Modeling Object Appearance Using Context-Conditioned Component Analysis

Daniyar Turmukhambetov, Neill D.F. Campbell, Simon J.D. Prince, Jan Kautz

http://visual.cs.ucl.ac.uk/pubs/ccca

Subspace Models

- Subspace models work well when images are aligned and have fixed structure, e.g. faces..

Factor Analysis of faces using 4 factors

Warping the source onto the target using correspondences from SiftFlow

- Objects with fixed structure can be aligned with a dense deformation field, e.g. Active Appearance Models or SiftFlow.
- Objects with varying structure cannot be readily aligned using existing methods, e.g. consider complex occlusions, varying numbers of parts, severe deformations, etc..

Context-Conditioned Component Analysis

The model for the jth pixel of the ith image is

$$x_{ij} = \mu\left[\mathbf{c}_{ij}, \boldsymbol{\theta}_\mu\right] + \sum_{f=1}^{F} \phi\left[\mathbf{c}_{ij}, \boldsymbol{\theta}_f\right] h_{if} + \epsilon_{ij}$$

where

$\mu\left[\mathbf{c}_{ij}, \boldsymbol{\theta}_\mu\right] =$ gives the mean value μ_{ij} at the pixel and is a function of the context $\mathbf{c}_{ij}$ at that pixel and a parameter vector $\boldsymbol{\theta}_\mu$

$F =$ number of components

$\phi\left[\mathbf{c}_{ij}, \boldsymbol{\theta}_f\right] =$ function that computes the component's value at the pixel; it is a function of the context $\mathbf{c}_{ij}$ and a parameter vector $\boldsymbol{\theta}_f$

$h_{if} =$ weight for the weighted sum of F function terms $\phi[\cdot,\cdot]$; the weights are the same for all pixels in each image

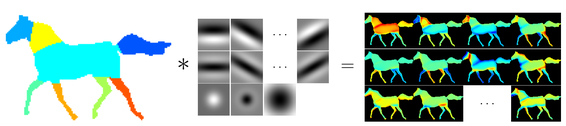
$\epsilon_{ij} =$ an independent stochastic noise variable distributed as $\text{Norm}_{\epsilon_{ij}}[0, \sigma^2]$

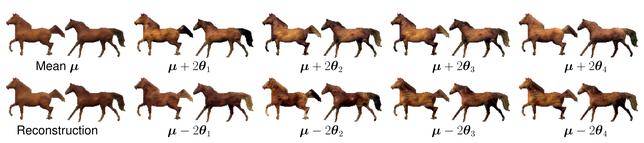
For tractability, we assume

$$\phi\left[\mathbf{c}_{ij}, \boldsymbol{\theta}_f\right] = \mathbf{a}\left[\mathbf{c}_{ij}\right]^T \boldsymbol{\theta}_f \ ,$$

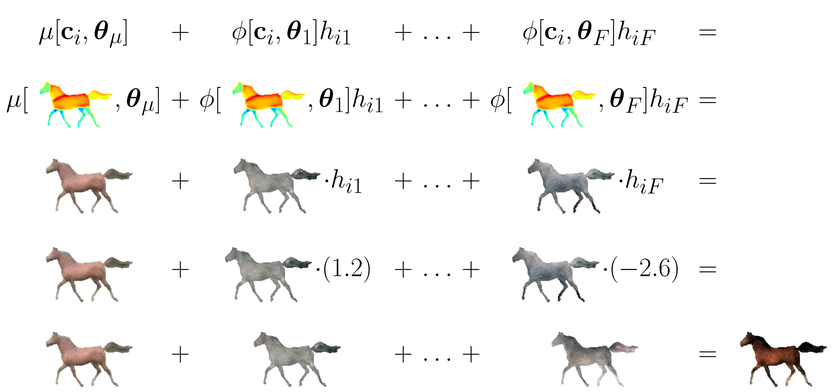
that the functions take the form of a pre-determined non-linear vector function $\mathbf{a}\left[\mathbf{c}_{ij}\right]$ of the context vectors which are linearly projected onto the parameter vector.

