INCORPORATING ELEMENTS FROM IMAGE RECOMMENDER SYSTEMS WITHIN A PERSONALIZED VIRTUAL TOUR FRAMEWORK

Asma Hanee Ariffin\textsuperscript{ab}, Yu-N Cheah\textsuperscript{b}, Mohd Helmy Abdul Wahab\textsuperscript{bc}

\textsuperscript{a}Faculty of Art, Computing and Creative Industry, Universiti Pendidikan Sultan Idris, Tanjung Malim, Perak, MALAYSIA
\textit{asma@fskik.upsi.edu.my}

\textsuperscript{b}School of Computer Sciences, Universiti Sains Malaysia, Penang, MALAYSIA
\textit{yncheah@cs.usm.my}

\textsuperscript{c}Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia, Batu Pahat, Johor, MALAYSIA
\textit{helmy@uthm.edu.my}

Abstract

A major vehicle that makes personalization possible is the recommender system that offers solution to overcome today’s world of information overflow. It provides specific suggestions based on user requirements, profiles or similar cases that have been previously handled by the system. This paper intends to illustrate the key elements summarized from previous image recommender systems, to be embedded within the suggested personalized virtual tour framework conducted in this research. This discussion will start by reviewing recommendation strategies that implement different ways of providing recommendations. As the focus of this paper is on image recommender systems, five key elements for better recommendations have been compiled based on a literature survey. Subsequently, these key elements were incorporated within a preliminary hybrid personalized virtual tour framework known as the ‘See What You Want, Feel What You See’ (SeeWYW, FeelWYS) model. Basically, the objective of this research on image recommendation is to implement a hybrid recommendation approach, which includes the socio-demographic and context-aware recommender engines. The socio-demographic recommender will deliver a suitable virtual route based on user demographic profiles. Following the suggested virtual route, a context-aware recommender will present a sequence of panorama that can be adapted based on user’s emotion as contextual information. It is hoped that these initial findings will provide insight on how to produce improved personalized and adaptive recommender system with good usability and good user feedback as well.

Keywords: recommendation strategies, image recommendation elements, personalized virtual tour framework

1. Introduction

A virtual tour can be described as a natural scenic tour that has been generated with a computer that incorporates multimedia elements to simulate existing locations and environments. The elements may involve sequences of videos and images, sound effects, music, narration or text intended to let the user experience the natural environment as in the real place. This virtual tour can be implemented via various techniques such as 3D models, images and video sequences or panorama where each technique is applied for different applications. In most cases, virtual tours have been developed to assist the tourism industry by providing virtual scenic tours of places of attraction. Virtual applications
may also act as supports for the entertainment affairs, business applications as well as teaching aids for analogical simulations and visual depictions of any topics of discussion. Creating a virtual tour can best be associated with three-dimensional (3-D) elements. However, conventional single or static images are normally too simple and require major enhancements via additional tools and methods, while 3-D computer graphic models often require complicated processing with suitable 3-D softwares. Therefore, utilizing panorama or panoramic images as a tool in a virtual tour can alternatively replace the previously mentioned techniques (Debevec, 1999).

In conjunction to a virtual tour, personalization may be defined as “the ability to provide contents and services tailored to individuals based on knowledge of their preferences and behaviour”. In addition, a recommendation system may be defined as a computer-based system that uses profile patterns from previous usage to provide relevant recommendations (Liang et al., 2008). A major vehicle that makes personalization possible is the recommendation system that matches potential products with customer preferences. However, most image-based virtual tours are focused on fixed environments with the assumption that all users have similar interests. Thus, this current research on image recommendation intends to deliver a suitable personalized virtual tour equipped with recommendation output based on user preferences.

As reported by the International Council of Museums (ICOM), the key features of an online interactive program are similar to a virtual tour, which consists of three main components;

- User preferences – concerning the type of user who is interacting with the system, type of knowledge that the user has and finally, what the user expects from the interaction.
- Item profile – primary concern is regarding the appropriate knowledge and context to be incorporated and classified to ease the matching process with user preferences.
- Suitable Interface with Recommendation Strategies - finally, to choose the best way to present the final output and adapt them with user feedback.

