Visual Perception Similarities to Improve the Quality of User Cold Start Recommendations

Crícia Z. Felício[1,3], Claudianne M. M. de Almeida[2,3], Guilherme Alves[3], Fabíola S. F. Pereira[3], Klérisson V. R. Paixão[3], and Sandra de Amo[3]

[1] Federal Institute of Triângulo Mineiro, Brazil
cricia@iftm.edu.br
[2] Federal Institute of Norte de Minas, Brazil
claudianne.almeida@ifnmg.edu.br
[3] Federal University of Uberlândia, Brazil
guilhermealves@mestrado.ufu.br, fabfernandes@comp.ufu.br,
klerisson@doutorado.ufu.br, deamo@ufu.br

Abstract. Recommender systems are well-know for taking advantage of available personal data to provide us information that best fit our interests. However, even after the explosion of social media on the web, hence personal information, we are still facing new users without any information. This problem is known as user cold start and is one of the most challenging problems in this field. We propose a novel approach, VP-Similarity, based on human visual attention for addressing this problem. Our algorithm computes visual perception's similarities among users to build a visual perception network. Then, this networked information is provided to recommender system to generate recommendations. Experimental results validated that VP-Similarity achieves high-quality ranking results for user cold start recommendation.

Keywords: Visual perception, Cold Start, Recommender system

1 Introduction

Traditional recommenders usually take advantage of users' previous choices and ratings to predict products (items) that would fulfill their wishes. When it comes to new users without rating records, however, the performance of these approaches fall a great deal. This is known as user cold start and it is one of the most challenging problems in this field [1]. Remarkably, with the emerging of social networks, social information has been largely explored to mitigate cold start problem in item recommendation [2–4]. However, it is important to notice that social information is not always available or it is costly to be accessed.

An emerging interesting source of information is *human visual perception*. Tracking users eyes movements to capture users' behaviors has been becoming increasingly tangible. Herein, we aim to explore the visual perception as an

* The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-34111-8_13.

important source of knowledge to address the user cold start problem. As a motivation example, let us imagine the domain of painting's recommendation. Consider a painting containing two main scenes, one describing a cat and other describing a child. Some people, looking at the painting, will focus their attention to the cat (U_{cat}). Others, to the child (U_{child}). Note that, in each group, there are people who like the painting and also the ones who do not like it. The reasons for which some people in U_{cat} or U_{child} do not like the painting are different, but we can affirm that people in U_{cat} (or U_{child}) are similar given their visual perception focuses. Now, imagine a new user u in the system, that has never expressed his preferences or tastes about paintings – u is a cold start user. Let us suppose that u focused on the cat. Then, we could assume that u has a lot in common with people from U_{cat} as regarding the way they perceive paintings. Thus, we hypothesize that it is reasonable to use people's preferences from U_{cat} to provide a recommendation to our new user.

In this work, we propose a method to identify similarities among users based on visual perception. We take advantage of the volume of existent approaches that address cold start problem through social network information [4, 3, 5] and innovate, by proposing to use a *Visual Perception Network*, inferred from the similarities on users visual perception. We evaluate our approach against the innovative Paintings dataset[4].

2 VP-Similarity Method

To adopt visual perception as contextual information for recommendation systems, first, we define the *VP-Similarity Method*. It is responsible to handle visual perception and infer visual perception similarities among users. Then, we incorporate visual perception network into social recommender systems and evaluate the performance in cold start scenario.

Definition 1 (Visual Fixation). *Let $\mathcal{I} = \{\mathcal{I}_1, ..., \mathcal{I}_m\}$ be a set of images. Let $\mathcal{U} = \{u_1, ..., u_n\}$ be a set of users. A visual fixation of an user u_j over an image $\mathcal{I}_k$ is a pair (p, f) where p is the position, represented by the pixels cluster centroid of that fixation, and f is the duration. We denominate $\mathcal{F}_{jk} = \{(p_1, f_1), ..., (p_z, f_z)\}$ the set of visual fixations of u_j over $\mathcal{I}_k$ (Fig. 1).*

