Exploration of Applying a Theory-Based User Classification Model to Inform Personalised Content-Based Image Retrieval System Design

Haiming Liu
Department of Computer Science and Technology
University of Bedfordshire
Park Square, Luton, LU1 3JU, UK
+44 (0) 1582489168
Haiming.liu@beds.ac.uk

Dawei Song
School of Computer Science and Technology
Tianjin University, China
The Computing Department
The Open University, UK
dawei.song@open.ac.uk
dwsong@tju.edu.cn

Paul Mulholland
Knowledge Media Institute
The Open University
Milton Keynes, MK7 6AA, UK
+44(0)1908654506
pual.mulholland@open.ac.uk

ABSTRACT

To better understand users and create more personalised search experiences, a number of user models have been developed, usually based on different theories or empirical data study. After developing the user models, it is important to effectively utilise them in the design, development and evaluation of search systems to improve users' overall search experiences. However there is a lack of research has been done on the utilisation of the user models especially theory-based models, because of the challenges on the utilization methodologies when applying the model to different search systems. This paper explores and states how to apply an Information Foraging Theory (IFT) based user classification model called ISE to effectively identify user's search characteristics and create user groups, based on an empirically-driven methodology for content-based image retrieval (CBIR) systems and how the preferences of different user types inform the personalized design of the CBIR systems.

Categories and Subject Descriptors

H.1.2 User/Machine Systems, H.3.3 Information Search and Retrieval, H.5.2 User Interfaces.

General Terms

Design, Human Factors, Theory.

Keywords

Personalised search, ISE user classification model, IFT, User model utilisation, CBIR.

1. INTRODUCTION

This paper proposes a methodology for applying an Information Foraging Theory (IFT) based user classification model [4] to classify users into different characteristic groups in order to understand different search preferences of the different user groups for providing them with personalised search experiences. The findings from a systematically structured analysis of the

users' interaction data, collected from an extensive empirical user study, further validate the user model and establish the preferences of the different user types for the design and development of personalised content-based image retrieval (CBIR) systems.

Users are a key element of all search systems. The users can be very different when they use a search system: some people are patient, but some are not; some people frequently change their mind on what they are looking for, but some do not; some people are easily satisfied with the result they get after a few rounds of search, but some are not [10]. It is also important to note that the user types are usually implicit, which can be reflected and characterised by the users' search behaviours during the interaction with the system [4]. Learning more useful information from users through user interaction data becomes vital to improve search system personalisation and better engage users during the search process [3].

Personalisation is an important strategy in web search engines to keep their users [2]. Extensive work has been done to provide personalised search for improving users' overall search effectiveness and experiences [1, 4, 9, 11, 13, 14, 15]. However, there still exist gaps and challenges on various aspects. For example, it is challenging to develop effective user models, and it is even more challenging to apply the user models to search systems to support personalised search, because applying different user models to different search systems requires different methodologies and different automation methods.

From the literature review, we have found that there is significant smaller number of user models that have been developed for CBIR than text search, when CBIR needs good user models for providing users a better interactive search experiences. The user classification model developed by Liu et al [4] based on Information Foraging Theory (IFT) is the only theory-based user model that we know which is developed for CBIR search using a deductive approach. We consider this model is a suitable candidate to be applied to interactive CBIR search systems for developing personalised search. From the literatures, we have also found that there is a lack of applications of the user models in real life search systems, especially CBIR systems. This may be caused by the challenges during the application, such as no effective methodologies for the application, etc. It is interesting to explore how to apply theory-based user models to interactive CBIR systems for improving the CBIR search personalisation.

2. RELATED WORK

Wu et al. (2014) [14] developed a learning-based framework for image search, which combines users' search behaviours analysis and current trend of search queries to suggest personalised trending queries and images. Xie et al. (2015) [15] applied machine learning techniques to model users' interests, which can be fed back to the search model for developing personalised search experiences. Shen et al. (2005) [9] proposed a method to infer users' interests from their search context, e.g. click through information, for personalised search. White and Drucker (2007) [11] identified two types of users, namely navigators and explorers, from a log-based study with a large number of users and over a long time period and the search interactions for the two user types are either very consistent or very evolving respectively. The findings suggested the personalised design of search tools should support effective interactions for different user types.

