Ten-million-order Human Database for World-wide Fashion Culture Analysis

Hirokatsu Kataoka, Kaori Abe*, Munetaka Minoguchi*, Akio Nakamura*, Yutaka Satoh
AIST, Tokyo Denki University (TDU)*

http://xpaperchallenge.org/cv
Dynamic Fashion Cultures in the world cities?

What kind of FashionDB do we need?

- Huge-scale fashion database
- GPS/time stamp are corresponded in a fashion snap
- Less noise with data refinement

The concept of dynamic fashion cultures
Fashion Culture DataBase (FCDB)

25M Fashion DB w/ geo-tagged and time-stamp

- Based on YFCC100M*
- On of the largest DB in context of fashion analysis (to our knowledge)

YFCC100M* = Yahoo Flickr Creative Commons 100M database [Thomee ACM2016]
Contributions

Discovering world-wide fashion trends w/ FCDB

(1) Fashion Culture DataBase (FCDB), which contains and is refined 25M images on Flickr

- Semi-automatic dataset collection with existing detector
- Data refinement with binary classification

(2) As the perspective of huge DB, we conduct inter-city similarity and temporal fashion trends in a simple way

- Bag-of-words (BoF) + StyleNet Vec (128-dim)
- Temporal subtraction between two consecutive BoF vectors
How to construct FCDB (1st step)

1) Collect images
   1)-1: Collect images on YFCC100M
   1)-2: Images are taken around 16 cities by GPS
2) Put bboxes on street fashion snaps
   2)-1: Human detection by Faster R-CNN
   2)-2: Crop the detected human regions

Result in the first step:

8,504,037 original images
76,532,519 clothing images
FCDB Refinement (2\textsuperscript{nd} step)

Human (fashion snap) or noise with binary classifier

1) Training configuration
   - Classifier by StyleNet feature and SVM
   - Train / Test : 2,886 / 2,886 (Total 5,772 images annotated by humans)

2) Refinement-by-Classification
   2)-1: Apply to all images in FCDB
   2)-2: Noise images are removed

Result in the second step:
- 8,504,037 original images
- 76,532,519 fashion snaps
- 1,981,812 original images
- 25,707,690 fashion snaps
FCDB Refinement (2nd step)

Human (fashion/snap)
or noise with binary classifier

1) Training configuration
   - Classifier by StyleNet feature and SVM
   - Train/Test: 2,886/2,886 (Total 5,772 images annotated by humans)

2) Refinement - by Classification
2) 1: Apply to all images in FCDB
2) 2: Noise images are removed

Result in the second step:
- 8,504,037 original images
- 76,532,519 fashion snaps
- 1,981,812 original images
- 25,707,690 fashion snaps

Noise images (Non-human/miss detection)
Human images (Fashion snaps)

Before refinement: noise images are included
After refinement: less noise images
Related fashion-oriented databases

- Ten-million-order images on SNS
- More diverse
  - Geo-tagged + Time-stamped images
  - Taken from daily life, rather than no-background

<table>
<thead>
<tr>
<th>Database</th>
<th>#images</th>
<th>#cate.</th>
<th>GPS?</th>
<th>Box?</th>
<th>Time Stamp?</th>
</tr>
</thead>
<tbody>
<tr>
<td>HipsterWars</td>
<td>1,893</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fashionista</td>
<td>158,235</td>
<td>53</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Fashion144k</td>
<td>144,169</td>
<td>N/A</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fashion14</td>
<td>13,126</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DeepFashion</td>
<td>800,000</td>
<td>1,050</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>FCDB (ours)</td>
<td>25,707,690</td>
<td>16</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Against to the previous databases, our database contains street snaps
City-scale fashion representation (1/2)

Fashion Style Distribution (FSD): aggregated fashion distribution at each city

1) Extracting StyleNet [Simo-Serra+, 2016] which has 128-dim vector
2) Feature quantization by Bag-of-Words (BoW)
   Step2-1: k-means clustering: the centroids define pseudo fashion style
   Step2-2: Vote for the most similar fashion style by distance from representative vectors of fashion styles

City-scale fashion representation (2/2)

Fashion Style Distribution (FSD): aggregated fashion distribution at each city

1) Extracting StyleNet [Simo-Serra+, 2016] which has 128-dim vector
2) Feature quantization by Bag-of-Words (BoW)
   Step2-1: k-means clustering: the centroids define pseudo fashion style
   Step2-2: Vote for the most similar fashion style by distance from representative vectors of fashion styles

