IDENTIFICATION OF FASHION TRENDS

Gunjan Batra

INTRODUCTION

- Fashion industry employs 4.2 million people and has a global annual revenue of 1.2 trillion dollars
- Over 2 million online blogs and stores devoted to fashion , millions of images readily available for computer vision applications
- Fashion is heavily driven by trends, similar aesthetics followed by people that appear/reappear cyclically e.g. Florals for spring and oranges for fall.
- Fashion trends have been analyzed by fashion designers and analysts but this is a novel topic in multimedia science.
- Fashion is an excellent domain for applying computer vision
- Understanding change in fashion is a good way to know society and time.
- New York City is one of the worlds fashion capitals and influences fashion trends.
- There is a need for fashion trends spotting!!!

APPROACHES PROPOSED

- Automatic Segmentation of Clothing for the Identification of Fashion Trends Using K-Means Clustering, 2013
- What are the fashion trends in New York? , 2014
- Runway to Realway : Visual Analysis of Fashion, 2015

AUTOMATIC SEGMENTATION OF CLOTHING FOR THE IDENTIFICATION OF FASHION TRENDS : METHODOLOGY

- This approach partitioned **individuals** from clothing, and also identified **individual pieces** of clothing.
- Combination of following algorithms:
  - Simple Segmentation Utilizing Global K-means Analysis
  - Segmentation through Combinatorial Computer Vision and Machine Learning Algorithms
  - Segmentation through Image-Specific Combinatorial Computer Vision and Machine Learning Algorithms
  - Segmentation using a combination of the algorithms
  - Clustering of Clothing Segments

AUTOMATIC SEGMENTATION OF CLOTHING FOR THE IDENTIFICATION OF FASHION TRENDS : DATASET

- **153 images** that were selected from a set of 2500 images of two large US clothing retail stores.
- Selected images had a **single individual** that displayed the face, torso and, at least, a portion of the leg and contained individuals wearing one to four pieces of clothing.
- **35 images were of men** and the remaining **118 were of women.**
- Image sizes ranged between 6 and 12 kilobytes, with approximate dimensions of 167 x 204.
- The types of clothes that individuals wore varied from **formal to casual**.

SEGMENTATION USING KMEANS

1. Images resized by factor of 2
2. Segmentation of Images performed using Global K Means Analysis (using Matlab) K=3
3. Segmented images subjected to face detection algorithm..face =layer of skin
4. Background and clothes layer were differentiated by determining range of pixel locations for each segment
5. Background , Clothes and Skin clearly segmented

Using this method, 102 of 153 images were segmented properly.

SEGMENTATION THROUGH COMBINATORIAL COMPUTER VISION AND MACHINE LEARNING ALGORITHMS

-Canny Edge Detection applied

-Dilation is applied

-Background mask is created for back projection

Face is located using Haar classification

Background and skin masks are subtracted from the original image, leaving behind the clothes.

Polygon approximation algorithm - To eliminates jagged contours along the clothing.

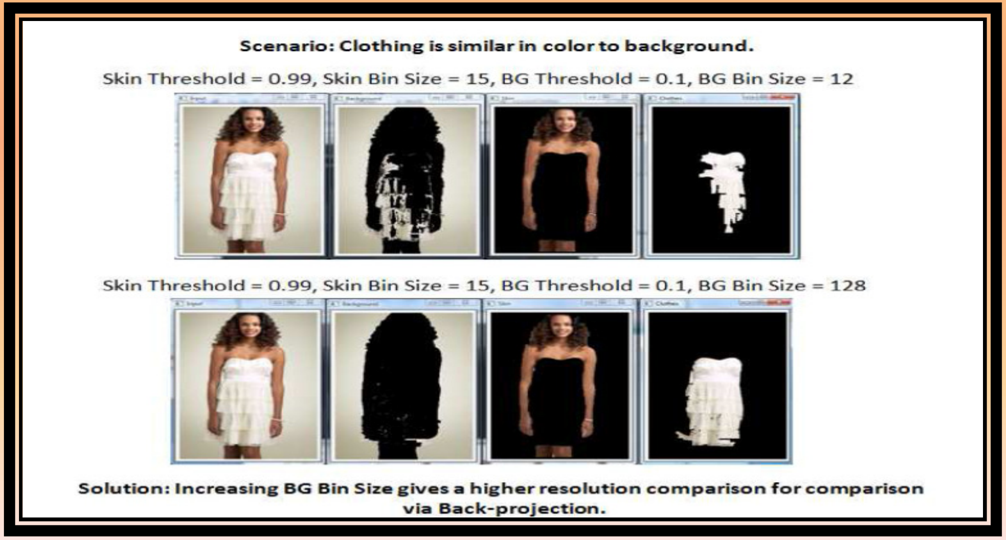
Images and masks are sent to the clustering algorithm.

