Cross-Language Peculiar Image Search Using Translation between Japanese and English

Shun Hattori
School of Computer Science
Tokyo University of Technology
1404-1 Katakura, Hachioji, Tokyo 192-0982, Japan
Email: hattori@cs.teu.ac.jp

***Abstract*—Most researches on Image Retrieval (IR) have aimed at clearing away noisy images and allowing users to retrieve only acceptable images for a target object specified by its object-name. We have become able to get enough acceptable images of a target object just by submitting its object-name to a conventional keyword-based Web image search engine. However, because the search results rarely include its uncommon images, we can often get only its common images and cannot easily get exhaustive knowledge about its appearance (look and feel). As next steps of IR, it is very important to discriminate between "Typical Images" and "Peculiar Images" in the acceptable images, and moreover, to collect many different kinds of peculiar images exhaustively. In other words, "Exhaustiveness" is one of the most important requirements in the next IR. As a solution to the 1st next step, my previous work has proposed a novel method to more precisely search the Web for peculiar images of a target object by its peculiar appearance descriptions (e.g., color-names) extracted from the Web and/or its peculiar image features (e.g., color-features) converted from them. This paper proposes a refined method equipped with cross-language (translation between Japanese and English) functions and validates its retrieval precision.**

***Keywords*-cross-language retrieval; content-based image retrieval (CBIR); text-based image retrieval (TBIR); machine translation; Web search; Web mining; text mining;**

I. INTRODUCTION

In recent years, the Web has had exploding Web images as well as Web documents (text), and various demands have arisen in retrieving Web images as well as Web documents to utilize this information more effectively. When a name of a target object is given by a user, the main goal of conventional keyword-based Web image search engines such as Google Image Search [1] and most researches on Image Retrieval (IR) is to allow the user to clear away noisy images and retrieve only the acceptable images for the target object-name, which just include the target object in their content, as precisely as possible. However, the acceptable images for the quite same object-name are of great variety. For instance, in different shooting environments such as angle, distance, or date, in different appearance varying among individuals of the same species such as color, shape, or size, with different background or surrounding objects. Therefore, we sometimes want to retrieve not only vague acceptable images of a target object but also its niche images, which meet some kind of additional requirements. One example of more niche image searches, when not only a target object-name and also impressional words as additional conditions are given, allows the user to get special images of the target object with the impression [2–4].

Another example of more niche demands, when only a name of a target object is given, is to search the Web for its "Typical Images" [5] which allow us to adequately figure out its typical appearance features and easily associate themselves with the correct object-name, and its "Peculiar Images" [6], [7] which include the target object with not common (or typical) but eccentric (or surprising) appearance features. For instance, most of us would uppermost associate "sunflower" with "yellow one", "cauliflower" with "white one", and "Tokyo tower" with "red/white one", while there also exist "red sunflower" or "black one" etc., "purple cauliflower" or "orange one" etc., and "blue Tokyo tower" or "green one" etc. When we exhaustively want to know all the appearances of a target object, information about its peculiar appearance features is very important as well as its common ones.

Conventional Web image search engines are mostly Text-Based Image Retrievals (TBIR) by using the filename, alternative text, and surrounding text of each Web image as clues. When such a text-based condition as a name of a target object is given by a user, they give the user the retrieval images which meet the text-based condition. It has become not difficult for us to get typical images as well as acceptable images of a target object just by submitting its object-name to a conventional keyword-based Web image search engine and browsing the top tens of the retrieval results, while peculiar images rarely appear in the top tens of the retrieval results.

As next steps of IR in the Web, it is very important to discriminate between "Typical Images" and "Peculiar Images" in the acceptable images, and moreover, to collect many different kinds of peculiar images as exhaustively as possible. In other words, "Exhaustiveness" is one of the most important requirements in the next-generation Web image searches as well as Web document searches [8].