Tokyo fashion trends

Paris fashion trends

Visual Words ⇔ Spontaneous Fashion Styles

Tokyo

Paris

Fashion-based city similarity graph with FSD

Nodes/edges show how much similar between cities

- A thick line indicates high similarity
- The fashion-based city similarity is affected by culture
Fashion-based city similarity graph with FSD

The graph shows how much similar between 2 cities

- A thick line indicates high similarity
- The fashion-based city similarity is affected by culture

North American countries are correlated
Fashion-based city similarity graph with FSD

The graph shows how much similar between 2 cities:

- A thick line indicates high similarity.
- The fashion-based city similarity is affected by culture.

NYC is similar to London & Barcelona.
Fashion-based city similarity graph with FSD

The graph shows how much similar between 2 cities

- A thick line indicates high similarity
- The fashion-based city similarity is affected by culture

Is Tokyo a unique city in the world?
Fashion-based city similarity graph with FSD

Randomly sampled images in Tokyo

Costume play? Character show?

New Delhi   Barcelona   London
San Francisco
Fashion-based city similarity graph with FSD

Randomly sampled images in Tokyo

Costume play? Character show?

OMG, is this fashion culture?
**Time series representation**

- **Fashion Trends Descriptor (FTD)**: temporal fashion trends detector

  Subtraction of temporally consecutive FSDs

  Increased/stable/decreased fashions can be detected

\[
\begin{align*}
\Delta v^t &= v^t - v^{t-1} \\
F^+(t) &= 4 \quad 5 \quad 7 \\
F^0(t) &= 1 \quad 6 \quad 9 \\
F^-(t) &= 2 \quad 3 \quad 8 \\
\end{align*}
\]

Avg. = \(TH^+\)  \(= TH^-\)
Yearly Fashion Trends Visualization

- Analysis of dynamic fashion cultures
- Yearly analysis in 2000 - 2015
- Visualizing appeared fashion styles

- **Boston**
  - **Sportswear** keeps *increasing since 2010*
    → High interest in sports
  - **Hockey** team won the Stanley Cup *in 2011*

- **Tokyo**
  - **Costumes** show *increase since 2011*
    → High interest in Japanese subculture

The result suggests that user's interest in the area, to detect a hot event
Yearly Fashion Trends Visualization

- Analysis of dynamic fashion cultures
- Yearly analysis in 2000 - 2015
- Visualizing appeared fashion styles
  - Boston
    - Sportswear keeps increasing since 2010 → High interest in sports
    - Hockey team won the Stanley Cup in 2011
  - Tokyo
    - Costumes show increase since 2011 → High interest in Japanese subculture

The result suggests that user's interest in the area, to detect a hot event
Yearly Fashion Trends Visualization

- Analysis of dynamic fashion cultures
- Yearly analysis in 2000 - 2015
- Visualizing appeared fashion styles

- Boston
  - Sportswear keeps increasing since 2010
    → High interest in sports
  - Hockey team won the Stanley Cup in 2011
- Tokyo
  - Costumes show increase since 2011
    → High interest in Japanese subculture

The character is newly appeared in 2013.

The result suggests that user's interest in the area, to detect a hot event.
Yearly Fashion Trends Visualization

• Analysis of dynamic fashion cultures
• Yearly analysis in 2000 - 2015
• Visualizing appeared fashion styles

➢ Boston
  ✓ Sportswear keeps increasing since 2010
    → High interest in sports
  ✓ Hockey team won the Stanley Cup in 2011

➢ Tokyo
  ✓ Costumes show increase since 2011
    → High interest in Japanese subculture

The character is newly appeared in 2013.

The result suggests that user's interest in the area, to detect a hot event
Summary

Proposal of **large-scale database** contributing to **fashion trends** analysis

1. **Proposal of Fashion Culture Database**

2. **fashion trends analysis method**
   - **Representation** of fashion trends
   - **Time series representation** of fashion trends

3. **Spatiotemporal analysis** of fashion trends
   - **Regionality** of fashion trends
   - **Temporal change** in clothing prevalence
Toward release database!?

We’re struggling to release the data!

• Commercial-free images are being collected
• Building a set focused on clear images