Using this method, 75 of 153 images were segmented properly.

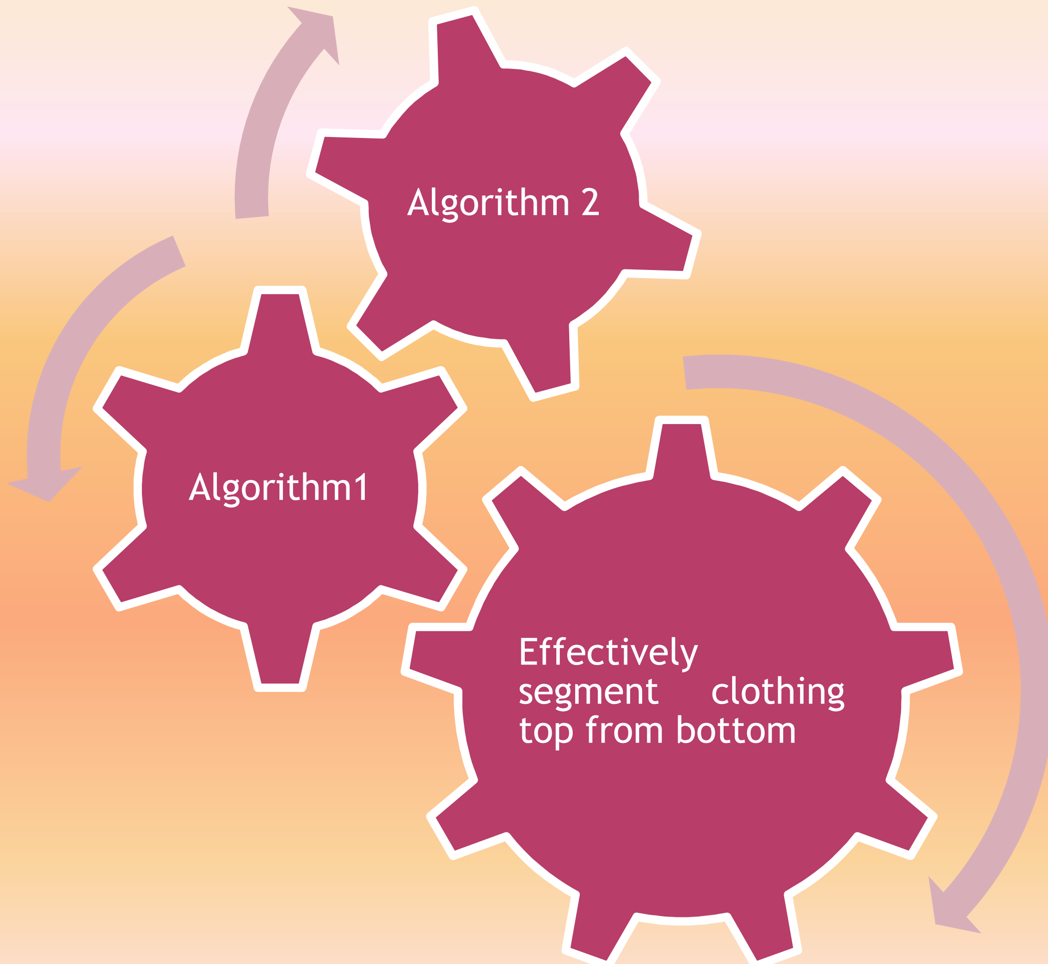
SEGMENTATION THROUGH IMAGE SPECIFIC COMBINATORIAL COMPUTER VISION AND MACHINE LEARNING ALGORITHM