As a solution to the 1st next step, my previous work [6] proposes a novel method to precisely search the Web for peculiar images of a target object whose name is given as a user's original query, by expanding the original query with its peculiar appearance descriptions (e.g., color-names) extracted from the Web by text mining techniques [9], [10] and/or its peculiar image features (e.g., color-features) converted from the Web-extracted peculiar color-names. In order to make the basic method more robust, this paper proposes a refined method equipped with cross-language (translation between Japanese and English) functions like [11], [12] and validates its retrieval precision (robustness).

The remainder of this paper is organized as follows. Section II explains my basic single-language method, and Section III proposes my cross-language method to search the Web for Peculiar Images. Section IV shows several experimental results. Finally, Section V concludes this paper.

II. Single-language Method

This section explains my basic method [6] to precisely search the Web for "Peculiar Images" of a target object whose name is given as a user's original query, by expanding the original query with its peculiar appearance descriptions (e.g., color-names) extracted from the Web by text mining techniques and/or its peculiar image features (e.g., color-features) converted from them. Figure 1 gives an overview of my basic (single-language) Peculiar Image Search.

Step 1. Peculiar Color-Name Extraction

When a name of a target object as an original query is given by a user, its peculiar color-names (as one kind of appearance descriptions) are extracted from exploding Web documents about the target object by text mining techniques.

The two kinds of lexico-syntactic patterns which consist of a color-name cn and the target object-name on are often used as follows:

1) "cn-colored on", such as "yellow-colored sunflower",
2) "on is cn", such as "sunflower is yellow".

The weight $\text{pcn}(cn, on)$ of Peculiar Color-Name extraction is assigned to each candidate cn for peculiar color-names of a target object-name on as follows:

$$\text{pcn}(cn, on) := \begin{cases} 0 & \text{if } \text{df}(\texttt{["}on \text{ is } cn\texttt{"]}) = 0, \\ \dfrac{\text{df}(\texttt{["}cn\text{-colored } on\texttt{"]})}{\text{df}(\texttt{["}on \text{ is } cn\texttt{"]})+1} & \text{otherwise.} \end{cases}$$

where $\text{df}(\texttt{["}q\texttt{"]})$ stands for the frequency of Web documents retrieved by submitting the phrase query $\texttt{["}q\texttt{"]}$ to Google Web Search [13].

Step 2. Color-Feature Conversion from Color-Name

The peculiar HSV color-features cf_p (as one kind of image features) of the target object are converted from its Web-extracted peculiar color-names cn_p by referring the conversion table [14] or [15] in each language.

Step 3. Query Expansion by Color-Name/Feature

Here, we have three kinds of clues to search the Web for peculiar images: not only a target object-name on (text-based condition) as an original query given by a user, but also its peculiar color-names cn_p (text-based condition) extracted from Web documents in the Step 1, and its peculiar color-features cf_p (content-based condition) converted from its peculiar color-names in the Step 2.

The original query (q0 = text: $\texttt{["}on\texttt{"]}$ & content: null) can be expanded by its peculiar color-names cn_p and/or its peculiar color-features cf_p as follows:

q1 = text: $\texttt{["}on\texttt{"]}$ & content: cf_p,
q2 = text: $\texttt{["}cn_p\text{-colored } on\texttt{"]}$ & content: null,
q3 = text: $\texttt{["}cn_p\text{-colored } on\texttt{"]}$ & content: cf_p.

Step 4. Image Ranking by Expanded Queries

First, the weight $\text{pis}_{q1}(i, on)$ of Peculiar Image Search based on the 1st type of expanded query (q1 = text: $\texttt{["}on\texttt{"]}$ & content: cf_p) is assigned to a Web image i for a target object-name on and is defined as

$$\text{pis}_{q1}(i, on) := \max_{\forall(cn_p, cf_p)} \Big\{ \text{pcn}(cn_p, on) \cdot \text{cont}(i, cf_p) \Big\},$$

$$\text{cont}(i, cf_p) := \sum_{\forall cf} \text{sim}(cf, cf_p) \cdot \text{prop}(cf, i),$$

where a Web image i is retrieved by submitting the text-based query $\texttt{["}on\texttt{"]}$ (e.g., $\texttt{["sunflower"]}$) to Google Image Search [1], $\forall(cn_p, cf_p)$ stands for not completely any pair but each pair of its Web-extracted peculiar color-name cn_p and its converted peculiar color-feature cf_p in the Step 2, $\text{sim}(cf, cf_p)$ stands for the similarity between color-features cf and cf_p in the HSV color space [16], and $\text{prop}(cf, i)$ stands for the proportion of a color-feature cf in a Web image i.

