2011

Labeling Images with Queries: A Recall-based Image Retrieval Game Approach

Jun Wang
Syracuse University, jwang72@syr.edu

Bei Yu
Syracuse University, byu@syr.edu

Recommended Citation
Jun Wang and Bei Yu (2011) Labeling Images with Queries: A Recall-based Image Retrieval Game Approach. Proceedings of the SIGIR 2011 Workshop on Crowdsourcing for Information Retrieval.

Labeling Images with Queries: A Recall-based Image Retrieval Game Approach

Jun Wang
School of Information Studies
Syracuse University
Syracuse, New York 13244
junwang4@gmail.com

Bei Yu
School of Information Studies
Syracuse University
Syracuse, New York 13244
byu@syr.edu

ABSTRACT

Human computation game-based image labeling has proven effective in collecting image descriptions for use in improving image searching, browsing, and accessibility. ESP and Phetch are two successful image description games: the ESP game is designed for collecting keyword descriptions while the Phetch game for sentence descriptions. In this paper we propose and implement an image retrieval game for collecting query-like descriptions. The idea is that a player is presented with a target image for a short time and then needs to construct a query based on his recall of the image, with the goal of making the target image ranked as high as possible by an image search engine. We have conducted preliminary analysis of the queries collected from players who were recruited through the Amazon Turk. Our result shows that the image retrieval game enabled us to collect queries that are comparable to the real queries reported in existing studies, and it can also be used to collect informative tags while avoiding undesired ones such as trivial color words.

Categories and Subject Descriptors

H.3.m [**Information Retrieval**]: Miscellaneous; H.5.3 [**HCI**]: Web-based interaction

General Terms

Design, experimentation, human factors

Keywords

Human computation, games with a purpose, image retrieval, image tagging, query analysis, user search behavior, ESP game

1. INTRODUCTION

Current image search engines on the web mainly use text-based clues to index images. Textual data such as filenames and surrounding text can be misleading or insufficient, making a large fraction of search results unrelated to the image queries. The best approaches available today for indexing or labeling images are still human-based: humans are far better than computers to understand and describe the content of images. However, manually labeling a vast amount of images is tedious and extremely costly.

Human computation games are the recently emerging approach to make tedious tasks enjoyable to do: tasks are solved by players as a side product of gameplay. A classical example of using games to label images is the ESP game, in which two randomly assigned players need to agree on a word to describe an image [12]. As a side product of gameplay, the words generated by the players can be used to label the images played. Another example is the Phetch game that is played by an image describer and multiple image seekers: the describer needs to generate an explanatory description of an image so that the seekers, when receiving the description, can translate it into a query to find the image [13]. As a side product of gameplay, the descriptions generated by the describer can be used to accurately describe the images played.

In this paper we propose and implement a single-player image retrieval game for collecting query-like descriptions. The key idea is that a player is presented with a target image for a short time and then needs to construct a query, based on his recall of the image, to find the target image. Using the image retrieval game to collect queries for images, we are interested in answering the following questions:

1) What are the characteristics of the user-generated image queries? Do they have parallels with earlier studies on image queries and user search behavior?

2) What are the characteristics of the tags aggregated from the queries? How are the tags similar to or different from those collected in the ESP game?

2. GAME DESIGN

Our image retrieval game is designed to be played by one player. In the game, a player is presented with an image that is only displayed for a very brief time and then fades away. After the image completely disappears, the player is asked to enter a query to a search engine to retrieve back the image. To encourage players to construct good queries, we used the following incentive mechanism. Suppose the rank position of the target image in the search results is r, the amount of points that a player can get is given by:

$$score(r) = \begin{cases} 10(k - r + 1) & \text{if } r \leq k \\ 0 & \text{otherwise} \end{cases}$$

Proceedings of the SIGIR 2011 Workshop on Crowdsourcing for Information Retrieval, July 28, 2011, Beijing, China.