This paper will be discussing the main elements compiled from previous image recommender systems that can be incorporated within an adaptive image-based virtual tour framework. Previous works will be described based on different elements of these three components; user preferences, item profile and interface with recommendation strategies.

This paper has been organized into 5 sections and this first section covers the introduction. An overview of recommendation strategies will be described in the next section. Section 3 will discuss the five key elements that have been summarized from a survey on image recommendations. The process of exploiting and fitting the elements within the personalized virtual tour will be described in section 4 and this paper will be concluded in section 5.

2. Recommendation Strategies: an Overview

According to Liang et al.(2008), Ramezani et al. (2008), Yujie et al. (2010), Kim et al. (2008), and Kwon (2003), recommendation strategies and information filtering can be described based on three types of filtering; collaborative, content-based and hybrid filtering.

2.1. Collaborative Filtering (CF) Recommenders

Collaborative filtering can be classified into two groups; user-based CF and item-based CF. A user-based CF algorithm uses the preferences of similar users in the same reference group as a basis for recommendation. On the other hand, an item-based CF algorithm “improves scalability by using
ratings to focus on the similarity among items rather than the users themselves” (Lee et al., 2008). CF can be further subdivided into two types of algorithms; i) the memory-based collaborative filtering, where it constructs a set of neighbors at runtime and uses the neighbors’ preferences to generate a recommendation and ii) the model-based collaborative filtering that learns an aggregated model of user behavior in advance and uses this model to generate recommendations.

2.2. Content-based (CB) Recommenders

Content-based (CB) filtering utilizes the features of an item to recommend similar items to those in which the user has expresses interest. The CB system uses “persistent user profiles”, also known as learning content-based. This system deduces user’s interests and preferences from inputs that have accumulated over time, using statistical and machine learning techniques by studying user interests based on a long-term profile (Ramezani et al., 2008). CB systems that use such ephemeral user profiles (usually in the form of queries) are typically known as search engines.

2.3 Hybrid Recommenders

A Hybrid Recommender (HR) combines two or more techniques to gain better results with fewer drawbacks (Ramezani et al., 2008). For example, the integration of CB and CF to form a HR has been reported to produce better results in some domains (Liang et al., 2008).

In addition, traditional categories such as the CB and CF have similar user interaction process that consists of a single-shot recommender where each user interaction is used to produce a recommendation independently, with no additional user feedback. However, both systems have their advantages and disadvantages as tabulated in Table 1.

<table>
<thead>
<tr>
<th>Type of Recommender Algorithm</th>
<th>User Interaction</th>
<th>Advantage</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collaborative Filtering (CF)</td>
<td>Single-shot recommender</td>
<td>Use only historical data</td>
<td>Cold-start problem</td>
</tr>
<tr>
<td>Content-Based Filtering (CB)</td>
<td>Single-shot recommender</td>
<td>No cold-start problem</td>
<td>Unable to provide the serendipitous recommendations that CF generates</td>
</tr>
<tr>
<td>Hybrid</td>
<td>Combines two or more techniques (eg. content + collaborative to produce good results in some domains)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2.4 Other Innovative Strategies and Context-Aware Recommenders

As the research trend in recommender systems is growing rapidly, other innovative strategies have emerged such as socio demographic, knowledge-based, utility-based and also psychological-based, which have all been proposed in Nunes (2010).

The most suitable strategy for this current research on the educational application domain would be the Socio-demographic Recommender. This strategy utilizes the demographic features contained in a user’s profile to suggest any recommendation. It does not require any historical user ratings. Similar user characteristics are clustered to form user stereotypes that can further be matched with the current user profile. Therefore, knowledge-based reasoning techniques such as the Case-based Reasoning that performs similar processes would be appropriate for this strategy.
The importance of context-aware computing has gained the attention of researchers and practitioners in many disciplines (Yujie et al., 2010). Although there is still no exact definition on context, Yujie et al. (2010) has pointed out the most frequently cited definition from Abowd et al. (1999);

“Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves.”