Next, the peculiarity of a Web image i for a target object-name on by the 2nd type of expanded query (q2 = text: $\texttt{["}cn_p\text{-colored } on\texttt{"]}$ & content: null) is defined as

$$\text{pis}_{q2}(i, on) := \max_{\forall cn_p} \left\{ \frac{\text{pcn}(cn_p, on)}{\text{rank}(i, on, cn_p)^2} \right\},$$

where $\forall cn_p$ stands for not completely any color-name but each Web-extracted peculiar color-name cn_p in the Step 1, and $\text{rank}(i, on, cn_p)$ stands for the rank of a Web image i in the retrieval results by submitting the text-based query $\texttt{["}cn_p\text{-colored } on\texttt{"]}$ to Google Image Search [1].

Last, the peculiarity of a Web image i for a target object-name on by the 3rd type of expanded query (q3 = text: $\texttt{["}cn_p\text{-colored } on\texttt{"]}$ & content: cf_p) is defined as

$$\text{pis}_{q3}(i, on) := \max_{\forall(cn_p, cf_p)} \left\{ \frac{\text{pcn}(cn_p, on) \cdot \text{cont}(i, cf_p)}{\text{rank}(i, on, cn_p)} \right\},$$

where $\forall(cn_p, cf_p)$ stands for not completely any pair but each pair of its Web-extracted peculiar color-name cn_p and its converted peculiar color-feature cf_p in the Step 2.

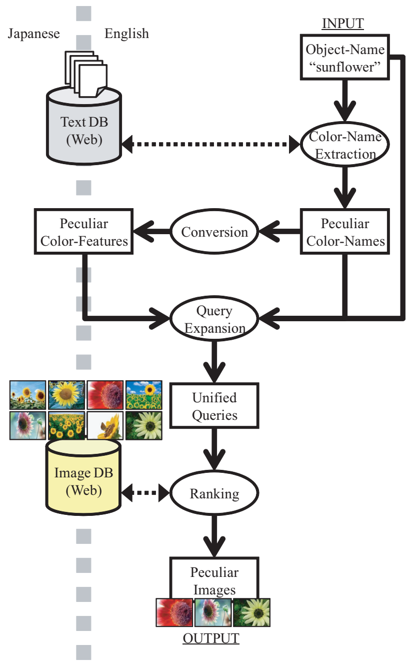

Figure 1. Single-Language Peculiar Image Search (in only English).

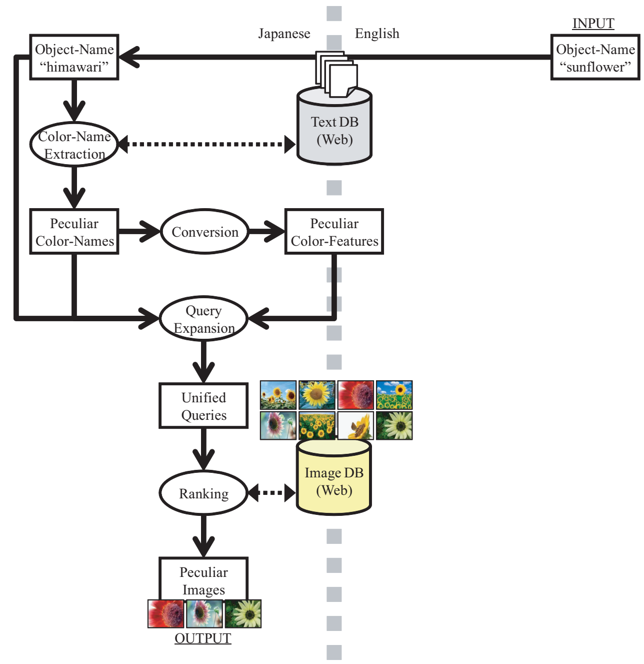

Figure 2. Cross-Language Peculiar Image Search to make a one-way trip (in English → Japanese).