This is a revised version of the original one which appeared in the workshop.

where k is the maximum number of images that a search engine can return in a search result page. In our case, we set the page limit $k = 10$. If the search engine has a large image collection, we can set k to be a larger number such as $k = 20$ so that the player will not find it too difficult to play the game.

Figure 1 gives two screenshots of the game interface. In the screenshot (a), the player entered a query "walk" based on his recall of a disappeared target image. Among the 27 images indexed with the term "walk" by the search engine, the target image was ranked 25th, beyond the page limit of $k = 10$, and so the player failed in this trial. Note that a player can click the disappeared image to take a peek at it again, but he only has a limited number of such "cheating" chances in a 3-minute game. In the screenshot (b), the player added one more word "shadow" to the previous query, and this time he succeeded — the target image was ranked 6th, within the limit of $k = 10$, and he got 50 points.

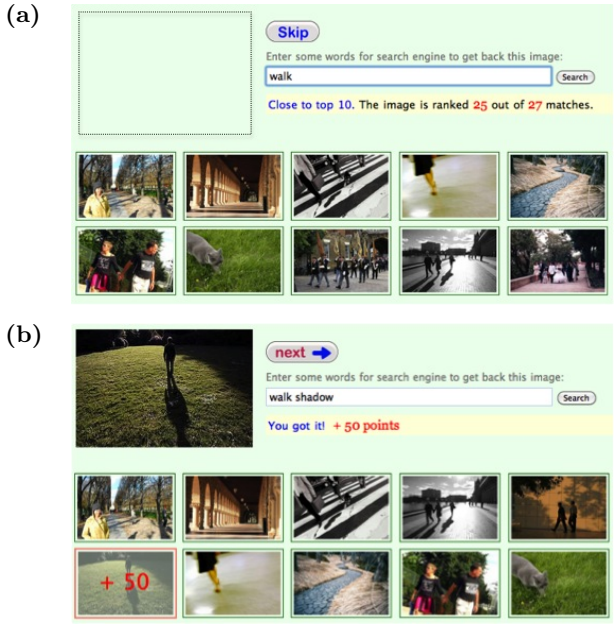


Figure 1: Screenshots of the game interface.

A unique feature of our image retrieval game is that the image is only presented to players very briefly and then fades away over time. A player is not allowed to enter a query until the image has completely disappeared. There are several reasons for this design. (1) Recall-based retrieval could be used to approximate the actual cognitive process of searching known images. Users usually generate queries based on an idea in their head about what they are searching for. (2) There is a capacity limit associated with our short-term working memory, meaning that our brain will only pick up the most salient features in a picture while ignoring others. This seems to meet our expectation of what constitutes good tags: those that can give people the sense of what is important in an image. By contrast, if we are allowed to tag an image while observing it, we may be easily distracted by many trivial visual features such as various colors appeared in an image. (3) Recall-based approach can make the game novel, interesting, as well as more challenging [14]. For example, one can increase the challenge level of the game by setting the display time of images to be shorter and shorter.

3. THE IMAGE SEARCH ENGINE

The image search engine is a crucial part of the game design for which the purpose is to encourage players to generate good queries. It is expected that for a good query, the target image should be ranked as high as possible by the search engine. Therefore, the image collection indexed by the search engine should be of a moderate size. If the collection is too small, it would be easy for the engine to return a target image in the top-k, and so the game could be boring; if it is too large, the chance for a player to find an image might be very low, and so he could easily get frustrated.

As in the Phetch game [13], we also used the ESP game dataset, which is publicly available at `http://www.cs.cmu.edu/~biglou/resources/`. The ESP dataset includes 100k images, each associated with on average 14.4 labels that we can use for indexing images. Another reason for us to use the ESP dataset is that the ESP tags can be used as a baseline for evaluating our game. For simplicity, we only used 1000 images and corresponding labels to build our image search engine. These 1000 images were randomly selected from the 6000 images whose thumbnails are around the size of 240x160, for the sake of nice layout of search results (see Figure 1). Then, out of the 1000 images, 101 images [1] that do not include any explicit words were chosen as the target images for our participants to play. We will show later that with even only 1000 images in the collection, for players who were from the US (suggesting it is very likely that they were native English speakers), only 37% of the queries they constructed can lead to successful searches.