5 critical values to improve image segmentation -

- Background probability threshold,
- Background histogram bin size,
- skin probability threshold,
- skin histogram bin size,
- background sample size.



SEGMENTATION USING A COMBINATION OF ABOVE ALGORITHMS



CLUSTERING OF CLOTHING SEGMENTS

The average standard deviation of the red, green, and blue channels was extracted to represent the complexity of the outfit.

R,G,B standard deviations were kept as separate entries in the feature vector

The algorithm could differentiate between different levels of complexity across the primary colors

H and S histograms were used as a feature vector to allow differentiation based on colors in the clothing.

The similarity clothing color (Represented as R, G, B) between tops and bottoms was used as a feature vector to understand contrasts.

Results from cluster analysis based on standard deviation of clothing using k=2

Figure 1- Low Complexity Clothing

Figure 2- Low Complexity Clothing

Results from cluster analysis based on standard deviation of color (histogram) using k=4

Figure 3- Dark Shades

Figure 4- Blue

Figure 5- Red

Figure 6-

Results from cluster analysis based on standard deviation of color (histogram) and color using k=4

Figure 7 – High Complexity & Dark

Figure 8 – Low Complexity & Red

Figure 9- Low Complexity & Light Shades

ANALYSIS/RESULT

- Automatic Image Segmentation
  - Global K –means yielded 67% accuracy
  - Combinatorial Computer vision and Machine Learning Algorithm was able to resolve clothing from 50% images.

    On adjustment of threshold and bin size it resolved all clothing
- Clustering Extracted Features successfully
- Fashion tends to steer in direction of simplicity
- Majority outfits tend to be more slim than wide

WHAT ARE THE FASHION TRENDS IN NEW YORK ? : METHODOLOGY

- The paper presents an algorithm that automatically discovers visual style elements representing fashion trends for a particular season with initial focus on New York Fashion Week (NYFW).
- The **catwalk model is detected** from the fashion show videos.
- With a combination of **Visual Style Element Initialization** and **Iterative Clustering**, patterns are identified that are *coherent*(frequently occurring within the fashion week event) and *unique* (for the fashion week).
- Color, Cut, head decorations and pattern are the four major elements under investigation

WHAT ARE THE FASHION TRENDS IN NEW YORK ? : DATASET

- Images were extracted from Fashion TV to build a fashion show image dataset of 3276 different catwalk models.
- These had been presented by nine fashion design companies for the past 10 seasons of NYFW

Season / Designer	Spring / Summer 2010	Autumn / Winter 2010	Spring / Summer 2011	Autumn / Winter 2011	Spring / Summer 2012	Autumn / Winter 2012	Spring / Summer 2013	Autumn / Winter 2013	Spring / Summer 2014	Autumn / Winter 2014
BCBG Max Azria	31	29	37	34	30	34	34	30	32	32
Calvin Klein	35	34	34	34	32	33	33	32	35	34
DKNY	46	40	37	34	13	40	48	40	47	42
Nicole Miller	39	37	37	34	39	36	39	44	39	39
Ralph Lauren	50	50	57	53	33	56	52	58	52	38
Rebecca Taylor	44	34	35	39	41	40	33	36	40	41
Richard Chai	26	27	29	28	26	28	27	26	31	27
Ruffian	33	34	36	32	30	31	13	32	31	23
Vera Wang	31	39	44	36	40	39	38	48	40	40
Total	335	324	346	324	284	337	317	346	347	316
	3,276									