III. Cross-Language Method

This section proposes a refined method equipped with cross-language (translation between Japanese and English) functions to make the basic method more robust. Figure 2 and 3 show my cross-language Peculiar Image Searches.

When an English object-name is given by a user, my proposed cross-language method in Figure 2 runs from English to Japanese language space as follows:

Step 0. translates the English object-name (e.g., "sunflower") into its Japanese one (e.g., "himawari"),

Step 1. extracts its Japanese peculiar color-names (e.g., "akairo" and "shiro") from the Web,

Step 2. converts its Japanese peculiar color-names into its peculiar color-features (e.g., ■:red and □:white),

Step 3-4. retrieves Web images by its Japanese object-name and its peculiar color-names and/or features.

Meanwhile, my proposed cross-language method in Figure 3 runs back and forth between English and Japanese language spaces as follows:

Step 0. translates its English object-name (e.g., "sunflower") into its Japanese one (e.g., "himawari"),

Step 1. extracts its Japanese peculiar color-names from the Web and translates them into its English peculiar ones (e.g., "red" and "white"),

Step 2. converts its English peculiar color-names into its peculiar color-features (e.g., ■:red and □:white),

Step 3-4. retrieves Web images by its English object-name and its peculiar color-names and/or features.

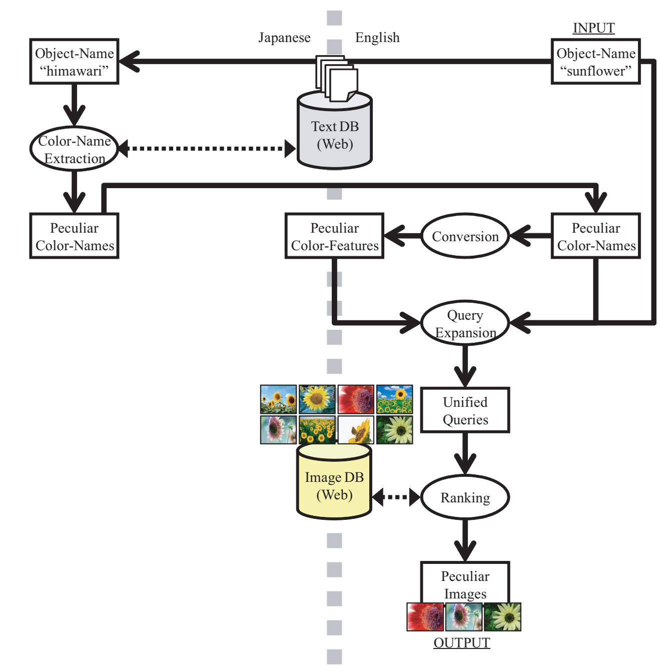

Figure 3. Cross-Language Peculiar Image Search to make a round trip (in English → Japanese → English).

IV. EXPERIMENT

This section shows several experimental results for the following eight kinds of target object-names from among four categories to validate my proposed cross-language methods to search the Web for their peculiar images more precisely than my previous single-language method and conventional keyword-based Web image search engines.

1) Plants:
   - Sunflower (typical-color: yellow)
   - Cauliflower (typical-color: white)
2) Landmarks:
   - Tokyo Tower (typical-color: red)
   - Nagoya Castle (typical-color: white)
3) Animals:
   - Praying Mantis (typical-color: green)
   - Cockroach (typical-color: brown)
4) Others:
   - Wii (typical-color: white)
   - Sapphire (typical-color: blue)

Table I shows each precision for the eight target object-names and the average precision of the top 20 and top 100 peculiar images searched by my proposed cross-language methods, my basic single-language methods, and Google Image Search [1]. The values listed in boldface are the best in each target object-name or the average. It shows that my cross-language EJE*q2 method gives the best performance.