For the information of the readers, here we show a naive algorithm that we used to compute the relevance between a query that includes m words, $q = \{w_1, \cdots, w_m\}$, and an image or picture that is indexed with n tags, $p = \{t_1, \cdots, t_n\}$. Denote by $v = \{v_1, \cdots, v_l\}$ the set of words shared by both p and q. Then the relevance score between p and q is given by

$$relevance(p, q) = \sum_{i \in \{1, \cdots, l\}} \log \frac{N}{f_i}$$

where f_i indicates the number of images that were indexed by word v_i and N the total number of images. The symbol $\log N/f_i$ gives the *idf* (inverted document frequency) of word v_i. Note that since the ESP dataset does not provide any information about how frequently a tag was used to label an image, we do not have the *tf* (term frequency) item in the above formula. When two images have the same relevance score, their rank order will be determined by the number of tags that they contain: the fewer tags an image is associated with, the higher the rank of the image. To enhance users' experience, the retrieval algorithm can be improved in many ways. For example, newly user-generated queries can be integrated into the relevance function, with caution that some players may enter queries aiming to cheat the system.

[1] We originally planned to select 100 target images, but later on realized that we actually ended up with 101 images.

4. PRELIMINARY RESULTS

In our preliminary study, we were interested in answering the following questions:

1) Will people enjoy playing the image retrieval game?
2) What are the characteristics of the user-generated image queries? Do they have parallels with earlier studies on image queries and user search behavior?
3) What are the characteristics of the tags aggregated from the queries? How are the tags similar to or different from those collected in the ESP game?

We recruited participants through Amazon Mechanical Turk. We set up two HITs there: one was for players in the US, and the other for players in India. The reason for recruiting two groups of players is that we also wanted to study the difference between native and nonnative speakers. Obviously, player's language skill can play an important role in the way they describe images or formulate queries. A comparison between native and nonnative speakers will be done in the future.

The game was designed for 3-minute play with an incentive of $0.15 for the Amazon Turk workers, who can play as many times as they liked. But since there was only one HIT for each worker, they would not get paid for extra plays. When a game was over, the player would receive a passcode for approval of his work. In addition to playing the game, a worker also needed to provide some demographic information, including their native language.

The description of the game that we posted on the Amazon Turk was: *"This is a study about how users generate queries so that a given image can be found or ranked in the top 10 by an image search engine."*

4.1 Enjoyability of the game

Within about 10 days, we recruited 373 players through the Amazon Turk. Here, a player refers to one who played at least once a whole 3-minute game. Among the 373 players, 342 were those who submitted their passcode for approval of work, and we call them *workers*, including 155 from India and 187 from the US. Actually, we were able to recruit the 155 India workers within three days, but it took us 12 days to recruit the 187 US workers.

Table 1: Distribution of players and play times.

Play times	1	2	3	4+	Sum
US players	147	31	6	3	187
	78.6%	16.6%	3.2%	1.6%	
India players	121	31	19	14	185
	65.4%	16.8%	10.3%	7.6%	
All players	269	62	25	17	373
	72.1%	16.6%	6.7%	4.6%	

Table 1 shows that, in general, about 28% players played at least two games. It is worth noting that there was an India player who played 17 games for over one hour. We also want to point out that a feature in our game design is that at the end of a 3-minute game, a player can see a short list of top players and recent game scores. We speculated that this feature, along with other design factors, played an important role in motivating people to play twice or more.

4.2 Query analysis

We wondered if the queries collected in the context of gameplay can demonstrate typical search behavior. A challenge in studying user search behavior is that it is expensive and time-consuming to collect large-scale query data in lab studies, in which the searcher's intent can be controlled [1, 2]. If our game can generate queries somewhat similar to the queries collected in labs, we may use it to collect large-scale query data at minimum or no cost.

