Modeling the Intransitive Pairwise Image Preference from Multiple Angles

Jun Chen, Chaokun Wang, Jianmin Wang

School of Software, Tsinghua University, Beijing 100084, P.R. China

Motivations

- Intransitive preference exists in pairwise comparisons.
- Images can be compared from more than one angle.
- Inconsistent image rankings from different angles.

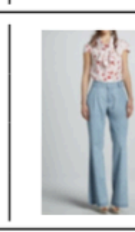
Contributions

- Multi-angle preference (MAP) models.
- Balance between the most/least liked angles.
- User/Image feature representations in multiple angles.

Intransitive Preference with Multiple Angles

- User rates one image differently under multiple (latent) angles.
- User compares two images according to ***the joint set of the most/least liked angles (highest/lowest rating)*** of the two images.
- Comparing image 1 and 2 with average rating, she uses the joint set of the most liked angles {TH, SO}. She rates image 1 as (3+5)/2=4, and rates image 2 as (5+2)/2=3.5, showing she prefers image 1 to 2.
- Similarly, in other pairwise comparisons, it has 1>2, 2>3, but 3>1.
- Due to different set of compared angles, intransitive preference happens.

Table 1: A user's angle-wise ratings on different images. *r* denotes the 5-scale rating. Images ①–③ are selected from [15]. Images ④–⑥ are selected from [11]. Abbreviations: CL-colors, SO-salience object, TH-theme, ST-sentiment, BS-body shape, TC-top clothing, BC-bottom clothing, SL-style.

No	Image	Angle	r	No	Image	Angle	r
①		CL SO TH ST	2 3 5 4	④		BS TC BC SL	1 5 3 4
②		CL SO TH ST	3 5 2 1	⑤		BS TC BC SL	1 4 5 1
③		CL SO TH ST	5 2 4 3	⑥		BS TC BC SL	3 3 3 5

Multi-Angle Preference (MAP) Models

- Represent each image/user with a latent feature matrix
- Define u's preference on i with row-wise inner product $\mathbf{p}_{uv_i} = \mathbf{U}_u \odot \mathbf{V}_i \in \mathcal{R}^D$
- The family of MAP models (argFmax/argFmin is the function to fetch the indices w.r.t the F max/min values in the vector):

Most-like rule model

$$r^{ML}_{v_i >_u v_j} = \frac{1}{F}\Big(\sum_{x \in \text{argFmax}\, \mathbf{p}_{uv_i}} (\mathbf{p}_{uv_i}[x] - \mathbf{p}_{uv_j}[x]) - \sum_{y \in \text{argFmax}\, \mathbf{p}_{uv_j}} (\mathbf{p}_{uv_j}[y] - \mathbf{p}_{uv_i}[y]) \Big),$$

Least-like rule model

$$r^{LL}_{v_i >_u v_j} = \frac{1}{F}\Big(\sum_{x \in \text{argFmin}\, \mathbf{p}_{uv_i}} (\mathbf{p}_{uv_i}[x] - \mathbf{p}_{uv_j}[x]) - \sum_{y \in \text{argFmin}\, \mathbf{p}_{uv_j}} (\mathbf{p}_{uv_j}[y] - \mathbf{p}_{uv_i}[y]) \Big).$$

Two-side balance rule model

$$r^{TS}_{v_i >_u v_j} = \eta r^{ML}_{v_i >_u v_j} + (1-\eta) r^{LL}_{v_i >_u v_j}, (0 \le \eta \le 1)$$

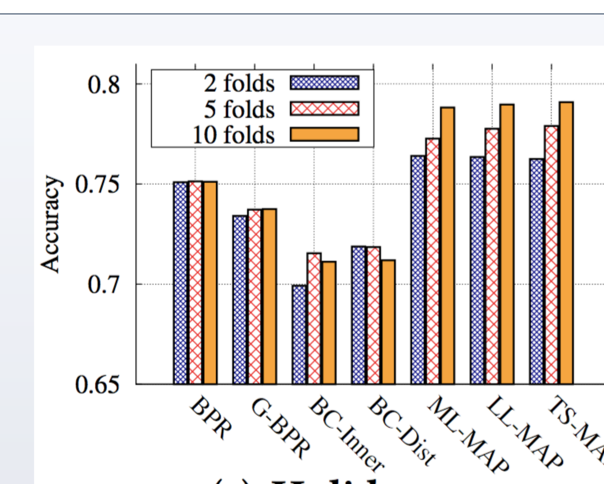
- Learning **U**, **V** with SGD: $\underset{\Theta}{\text{argmin}}\, L = -\sum_{(u, v_i, v_j) \in \mathcal{D}} \ln \sigma(r^{\diamond}_{v_i >_u v_j}) + \frac{\lambda}{2}\Theta^2$