The above studies all employed an inductive approach, in which the user models were learnt through analysing pattern in the user interaction data and then applied to the search systems. As users' information goal and search strategy often evolve, the users' interaction with the search is exploratory in nature [4, 12]. With increasing amount of constantly evolving interaction data, the data pattern can become unclear. Thus the inductive approach can become complicated and computationally expensive. To tackle this problem, the deductive approach has recently be considered, which is to develop and validate user models based on established user behaviour theories, and then apply the validated user models to search systems. For example, Wilson et al. (2009) [13] presented a formative inspection framework based on two established user models, namely the ISS conditions from episodic model and the levels of search strategies in strategic model. The framework was applied to evaluate three search interfaces. Collins-Thompson et al. (2011) [1] investigated personalising web search results based on reading levels. They firstly estimated the user' reading proficiency level and the results' difficulty levels, and then re-rank the results based on the difference between the two. Liu et al. (2010) [4] proposed and verified a user classification model, namely the ISE model, for categorizing different user types based on IFT [8] for interactive CBIR. The different user types can then be applied to inform personalised CBIR search system design as well as to improve the overall search experience.

The existing work described above has demonstrated how the user interaction data can be effectively applied to develop personalised search, such as users' search history/log [9, 11], users' search behaviour [4, 14], users' profile data [1, 15]. Many of the existing studies focused on improving the search accuracy and efficiency, while a few of them focused on improving the design of the search system and the users' overall search experiences [1, 4, 11, 13]. All the studies above had evidenced the user models can be generated based on the users' interaction data or established theories or both. The user models are the key element to learn different user preferences for different user types [1, 4, 11], and to design and develop effective personalised search systems [4, 13]. Whilst the theory-based user models [1, 4, 13] have been developed for personalised search, there has been lack of a methodology for systematic empirical studies with real search systems in order to utilise the user models for better understanding the users. Among all of the studies, the information foraging theory based ISE model [4] is particularly developed for CBIR. In this paper, we are motivated to explore the possibilities and challenges of applying the ISE user model to classify different user type base on their interaction data for personalised CBIR search.

3. AN INTRODUCTION TO THE ISE MODEL

Information Foraging Theory (IFT) [8] suggests how humans seek information is like how wild animals seek food. Animals' general food seeking behaviour is that they first find a patch of food (scents model); next they select what to eat from the patch (diet model), and then they decide when to hunt elsewhere (patch model). White and Roth (2009) [12] suggested users' exploratory search behaviour is similar to IFT in the respect of users try to find and judge an optimal result for their information goal, and they apply different search strategies, and they have different opinions when they decide which result to use. CBIR search is rather exploratory because users do not have a clear idea on what exactly they will found from the search results; therefore their search goal evolves during CBIR search more than when they carry out a keyword-based search.

Liu et al. (2010) [4] proposed and verified a new user classification model called ISE model based on IFT, which includes three criteria: Information goals (I), Search strategies (S) and Evaluation thresholds (E). Each criterion categorizes users into two different user characteristics[1]: I - fixed information goal or evolving information goal, corresponding to the Information Scent model in IFT; S - risky search strategy or cautious search strategy, corresponding to the Information Patch model of IFT; E - weak evaluation threshold or precise evaluation threshold, corresponding to the Information Diet model of IFT.

Table 1 shows the definitions of the six characteristics in ISE model based on IFT. Table 2 shows the operational definitions of the six characteristics in the ISE model in an interactive CBIR search context.

Table 1: Definitions of the six user characteristics

Criteria	Characteristic	Definition
Information goal	Fixed	Searchers with fixed information goal know what they are looking for.
	Evolving	Searchers with evolving information goal are not sure what they are looking for.
Search strategy	Cautious	Searchers with cautious search strategy move slowly between patches.
	Risky	Searchers with risky search strategy move quickly between patches.
Evaluation threshold	Weak	Searchers with weak evaluation threshold are lenient on selecting the result.
	Precise	Searchers with precise evaluation threshold are strict on selecting the result.

---

[1] There is in total six characteristics in the ISE model.

Table 2: Operational definitions of the six characteristics

Characteristics	Operational definition
Fixed	1.Use small number of jump query transition; 2.Use small number of history functionality; 3. Find the best result image early.
Evolving	1.Use large number of jump query transition; 2. Use large number of history functionality; 3. Find the best result image late.
Cautious	1. View large number of result pages; 2. Spend long time per search iteration; 3. Select result at the end of the search.
Risky	1. View small number of result pages; 2. Spend short time per search iteration; 3. Select result while searching.
Weak	1. Select large number of results; 2. Select large number of feedbacks; 3. Search small number of iterations.
Precise	1. Use lots of subset query transition; 2. Use many times ranking functionality; 3. Search large number of iterations.