The traditional techniques of recommender systems are generally based on a two dimensional matrix representation, known as the N*M matrix (N users and M items). However, when users are involved in making the decision on which items to choose, they will somehow be affected by other factors, depending on the specific conditions of the decision that has to be made. This is where contextual information plays an important part in recommender systems (Adomavicius et al., 2005; Adomavicius et al., 2011).

3. Key Elements on Image Recommendation

A survey on the classifications of recommender systems conducted by Park et al. (2012), discussed seven articles on recommender systems which are related with images. Hence, this section will discuss the five key elements extracted from this survey as a guideline for further research.

3.1 Hybrid Recommendation Strategies

Hybrid recommendation strategies are most likely the best choice as they resolve the shortcomings of CF and CB (refer Table 1). Kim et al. (2004) was the first to introduce this hybrid system for cell phone wallpaper image recommendation through Visual Content Recommender System (VISCORS) to reduce a customer’s search effort. The technique was focused on visual content characteristics to tackle the semantic gap problem. The long list of recommended images were shortlisted by CF, followed by the Content-based Image Retrieval (CBIR) algorithm that utilizes preference feedback applied for further search. The hit image was then labeled with binary weight, recommending the user to buy the preferred image.

In addition, in another work by Kim et al. (2008), the combination of CBIR and CF has also been proven to overcome the shortcomings in recommending image files in a decentralized domain such as the peer-to-peer (P2P) architecture. CB and CF are represented as personal agents, with the purpose of communicating with each other in order to decide who should receive recommendations and who should be kept as a neighbor, as well as what contents are of relevance to the host peer. Based on a conducted experiment using a simulator, the proposed model was able to offer qualified recommendations by minimizing the effort and time of users.

Other previous works such as the Visual document Context Aware Filtering (VCAF) Model (Boutemedjet et al., 2008) and the Context-Aware Multimedia Agent (CAMA) (Kwon, 2003) implemented context-aware recommendation strategies where contextual information is used in addition to user and item profiles. This approach can also offer more adaptive recommendations as user preferences may differ depending on different situations.
3.2 Consumer Psychology

During the decision-making process, users tend to be influenced by their own affective states or current situation. Therefore, the consumer psychology is an additional field that requires due attention in order to attain better user satisfactions. As VISCORS was the early approach on image recommendation, due to the advances in the field of research, it has some drawbacks that were soon investigated and enhanced. A recent research has discovered the VISCORS’ inability to compute the correlation between two users with similar preferences but with different context, as the users have no history on visual document rating (Boutemedjet, 2008). Hence, a new framework for context-aware recommendation of visual documents named the VCAF model was proposed. This new proposed model was inspired by consumer psychology based on two factors;

i) Visual content and metadata - that can be described by visual features (may be local, global, low level or of semantic nature) and keywords (can be automatically or semi-automatically extracted by annotation or recognition process).

ii) User preferences - which can be predicted by considering the user’s conformity (based on history of similar users), external environment (such as context) and the user’s novelty-seeking and variety-seeking behaviours (by using visual and semantic similarities).

3.3 Suitable Multimedia-based Ontology

In contrast with the VCAF, CAMA is an ontology-based recommender system that utilizes user preferences as ontology (Kwon, 2003). It uses an agent-based methodology to identify semantics of hypermedia images to derive contextual information and prior recommending specific web services that the user might be interested in. To identify and interpret images, the ontology is combined with extended Attributed Relational Graph (ARG) in which ARG acts as an image identifier. However, it lacks the incorporated information regarding mathematical models that play crucial roles in solving decision-making problems. Ontology as an image interpreter is used to understand what the items that appear in a semantic web actually mean, and also to represent user preferences. To increase the security level of a user’s private information, CAMA consists of two ontologies which are the public ontology and the private ontology.