From now on we only report the results of our preliminary analysis of the queries generated by the *US players*, for the purpose of making the results comparable to earlier studies. But for the information of the readers, we want to mention that the queries from our India players were significantly shorter than those from the US players, indicating the effect of language skill. Analysis of the difference between native and nonnative speakers remains to be future work.

Firstly, we report some general statistics of the collected image queries, and show that our result largely agrees with earlier studies by Goodrum & Spink [5] and Jorgensen & Jorgensen [8]. Overall, our game has generated 4,500 queries from 201 US players (including 14 players who did not finish the whole 3-minute game), and 37% queries lead to successful searches (i.e., a target image can be ranked by the search engine in the top 10). The average number of queries constructed for each target image is 1.86, and the average number of words used in each query is 2.56. The comparison given in Table 2 shows that our result agrees well with earlier studies on analyzing image queries.

Table 2: Statistics of image queries

Source	Users	Query tries[a]	Query length
present study	game players	1.86	2.55 words
[5]	Excite users	3.36	2.74 or less[b]
[8]	professionals	2.10	1.87 terms[c]

[a] *Query tries* refers to the number of queries attempted by a player in an image search task.
[b] One or more words such as "picture" and "jpeg" were used for specifying images and image formats. Therefore, the number of words per query, 3.74, as reported in [5], was actually 2.74 or less [8].
[c] Some terms such as "New York" include two or more words.

The number of queries attempted in our search task, 1.86, compared with other studies, is relatively small. This can be explained by the observation that 27% players, when failed in the initial search, just gave up the current search task and moved on to the next task, and as a result we have a relatively smaller number of queries per search task. We speculated that this would only happen in game settings since for some players their goal was to achieve high points. Actually it was the case in our data that a player, on average, would have less and less chance to find an image if they failed in previous searches.

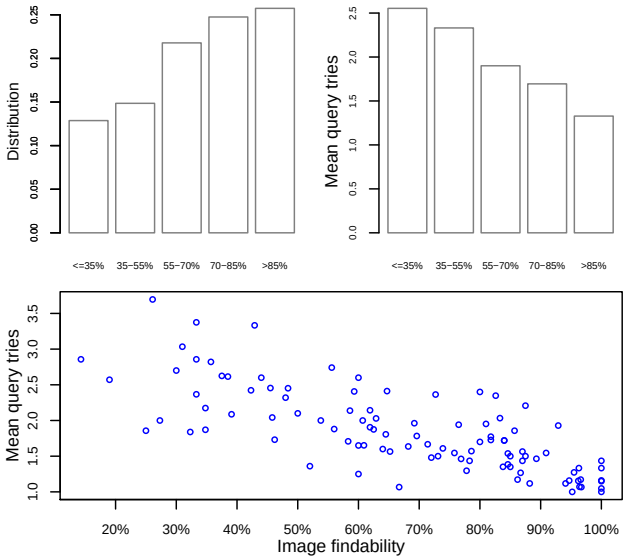


Figure 2: The number of queries attempted in an image search task decreases as the findability of the image increases.

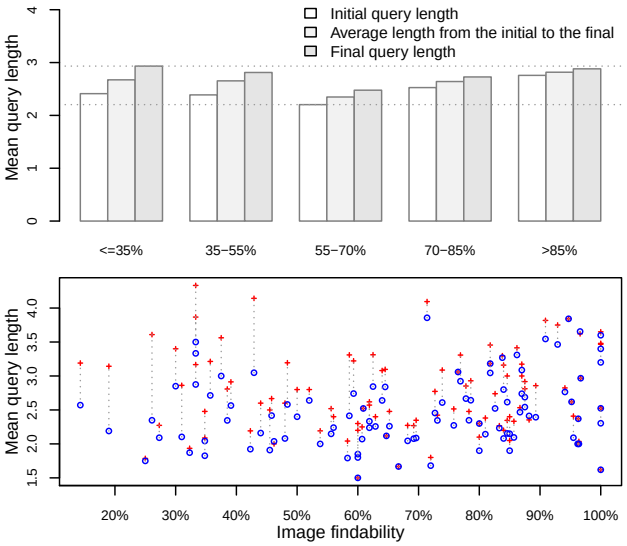


Figure 3: The number of words in queries tends to increase as an image becomes hard to find. Blue-circled points correspond to *initial queries*, and the red-crossed correspond to *final queries*.