Information Goal (I) - At the beginning of a search, the searchers might or might not have a clear information goal (idea on what they are looking for) to start the search. In IFT terms, the searchers might or might not get strong information scent from reading the task based on their information environment (knowledge). Thus, the searchers can be categorised into two types based on the information scent concepts: One type with fixed information goal and the other with evolving information goal. According to the information scent concepts, if the searchers have a fixed information goal, they tend to focus on what they are looking for and likely make consistent decisions at every stage. On the other hand, if the searchers have an evolving information goal, their search will be more exploratory. They tend to walk around and learn from the data before they make a decision.

Search Strategy (S) - When searchers start a search, they often submit the first query, which can be seen as an initial effort to find the first information patch, and then they might or might not walk around within the patch and evaluate what they have found before they provide feedback to refine or reformulate the query to start a new search (we can consider this as looking for a new patch). In IFT terms, the searchers can decide whether they would like to do between or within patch activities based on their search strategy. Thus, the searchers can be categorized into two types based on the information patch model: One type has cautious search strategy and the other has risky search strategy. According to the information patch concept, the searchers with a cautious search strategy tend to perform more within-patch activities, which means they will carefully search through the current patch before they go to the next patch (e.g. reformulating the query to start a new search); the searchers with risky search strategy, on the other hand, tend to be more adventurous and perform more across-patch activities, which means that they will skip over the current patch and move to next patch quickly.

Evaluation Threshold (E) - When searchers select the result images for completing the tasks, they need to decide which images to choose from the result list. In IFT terms, some foragers like easy-to-catch prey, but others like hard-to-catch prey. Thus, the searchers can be categorized into two types based on the information diet concepts: One type with weak evaluation

threshold and the other with precise evaluation threshold. According to the information diet concepts, the searchers with weak evaluation threshold will be likely to go for easy-to-catch information although the information maybe just slightly relevant to the their information goal; the searchers with precise evaluation threshold will instead go for hard-to-catch information: For example, they will not select the information unless they consider it highly relevant to their information goal.

4. EMPIRICAL STUDY SET UP AND DESCRIPTION OF DATA

The CBIR search system that we apply the ISE model to is called uInteract (Figure 1). The uInteract system is developed based on a four-factor user interaction model [5] for user-centred effective and interactive CBIR search. The user interaction data is collected from a user study that involves 50 subjects and four search tasks with different complexity levels. The users were a mixture of males and females, undergraduate and postgraduate students, and academic staff from a variety of departments.

The users were given search tasks with different complexity levels to complete within a time constraint using the four interactive CBIR systems in a random order, and provided feedback on their search experiences through five point likert scale questions and open questions in questionnaires. The user's actual search results for all the tasks were also collected. The screen capture of the whole search process was recorded with video and audio input. The tasks are simulated natural life tasks, such as "*Imagine you are a graphic designer with responsibility for the design of leaflet on the newly built sport stadium for the local council. The leaflet is intended to raise interest among the general public and encourage people to use the stadium and to watch the sports in the stadium. Your task is to find 3-5 images, from a large collection of images, to include in the leaflet. The images should represent the kind of sports you think can be held in the stadium.*", which allowed the users to develop their own interpretation of the task description, use their own judgement for selecting relevant images as feedback and result, and decide when to use different functionalities of the interface to support their search.
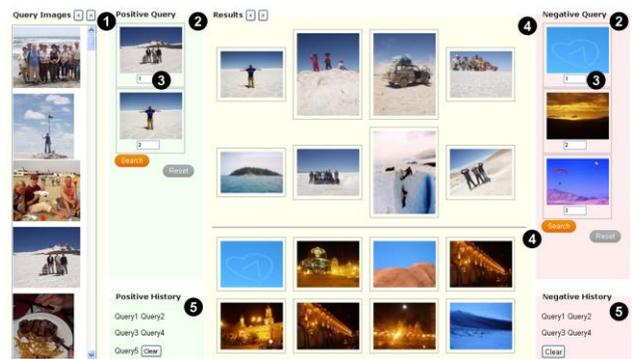
Figure 1: Interface 4 (S4) - The uInteract interface. *The keys' description: (1) Query images panel provides a browsing functionality that facilitates the selection of the initial query images. (2) Users can provide both positive and negative examples to a search query in the positive and negative query panel, and further expand or remove images from that query. (3) By allowing the user to override the system-generated scores (integer 1-20) of positive and negative query images, users can*

directly influence the relevance level of the feedback. (4) The displaying of the results in results shows not only the best matches but also the worst matches. This functionality can enable users to gain a better understanding of the data set they are searching. (5) Combining both positive and negative query history functionality has not previously been undertaken in CBIR. The query history not only provides users with the ability to reuse their previous queries, but also enables them to expand future search queries by taking previous queries into account.

