A Novel Commodity Image Recognition and Recommendation System for Mobile Shopping & Environment Protection

Haohan Wang[1], Ming Xu[2]
Department of Computer
Beijing University of Posts and Telecommunications
Beijing, China
[1]WangHaohan.of.BUPT@gmail.com
[2]xum@bupt.edu.cn

Zhen Chen[3], Feng Xie[4]
[3]Research Institute of Information Technology (RIIT)
[4]Department of Automation
Tsinghua University
Beijing, China
{xief10, zhenchen}@tsinghua.edu.cn

Abstract—**With the developing of Low-Carbon Economy and increasing attention on environment-friendly projects, we introduce the system for mobile shopping to make lives greener and more convenient. It allows the customers to perform the purchase items through taking photos of the commodity and the system recommends customers to choose the similar items but with lower harm to environment to buy. For the system to better work on massive data of the millions of commodities, we introduce two algorithms for image recognition and similar commodity recommendation respectively. We perform an experiment to test these two algorithms and the results show that this system performance is promising.**

Keywords-mobile shopping; image recognition; recommendation; algorithm; environment protection;

I. INTRODUCTION

Shopping patterns change as the society develops: in 1990s, people went to shops and wait in queue for daily supplies; in 2000s, people walked in the supermarket and bought items at low prices; now people prefer to buy things online and make the commodities delivered. Recently, Amazon① has launched an application on mobile phone allowing people to purchase a commodity through taking photos of it. Though it only allows four kinds of commodities to be bought in this way, this technology will lead us into a new stage of shopping.

However, as the Low-Carbon Economy will be the premier option of future economy, whether a shopping pattern is convenient is not the only thing the society today concerns, the contribution to environment protection of a shopping pattern should also be considered. At the forefront of science and technology, reference [1] has emphasized greatly on energy-saving calculations and environment-friendly projects, which brings a new era of blending the thought of environment-protection into the development of science and technology.

Under this circumstance, our system is presented for environment protection with convenient mobile phone shopping. The system is developed with the shopping pattern of taking photos of commodities and paying online. Whenever and wherever customers take photos of a commodity and send to the server. They can get the detailed information of it and the most similar ten commodities with lower harm to environment. Then they can choose to buy one of the commodities or cancel the purchase. Paying and delivering are performed after the customer confirms. This system is based on the image recognition and recommendation algorithms and should be implemented on the cloud.

Firstly, for the image recognition, an algorithm applicable to our system is necessary. However, in the extensive overview of invariant feature detectors dealing with global or local features for image recognition provided by [2]: Global features, like descriptors based on Principal Component Analysis (PCA) [3], requiring the pre-processing step to discriminate background and foreground, deal with overall composition of the image; Local features, like SURF [4] express the properties of the objects' shape. There are also combined ways for image recognition introduced: Reference [5] mentioned a two-level process combining CHEF with KNN and SURF with SVM. Reference [6] has introduced a combination of PCA and SURF. However, our system cannot be satisfied with the above, considering the following problems:

- There are millions of pictures to deal with; each query-by-picture should be quickly allocated into its category.
- The influence of the background should be minimized, since costumers may take photos anywhere.
- The validity of the recognition is determined by correctness of distinguishing slight difference because some commodities can only be distinguished by labels.

The application Amazon launched can solve these problems, but with only four kinds of commodities. To deal with other hundreds of kinds of commodities, we introduce a pattern of two-level image recognition with weighted features.

Secondly, a method for recommending the environment-friendly commodities related to Carbon Footprint is presented. Our recommendation is different from most existing algorithms [7-9], which are based on a group of similar customers or aimed at predicting the future preference. Our algorithm is presented for the following problems:

- The information of commodities is stored with data sparseness and attributes redundancy problems.

①http://techcrunch.com/2010/11/22/amazons-new-iphone-app-offers-in-store-price-comparisons-one-click-purchases/

- The recommendation should be based on the past reference of the specific customer and recommend choices to replace the current choice, not recommend the future choice.
- There are huge amounts of information to deal with and customers would not like to wait long.

For the first two problems, there are clustering algorithms to deal with, [10, 11] but clustering process is too complex to be performed on massive information, even when the information is already classified into small groups. Netflix① has employed a data storage method that can reduce the complexity, but commodities are no like movies. Considering the need of our system, we present an algorithm based on commodity attributes and customer's preference to recommend new choices.