Figure : No. of fashion styles in the dataset

CATWALK MODEL DETECTION

- This step identifies the catwalk model who demonstrates clothing and accessories on the stage at the fashion shows.
- The presence of faces is tracked over all video frames of a fashion show video of a particular designer who has exhibition at a season of NYFW
- It is detected if the faces belong to same person
- Also a constraint for frontal faces is used as models show their front faces and audiences show side faces.
- The detected faces are used to obtain models full body image(with clean background) using an algorithm that estimates the spatial layout of humans via pictorial structure model
- The set containing full body image of catwalk model is represented as $D_i = \{v_i^{1,1}, v_i^{1,2}, \ldots, v_i^{1,G}, v_i^{2,1}, v_i^{2,2}, \ldots v_i^{E,G}\}$
- Where $v_v^{w,x}$ denotes the fashion show video of designer x who has exhibition at NYWF season w, E is the no. of NYFW events and G is the number of designers

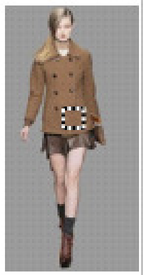
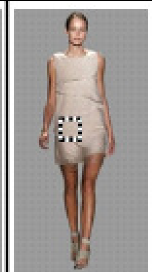
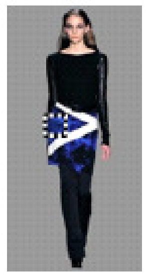
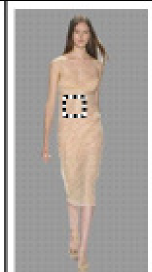
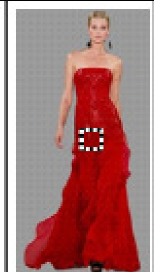
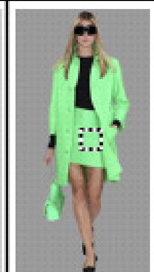
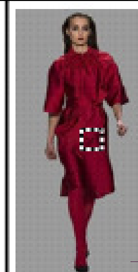
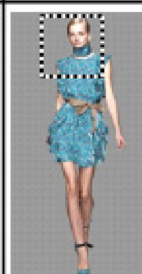
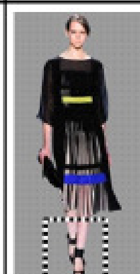
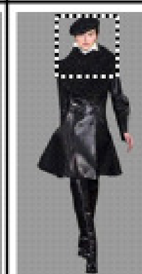
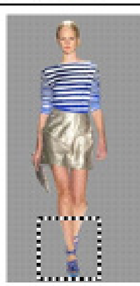
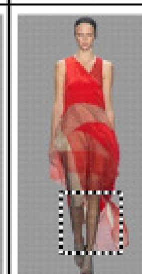
VISUAL STYLE ELEMENT INITIALIZATION

- The collection of full body image of models is divided into 2 sets
  - Positive Set, D^+ - Containing catwalk models from the selected fashion week event whose visual style elements of the fashion trends are to be discovered, e.g. catwalk models at NYFW S/S 2014
  - Negative Set, D^- - Containing catwalk models from other NYFW events, e.g. all catwalk models at NYFW except from NYFW S/S 2014).
- The full body images of catwalk model from both the sets are partitioned into non overlapping square patches.
- The 4 elements of fashion trends are captured using color and texture information.
  - Color – HSV color histogram, color correlogram, color moments
  - Texture – HOG, Haar wavelet transform, 2D Gabor wavelets
- Each patch is represented using 3 features
  - Color, texture and both color and texture
- Candidate patches that might represent the visual style elements are determined by finding out nearest neighbors of the patches in entire data set D_i ($D^{+ \text{ and}}$ D^-) and filtering out the ones who have more than 10 nearest neighbors in top 20 NN exist in D^+
- Standard k-means clustering is used to cluster similar style elements
- Groups of similar style elements are obtained as initial cluster