A substantial set of real user interaction data was extracted from the screen capture of the user study. There are in total 50 users' screen captures. Every screen capture is about two hours long with both audio and video input. We extracted totally 48 unique interaction features, which can be categorised into six groups:

- **Time and iteration**: Time to complete each iteration, time to complete task, time to find the best result, number of iterations/queries per task;

- **Results page**: Number of result pages viewed, page result selected from, page found the best result[2], page positive feedback selected from, page negative feedback selected from;

- **Image**: Number of images used per query (positive and negative query), number of feedback images selected (positive and negative query), number of results selected;

- **Functionality used**: Number of times positive/negative ranking used, number of times positive/negative history used;

- **Select result strategy**: Some users select result while searching, and some users select result at the end of the search;

- **Query transitions**: We adapted the five query transitions (Table 3) for both positive and negative queries from Mulholland et al. (2008)'s work [6].

Table 3: Five types of query transition

Query transition	Description
Repeat	Consecutive positive or negative query contains identical images.
Subset	The next positive or negative query contains a subset of the query images.
Superset	The next positive or negative query contains all the previous images plus one or more additional images.
Overlap	The next positive or negative query contains some but not all of the previous images plus one or more additional images.
Jump	There is no intersection between the images used in consecutive positive or negative queries.

5. A METHODOLOGY FOR UTILISING ISE MODEL TO IDENTIFY USER TYPES AND PREFERENCES

In this section, we propose a methodology to utilise the ISE model based on the user interaction data that we have collected for the

---

[2] The best result here is judged based on the ratings of five independent raters [6].

user study. We first map the definitions and operational definitions of the six characteristics in the ISE model (see Table 1 and Table 2) to the user interaction features extracted from the screen captures of the user study, in order to find out the type[3] of each individual user. Further we put the 50 users into different groups[4] based on their user types to identify the different search preferences of different user types. The methodology is detailed as follows:

5.1 Identifying general User Types for Each Individual User

The following steps describe how we identify user type for each individual user based on the ISE model through analysing the users' interaction data obtained from the screen capture of the user study.

Step 1 - Find the interaction features for the 3 operational definitions of each of the 6 user characteristics. We select three interaction features from Table 3, corresponding to the three operational definitions in Table 2 for each characteristic. For instance, the operational definitions of Risky character are: (1) view small number of result pages; (2) spend a short time per search iteration; (3) select results while searching. Correspondingly, the relevant interaction features are (1) NumResultPageViewed; (2) TimePerIteration; (3) SelectResultStrategy. Table 4 shows the 3 interaction features we chose for each of the 6 characteristics in the ISE model[5].

Table 4: Interaction features that support the operational definitions of the six characteristics

Characteristic	Interaction features
Risky	-Num_ResultPageViewed; -TimePerIteration; SelectResultWhileSearching
Fixed	-Num_JumpQuery; -Num_History; -TimeFindBestResultImage
Weak	+Num_ResultSelected; +Num_RFSelected; -Num_Iteration_Query
Cautious	+Num_ResultPageViewed; +TimePerIteration; SelectResultAtTheEnd
Evolving	+Num_JumpQuery; +Num_History; +TimeFindBestResultImage
Precise	+Num_Ranking; +Num_SubsetQuery; +Num_Iteration_Query

Step 2 - Identify the characteristics of every user for each task. We calculate the mean value for every interaction feature for each task across all the users. We then suggest the characteristic of a user based on whether or not the value of interaction feature for the user is bigger than the mean value. For the interaction features with binary values such as *SelectResultStrategy*, we do not calculate the mean value, but judge the characteristics of the user based on the data itself. The final characteristic of a user for a specific task is determined by all three interaction features. For example, a user will be identified as a risky/cautious user when he/she shows a risky/cautious characteristic with regard to all the 3 interaction features, otherwise the user will be classified as

---

[3] A user type could include more than one characteristic from different classification criteria of the ISE model.

[4] A user group contains users with the same user type.

[5] "+" means more and "−" means less.