Secondly, we report our finding regarding how difficult search tasks affect search behavior, and show that our result is in line with and complements the finding reported in a recent study conducted by Aula et al. [2]. In [2], the difficulty of a search task was measured as the proportion of users who reported success for the task. In our study, similarly, the difficulty of finding a specific image is measured as the proportion of successful searches —the success could happen at the initial query attempt or a later refined one. We use proportion of *successful searches* instead of *successful users* because in our case one user may encounter the same image search task during different games. For convenience, we will use *image findability* to refer to the inverse of task difficulty.

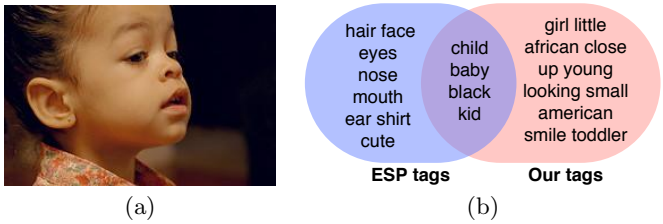
Figure 2 gives the mean number of queries attempted per image (query tries) as a function of image findability. It shows that the more findable an image, the fewer queries a player would issue to search for the image. Each data point in the figure corresponds to one image; totally we have 101 images or data points. For example, the left most point represents the image given in Figure 4(a). It was played by 21 players (each played once), and only 3 of them succeeded in finding it, resulting in a 14.3% success ratio or findability. On average, these players tried 2.86 queries per image. For detailed queries performed by the players on this image, see Table 4 in the appendix.

Figure 3 gives the mean number of words per query (query length) as a function of image findability. It shows that for hard-to-find images, the query length tends to increase, indicating users tended to add more descriptions in their image queries. This phenomenon is related to the design of the search interface in which the previous query is kept in the search box until a new search task starts or the player modifies it. It is also worth noting that the highly findable images —those whose findability is over 85%— are associated with relatively longer initial queries. This relationship suggests that easy-to-find images seem to be able to trigger players to come up with more words in their mind, resulting in long initial queries.

These preliminary results, though not comprehensive, suggest that our game can generate queries that are comparable to the real queries reported in earlier studies, and has the potential to be used in the future to collect large-scale query data at minimum or no cost.

4.3 Comparison to the ESP tags

As shown above, the queries collected using the image retrieval game can provide information about user search behavior. The queries can also be used to acquire useful tags for annotating images. For example, if queries constructed by two or more different players agree on a word for an image, then the word could be a useful tag for the image.



(a) (b)

Figure 4: An example

By comparing the tags from our game to those from the ESP game, we show that our game can be used to collect high-quality tags while avoiding undesired ones. Firstly, we show how our tags are similar to or different from the ESP tags through an example. The image used for the example is given, again, in Figure 4(a). The queries generated by the US players for this image are given in Table 4 in the

appendix. Totally 15 words were used by at least two players in their queries, as shown in the following table along with their frequency:

child 12	girl 11	little 6	black 6	up 4
toddler 4	close 3	african 3	young 3	baby 3
kid 3	looking 2	small 2	american 2	smile 2

For the ESP game, it generated 12 tags[2], as shown by the words in the left circle in Figure 4(b). The Venn diagram in Figure 4(b) clearly illustrates the difference between our game and the ESP game. Half of the ESP tags are about body parts, which are usually not desired. By contrast, our game did not elicit such undesired words, and in particular, it was able to generate many words that were relatively more informative, e.g. "toddler" and "african american."

That our game outperformed the ESP game in terms of the ability of obtaining *informative words* can be explained as follows. The optimal strategy for a player to take in the ESP game is to enter general or obvious words because it is far easier for two people to agree on general words than to agree on specific words [11]. By contrast, the best strategy to use in our game is to enter informative words in their queries so that a specific image can be selected out from within an image collection by a search engine.