3.4 Time-based Implicit Feedback

Collaborative Filtering (CF) mainly depends on its user explicit feedback. Somehow, this method has created some problems when the user rating patterns are formed in too many variations and hard to be classified. In certain cases, especially concerning the mobile e-commerce environment, explicit feedback by ratings preferences is difficult to collect as the user would be required to pay for some services. Therefore, Lee et al. (2009) has presented a time-based CF with implicit feedback to recommending character images (wallpapers) in a mobile e-commerce environment, which is similar to VISCORS (Kim et al., 2004). Lee et al. (2008) has reported some modifications of the conventional CF by constructing a pseudo rating data from implicit feedback data by incorporating temporal information (the time when the item was launched and the time when the user purchased an item) in order to achieve better recommendation accuracy. The results presented by Lee et al. (2008, 2009) have indicated that the implicit feedback may replace the difficulties of collecting explicit feedback and it may also provide better opportunity to any e-commerce environment.
3.5 Social Tagging Systems

Since CF is based on the preferences of similar users when recommending items, another innovative way of improving CF is by using the social tagging system (Zheng et al., 2010). Marinho et al. (2011) has defined it as such: “Tags are a way of grouping content by category to make them easy to view by topic”. Zheng et al. (2010) has used seven CF algorithms when recommending groups for Flickr users. Results have indicated that using a tag to represent an image in the recommendation algorithms can assist in delivering a more precise recommendation.

On the other hand, the use of CB as a social tagging recommender system is one way to overcome the cold start-semantic gap issue (Marinho et al., 2011). The semantic gap issue in image-based recommendation can be handled by performing a combination of tags with low level image features which are utilized as input for recommender systems. Nevertheless, real data for Flickr has demonstrated that the classifier (One Class SupportVector Machine (SVM)) was outperformed during training within the features vector instead of training alone with the tag feature or low level image vectors, respectively.

All these five elements will be incorporated into the proposed virtual tour recommender model as discussed in this current research.

4. Exploiting the key elements for the Virtual Tour Recommender model

The proposed framework is as illustrated in Figure 1: See What You Want, Feel What You See (SeeWYW, FeelWYS) model. This current model was inspired by the Case-Based Reasoning (CBR) cycle (Aamodt and Plaza, 1994).

As illustrated in Figure 1, the basic four main phases in CBR may be described as follows:

- **Retrieve** – Knowledge is represented as cases in CBR. Therefore, the CBR starts by retrieving the most similar preceding cases from the case library that have the highest similarity value as compared to the new case assigned by the current user.
- **Reuse** – The cycle continues by adapting the retrieved cases with other general knowledge to suit them with the current user’s problem. A combination of several cases and additional knowledge may produce the best solution to solve the problem.
- **Revise** – The solved case is presented to the user to generate a validation of the solution. If the validation process leads to an unsatisfactory result, the solution will be revised.
- **Retain** – The final solution with good results and validation is retained in the case library for future use.

The ‘SeeWYW, FeelWYS’ model represents the above phases which are tailored to the chosen application domain; virtual tour plan for education. The Previous Cases will be represented as domain ontology (for item) and user ontology (for previous user profiles). In addition, the CBR cycle within this model will be integrated with the Rule-based Reasoning technique in the third cycle (Retain) which also combines two recommendation strategies, the Socio-Demographic and Context-Aware Recommender.
Based on Figure 1, the ‘SeeWYW, FeelWYS’ model involves two main modules which are the Recommend module (See What You Want), followed by the Re-recommend module (Feel What You See). The findings from the first module will be compared within the second module to present an empathic recommendation to users. This work will be emphasized on recommending virtual museum tour guiding plans based on Similar Cases as compared to the current User Profile. When the user is satisfied with the recommended route (presented as the Overall Visiting Route), the second module will commence by presenting the panorama of the first suitable museum gallery (Panorama of Gallery). Then, the tour will follow the recommended route from the first module until the last gallery is visited, only if the user’s emotion matches with the emotion’s of the states of the gallery. If changes are detected, the re-recommended plan will create a new pattern of visiting route.