Definition 2 (Visual Perception). *Let the images in $\mathcal{I}$ be divided in r equal parts $Q = \{q_1, ..., q_r\}$ as illustrated in Fig. 2. From the positions and durations described in the set of visual fixations $\mathcal{F}_{jk}$, we call v_s the percentage of time that u_j fixed to $\mathcal{I}_k$ in each quadrant q_s, for $1 \leq s \leq r$. The visual perception of an user u_j over an image $\mathcal{I}_k$ is defined as the vector $\mathcal{P}_{jk} = (v_1, ..., v_r)$. Finally, the visual perception of u_j over all images $\mathcal{I}$ is represented by the concatenation of all visual perceptions vectors from u_j: $\mathcal{P}_j = \mathcal{P}_{j1} \parallel ... \parallel \mathcal{P}_{jm}$.*

Table 1 shows an example of visual perception. There are visual perceptions from 7 users over 2 images. Images are divided in 4 parts. For each image, we have

[4] Publicly available at: `www.lsi.facom.ufu.br/datasets`

the percentage of time each user fixed his visual attention in a corresponding image part.

We denominate *VP-similarity score* between two users u_1 and u_2 as the distance between their respective visual perceptions vectors $\mathcal{P}_1$ and $\mathcal{P}_2$. This distance is defined by the function $l(u_1, u_2)$, where $l : \mathcal{P}_1 \times \mathcal{P}_2 \rightarrow \mathbb{R}$ and $l(u_1, u_2)$ can assume any classic similarity function like Euclidean distance, cosine similarity or Pearson distance correlation. By abuse of notation, we will write $l(u_1, u_2)$ as $l_{1,2}$. Table 1 highlights the VP-similarity score between u_4 and u_5, considering l as cosine similarity is 0.76 (* has been assumed as 0).

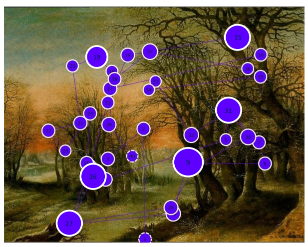

Fig. 1: Gaze positions and fixation length captured of an user.

Fig. 2: Example of a painting split in four equal parts.

Table 1: Users visual perception over two images of paintings dataset.

	$\mathcal{I}_1$				$\mathcal{I}_2$			
	q_1	q_2	q_3	q_4	q_1	q_2	q_3	q_4
u_1	0.50	0.10	0.40	0.00	*	*	*	*
u_2	0.60	0.20	0.10	0.10	0.10	0.70	0.10	0.10
u_3	0.40	0.40	0.20	0.00	0.00	0.90	0.10	0.00
C_1	**0.50**	**0.23**	**0.23**	**0.03**	**0.05**	**0.80**	**0.10**	**0.05**
u_4	*	*	*	*	0.75	0.08	0.05	0.12
u_5	0.05	0.25	0.20	0.50	0.70	0.05	0.10	0.15
u_6	0.15	0.20	0.25	0.40	0.82	0.02	0.10	0.06
u_7	*	*	*	*	*	*	*	*
C_2	**0.10**	**0.23**	**0.23**	**0.45**	**0.76**	**0.05**	**0.08**	**0.11**

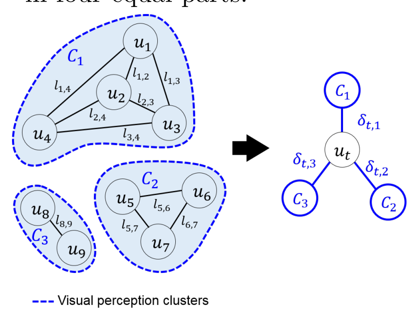

Fig. 3: Inferred visual perception networks (left) and selection of visual perception cluster for target user u_t (right).

Inferring Visual Perception Network. Our hypothesis is that users with similar visual perceptions are a good source for new user recommendation. Thus, we propose to cluster users according to their VP-similarity scores. In this work, we use *K-means* as classical clustering algorithm. Inside each resultant cluster we have a *Visual Perception Network*, defined as a complete graph, where nodes are users and the edges are labeled with respective VP-similarity scores. This process is shown in the left side of Fig. 3.