Table 5: An example of how to identify each user's type for each task

Task	Num_Rresult PageViewed	Characteristic	TimePerI teration	Characteristic	SelectResul tStrategy	Characteristic	Summary
1	37	Cautious	00:02:30	Cautious	End	Cautious	**Cautious**
1	10	Risky	00:00:52	Risky	View	Risky	**Risky**
1	22	Cautious	00:00:54	Risky	View	Risky	
1	24	Cautious	00:00:42	Risky	View	Risky	
1	6	Risky	00:00:30	Risky	View	Risky	**Risky**
1	9	Risky	00:01:09	Risky	View	Risky	**Risky**
1	43	Cautious	00:02:15	Cautious	End	Cautious	**Cautious**
1	22	Cautious	00:02:40	Cautious	View	Risky	
1	45	Cautious	00:02:51	Cautious	End	Cautious	**Cautious**
1	20	Risky	00:02:47	Cautious	View	Risky	

neither risky nor cautious (Table 5). We apply the same methodology to identify fixed/evolving characteristic and weak/precise characteristic.

Step 3 - Identify the final characteristics for every user. After checking every user's characteristic for each task, we need to summarise the user's overall characteristics. From previous steps, we find that a user can have completely different characteristics when doing the four different tasks. Further we also find that in reality there are five characteristics instead of two characteristics (Table 2) in each criterion based on our data, such as *risky*, *cautious*, mixed risky and cautious (*mixedRC*), neither risky nor cautious (*noneRC*) and not sure whether it is risky and/or cautious (*undefinedRC*) for the search strategy criterion. When we put the user's characteristics from the four tasks together, it is difficult to determine the user's final characteristic. Therefore, we propose the following approach to decide the type of a user (an example is shown in Table 6):

- Risky (R) user: ≥ 2 tasks shown risky and 0 tasks shown cautious;

- Cautious (C) user: ≥ 2 tasks shown cautious and 0 tasks shown risky;

- MixRC user: > 0 tasks shown risky and > 0 tasks shown cautious;

- NoneRC user: 0 tasks shown cautious and 0 tasks shown risky;

- UndefinedRC user: Does not match 1 - 4;

- Fixed (F) user: ≥ 2 tasks shown fixed and 0 tasks shown evolving;

- Evolving (E) user: ≥ 2 tasks shown evolving and 0 tasks shown fixed;

- MixFE user: > 0 tasks shown fixed and > 0 tasks shown evolving;

- NoneFE user: 0 tasks shown fixed and 0 tasks shown evolving;

- UndefinedFE user: Does not match 6 - 9;

- Weak (W) user: ≥ 2 tasks shown weak and 0 tasks shown precise;

- Precise (P) user: ≥ 2 tasks shown precise and 0 tasks shown weak;

- MixWP user: > 0 tasks shown weak and > 0 tasks shown precise;

- NoneWP user: 0 tasks shown weak and 0 tasks shown precise;

- UndefinedWP user: Does not match 11 - 14.

Table 6: An example of how to identify user types for each user across four tasks

User Type	Risky / Cautious			
	Task1	Task2	Task3	Task4
NoneRC				
Riksy		Risky	Risky	Risky
UndefinedRC		Risky		
UndefinedRC				Cautious
Risky		Risky		Risky
UndefinedRC				Risky
Risky	Risky		Risky	Risky
Risky	Risky			Risky
UndefinedRC	Cautious			
MixedRC	Cautious	Risky		

Table 7: The number of users assigned into three characteristic cross tables

Search strategy / Information goal	Risky	Cautious	MixedRC	NoneRC	UndefinedRC	Total
Fixed	2	2		4	5	13
Evolving						
MixedFE	2		1	1	3	7
NoneFE	5	1	2	3	5	16
UndefinedFE	3	1		3	7	14
Total	12	4	3	11	20	50

$\chi 2$ test = 0.9950

Search strategy / Evaluation threshold	Risky	Cautious	MixedRC	NoneRC	UndefinedRC	Total
Weak					1	1
Precise						
MixedWP						
NoneWP	9	4	3	10	12	38
UndefinedWP	3			1	7	11
Total	12	4	3	11	20	50

$\chi2$ test = 0.9965

Information goal / Evaluation threshold	Fixed	Evolving	MixedFE	NoneFE	UndefinedFE	Total
Weak				1		1
Precise						
MixedWP						
NoneWP	13		5	12	8	38
UndefinedWP			2	3	6	11
Total	13		7	16	14	50

$\chi2$ test = 0.9871

5.2 Grouping Individual Users into Types

The When each user has more than one characteristics, the following steps suggest how we can categorize the 50 users into different characteristic groups.