Secondly, we present a preliminary result of an overall comparison analysis on the 101 images played. The average number of tags acquired for each image is 12.5, a number that will gradually increase as the game is played by more people. Correspondingly, the average number of tags per image from the ESP game is 17.5.

Table 3: Top 15 words that were significantly used in one group (Ours or ESP) but not in the other group.

Ours	p-value	ESP	p-value
and	2e-07	cloud	0.0007
in	1e-06	blue	0.0009
on	1e-05	brown	0.001
with	6e-05	grey	0.003
of	2e-04	hair	0.006
a	0.01	gray	0.007
young	0.01	nose	0.02
little	0.03	sun	0.02
the	0.03	eyes	0.02
at	0.05	tree	0.03
african	0.05	yellow	0.04
playing	0.07	floor	0.06
walking	0.07	orange	0.07
nature	0.07	light	0.08
smiling	0.08	pink	0.09

Table 3 shows which words were significantly used in our game but not by the ESP game and the other way. The p-value of each word is calculated using Pearson's χ^2 test. For example, for word "blue", the ESP game has 46 images associated with it, and our game has 11 images (which are also included in the 46 images in this case). Considering that the average number of tags per image in our game is 12.5 and it is 17.5 for the ESP game, we scaled down the number 46 by the ratio of 12.5/17.5 to 33. Alternatively, if we scale up the number 11 to 15 the other way, we will have much smaller p-value but the order of those words in the list is essentially the same. The p-value of 0.0009 gives how the "observed" distribution 33:11 differs significantly from what would be expected to occur *by chance*, by which we mean that half of the 33+11 images associated with "blue" will be from the ESP and half from our game.

From the comparison given in the above table, we can see that our game is good at collecting function words (e.g., "and" and "in"), adjective words (e.g., "young" and "little"), and gerund forms of verbs (e.g., "playing" and "walking"). By contrast, the ESP can elicit from players many undesired words such as color words (e.g. "blue" and "brown"), common backgrounds and scenes (e.g., "cloud" and "tree"), and body parts (e.g., "hair" and "nose").

It is easy to understand that our game can avoid eliciting undesired tags from players because these tags are neither informative nor salient features of an image. However, it is interesting to see that many players also used function words in their queries, despite function words provide no information for retrieving images. An explanation is that people have the tendency to use function words to help organize their thoughts. This may imply that function words, especially those signaling the structural relationships among components of an image, could be used by image search engines to improve search performance.

It is also interesting to discover that many queries generated in our game include the "-ing" or gerund form of verbs. Correspondingly, the ESP game has relatively much fewer gerund forms. This could be related to the way that people describe activities shown in an image. For example, here are two queries generated by different players to describe a picture: "couple *gazing* into each others eyes," and "man and woman *looking* at each other." See Table 4 and 5 in the appendix for more such cases. In the future it will be interesting to study whether actual image seekers also tend to use the gerund form of verbs in their queries.

5. RELATED WORK

Since Luis von Ahn's pioneering work of labeling images with a computer game [12], many human computation games, also called *games with a purpose*, have been proposed to harness the intelligence of crowds in an enjoyable way. Here we address two games that are closely related to our work. Some other image labeling games (e.g. KissKissBan [6]) are also related, but they are not of particular interest regarding the purpose of the present study. For a comprehensive introduction of human computation games, check the latest book by Law & von Ahn [9].

The first one that is of particular interest is the Phetch game [13] that we mentioned above. Phetch is a game played by one image describer and 2 to 4 image seekers. The describer is given an image and broadcasts his description of the image to all the seekers. The seekers then need to translate the description into a query so that they can find the image using an image search engine. The describer and the first

[2]The obvious word "girl" is missing in the ESP tag list, suggesting it is a taboo or off-limits word [12].

seeker to find the image get points. Different from the ESP game, the design purpose of Phetch is to collect accurate descriptions for images. In contrast to the Phetch game, in our game, the seekers see the image before seeking, and have to construct queries based on their recall of the image just seen. When people are searching for an image, especially for a known item, they might have some clue in their head about what they are searching for. Our game was designed to more closely approximate that model of search. In addition, despite the Phetch game can be used to collect queries [13], there has been no further study on analyzing the queries and descriptions collected in the Phetch game.