We define as *cluster consensual vector* the vector containing the averages of all visual perceptions from users inside the cluster. Table 1 illustrates two clusters C_1 and C_2 and their respective consensual vectors. This notion is specially important to perform a recommendation: when a target new user u_t is added to the system, our VP-Similarity method generates his visual perception $\mathcal{P}_t$ of u_t. Also, the VP-similarity score between u_t and each cluster C_k is computed. We denote $\delta_{t,k}$ as the VP-similarity score between an user u_t and a cluster C_k. This notation is similar to l, previously defined. The goal is to determine the most similar cluster concerning the target user and use the respective Visual Perception Network to perform recommendations. Thus, VP-Similarity finds the nearest visual perception cluster for a target user.

3 Experimental Setup

Dataset. In order to obtain visual perception information and ratings over items, we recruited 194 volunteers for rating 200 paintings, which were randomly chosen between 605 paintings public available at *Ciudad de la Pintura* (*http://pintura.aut.org*). For each volunteer, an eye tracker device captures eye movements on each painting displayed on the 22' monitor with image resolution of 500 x 700 pixels. The volunteer should rate each painting in a 1-5 scale according to its preference. The dataset is composed by 194 users, 605 items, 38,753 ratings, 67% of sparsity and 28,992 links among users.

Recommendation Baseline. We provide visual perception to social recommender systems based on matrix factorization. We compare the effectiveness of this approach among the following state-of-the-art recommenders:

- *Global Average:* A standard "popular" baseline, which recommends using the global average rate for an item.
- *SoRec:* A matrix factorization approach that combining social information and ratings to build a recommendation model[3].
- *SocialMF:* It is a model-based matrix factorization approach for recommendation in social networks with trust propagation mechanism [4].
- *TrustMF:* This approach combines a truster and trustee model mixing information of both, the users who trust the target user and those who are trusted by the user to build a recommendation model [5].

Parameter Settings. We use LibRec [6] library implementation of SoRec, SocialMF, TrustMF and GlobalAverage methods with default parameters. For social matrix factorization approaches the experiments were executed with 10 latent factors and number of interactions equal to 100.

Evaluation Protocol. We simulate a cold start scenario, using an evaluation protocol called **1-rating protocol**. To do that we employ the *"All but One"* protocol. Instead give as much as possible user's ratings for training, we provide only 1 rate from the target user per iteration. Note that all linked users' preferences in VP-Network are provided to build the recommendation model. We validate the model over the target user's remain rates.

Table 2: nDCG results for cold start scenario using 1-rating protocol.

Rank size	**SocialMF**	**SoRec**	**TrustMF**	**GlobalAvg**
5	0.6655 ± 0.1731	**0.8330** ± 0.1273	0.6281 ± 0.1468	0.7167 ± 0.1532
10	0.6726 ± 0.1550	**0.8297** ± 0.1103	0.6330 ± 0.1302	0.6954 ± 0.1305
15	0.6724 ± 0.1457	**0.8259** ± 0.1006	0.6346 ± 0.1249	0.6812 ± 0.1178
20	0.6749 ± 0.1379	**0.8218** ± 0.0988	0.6409 ± 0.1183	0.6656 ± 0.1142

4 Results and Discussions

We assess the effectiveness of our proposed Visual Perception Network model for item recommendation. In particular, we aim to answer the follow research question: ***How effective is visual perception networks for item recommendation?*** We also evaluate the prediction quality of visual perception approaches among the state-of-art recommenders presented in section 3. Table 2 shows the result of this comparison in terms of ***nDCG*** rank size of 5, 10, 15, and 20 for items recommended in our Paintings dataset.

We note that the experimental results show the superiority of SoRec over the others social approaches and the baseline. In particular, its performance might be explained because it work better in cold start scenario, since all social recommenders use the same VP-Network and are tested over the same conditions. The SoRec result attests the effectiveness of apply visual perception to leading with cold start users recommendation in contrast to others social approaches that have similar or inferior results than the baseline method.