Learning algorithm

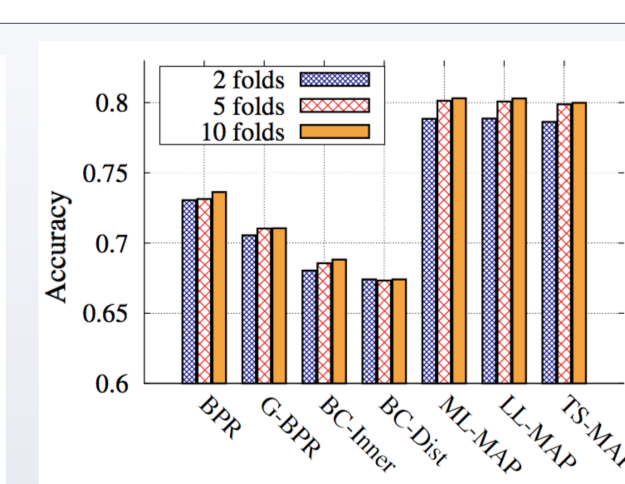
```
initialize U, V ~ Gaussian(0, λI).
for i ∈ {1, .., T} do
  shuffle D
  for (u, a, b) ∈ D do
    compute argFmax p_{uv_i}, argFmax p_{uv_j}.
    compute argFmin p_{uv_i}, argFmin p_{uv_j}.
    compute r^{ML}_{v_i >_u v_j} using Eq. (2).
    compute r^{LL}_{v_i >_u v_j} using Eq. (3).
    compute r^{TS}_{v_i >_u v_j} using Eq. (4).
    for x ∈ {1, ..., D} do
      u ← (1 − αλ)U_u[x,:] + α(1 − σ(r^{TS}_{v_i >_u v_j})) ∂r^{TS}_{v_i >_u v_j} / ∂U_u[x,:]
      v_i ← (1 − αλ)V_i[x,:] + α(1 − σ(r^{TS}_{v_i >_u v_j})) ∂r^{TS}_{v_i >_u v_j} / ∂V_i[x,:]
      v_j ← (1 − αλ)V_j[x,:] + α(1 − σ(r^{TS}_{v_i >_u v_j})) ∂r^{TS}_{v_i >_u v_j} / ∂V_j[x,:]
    end for
    U_u[x,:], V_i[x,:], V_j[x,:] ← u, v_i, v_j
  end for
end for
return U, V
```

Experiments

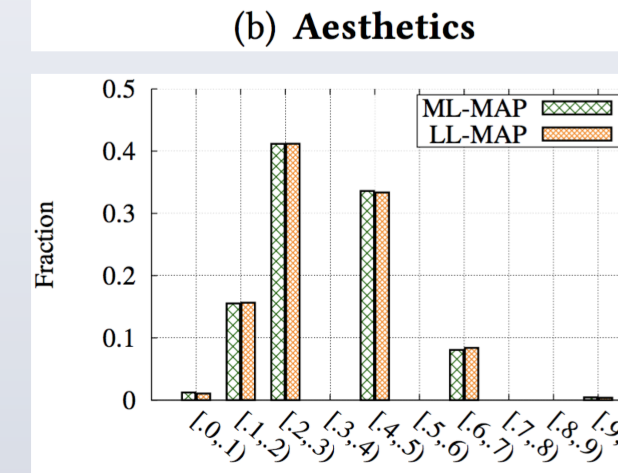
- Datasets: Holidays (scenery images), Aesthetics (fashion images).
- The accuracy of correct prediction of pairwise image comparison.
- Superior accuracy performance of MAP models versus baselines.
- The models of intransitive reciprocal relations (BC-Inner, BC-Dist) perform badly on datasets with not enough intransitive relations.
- The most/least liked angles for different images are usually different since most of their overlap is less than 0.5. (That is a reason why personalization is important)



(a) **Holidays**



(b) **Aesthetics**



ACM International Conference on Multimedia (MM-17, Mountain View, USA).
Contact: chenjun082@gmail.com, chaokun@tsinghua.edu.cn.