ITERATIVE CLUSTERING

- Performed on the initial visual elements obtained from the previous step to improve the stylistic coherence and uniqueness of patches.
- A linear SVM classifier is trained for each cluster, treating patches within the cluster as positive examples and all negative-set patches as negative examples.
- During iteration SVM faces the Overfitting problem.
- Solution - Cross Validation is applied at each round
- After final iteration, top ten detectors of SVM represent fashion trends at a certain season of NYFW

RESULTS

	S/S 2010	A/W 2010	S/S 2011	A/W 2011	S/S 2012	A/W 2012	S/S 2013	A/W 2013	S/S 2014	A/W 2014
Color Features										
Texture Features										
Color + Texture Features										

RESULTS

- The fashion trends identified are both coherent and unique
- Season S/S 2014 - Pastel green is identified as the most representative color trend for the investigated nine designers.
- Season A/W 2010 - High bun together with a neck scarf is identified as the most representative fashion trend in terms of texture information.
- Based on high average accuracy of fashion trend detector for each season, the framework is very effective
- NYFW with higher average accuracy has more coherent and unique fashion styles compared to fashion week with lower average accuracy

  E.g. the low average accuracy of A/W 2012 implies the fact that the designers had less fashion elements in common compared to the other seasons).
- The highest average accuracy among three types of features (color, texture, color+texture) for each fashion week shown in Table.

  From S/S 2010 to S/S 2013, the NYFW designers had put more concern in designing new textures rather than employing new color, while for S/S 2014 and A/W 2014 they had put more efforts in finding new colors and color schemes.

		Extracted Features		
		Color	Texture	Color + Texture
NY Fashion Week	Spring/Summer 2010	80.33 %	90.67 %	88.00 %
	Autumn/Winter 2010	64.67 %	91.00 %	88.67 %
	Spring/Summer 2011	87.00 %	92.00 %	92.33 %
	Autumn/Winter 2011	67.00 %	74.00 %	72.33 %
	Spring/Summer 2012	90.67 %	88.33 %	89.00 %
	Autumn/Winter 2012	50.33 %	67.67 %	60.33%
	Spring/Summer 2013	67.67 %	84.67 %	83.67 %
	Autumn/Winter 2013	66.33 %	65.67 %	68.00 %
	Spring/Summer 2014	94.67 %	90.67 %	92.33 %
	Autumn/Winter 2014	81.00 %	76.00 %	71.67 %