Inference





C-CCA of horses using 4 components



Extra Information

C-CCA Learning Procedure Requires: *RGB values* $\{\tilde{\mathbf{x}}_i^*\}$, *Context vectors* $\{\mathbf{c}_{ij}\}$

$\forall i : \mathbf{h}_i \leftarrow$ Random sample of $\text{Norm}_{\mathbf{h}_i}[0, \mathbf{I}]$

$\forall i : \mathbf{R}_i \leftarrow \mathbf{I};\ \mathbf{t}_i \leftarrow \vec{0}$

$\sigma^2 \leftarrow 1$

$\forall i, j : \mathbf{c}_{ij} \leftarrow (\mathbf{c}_{ij} - \texttt{mean}[\mathbf{c}]) \,/\, \texttt{std}[\mathbf{c}]$

$\{\mathbf{z}_m\}_{m=1}^{M} \leftarrow$ M random samples of $\{\mathbf{c}_{ij}\}$

$\forall i, j : \mathbf{a}[\mathbf{c}_{ij}] \leftarrow \texttt{kNN}\left[\mathbf{c}_{ij}, \{\mathbf{z}_m\}\right]$

for number of iterations **do**

$\forall i, j : \mathbf{x}_{ij}^* \leftarrow \mathbf{R}_i^{-1}(\tilde{\mathbf{x}}_{ij}^* - \mathbf{t}_i)$

$\boldsymbol{\theta} \leftarrow \text{argmax}_{\boldsymbol{\theta}_\star} \sum_{i=1}^{I} \log\left[Pr(\mathbf{x}_i|\mathbf{h}_i, \boldsymbol{\theta}_\star, \sigma^2)\right] = \text{argmax}_{\boldsymbol{\theta}_\star} \sum_{ij} \left(x_{ij} - \sum_{f=0}^{F} \mathbf{a}\left[\mathbf{c}_{ij}\right]^T \boldsymbol{\theta}_f h_{if}\right)^2$

$\sigma^2 \leftarrow \text{argmax}_{\sigma^2} \sum_{i=1}^{I} \log\left[Pr(\mathbf{x}_i|\mathbf{h}_i, \boldsymbol{\theta}_\star, \sigma^2) Pr(\mathbf{h}_i)\right] = \frac{1}{IJ} \sum_{i=1}^{I} (\mathbf{x}_i - \boldsymbol{\mu}_i - \boldsymbol{\Phi}_i \mathbf{h}_i)^T (\mathbf{x}_i - \boldsymbol{\mu}_i - \boldsymbol{\Phi}_i \mathbf{h}_i)$

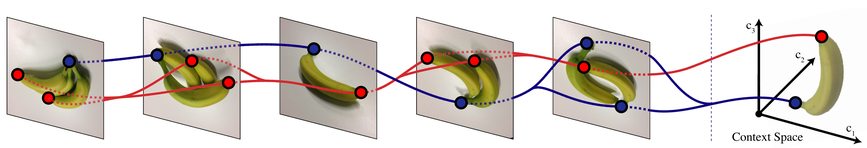
$\forall i \ : \ \text{argmax}_{\mathbf{h}_i} \log\left[Pr(\mathbf{x}_i|\mathbf{h}_i, \boldsymbol{\theta}_\star, \sigma^2) Pr(\mathbf{h}_i)\right] = (\boldsymbol{\Phi}_i^T \boldsymbol{\Phi}_i + \sigma^2 \mathbf{I})^{-1} \boldsymbol{\Phi}_i^T (\mathbf{x}_i - \boldsymbol{\mu}_i)$

$\mathbf{h}_i$

$\forall i : \mathbf{R}_i, \mathbf{t}_i \leftarrow \text{argmin}_{\mathbf{R}_i, \mathbf{t}_i} \sum_j \left(\tilde{\mathbf{x}}_{ij}^* - \mathbf{R}_i \mathbf{y}_{ij}^* - \mathbf{t}_i\right)^2$

end for

return $\{\mathbf{z}_m\}, \boldsymbol{\theta}, \{\mathbf{h}_i\}, \sigma^2, \{\mathbf{R}_i\}, \{\mathbf{t}_i\}$