The next section continues with discussions on how to incorporate key elements as described in this section in order to plan a recommender system which will be more user-centered with adaptable features.
4.1 Hybrid Recommendation Strategies - Socio-demographic with Context-Aware Recommender

As previously mentioned, hybrid recommender strategies are suitable to compliment each other, based on the specific chosen domain. At first, the early assumption is that an image-based recommendation leans more towards content-based filtering approach or a simple hybrid of content-collaborative recommender system. However, Ramezani et al. (2008) argued that both of them cannot be directly extended to support personalized image recommendation because of the huge diversity and the time-varying properties of the user’s preferences. It is also very hard, if not impossible, to precisely learn users’ preferences from a few relevant judgments. Therefore, this study will narrow down into innovative strategies, which are inspired by these basic CB and CF strategies. The SeeWYW, FeelWYS model is a preliminary methodology to justify the flow of process to combine two recommendation strategies, which are the Socio-demographic Recommender in the first module, and the Context-Aware Recommender in the second module. With the intention of attaining a user-centric approach, this model integrates two reasoning techniques to suit the hybrid strategies with the application domain. The Rule-Based Reasoning will be embedded in the retain phase in CBR where the CBR cycle will be the main reasoning tool adapted in this model.

4.2 Consumer Psychology – Current User Emotion as Contextual Information

In the second module named the Re-recommend module (Feel What You See), a context-aware recommender was used to adapt the item (Panorama of Gallery) that are presented to the user. As discussed in Section 2.4, the context-aware recommender is not only based on users and items, this recommendation algorithm also focuses on other specific side factors. In the Re-recommend module, such side factor is referred to as the Current User Emotion, which pertains to consumer psychology in the decision making process, as suggested in Section 3.2. Since this model is focused on education, it adapts the user emotion card, inspired by Sylvie et al. (2009), as a tool to capture the user’s emotional state.

4.3 Suitable Multimedia-based Ontology – User Cases Ontology

The process of representing the elements into ontology may form sets of knowledge that can be easier to understand and shared. Yun et al. (2008) and Mihaela et al. (2012) claimed that ontology and CBR can work together to assist in measuring the semantic distances between similar cases and situations. Therefore, the See WYW, Feel WYS model utilizes the domain ontology and user ontology related to items and users of previous cases to be matched with the current user profiles in the early phase of the Socio-Demographic Recommender.

4.4 Time-based Implicit Feedback - User Feedback

It is a crucial step to capture user feedback in order to evaluate the proposed recommendation so that it can be enhanced and stored for future services. A similar process also takes place in the last CBR cycle, which is the Retain phase. In addition to the individual user feedback, this model moves a step forward into group adaptation, as discussed by Masthoff (2008) and Lara et al. (2013), on the importance of group adaptation and group recommender systems as a new research area. By evaluating user emotion feedbacks, this model will retain the current user details for future use.
4.5 Tagging Systems – Product-Emotion Space

Tags are useful as a way to categorize user input to make it easier when delivering a recommendation. This current model has adapted a similar practice from Desmet (2005) that maps the Product-Emotion Space to show the relationships between user emotions towards certain models of vehicle. Based on previous user cases, the Overall Visiting Route will display the item, tagged with emotions from previous users with similar profiles.

5. Conclusion

This paper discusses a brief introduction on recommendation strategies that have been the inspiration for a few recommending strategies for the current recommender system. The five key elements that are required for incorporation into the proposed model have also been highlighted. This introduction will hopefully give a better picture on suitable recommendation algorithms, as well as further processes that are suitable when applying the proposed framework.

In the future, this work will move towards implementing specific and suitable techniques based on the general flow of processes in this ‘See WYW, Feel WYS’ model. With suitable hybrid of reasoning techniques and recommender strategies analyzed with embedded chosen contextual information, it is hoped that this future personalized image recommendation model will produce not only good recommendation accuracy, but may also be well perceived in terms of usability and satisfaction to the targeted users.

Acknowledgment

This research was supported by the Universiti Sains Malaysia-Research University-Postgraduates Research Grant Scheme (USM-RU-PRGS 1001/PKOMP/846055). Thank you to the administrative staff of Museum and Gallery Tuanku Fauziah (MGTF), Universiti Sains Malaysia, Penang, for their assistance, useful ideas and comments.

References