Step 1 - Put the users into characteristic cross tables. Firstly, we make a cross table for each pair of characteristic criteria: For instance, one cross table between search strategy and information goal; one between search strategy and evaluation threshold; and one between information goal and evaluation threshold. Secondly, we assign the 50 users into every cross table. Each cell indicates the number of users identified as the combined characteristics. Table 7 shows the three assigned cross tables and the numbers are the number of people who match the characteristic. The $\chi2$ test is applied to test the independence between the five row characteristics and the five column characteristics. The results show there is no significant correlation for any pair of categorical variables, which suggesting that we can analyse the row and/or column characteristics independently.

Step 2 - Group the users into the types. From the middle and bottom sub-tables of Table 7, we can see that there is insufficient variance of characteristics in the evaluation threshold criterion. For instance, only one user shows weak characteristic and the rest of the users' carries noneWP and undefinedWP characteristics, so we decide we are not going to analyse the characteristics in this criterion any further. We then focus on the other two criteria: Search strategy and information goal. The columns of the top sub-table show that there are 12 risky users, 4 cautious users, 3 mixedRC users, 11 noneRC users and 20 undefinedRC users. The rows of the top table shows 13 users with fixed goals, 7 users with mixed fixed and evolving goals, 16 users with neither fixed nor evolving goals, 14 users with undefined fixed or evolving goals.

6. DATA ANALYSIS RESULTS FOR DIFFERENT USER TYPES

In this section we report the analysis method and results of the questionnaire answers that was obtained from the user study. The analysis will be done to each user types that we identified using the ISE model in the above section, in an effort to find out what the search preferences are for the different user types.

We firstly assigned the 50 users to the five characteristics groups of the search strategy criterion, namely Risky, cautious, mixedRC, noneRC and undefinedRC. Then we link the users' answers to the questionnaire questions with the five characteristics. We repeat the same process for the five characteristics of the information goal category, namely Fixed, evolving, mixedFE, noneFE and undefinedFE. After matching the users' questionnaire answers to the different user characteristics (type), we carry out a content analysis [7] to the answers for each of the user types. The content analysis shows that for each user type the answers can be classified into three categories, namely expected image search tool, search experience and suggestions to the search systems.

6.1 Search preferences of the user types based on the Search Strategy Criterion

Risky Users: The users with risky search strategy prefer accurate and diverse results and care less about the data source quality and where they search from. They prefer rich functionalities to support different search aspects, so that they can find the good results quickly and easily. They judge the effectiveness of the system depending on the tasks they are performing. They tend to think the system is good when they perform an easy task using the system and get the good results fairly quickly; otherwise, they think the system is bad. As our evaluation systems support multiple images for each query, the users with risky search strategy feel the search accuracy drops with more image examples in the query. This might be because they are likely to provide image query examples using variety of judgement, such as colour, shape and semantic relevance, which is not supported by our colour only based evaluation systems. However, a risky user could perform quite well if s/he gets the supportive functions needed. This is why risky users provided many useful suggestions about improving the usability of the evaluation systems, such as adding an egg timer, image zooming, and dragging and dropping function, etc.

Cautious Users: The users with cautious search strategy are another group that showed a clear pattern. Like the risky users, they hope the search system is accurate and with rich functionalities to support different search aspects. They do not care much about the search speed. This might be because cautious people are usually patient. They are more satisfied with the search results and the evaluation systems than risky users.

They did not need to use all of the provided functionalities on completing some tasks although they are more likely to think the functions could be useful. The difference between risky users and cautious users is that the cautious users feel that using more image examples in a query improves the search results. This might be because cautious users are more likely to be careful with query refinement and they understand the nature of the colour-based evaluation systems and the tasks better than risky users. Half of the cautious user population think the tasks are interesting and clear. As they are satisfied with the search performance, they suggested minor improvement that could be made to the evaluation systems, such as better graphic design for the interfaces. They strongly suggest combining keyword-based and content-based search. This might also because it is hard for cautious users to change their search strategy completely. They like the content-based search strategy, but they also want to keep their normal keyword-based search strategy.

MixedRC Users: Some comments provided by the mixedRC users are similar to the comments from the risky and cautious users. For instance, they do not care about the speed of the system, they like rich functionalities on the image search system, they are satisfied with the search results and the evaluation systems, they think the negative query and query scoring functionalities are useful and have many suggestions on improving the evaluation systems. However, they comment on some things that the risky and cautious users have not mentioned. For instance, they strongly believe the image search tool is easy to use, they think the content-based image search strategy is better than a keyword-based search strategy, and they would like to see the negative results become optional.

