

Do you have a Pop face? Here is a Pop song.

Using profile pictures to mitigate the cold-start problem in Music Recommender Systems

Eugenio Tacchini
Università Cattolica di
Piacenza
eugenio.tacchini@unicatt.it

Ramon Morros
Universitat Politècnica de
Catalunya
ramon.morros@upc.edu

Veronica Vilaplana
Universitat Politècnica de
Catalunya
veronica.vilaplana@upc.edu

Enrique Sañoso
Universitat Politècnica de
Catalunya
enriquesv19@gmail.com

ABSTRACT

When a new user registers to a recommender system service, the system does not know her taste and cannot propose meaningful suggestions (cold-start problem). This preliminary work attempts to mitigate the cold-start problem using the profile picture of the user as a sole information, following the intuition that a correspondence may exist between the pictures that people use to represent themselves and their taste. We proved that, at least in the small music community we used for our experiments, our method can improve the precision of both a classifier and a Top-N music recommender system in a cold-start condition.

Categories and Subject Descriptors

H.3.3 [Information Search and Retrieval]

Keywords

Top-N recommendations, cold-start, evaluation, pictures.

1. INTRODUCTION

One of the limits of Collaborative Filtering (CF) Recommender Systems (RSs) is the "cold start problem": when a new user registers to a RS, the system does not know her taste and cannot propose meaningful suggestions until the user provides some feedback. Nowadays it is very common to provide a profile picture when we register to a Web site (including web-based RSs), furthermore in many cases users register using Social Network accounts, which allows in turn to access their profile pictures. This work analyzes users' profile pictures to provide hints about their musical taste and thus provide better recommendations since the registration, without additional input.

2. RELATED WORK

The cold-start problem hides two different subproblems: the new user problem (users need to rate some items before getting meaningful suggestions) and the new item problem (items need to be rated by some users before being suggested). For the new user problem, a solution is to fill the missing ratings with default values such as, for each item, the average rating received by other users [1]. Other approaches involve the segmentation of the users in homogeneous classes and the suggestion of items suitable for a specific class. Some classification criteria in literature are: data coming from questionnaires and demographic data [2,3]. An alternative approach [4] relies on the detection of communities through Social Networks analysis: for a new user, the RS suggests the items typically liked by the community she belongs to.

We focus just on the new user problem in the music domain, assuming we do not have any information about user preferences and relying exclusively on her profile picture. To our knowledge, there is no work that has attempted to use profile pictures to guess preferences. This method is not proposed as a substitution for existing approaches; it can be used when the RS does not have information about a user and can be combined with other methods in order to increase the accuracy in a cold-start situation.

3. METHODOLOGY

3.1 Background, goals and method

Our dataset comes from Last.FM: through the Last.fm APIs we can retrieve users' profile pictures and listening logs. Last.FM also allows users to create and join *Groups*; a group is a place where people talk about a topic, in particular, there are groups related to music genres e.g. the *Pop* group or the *Jazz* group. We assumed that a user joins a group if she is interested in that specific genre. Our experiments have been performed with users belonging to three quite different genre groups: Pop, Black Metal and Jazz (P, M and J). The three genres traditionally have a different audience: their fans tend to have not only a different music taste but also different styles so we thought that a dataset coming from M, P and J was a good starting point to test our intuition. We retrieved for each user her profile picture and listening logs, we built a playcount matrix $M(n,m)$ ($n=3,000$ users, $m=48,868$ artists) and for each user-artist pair we stored the number of times that user listened to that artist. The users dataset, together with some details and examples, has been released here [5]. Starting from $M(n,m)$ we built a preferences binary matrix $P(n,m)$ that represents, for each user, the artists she liked. P was computed using an approach similar to the one used in [6]; we assumed a user liked an artist if she listened to the artist more than five times.

We used the dataset for two different goals. The first one was a classification problem: given the profile picture of a user (without other information) and the information related to all the other users (profile pictures and groups they belong to), can we predict to which group the user belongs to (M, P or J)? To guess the group a user belongs to, we used a k-nearest neighbors (kNN) approach. The nearest *neighbors* of a user in this context were the ones having the most similar profile pictures (see section 3.2). The prediction was based on the groups the *picture-neighbors* belonged to: if most of the picture-neighbors of user U_x belonged to M, we predicted U_x belonged to M as well.

The second goal was related to a Top-N recommendation problem in a cold-start situation: a user U_x has just subscribed to a RS service and we want to suggest N artists. If we do not know anything about her taste we will end up suggesting random artists or the most popular artists. Given the profile picture of U_x and the profile pictures and preferences (P) of all the other users, can we

provide to U_x a meaningful Top-N artists recommendation list? To exploit the information from the user's picture, we mimicked a CF user-based technique using, as a similarity measure between two users, the similarity between their profile pictures. Given a user U_x , we selected her k nearest *picture-neighbors*, we computed the list of the N most appreciated artists by the *picture-neighborhood* and we suggested them to U_x .

3.2 Image analysis

Visual inspection of several profile pictures shows that the content of the pictures is very heterogeneous. Some people use pictures of their faces, while other use images of objects, places, logos, cartoons, etc. However, we expect that users with similar musical tastes select profile images that are related in some way.