The second one is the Page Hunt game, an information retrieval game [10]. In Page Hunt, a player is given a web page and needs to generate a query that can make the page ranked in the top-k list by a real search engine like Bing. Similarly, in our game, the goal of a player is to search for a given image. A design defect in Page Hunt seems to be that if a player simply enters a quoted, sufficiently long sequence of words occurred in the web page that he is searching for, he would easily get away with a high score. If this is the case (we could not test it any more because the game site has been down), the queries generated would be of much less value. To address the issue, one may take the approach used in the present paper: a page is only presented for a short time and the player is not allowed to enter queries until the page has completely disappeared. In this design, a player needs to generate queries based on his recall of the page instead of simply adopting words from the page.

6. CONCLUSION

In this paper, we proposed to use a single-player image retrieval game to collect queries for images. We have conducted preliminary analysis of the queries collected from participants through Amazon Mechanical Turk. We have shown that the game enabled us to collect queries that are comparable to the real queries reported in earlier studies, and it can also be used to collect high-quality tags while avoiding undesired ones such as trivial color words, common backgrounds, and body parts.

A limitation of the present study is that we have not yet investigated whether the outcome that our game was able to avoid undesired tags was caused mainly by the recall-based approach or by the image retrieval game, or by both. Our guess is that they both had an effect, but further work needs to do to clarify it. If the recall-based approach also plays an important role, it may be used in the ESP game to improve its quality [4, 11].

Future work also includes exploring in detail the strategies used by successful or unsuccessful searchers to formulate and reformulate their queries, and how search engines can be designed to help users to learn better search strategies [1, 3, 7, 15]. We will also study the difference of search behaviors between native and nonnative speakers.

Acknowledgement

We want to thank those Amazon Mechanical Turk players for their anonymous but indispensable contribution. We also thank the anonymous SIGIR 2011 Workshop reviewers for their very helpful and insightful comments and suggestions.

7. REFERENCES

[1] M. Ageev, Q. Guo, D. Lagun, and E. Agichtein. Find it if you can: a game for modeling different types of web search success using interaction data. In *SIGIR '11*, pages 345–354, New York, NY, USA, 2011. ACM.

[2] A. Aula, R. M. Khan, and Z. Guan. How does search behavior change as search becomes more difficult? In *CHI'10*, pages 35–44, 2010.

[3] A. Aula and K. Nordhausen. Modeling successful performance in web searching. *Journal of the American Society for Information Science and Technology*, 57(12):1678–1693, 2006.

[4] D. H.-L. Goh, R. P. Ang, C. S. Lee, and A. Y. Chua. Fight or unite: Investigating game genres for image tagging. *Journal of the American Society for Information Science and Technology*, 62(7):1311–1324, 2011.

[5] A. Goodrum and A. Spink. Image searching on the excite web search engine. *Information Processing & Management*, 37(2):295–311, 2001.

[6] C.-J. Ho, T.-H. Chang, J.-C. Lee, J. Y.-j. Hsu, and K.-T. Chen. KissKissBan: a competitive human computation game for image annotation. In *Proceedings of the ACM SIGKDD Workshop on Human Computation*, HCOMP '09, pages 11–14, New York, NY, USA, 2009. ACM.

[7] J. Huang and E. Efthimiadis. Analyzing and evaluating query reformulation strategies in web search logs. In *CIKM '09*, pages 77–86. ACM, 2009.

[8] C. Jörgensen and P. Jörgensen. Image querying by image professionals. *Journal of the American Society for Information Science and Technology*, 56(12):1346–1359, 2005.

[9] E. Law and L. von Ahn. *Human Computation*. Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool Publishers, June 2011.