II. SYSTEM OVERVIEW

As Fig. 1 shows: Mobile phone serves as a device to take the pictures, send images to server and receive the data returned from the server through WIFI or wireless network. It is a device to communicate with the server for the customer. The server is implemented on the cloud and it consists of two main processors, which are used for image recognition (*Image Recognition Processor, IRP*) and recommendation (*Recommendation Processor, RP*). There are also three main storage unit for the information of image feature, commodity and Users (*Image Feature Storage, IFS, Commodity Information Storage, CIS & User Information Storage, UIS*).

When the Cloud Server receives a query, it sends the query to *IRP*, Image Recognition process is called, loading information from *IFS*. Then the result is sent to *RP*, which loads information from *CIS* and *UIS*. The final result is sent to the customer through the Center of the Cloud Server. After the customer purchases or rejects, the Center will modify the information in *IFS* and *UIS* for improving the validity of image recognition and recommendation for future use.

III. DETAILED PROCESS

A. Two-level Image Recognition with Weighted Features

First, to overcome the difficulties of massive pictures, we implement a two-level process of recognition named *TIRWF*. The first level serves as the classification for overall features of the image related to a subset of samples in each category and the second level search is focused on the slight differences of each picture related to all the samples in one category.

1) First Level: In order to recognize the images by overall features and classify them fast enough, we implement the fast image classification algorithm introduced in [12]. However, considering that the costumers will take photos with any kind of background, we improve this algorithm with weighted features. The information is stored in vertically decomposed data and we enlarge the bit vectors storing feature attribute eight more bits, one for recording whether being chosen as the category, the other seven for the weight. The new algorithm, especially the modified part, is listed below:

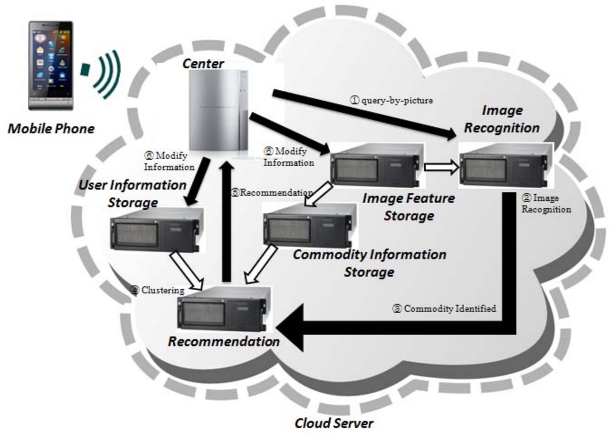

Figure 1. Architecture and Work flow of the system

a) Prepocessing

Extract the data information from image a.

b) Selecting a Category

- Determine the k nearest neighbors of Image a with:

$$d(x,a) = \sqrt{\sum_{i=1}^{D} (x_i - a_i)^2 x_{i\_weight}}$$

where, x_i stands for the features of Category x. a_i stands for the features of Image a; x_{i_weight} stands for the weight of feature x_i in Category x.

- Using k nearest neighbors, vote a category for a.
- Label the nearest k features x_i of the chosen category x by $x_{i\_label} = 1$.

c) Modifying the feature

If the costumer replies the recognition result is incorrect, which means the labeled features in selected category are unlikely to be useful for recognition. For better future use, modify the weight of labeled features. For each $x_{i\_label} = 1$ using:

$$x_{i\_weight} = \frac{\pi}{3} + \left(\frac{1}{2} - \frac{\pi}{12}\right) \times ((\sin(6(x'_{i\_weight} - 1) - \frac{\pi}{2})) + 1)$$

where x_{i_weight} means the present weight of x_i before modifying.

With this formula, the weight of the feature firstly increases slowly, then after several times of mismatch, it increases faster, finally approaches slowly to a threshold, approximately 1.5236.

For the details about the unmodified part of the fast image classification algorithm, refer to [12].

2) Second Level: The second level is applied to distinguish the slight differences of pictures with SURF feature. For more efficiency and accuracy of the recognition, this level only considers the discriminative features of the samples.

Recognize the Image a using SVM Classification through the features of Image a, computing with the discriminative features d_x of category x. Vote a recognition result for a. The speed and efficiency of recognition depends on the extraction of discriminative features, which will be discussed in next part.

