Learning the Latent Look: Unsupervised Discovery of a Style-Coherent Embedding from Fashion Images University of Texas at Austin

Motivation

How to define similarity for fashion images?

Query













Proposed Method

- **Discover styles from** unlabeled images
- II. Style-coherent representation





1. Discovered styles













Localized

upper color white pattern plain

outer layer color orange color white pattern printed

Localized attribute recognition

background sunglasses skin hair boots T-shirt bag belt blazer blouse leggings pants shoes



Color segmentation segmentation

Style modeling by Polylingual Latent Dirichlet Allocation (PolyLDA)

attribute distribution $\varphi_{k} \sim \text{Dir}(\beta)$ style composition $\boldsymbol{\theta_i} \sim \text{Dir}(\alpha)$ style assignment $z_{ij} \sim \text{Multinomial}(\boldsymbol{\theta_i})$ observed attribute $x_{ij} \sim \text{Multinomial}(\varphi_{z_{ij}})$



Shared across body regions

LDA: decouples different regions

•	U			
opic 1	Topic 2		Тор	
hirt collar	L	skirt	Ο	deco
leco button	L	skirt short	Ο	patte
outtoned	L	skirt full	Ο	blaz
leco button	L	skirt pleat	U	butt
leeve long	L	skirt high-rise	U	shir
attern plaid	G	pattern plain	L	leng
attern plain	G	front pullover	L	shap
	G	deco button	G	deco
		_	G	pant







Approach

aggregate

lower

pants color blue pants pattern plain pants material denim pants length short

global

jacket pants color orange color blue color white pattern printed pattern plain material denim decoration belt





PolyLDA: across whole body

Topic 2

pic 1 o button tern plain zer toned rt collar gth long pe straight o button G jacket

O length short sleeve long pullover shirt collar color white skirt short skirt full pattern plain I length short G sweater



- U :upper body L :lower body
- O :outer layer III:hosiery
- G :global

Baselines:

Attribute: Indicator vectors

Discovered styles align with human defined style labels

Our representation retrieves most stylecoherent, diverse outfits



Given a gallery of photos ne Street Style Community CHICTOPIA











Wei-Lin Hsiao

Kristen Grauman



http://vision.cs.utexas.edu/projects/StyleEmbedding/

Experiments

StyleNet [Simo-Serra et al. 16]: clothing feature trained on Fashion144K **ResNet** [He et al. 16]: last layer of (ImageNet pretrained) ResNet

Attr-ResNet: ResNet-50 fine-tuned on our attributes



Summarize by dominant styles