				C-CCA					PPCA		
Dataset	I	Test I	F	K	M	Iter time (min)	Training Set Mean Error	Test Set Mean Error	M	Training Set Mean Error	Test Set Mean Error
Horses	200	95	24	16	2000	9	0.0258 ± 0.0079	0.0547 ± 0.0432	11459	0.0268 ± 0.0092	0.0886 ± 0.0661
Cats	450	117	16	16	4500	27	0.0280 ± 0.0119	0.0511 ± 0.0351	44015	0.0315 ± 0.0131	0.0548 ± 0.0371
Elephants	200	75	20	16	3000	10	0.0120 ± 0.0031	0.0346 ± 0.0163	11803	0.0169 ± 0.0048	0.0470 ± 0.0220
Facades	80	24	20	16	2000	9	0.0333 ± 0.0094	0.0593 ± 0.0264	23629	0.0350 ± 0.0198	0.0889 ± 0.0330

Parameters and Mean Per-pixel Reconstruction Error (mean squared error in RGB colorspace) for C-CCA and PPCA.

Results

Appearance Transfer

Subset of images from the training set

Reconstruction using fitted model

Reconstruction using the fitted model with fixed appearance parameters from the first (leftmost) image

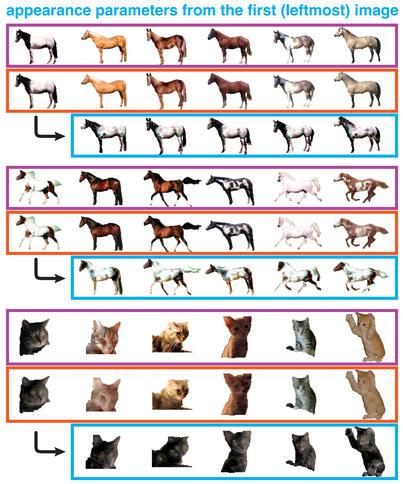
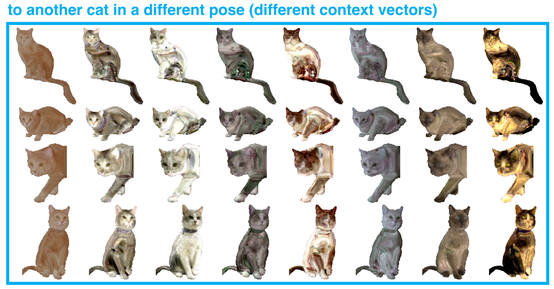


Image from the training set

Reconstruction using appearance parameters of the above cat transferred to another cat in a different pose (different context vectors)



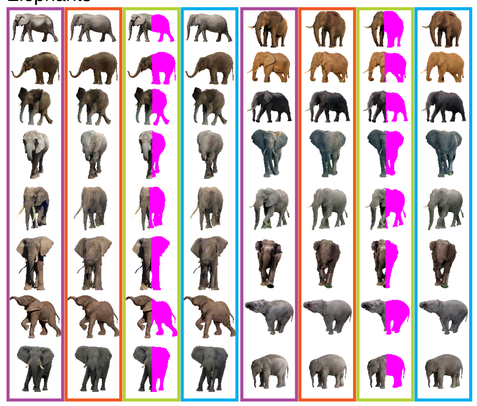
Structured Inpainting

Image from test set | Reconstruction using fitted model | Input for inpainting | Inpainted result

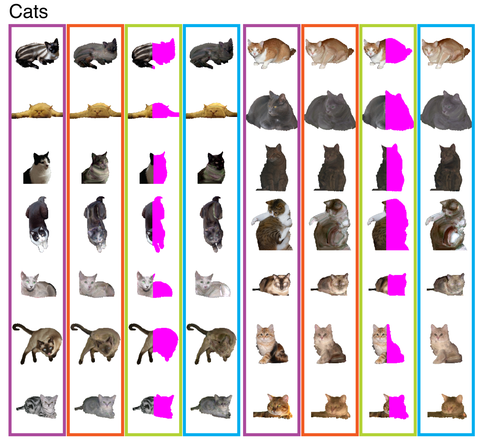
(context vectors known)

Horses

Elephants



Cats



Facades