Figures 4 and 5 show the top k average precision of my proposed cross-language methods, my basic single-language methods, and Google Image Search. They also show that my cross-language EJE*q2 method is superior to all the others, and that my cross-language EJE*qX methods to make a round trip from English to Japanese are the best, my cross-language EJ*qX methods to go from English to Japanese (and not to come back) are the second-best (better), and my basic single-language E*qX methods are the worst.

Figures 6 to 14 show the top 20 search results for each target object-name to compare between Google Image Search, my basic single-language method, and my cross-language method. Figure 13 shows that my previous method sometimes ($37.5\% = 3/8$) returns none as the search results.

Table I
CROSS-LANGUAGE EFFECT ON TOP 20 & TOP 100 PRECISION OF PECULIAR IMAGE SEARCHES.

		E only Eng		EJ Eng → Jap		EJE Eng → Jap → Eng	
q0 Google Image	Sunflower	0/20	2/100				
	Cauliflower	6/20	40/100				
	Tokyo Tower	0/20	7/100				
	Nagoya Castle	0/20	1/100				
	Praying Mantis	1/20	4/100				
	Cockroach	6/20	14/100				
	Wii	5/20	17/100				
	Sapphire	3/20	8/100				
	(Avg.)	2.6/20	11.6/100				
q1	Sunflower	1/20	2/100	1/20	9/100	0/20	2/100
	Cauliflower	2/20	40/100	8/20	40/100	0/20	40/100
	Tokyo Tower	0/20	7/100	5/20	12/100	3/20	7/100
	Nagoya Castle	0/20	0/100	0/20	0/100	0/20	1/100
	Praying Mantis	2/20	4/100	2/20	8/100	0/20	4/100
	Cockroach	3/20	14/100	3/20	23/100	0/20	14/100
	Wii	0/20	17/100	5/20	13/100	6/20	17/100
	Sapphire	5/20	8/100	9/20	40/100	4/20	8/100
	(Avg.)	1.6/20	11.5/100	4.1/20	18.1/100	1.6/20	11.6/100
q2	Sunflower	11/20	37/100	9/20	**54/100**	6/20	29/100
	Cauliflower	5/20	20/100	14/20	61/100	14/20	**62/100**
	Tokyo Tower	0/20	0/100	9/20	40/100	**13/20**	**43/100**
	Nagoya Castle	0/20	0/100	**4/20**	**7/100**	0/20	0/100
	Praying Mantis	2/20	3/100	6/20	15/100	**9/20**	**24/100**
	Cockroach	0/20	0/100	8/20	12/100	**12/20**	**43/100**
	Wii	5/20	18/100	2/20	11/100	**16/20**	61/100
	Sapphire	13/20	48/100	11/20	66/100	18/20	**81/100**
	(Avg.)	4.5/20	15.8/100	7.9/20	33.2/100	**11.0/20**	**42.9/100**
q3	Sunflower	7/20	36/100	**12/20**	50/100	2/20	18/100
	Cauliflower	5/20	13/100	13/20	51/100	**16/20**	48/100
	Tokyo Tower	0/20	0/100	1/20	29/100	7/20	20/100
	Nagoya Castle	0/20	0/100	**4/20**	6/100	0/20	0/100
	Praying Mantis	2/20	3/100	3/20	6/100	3/20	14/100
	Cockroach	0/20	0/100	4/20	11/100	7/20	26/100
	Wii	9/20	44/100	4/20	5/100	**16/20**	**72/100**
	Sapphire	14/20	62/100	16/20	64/100	**20/20**	79/100
	(Avg.)	4.6/20	19.8/100	7.1/20	27.8/100	8.9/20	34.6/100