To compute image similarities, images can be described using different types of visual information such as color, texture, shape of objects or similar characteristics. The MPEG-7 Visual Standard [7] specifies several content-based descriptors which can be used to efficiently identify, filter or browse images or video. The experiments have been performed using several MPEG7 image descriptors [7] (Dominant Colors, Color Layout, Color Structure and Edge Histogram). For space reasons we only present results obtained with Color Structure (CS), the best performing one. CS captures information about both color content and spatial arrangement of this color content. It is a histogram counting the number of times a color is present in a windowed neighborhood, as this window progresses over the image rows and columns. This enables it to distinguish, for example, between an image in which pixels of each color are distributed uniformly and an image in which the same colors occur in the same proportions, but are located in distinct blocks. The matching function used to compare the CS of two images is the L1 metric. We then convert distance into similarity multiplying distance values by -1.

4. RESULTS AND FUTURE WORK

To evaluate our method we used a *leave-one-out* approach. In the classification experiment, for each of the n users, we alternatively hid the group she belonged to and we tried to predict it according to her picture-neighbors. Fig. 1 (top) shows the results of the experiment: the precision at various levels of k (number of neighbors). Our method (*FaceBasedClassifier*) already overcomes *RandomClassifier* at $k=1$ and reaches the maximum value at $k=281$ where the precision is 0.467 (46.70% correct predictions); since M , P and J were composed by the same number of users, we assume, for *RandomClassifier*, a precision of 0.333; therefore our method overcomes *RandomClassifier*, at $k=281$, by 40.24%.

In the recommendation experiment, for each of the n users, we alternatively hid her preferences and we tried to predict them using her picture-neighbors. We used, as an accuracy metric, the precision, defined as number of *true positive* divided by the sum of *true positive* plus *false positive*, where a *true positive* here is an artist correctly guessed. Precision is a typical metric used for off-line evaluation of a Top-N task ([1]). We experimented with $N=20$ and we tested the precision at various level of k (number of neighbors), the final precision is the average of the users' precision and we compared it with the precision provided by *SuggestRandom* and *SuggestPopular*. The *SuggestRandom* method suggested, for each user, a different set of artists, randomly extracted among all the m available. The *SuggestPopular* method computed in advance the 20 most appreciated artists in the community of 3,000 users and suggested

those artists. Fig. 1 (bottom) shows the results: the random approach performs very poorly (precision 0.003). On our *FaceBasedRecommender* method, as expected for a kNN approach, the precision increases as k increases and at some point starts decreasing until it reaches the value of *SuggestPopular*. At $k = 250$ reaches its maximum: 0.2883, overcoming the precision provided by *SuggestPopular* by 10.01%.

This preliminary work shows that profile pictures can be used to mitigate the cold-start problem. Our hypothesis have been tested with both a classification and a Top-N recommendation experiment. As future work, we will explore ways to explicitly model correlations between musical taste and pictures using KCCA [8]. Another line of research will be improving the description of the image content by combining color and texture information using Bag of Features [9] based on color SIFT descriptors. Also, we will make experiments with more users, from different genre groups.

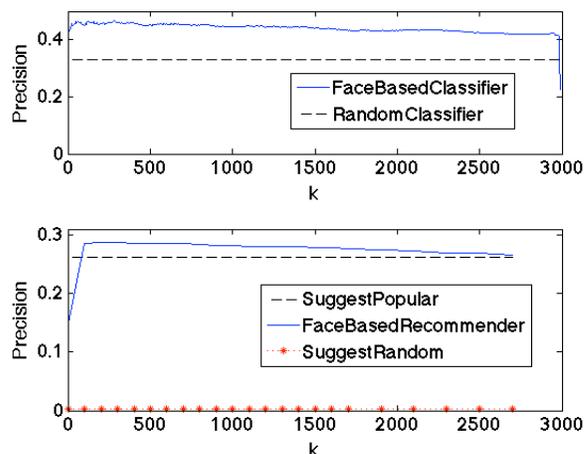


Figure 1: Classification and recommendation precision

5. REFERENCES

- [1] Ricci, F. et al. (2011). Recommender systems handbook. New York: Springer. (Chapter 4 and 8)
- [2] Park, S.T. et al. Pairwise preference regression for cold-start recommendation. *Proc of the 3rd ACM conference on Recommender system*. 2009.
- [3] Lika, B. et al. (2014). Facing the cold start problem in recommender systems. *Expert Systems with Applications*, 41(4), 2065-2073.
- [4] Sahebi, S. et al. "Community-based recommendations: a solution to the cold start problem.", *RSWEB*. 2011.
- [5] Dataset and some more details: <http://ds.dreamhosters.com/>
- [6] Tacchini, E. (2012), Serendipitous Mentorship in Music Recommender Systems. (Ph.D. Thesis).
- [7] B. S. Manjunath et al. Introduction to MPEG-7, Multimedia Content Description Interface, J. Wiley and Sons, Ltd., 2002.
- [8] D. R. Hardoon et al. Canonical Correlation Analysis: An Overview with Application to Learning Methods. *Neural Computation*, 16(12), 2004.
- [9] G. Csurka et al. Visual categorization with bags of keypoints. In *Workshop on Statistical Learning in Computer Vision, ECCV*, volume 1, page 22. Citeseer, 2004.