Hi, Magic Closet, Tell me what to wear

Si LIU, Jiashi FENG, Zheng SONG, Tianzhu ZHANG, Changsheng XU, Hanqing LU, Shuicheng YAN
National University of Singapore
Institute of Automation, Chinese Academy of Science
Advanced Digital Science Center

HEADACHED QUESTIONS FOR FEMALE

WHAT TO WEAR

I am going on a conference, what should I wear?

MAGIC CLOSET: CLOTHING SUGGESTION

MAGIC CLOSET – OCCASION ORIENTED CLOTHING RECOMMENDATION SYSTEM

HOW TO PAIR?
MAGIC CLOSET: CLOTHING PAIRING

RECOMMENDATION CRITERION

- Criterion 1: Wear Properly (occasion-oriented)
  - Occasion1: Wedding
  - Occasion2: Funeral

WHAT TO WEAR (WoW) DATASET

CLOTHING ATTRIBUTES

- Clothing Category Attributes:
  - Upper body
  - Lower body
  - Set
  - Outwear/Casual
  - Fashion/Hoodies

- Detailed Attribute:

DETAILED ATTRIBUTES

- We manually define 6 kinds of clothing attributes based on a comprehensive study of many online shopping websites.
HUMAN POSE ESTIMATION
- Use the annotated key points in human photos
- Train one human upper-body detector and one human lower-body detector[1]

LATENT SVM RECOMMENDATION MODEL POTENTIAL FUNCTION
- Potential Function

\[ \phi(x, a) = \sum_{o \in \mathcal{O}} \psi(x, o) \]

- Feature vs. Occasion Potential:
  \[ \psi(x, o) = \frac{1}{1 + e^{-\beta \cdot \phi(x, o)}} \]

- Attribute vs. Occasion Potential:
  \[ \psi(a, o) = \frac{1}{1 + e^{-\beta \cdot \phi(a, o)}} \]

LATENT SVM RECOMMENDATION MODEL FORMULATION
- Formulation

\[ \min_{w, \phi} \frac{1}{2} \|w\|^2 + \frac{1}{N} \sum_{i=1}^N \ell(y_i, \hat{y}_i) \]

s.t. \[ \max_{a_{k-1}} \max_{a_k} w^T \Phi(x, a_{k-1}, a_k) - \max_{a_{k-1}} \max_{a_k} \Phi(x, a_{k-1}, a_k) \geq \Delta_0(\phi(a_{k-1}, a_k)) - \ell(\phi(a_{k-1}, a_k)), y_{i}, y_{i} \in \mathcal{C} \]

where \[ \Delta_0(\phi(a_{k-1}, a_k)) = \begin{cases} 1 & \text{if } \phi(a_{k-1}, a_k) > 0 \\ 0 & \text{otherwise} \end{cases} \]

LATENT SVM RECOMMENDATION MODEL INFERENCE
- Given the parameter model \( w \), we need to solve the following inference problem during recommendation:

\[ \{a^*_j, a^*_k\} = \arg \max_{a_j, a_k} w^T \Phi(x, a_j, a_k, o) \]

- Clothing Suggestion Criterion:

\[ \hat{x}^* = \arg \max_{x} \left\{ \max_{a_j, a_k} w^T \Phi(x, a_j, a_k, o) \right\} \]

Given an occasion, we recommend the clothing pair with the largest potential value.

LATENT SVM RECOMMENDATION MODEL INFERENCE
- Clothing Pairing Criterion:

\[ x^* = \arg \max_{x} \left\{ \max_{a_j, a_k} w^T \Phi(x, a_j, a_k, o) \right\} \]

Given an occasion and a reference coat \( x_0 \), we recommend the \( x \) corresponding trousers/skirts with the largest potential value.
RECOMMENDATION MODEL ANALYSIS: Attribute Estimation

(1) Occasion-Attribute Matching Rules
- Wedding
- Sleeveless
- Silk
- Dress

The top-3 most representative attributes:
- Funereal
- Black
- Plain
- Long

Wear Properly

EXPERIMENTAL SETTING
- Evaluate Criterion: Normalized Discounted Cumulative Gain (NDCG):
  - Feature-occasion SVM: predicting occasion from visual features directly without considering attributes.
  - Feature-attribute-occasion SVM: the 1st layer SVM linearly maps visual features to attribute values; the 2nd SVM is trained on attribute confidence vectors to predict occasion labels.

EXP1: Clothing Suggestion Results
- The feature-occasion SVM consistently outperforms the feature-attribute-occasion SVM.
- The proposed latent SVM model outperforms the two baseline models significantly.
EXP1: CLOTHING SUGGESTION EXEMPLAR RESULTS

EXP2: CLOTHING PAIRING EXEMPLAR RESULTS

EXP2: CLOTHING PAIRING RESULTS

RELATED WORK – ACADEMIC

RELATED WORK – INDUSTRY

CONCLUSION: Magic Closet

The feature-attribute-occasion SVM significantly performs better than the feature-occasion SVM (= SVM_U + SVM_L).

The proposed latent SVM model outperforms the two baseline models significantly.

Few researches in the clothing related study:
- Street-to-shop clothing retrieval system[1]: Snap a street fashion, and search similar clothing from online shops.
- Clothing Parsing[2]: Given an image, segment it into several semantic regions, such as trouser, shirts, etc.

We are the only work focusing on clothing recommendation task.


Many fashion sharing websites, where girls learn how to wear aesthetically by looking at others.

The first to explore two applications: clothing suggestion and clothing pairing.

Construct a large dataset, called What-to-Wear (WoW), containing 24,417 clothing with complete attribute and occasion annotations.

Apply a latent SVM based framework to learn the occasion-oriented clothing recommendation model and simultaneously mine the clothing matching rules.

Great application potential in social media (from social media, and used for social media).

The feature-attribute-occasion SVM significantly performs better than the feature-occasion SVM (= SVM_U + SVM_L).

The proposed latent SVM model outperforms the two baseline models significantly.

Few researches in the clothing related study:
- Street-to-shop clothing retrieval system[1]: Snap a street fashion, and search similar clothing from online shops.
- Clothing Parsing[2]: Given an image, segment it into several semantic regions, such as trouser, shirts, etc.

We are the only work focusing on clothing recommendation task.


Many fashion sharing websites, where girls learn how to wear aesthetically by looking at others.

The first to explore two applications: clothing suggestion and clothing pairing.

Construct a large dataset, called What-to-Wear (WoW), containing 24,417 clothing with complete attribute and occasion annotations.

Apply a latent SVM based framework to learn the occasion-oriented clothing recommendation model and simultaneously mine the clothing matching rules.

Great application potential in social media (from social media, and used for social media).
FUTURE DIRECTIONS

- Inaccurate body detection and parsing/segmentation
- Some attribute predictions are not reliable yet very important, such as materials
- Personalized clothing recommendation
- Complex clothing for Winter
- Beyond clothing: other wearing and carrying

THANK YOU & QA