *Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments*. Tech. rep. 07-49. University of Massachusetts, Amherst, Oct. 2007.
-  Paul Upchurch et al. "Deep Feature Interpolation for Image Content Changes". In: *CoRR* abs/1611.05507 (2016). URL: <http://arxiv.org/abs/1611.05507>.