5 Related Work

Solutions for user cold start problem. Generally, state-of-the-art works explore the same directions of our proposal: using contextual information to mitigate user cold start problem. The *similarity-networks* and *kinds of contextual information* to handle that are varied. Remarking on the use of a similarity-based networks, the work [7] uses a network among the users based on trust. It is not a social or a visual perception network. It is a method that generates a trust network based on ratings and users' similarities. A substantial body of literature explores *hybrid recommender systems* as a way to provide accurately recommendations on user cold start scenario. Those systems take advantage of different kinds of contextual information: [2] (preference and social data), [8] (item and preference data), [4] (demographic and social data).

Visual Perception as Implicit User Feedback. Few works investigated the use of eye tracker for recommending task [9, 10]. In these works, the user preference is not measured from ratings but from the way he/she looks at different images, such as eye fixation time and location. In [9] human visual perception data are adopted to build a gaze-based classifier for the image preference mining. User visual perception and preference data have been taken as a knowledge source to recommend images. However, while [9] proposed a classifier, our method is cluster-based. In [10] a content-based filtering enhanced by human visual attention was applied to clothing recommendation. This approach is specific for clothes domain and relays on visual attention similarity combined with the measures conventionally used in content-based image recommendation systems.

6 Final Remarks

In this paper, we have devised and evaluated VP-Similarity, a method whose ultimate goal is to incorporate visual perception in social recommender systems to deal with the user cold start problem. Our main contributions were: (1) A method to compute visual perception similarities among users; and (2) evaluating the use of visual perception similarities in social recommender systems to leading with user cold start problem. We have several ideas to extend this work in the future. First, it is worth exploring other datasets. Second, work to extend matrix factorization methods, and hybrid models approaches using visual perception similarities. Finally, we intend to evaluate if visual perception could be complementary to explicit preferences for general users recommendation.

References

1. Sahebi, S., Cohen, W.: Community-based recommendations: a solution to the cold start problem. In: Workshop on Recommender Systems and the Social Web (RSWEB), held in conjunction with ACM RecSys'11. (2011)
2. Felício, C.Z., Paixão, K.V.R., Alves, G., de Amo, S.: Social prefrec framework:leveraging recommender systems based on social information. In: Proc. of the 3rd Symposium on Knowledge Discovery, Mining and Learning. (2015) 66–73
3. Ma, H., Yang, H., Lyu, M.R., King, I.: Sorec: Social recommendation using probabilistic matrix factorization. In: Proc. of the 17th ACM Conference on Information and Knowledge Management. CIKM '08, ACM (2008) 931–940
4. Jamali, M., Ester, M.: A matrix factorization technique with trust propagation for recommendation in social networks. In: Proc. of the Fourth ACM Conference on Recommender Systems. RecSys '10, ACM (2010) 135–142
5. Yang, B., Lei, Y., Liu, D., Liu, J.: Social collaborative filtering by trust. In: Proc. of the 23rd Int. Joint Conference on Artificial Intelligence. IJCAI (2013) 2747–2753
6. Guo, G., Zhang, J., Sun, Z., Yorke-Smith, N.: Librec: A java library for recommender systems. In: 23rd User Modeling, Adaptation, and Personalization. (2015)
7. Victor, P., Cornelis, C., Teredesai, A.M., De Cock, M.: Whom should i trust?: The impact of key figures on cold start recommendations. In: Proc. of the 2008 ACM Symposium on Applied Computing. SAC '08, ACM (2008) 2014–2018
8. Amo, S., Oliveira, C.G.: Towards a Tunable Framework for Recommendation Systems Based on Pairwise Preference Mining Algorithms. In: Advances in Artificial Intelligence: 27th Canadian Conference on AI, May. Springer (2014) 282–288
9. Sugano, Y., Ozaki, Y., Kasai, H., Ogaki, K., Sato, Y.: Image preference estimation with a data-driven approach: A comparative study between gaze and image features. Journal of Eye Movement Research **7**(3) (2014)
10. Viriato de Melo, E., Nogueira, E.A., Guliato, D.: Content-based filtering enhanced by human visual attention applied to clothing recommendation. In: Tools with Artificial Intelligence (ICTAI), IEEE 27th Int. Conference on. (2015) 644–651