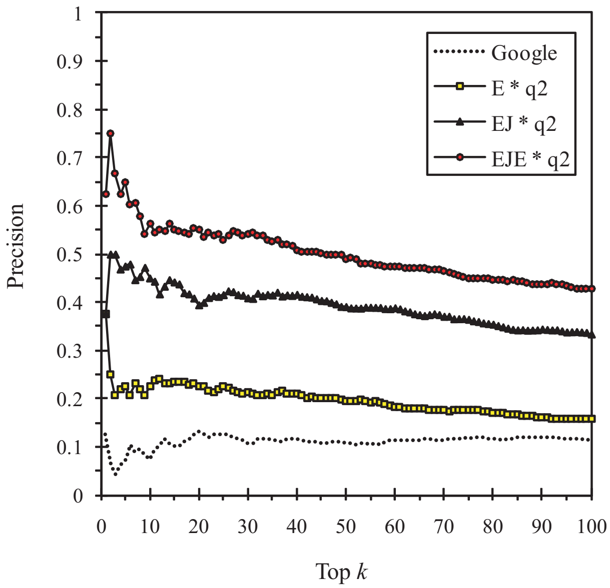

Figure 4. Top k Average Precision of Google Image Search vs. Peculiar Image Searches (method: X*q2).

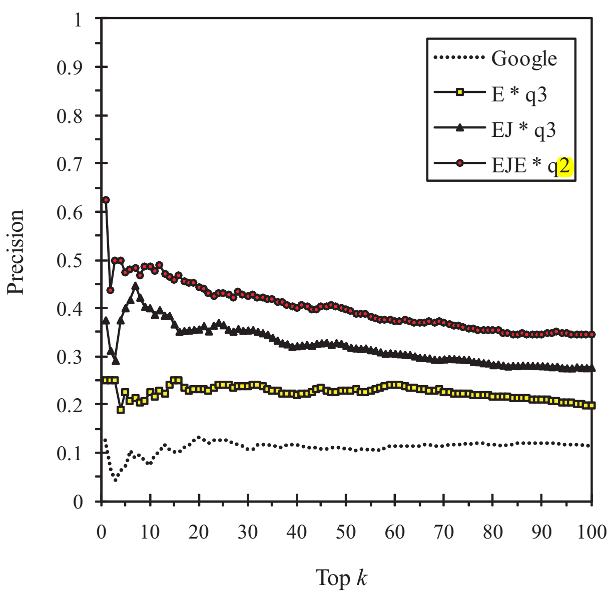

Figure 5. Top k Average Precision of Google Image Search vs. Peculiar Image Searches (method: X*q3).

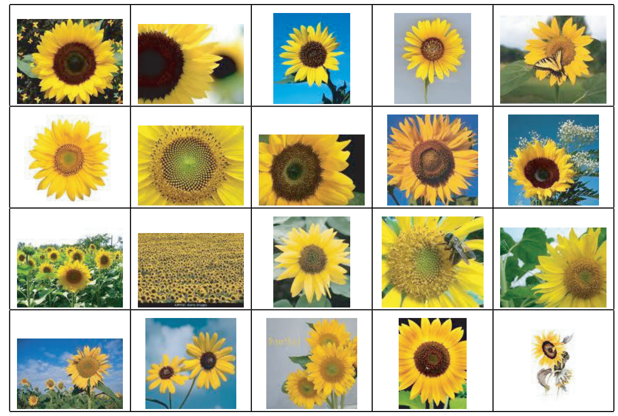

Figure 6. Top 20 results of Google Image Search (method: q0, object-name on = "Sunflower").

Figure 7. Top 20 results of Single-Language Peculiar Image Search (method: E*q2, object-name on = "Sunflower").

Figure 8. Top 20 results of Cross-Language Peculiar Image Search (method: EJ*q3, object-name on = "Sunflower").

Figure 9. Top 20 results of Google Image Search (method: q0, object-name on = "Cauliflower").