Figure : The average accuracy of fashion trend element detector

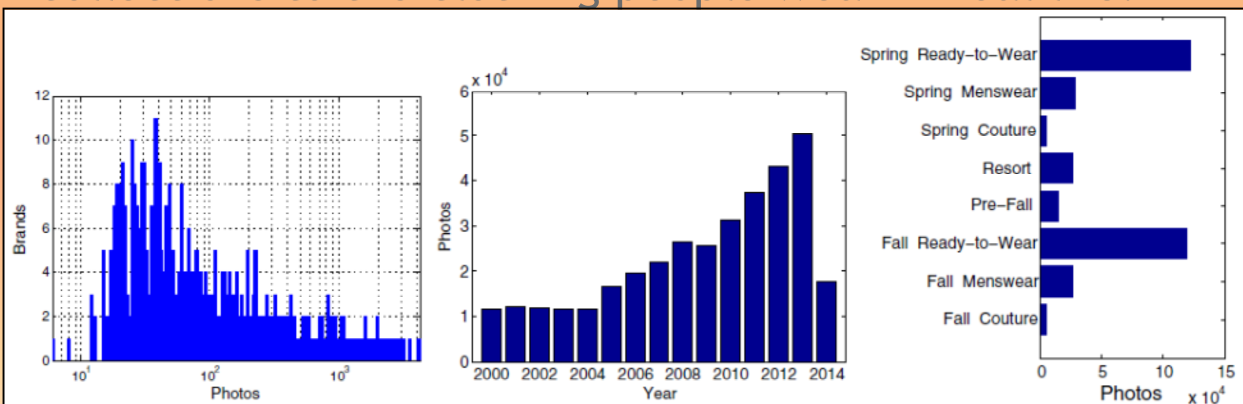
RUNWAY TO REALWAY : VISUAL ANALYSIS OF FASHION : METHODOLOGY

- Visual Representation
- Human Judgment
- Learning to compare outfits

RUNWAY TO REALWAY : VISUAL ANALYSIS OF FASHION : DATASET

- Runway Dataset –
  - Contains 348,598 runway fashion photos collected from *style.com*
  - 9,328 fashion show collections over 15 years from 2000 to 2014.
  - Collected together with the metadata which includes season, brand name, category, date of event, and a short text description.
- Paper Doll dataset –
  - Contains 339,797 images collected from the Chictopia social network.

    The combination of these two datasets allows us to make initial steps toward analyzing how fashion trends transfer from runway collections to the clothing people wear in real life.



(a) Photos per brand (b) Photos per year (c) Photos per season

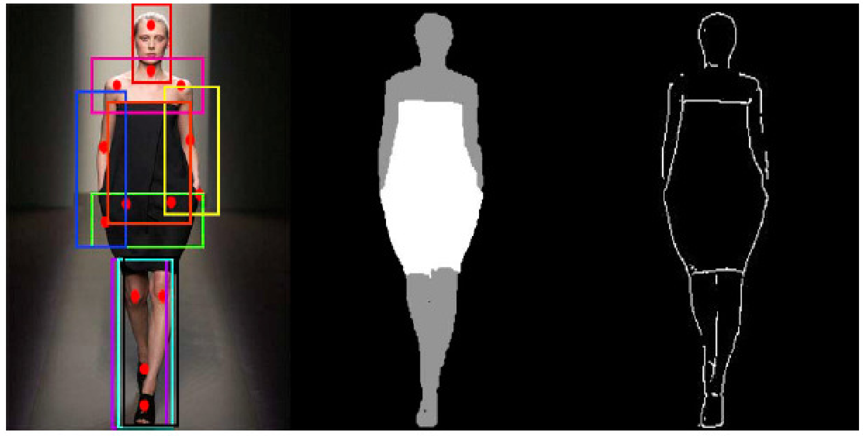
Figure 1 : Style dataset statistics

RUNWAY TO REALWAY : VISUAL ANALYSIS OF FASHION : DATASET

Figure 2 : Example images from our new Runway Dataset with estimated clothing parses.

VISUAL REPRESENTATION

Figure 4: (a) Pose estimation and corresponding bounding boxes (b) Foreground binary mask and (c) Binary edge maps



Preprocessing

Resize each person detection window to 320 × 160 pixels

Extract 9 sub-regions for the head, chest, torso, left/right arm, hip, left/right/between legs, as shown in Fig.

Measure Features : Color

For **each sub region :**

Two 512 dimensional histograms in RGB and Lab color spaces are extracted from the pixels that are parsed as foreground (clothes, skin, hair),

Measure Features : Texture

For each sub region : Two bag of words histograms are extracted from the pixels parsed as foreground. 1. Histogram of MR8 responses quantized into 256 visual words

2. Histogram of HOG descriptors quantized into 1000 words.

VISUAL REPRESENTATION

Measure Features: Shape

For each sub region, Two features are extracted : 1. Binary mask of the foreground pixels estimated by the clothing parsing algorithm

2. An edge map of the foreground region estimated using structured forest efficient edge detection.

Measure Features: Clothing Parse Estimates

For each sub region ,

1. The individual item masks are extracted for all the 56 different clothing categories (e.g. dress, shirt, shoes, etc)

2. 56-dimensional descriptor of the percentage of each item present in the sub-region is formed.

Measure Features: Global Style Descriptor

The style descriptor is computed over the entire body.

It also includes RGB, Lab, MR8, Gradients, HOG, Boundary Distance and Pose Distance.