①https://signup.netflix.com/global

3) Initialization: the algorithm adopts a two level recognition with two classification processes (KNN & SVM). The preprocessing and training procedure of each classification has been discussed in detail in [5][6][8]. The category of the pictures should be intialized by a clustering process. The data should be clustered by the similiarity of image, not utility. After clustering, the training process is supposed to start with easily recognized pictures in order to assign a correct weight for each feature. Moreover, a new way to select discriminative features is presented as following:

For all the pictures of category x, calculate $D_{i\_l\_j\_k}$ and t as threshold with:

$$D_{i\_l\_j\_k} = \left| f_{i\_l} - f_{j\_k} \right|, \qquad i \neq j$$

where f stands for features.

$$t = \frac{avg(D_{i\_l\_j\_k})}{e^{(p-1)}}$$

where avg means to calculate the average; p means the proportion of common features to all features. Each category may have a different proportion. This should be assigned manually for each category for a more precise calculation of t.

For all the features, if $D_i<t$, this feature is deleted; if not, it is labeled as discriminative feature. Label each feature for each category and the initialization process is finished. Though this process is complex, it does not affect the efficiency of the system since it is only computed once.

B. Recommendation Based on Carbon Footprint

After *IRP* recognizes the commodity, the *Center* called *RP* to recommend similar commodities, but with lower Carbon Cost. This function is aimed at providing the costumer more accesses to environment-friendly commodities that are probably out of their sight before.

1) Data stored in CIS: To complete the recommendation, some necessary information of each commodity should be stored in *CIS*. The information includes the parameters about appearance, performance, price, manufacturers etc. The harm of each commodity to the environment can be stored in the form of Carbon Footprint. All these data should be stored group by group according to its utility.

In one group of the data, all the attributes of each commodity should be transformed to digital form. In the initialization process, calculate the average of each attribute k, avg_k and the scope of each attribute k, $scope_k$. Then, we calculate the storing form of each attribute k for item i with:

$$sfa_{i\_k} = \frac{attribute_{i\_k} - avg_k}{\frac{1}{2} \times scope_k}$$

where sfa_{i_k} means storing-form-attribute of attribute k of commodity i and $attribute_{i\_k}$ stands for the original attribute k of commodity i.

In this form, all the data of different attributes are uniformed, so the intra-differences among attributes are eliminated. This process is not involved with Carbon Footprint and the missing attribute is not calculated.

Besides the above, a successful recommendation needs the information of each user's preference from *UIS*.

2) Recommendation Process: In the recommendation process, the system calculates the distance (dis_i) between the recognized commodity and each commodity i of the group as following:

$$dis_i = \frac{\sum_k sfa_{i\_k} \cdot \frac{1}{pre_m}}{num}, \qquad k \in m$$

where pre_m stands for the user's preference of the attribute of the kind m (It could be color, shape, performance, price etc. k□m means k is the attribute of the kind m) and num stands for number of attributes.

Select a certain number (implemented in the system as one percent of the number of the commodities in this group) of commodities with the least dis_i. These commodities are chosen as the similar ones. Recommend the 10 with the least Carbon Footprint to recommend to the customer. Because of the information of the commodities does not frequently change, the time complexity of this recommendation algorithm is approximately:

$$O(c \cdot a + \frac{c}{10} \log \frac{c}{10})$$

where c is the number of the commodities of this group. a is the number of the attributes.

3) Preference Recording: If the customer selects one recommended item to buy, calculate dis_k attribute k between the selected commodity and recognized commodity. The attribute with the least distance is supposed to reflect that the customer cares more of attribute of this kind, so the preference of the kind m and other kinds are updated as:

$$pre_m = pre_m^{'} + \frac{1}{k} \times \sum_{n \neq m} pre_n^{'}$$

$$pre_n = \frac{1}{k} \times pre_n^{'}, \qquad n \neq m$$

where $pre^{'}_m$ is the present preference before updating. pre_n is the preference of kind n, k is the number of attribute kinds

Generally, pre_m should be initialized as 1. Considering the image recognition process may not distinguish the commodities clearly, especially for the similar ones that are difficult for recognition, like laptops or cellphones, the pre_m for appearance is initialized as 1.3.