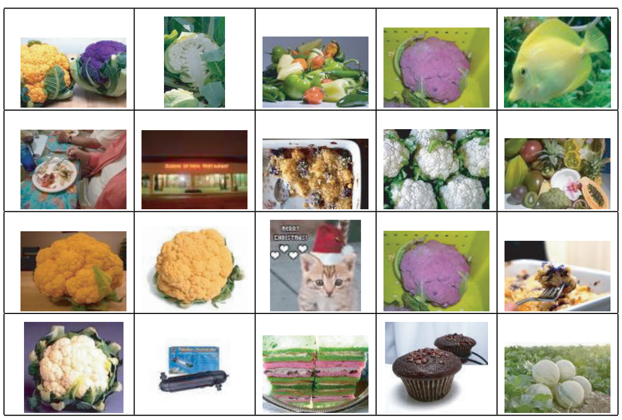

Figure 10. Top 20 results of Single-Language Peculiar Image Search (method: E*q2, object-name on = "Cauliflower").

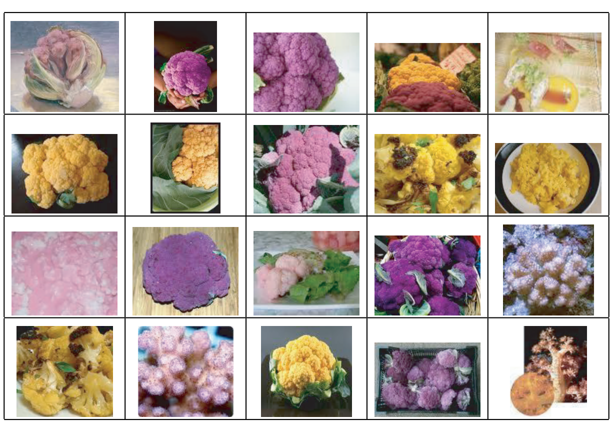

Figure 11. Top20 results of Cross-Language Peculiar Image Search (method: EJE*q3, object-name on = "Cauliflower").

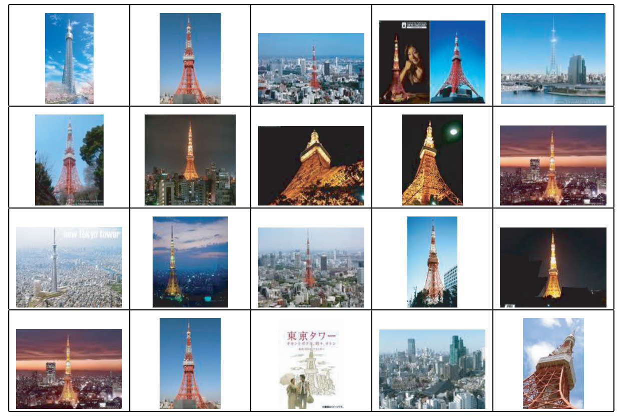

Figure 12. Top 20 results of Google Image Search (method: q0, object-name on = "Tokyo Tower").

No Image	No Image	No Image	No Image	No Image
No Image	No Image	No Image	No Image	No Image
No Image	No Image	No Image	No Image	No Image
No Image	No Image	No Image	No Image	No Image

Figure 13. Top 20 results of Single-Language Peculiar Image Search (method: E*q2, object-name on = "Tokyo Tower").

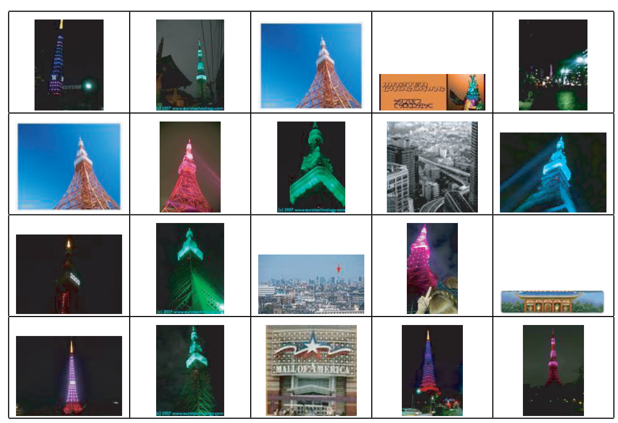

Figure 14. Top 20 results of Cross-Language Peculiar Image Search (method: EJE*q2, object-name on = "Tokyo Tower").