HUMAN JUDGMENTS

- Amazon Mechanical Turk (MTurk) is used to collect human judgment of "similarity" between images of clothing outfits.
- Labelers were asked to select the most similar outfit from a set of 5 pictures, or none in case there were no similar outfits.
- The 5 outfits were selected basis the cosine similarity using each individual feature in isolation (e.g., color, texture), or to an equally weighted combination of all features.
- Human judgments were obtained for 2000 random query images from the runway dataset.
- The 5 similar candidates were chosen for each query under two scenarios; by selecting the images from the runway dataset, and selecting the images from the Paper Doll, Realway dataset. The similarity annotations were collected for each query image and scenario from 5 labelers.
- The majority of labelers agreed for 74.8% of the queries in the within-runway scenario and for 73.9% of the runway-to-realway queries.

LEARNING TO COMPARE OUTFITS

A linear SVM was trained to classify a pair of pictures as similar or dissimilar.

Three strategies of Majority, Unanimity and Some human judgments were converted to positive/negative labels for training.

- Majority meant that if an image-pair is marked as positive when the pair gets the majority of labeler clicks.
- Unanimity meant that any query-image pair for which all-five labelers agree on the best match are treated as positive.
- Some meant that image pairs marked by any of the five labelers are treated as positive.

Any query for which all the labelers clicked “none” is used to form 5 negative pairs with each of its 5 potentially similar images.

QUALITATIVE EVALUATION

- Towards, the left, the first image is the query image i.e. a runway outfit and the retrieved results (images 2-5) are also sampled from runway collections.
- Towards the right, the query image is a runway image and the images 2-5 images are the retrieved outfits from the Realway Chictopia dataset.
- In both cases the majority training approach is used.
- Outfits retrieved both from the runway and from the realway images look similar in terms of:
  - color
  - Shape
  - pattern
  - visual characteristics such as neckline shape

QUALITATIVE EVALUATION

Retrieved similar outfits for example query runway outfits(red boxes) using learned similarity

Runway to Runway

Runway to Realway

QUANTITATIVE EVALUATION

- Performance of the model is evaluated by area under the precision recall curve (AUC)
- Learned Similarity models agree with human similarity judgments quite a lot
- The results show 73-76% AUC in Runway to Runway and 53-55% AUC in Runway to Realway
- Also performed comparison with Style descriptor – Increase in performance of 10%
  - 62-66%AUC in Runway to Runway
  - 42-45%AUC in Runway to Realway

AUC for predicting outfit similarity

	Runway to Runway		Runway to Realway	
Method	our feature	Style descriptor	our feature	Style descriptor
Majority	0.76 ± 0.11	0.66 ± 0.11	0.54 ± 0.03	0.45 ± 0.02
Unanimity	0.73 ± 0.08	0.62 ± 0.07	0.53 ± 0.02	0.42 ± 0.01
Some	0.73 ± 0.14	0.63 ± 0.12	0.55 ± 0.01	0.43 ± 0.03

COMPARISON / CONCLUSION

	Automatic Segmentation of Clothing for Identification of Fashion Trends	What are the Fashion Trends in New York ?	Runway to Realway : Visual Analysis of Fashion
Year of Publication	2013	2014	2015
Outcome	Clustering based on color complexity and outfit shape	Algorithm for fashion trend detection	Approach to studying fashion on runway and real world setting
Classification/ Clustering Used	K-Means Clustering	K-Means Clustering Simple Vector Machine	Simple Vector Machine
Data Set	153 images from 2 large US clothing Retail store	3276 images extracted from Fashion TV	348,598 Runway fashion photos
Feature Measures Used	Standard deviation of clothing, Standard deviation of color(histogram), Std dev of color(histogram) and color	Color, Texture, Color+Texture	Color, Texture, Shape, Parse, Style Descriptor

FUTURE WORK

- Similarity models to mine large datasets for similar styles/ trends and how street fashion is influenced by runway styles
- Develop clothing recommender systems
  - To pair clothing,
  - For events
  - Product suggestions
- Study of Fashion Trends on the social network in real time

THANK YOU!

QUESTIONS?