IV. Experiment

The experiment program is compiled in C#, modified from the original version of SURF in GoogleCode①. The client

①http://code.google.com/p/opensurf1/

Figure 2.Photo Shopping & Recommendation on WP7 Emulator

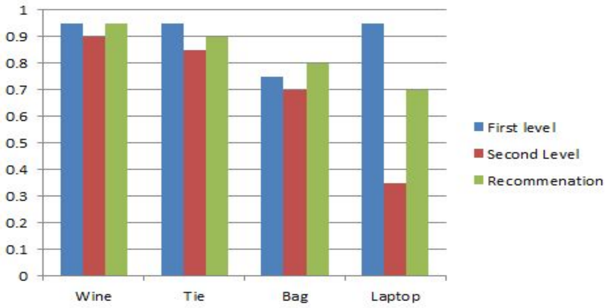

Figure 3. Results of the Experiments

prototype of our system is implemented on Windows Phone 7 Emulator on Windows 7, as showed in Fig. 2. The experiment is performed on four kinds of commodities: 1) wine: the labels make them easily to recognize; 2) ties: the photos are always shot with a big background; 3) bags: with varieties of styles; 4) laptops: even nearly impossible for human to distinguish. Because of the lack of related data sets, images and detailed information of the database are mainly from the internet, especially from online sale websites. The information of Carbon Footprint is assigned by us. The test sets are the pictures we take around us. Each set consists of twenty photos. The results, mainly reflecting the precision of the recognition, are showed in three phases: the two levels of image recognition and the recommendation. When an image is recognized wrongly, but showed in the list of recommendation, it is calculated as successfully recognized in Recommendation.

Fig.3 shows that the algorithm shows a good precision for most commodities. Even when the recognition gives a wrong result, the system can list it in the recommendation process.

V. Conclusion and Future Work

We have introduced a system for mobile shopping. In the system customers can take photos of an item and after the server recognize the item, they can choose to buy it or buy the similar but more environmentally friendly one. For the system to better work on the massive data of each commodity, we introduce two algorithms for image recognition and recommendation respectively. Experiments are performed to show that our algorithms have a relatively good precision.

In the future, the precision of the algorithm should be improved in order to recognize the labels of a computer. Future generations will be more likely to depend on mobile phones, they especially like to purchase through mobile phones. As a result, our system, though probably inconvenient for the first use, has a promising future.

Acknowldgment

This work is supported by Ministry of Science and Technology of China under National 973 Basic Research Program Grant No. 2011CD302600, Grant No. 2011CB302805, Grant No. 2011CB302601, and Grant No. 2012CB315800.This work is also supported by China NSFC A3 Program (No.61161140320).

References

[1] C. Lin , Y, Tian , M Yao, “Green network and green evaluation: mechanism, modeling and evaluation. In Chinese Journal of Computers, vol. IV, 2011.

[2] T. Tuytelaars, , K. Mikolajczyk, “Local invariant feature detectors: A survey,” in Foundations and Trends in Computer Graphics and Vision, vol. III, 2007, pp.177-280

[3] M. A. Turkand and A. P.Pentland, “Face recognition using eigenfaces.” in Proceedings of the Conference on Computer Vision and Pattern Recognition, 1991, pp.586-591.

[4] H. Bay, T. Tuytelaars, L. V. Gool, “Surf: Speeded up robust features.” in Computer Vision-ECCV 2006, pp.404-417.

[5] Y. Li, J. Lim, K. L. Lim and Y. Y, “Showroom introduction using mobile phone based on Scene Image Recognition.” in IEEE, ICME, 2009

[6] F. Temmermans, B. Jansen, R.. Deklerck, P. Schelkens, and J. Cornelis, “The mobile museum guide: artwork recognition with Eigenpaintings and SURF.” in WIAMIS 2011.

[7] S. Gong, “A collaborative filtering recommendation algotithm based on user clustring and item clustering”, in Journal of Software, vol. V, 2010, pp.754-752

[8] E. H. Lu, W. Lee and V. S. Tseng, “A framework for personal mobile commerce pattern mining and prediction” in IEEE Transaction on Knowledge and Data Engineering, 2011

[9] S. Gong, “An effective collaborative recommendation algorithm based on item clustering” in Springer-Verlag Berlin Heidelberg, 2010

[10] H. Kriegel, P. Kroger and A. Zimek, “Clustering high-dimensional data: a survey on subspace clustering, and colleration clustering” in ACM Transaction on Knowledge Discovery from Data, vol. III, 2009

[11] L. Bai, J. Liang, C. Dang and F. Cao, “A novel attribute weighting algorithm for clustering high-dimensional categorical data” in Pattern Recognition, 2011, pp. 2843-2861.

[12] T. Abidin, A. Dong, H. Li, and W. Perrizo, “Efficient image classification on vertically decomposed data.” in IEEE, ICDEW (2006)