[10] H. Ma, R. Chandrasekar, C. Quirk, and A. Gupta. Improving search engines using human computation games. In *CIKM '09*, pages 275–284. ACM, 2009.

[11] S. Robertson, M. Vojnovic, and I. Weber. Rethinking the ESP game. In *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems*, CHI EA '09, pages 3937–3942, New York, NY, USA, 2009. ACM.

[12] L. von Ahn and L. Dabbish. Labeling images with a computer game. In *CHI'04*, pages 319–326, 2004.

[13] L. von Ahn, S. Ginosar, M. Kedia, R. Liu, and M. Blum. Improving accessibility of the web with a computer game. In *CHI'06*, pages 79–82, 2006.

[14] J. Wang and B. Yu. Sentence recall game: A novel tool for collecting data to discover language usage patterns. In *Proceedings of the ACM SIGKDD Workshop on Human Computation*, HCOMP '10, pages 56–59, New York, NY, USA, 2010. ACM.

[15] R. W. White and D. Morris. Investigating the querying and browsing behavior of advanced search engine users. In *SIGIR'07*, pages 255–262, 2007.

Appendix

Table 4: Queries for the image in Figure 4(a)

Player	Query
1	child child family
2	little girl close up
3	child child girl child girl foreign child girl foreign african child girl foreign african dark child girl foreign african dark toddler
4	baby black
5	girl black girl black child girl black child smile
6	little child close up little kid little kid in dress
7	child face
8	toddler, African American, baby
9	child young boy baby child
10	black child girl little girl black child
11	young girl singing young hispanic girl
12	girl asian young little looking girl asian young girl indian young
13	child toddler
14	child wonder child wonder small child wonder small close up child wonder small african child wonder small african american
15	toddler, girl,
16	child kid smile
17	little girl child little girl black girl little black girl
18	child kid question quizzical
19	little girl floral shirt
20	small girl kimono small black girl kimono kimono cute girl kimono small girl
21	girl looking hair up bun collar girl looking hair up bun collar black girl looking hair up bun collar black portrait girl looking hair up bun collar black portraite girl looking hair up bun collar black portrait mouth open

↑ ↗

Table 5: Queries for the image at the left-bottom of this page. "Bagpipe(s)" were not indexed for the image.

1	bagpipes bagpipes parade
2	black and white bagpipe player
3	Black and white of man playing bagpipes Black and white photo of man playing bagpipes Black and white photo of heavy man playing bagpipes Black and white photo of heavy man playing bagpipes with others in the background Black and white photo of heavy older man playing bagpipes with others in the background close up of man with bagpipes close up of man playing bagpipes
4	bagpipes man parade suit
5	man bagpipe bagpipes scottish scotish
6	black and white bagpipe black and white bagpipes march
7	march parade scottish irish
8	bagpope bagpipe bagpipe man bagpipe bag pipe bag \pipe bag pipe man man instrument
9	bagpipes
10	bagpipe bagpipes scottish
11	men with bagpipes and microphone
12	bagpipes, man, black and white, photograph, Scotland
13	men with bagpipes and microphone mad men with bagpipes and microphone mad men
14	pipe bagpipe smoking
15	bagpipes bagpipe bagpipe scott bagpipe scott playing bagpipe players bagpipe players music bagpipe players music scottish
16	man bagpipe flag man bagpipe flag black and white man bagpipe flag black and white hat man bagpipe flag black and white hat army man bagpipe flag black and white hat man bagpipe flag black white hat man bagpipe flag black white hat music
17	bagpipe bagpiper bagpipers
18	man bagpipe man bagpipe marching
19	black and white bagpipe black and white bagpipe old man black and white bagpipe old man and kid black and white bag pipes old man and kid
20	man, pipes, black and white man, pipes, black and white, scottish, bagpipes
21	bagpipe playing bagpipes bagpipe parade
22	bagpipers bagpipers black and white black and white bagpipers black and white photos of outdoor bagpipers b/w photos of outdoor bagpipers photos of men outdoor bagpipers photos of men outdoors playing bagpipes