NoneRC Users: The noneRC users like fast and accurate systems and prefer rich functionalities to support different search aspects. They think the tasks are interesting and clear. They do not think the history functionality is useful at all. They find that it is hard to decide the relevant results for the tasks. Their initial search idea changes during the searching but they think the system supports the change well. They strongly suggest to improve the usability of the query history and query image scoring functionalities by showing thumbnail images in query history section, ranking the query images by slide bar or dragging and dropping to a different position in a query. They do not have many comments on the negative functionalities.

UndefinedRC Users: The undefinedRC users like an accurate and rich functionality search system. They are more likely to think the negative query functionality is useful in the evaluation systems. They also like the negative result functionality because they think they get to know the data collection quality better by seeing the negative results in the result panel. They are more likely to think the query history functionality is useful. Like the risky user, they also feel the search accuracy drops with more image examples included in a query. They feel the functionalities are more useful when they perform more difficult tasks. Their initial search idea changes during the search and the evaluation systems support the change well. They think the usability of the functionalities can be improved by showing query history automatically rather than having to press reset, showing diverse negative results rather than based on colour only, showing page number, etc.

6.2 Search preferences of the user types based on the Information Goal criterion

Users with Fixed Goal: The users with fixed information goal prefer accurate systems and rich functionalities. They are satisfied with the search results, and they are basically satisfied with the evaluation systems, but they think that the usability of the interface needs to improve. They have a clear information goal in mind before starting the search, and the goal does not change during the search. They feel the search results get increasingly better with every query refinement. They find that it is easy to make decisions on results selection. Whilst they prefer the content-based search, they also like keyword-based search, thus they suggest combining content-based and keyword-based image search. They like all the functionalities provided, but again they think the usability of some functionalities can be improved, such as ranking query images by a scale bar or dragging and dropping, showing image thumbnails in the query history section, starting with keyword-based search, etc.

MixedFE Users: The MixedFE users have less expectations of the system accuracy than the uses with other characteristics, but they have the same expectation with the other types of users on rich functionalities to support different search aspects. They think the tasks are interesting and clear. They think all the provided functionalities are useful especially the query image scoring and query history functionality. They also find that the functionalities are more useful when they perform more difficult tasks. Whilst they prefer the content-based search, they also like keyword-based search and the combination of keyword-based and content-based search. As they tried many functionalities for completing the tasks, they provide lots of suggestions on improving the functions, such as, ranking query images by a slide bar, showing image thumbnails in query history section, providing a colour histogram or pie chart for selecting negative colour examples, etc.

NoneFE Users: Comparing to the users with other characteristics, the NoneFE users like good a quality and large data source, a fast and easy to use system, accurate search results and rich functions to support different search aspects. These users are satisfied with both the search results and the evaluation systems. They think the tasks are fairly interesting and clear. They do not know what they are looking for before they start search. They think the negative query is extremely useful. They also like the query image scoring function. They sometimes find it is hard to decide the relevance of the results for the tasks. Their idea changes during the search and the systems support the changes very well. They also suggest making the negative results optional. As with other types of users, they think the usability of the interface can be improved by providing drag and drop and image zooming functionality.

UndefinedFE Users: The UndefinedFE users believe that the image search tools are accurate and easy to use, and have rich functionalities. They are more satisfied with the search experience with the systems than with the search results because they judge the system accuracy based on the complexity of the tasks. They are satisfied with the search results when they perform easier tasks, and they are not satisfied with the search results when they perform harder tasks. They think the positive and negative feedback functions are useful. They suggested improvements to the interfaces of the evaluation systems, such as adding drag and

drop, and image zoom functionality and providing diverse negative results rather than based solely on colour only, etc.

In this section, we categorized the 50 users into different groups based on the ISE user classification model. We have found that only the users grouped based on the search strategy and information goal criteria are relatively evenly spread to every characteristic, so we decided to discuss the characteristics of these two criteria only. After grouping the 50 subjects into the characteristics, we have extracted some qualitative data from the questionnaires. Through carrying out a content analysis on the qualitative data of the different user types, we have found clear evidence concerning users' different search preferences in terms of expectation of image search tools, search experience, and suggestions to improve the search systems.

7. CONCLUSION AND FUTURE WORK

In this paper, we reviewed the importance, challenges and gaps for developing and utilising user models in an effort to provide personalised search and to improve users' search experiences. A user classification model based on IFT called ISE model was introduced and applied to a CBIR search system called uInteract. An empirical user study set up was described and carried out to collect user interaction data. A methodology was proposed for the utilisation of the ISE model. We successfully applied the ISE model based on the methodology to identify different user types based on extensive analysis of the user interaction data collected from the user study. These user types were then applied to investigate the users' search preferences for user-centred and personalised CBIR search system design and development.

The key contributions of this work are as follows: we proposed a new methodology for the utilisation of the theory-based ISE user model, so that different user types can be effectively identified; we established the search preferences of different user types through successfully applying the ISE model to the uInteract CBIR search system based on the proposed methodology, which provided useful insights for effective design and development of personalised CBIR search systems. Although our work focused on the CBIR search systems, we believe that the methodology and the findings could be adapted to text-based search systems.

In the future we plan to design and evaluate the personalised CBIR search interfaces for different user types based on the findings of this work. We would like to test the proposed methodology to different user data. For example, the text-based search users' log data. We would also like to automate the user model utilisation process so that the user models can be embedded into the current search models.

8. REFERENCES

[1] Collins-Thompson, Kevyn, Bennett, Paul N., White, Ryen W., Chica, Sebastian de la and Sontag, David. (2011) Personalizing web search results by reading level. In Proceedings of the 20th ACM international conference on Information and knowledge management (CIKM2011), ACM, 403-412.

[2] Keenoy, Kevin and Levene, Mark (2003) Personalisation of web search, Proceedings of the 2003 international conference on Intelligent Techniques for Web Personalization, Mexico

[3] Lew, Michael S., Sebe, Nicu, Djeraba, Chabane and Jain, Ramesh (2006). Content-based multimedia information retrieval: State of the art and challenges. ACM Transactions on Multimedia Computing, Communications and Applications 2 (1), 1–19.

[4] Liu, Haiming, Mulholland, Paul, Song, Dawei, Uren, Victoria and Rüger, Stefan (2010) Applying Information Foraging Theory to Understand User Interaction with Content-based Image Retrieval, The Information Interaction in Context conference (IIiX2010), pp. 135-144, ACM

[5] Liu, Haiming, Uren, Victoria, Song, Dawei and Rüger, Stefan (2009a). A four-factor user interaction model for content-based image retrieval. In Proceeding of the 2nd International Conference on the Theory of Information Retrieval (ICTIR), pp 297–304.

[6] Mulholland, Paul, Zdrahal, Zdenek and Collins, Trevor (2008). Investigating the effects of exploratory semantic search on the use of a museum archive. In Proceeding of the IEEE 2008 International Conference on Distributed Human-Machine Systems.

[7] Neuendorf, Kimberly A. (2002) : The Content Analysis Guidebook Thousand Oaks, CA: Sage Publications.

[8] Pirolli, Peter and Card, Stuart K. (1999). Information foraging. Psychological Review 106, 643–675.

[9] Shen, Xuehua, Tan, Bin and Zhai, ChengXiang. (2005) Implicit user modeling for personalized search. In Proceedings of the 14th ACM international conference on Information and knowledge management (CIKM '05). ACM, 824-831.

[10] Urban, Jana, Jose, Joemon M. and van Rijsbergen, Keith (2006). An adaptive technique for content-based image retrieval. Multimedia Tools and Applications 31, 1–28.

[11] White, Ryen W. and Drucker, Steven M. (2007) Investigating behavioral variability in web search. In Proceedings of the 16th international conference on World Wide Web (WWW2007). ACM, 21-30.

[12] White, Ryen W. and Roth, Resa A. (2009). Exploratory search: Beyond the query - response paradigm. Morgan and Claypool Publishers.

[13] Wilson, M. L., schraefel, m.c. and White, R. W. (2009), Evaluating advanced search interfaces using established information-seeking models. J. Am. Soc. Inf. Sci. (JASIST), 60: 1407–1422. doi: 10.1002/asi.21080

[14] Wu, Chun-Che, Mei, Tao, Hsu, Winston H. and Rui, Yong. (2014). Learning to personalize trending image search suggestion. In Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval (SIGIR '14). ACM, New York, NY, USA, 727-736.

[15] Xie, P., Pei, Y., Xie, Y. and Xing, E.P. (2015) Mining User Interests from Personal Photos, In Proceeding of the 29th AAAI Conference on Artificial Intelligence (AAAI 2015).