V. Conclusion

As next steps of Image Retrieval (IR), it is very important to discriminate between "Typical Images" and "Peculiar Images" in the acceptable images, and moreover, to collect many different kinds of peculiar images exhaustively. In other words, "Exhaustiveness" is one of the most important requirements in the next IR. But it is difficult to find clusters which consist of not noisy but peculiar images only by clustering based on image content features. As a solution to the 1st next step, my previous work proposed a basic method to precisely search the Web for peculiar images of a target object by its peculiar appearance descriptions (e.g., color-names) extracted from the Web and/or its peculiar image features (e.g., color-features) converted from them.

To make the basic method more robust, this paper has proposed a refined method equipped with cross-language (translation between Japanese and English) functions. And several experimental results have validated the retrieval precision (robustness) of my cross-language methods by comparing with such a conventional keyword-based Web image search engine as Google Image Search [1] and my basic single-language method [6]. My proposed cross-language Peculiar Image Search has been about twice as precise as my basic Peculiar Image Search, and about quadrice as precise as Google Image Search, when an English object-names is given as a user's original query for peculiar images.

In the future, I try to utilize the other appearance descriptions (e.g., shape and texture) besides color-names and the other image features besides color-features in my basic single-language and my proposed cross-language Peculiar Image Searches. In addition, I also try to evaluate my proposed cross-language Peculiar Image Searches with translation between English and the other languages besides Japanese, or between Japanese and the other languages besides English.

Acknowledgment

This work was supported in part by Open Research Center Project "Tangible Software Engineering Education" [1] for Private Universities: matching fund subsidy from MEXT (Project Leader: Taichi Nakamura, 2008-2011).

References

[1] Google Image Search, http://images.google.com/ (2010).

[2] Inder, R., Bianchi-Berthouze, N., and Kato, T.: "K-DIME: A Software Framework for Kansei Filtering of Internet Material," Proc. of IEEE International Conference on Systems, Man and Cybernetics (SMC'99), Vol.6, pp.241–246 (1999).

[3] Kurita, T., Kato, T., Fukuda, I., and Sakakura, A.: "Sense Retrieval on a Image Database of Full Color Paintings," Transactions of Information Processing Society of Japan (IPSJ), Vol.33, No.11, pp.1373–1383 (1992).

[4] Kimoto, H.: "An Image Retrieval System Using Impressional Words and the Evaluation of the System," Transactions of Information Processing Society of Japan (IPSJ), Vol.40, No.3, pp.886–898 (1999).

[5] Hattori, S. and Tanaka, K.: "Search the Web for Typical Images based on Extracting Color-names from the Web and Converting them to Color-Features," Letters of Database Society of Japan (DBSJ), Vol.6, No.4, pp.9–12 (2008).

[6] Hattori, S. and Tanaka, K.: "Search the Web for Peculiar Images by Converting Web-extracted Peculiar Color-Names into Color-Features," IPSJ Transactions on Databases, Vol.3, No.1 (TOD45), pp.49–63 (2010).

[7] Hattori, S.: "Peculiar Image Search by Web-extracted Appearance Descriptions," Proceedings of the 2nd International Conference on Soft Computing and Pattern Recognition (SoCPaR'10), pp.127–132 (2010).

[8] Yamamoto, T., Nakamura, S., and Tanaka, K.: "Rerank-By-Example: Browsing Web Search Results Exhaustively Based on Edit Operations," Letters of Database Society of Japan (DBSJ), Vol.6, No.2, pp.57–60 (2007).

[9] Hattori, S., Tezuka, T., and Tanaka, K.: "Extracting Visual Descriptions of Geographic Features from the Web as the Linguistic Alternatives to Their Images in Digital Documents," IPSJ Transactions on Databases, Vol.48, No.SIG11 (TOD34), pp.69–82 (2007).

[10] Hattori, S., Tezuka, T., and Tanaka, K.: "Mining the Web for Appearance Description," Proc. of the 18th International Conference on Database and Expert Systems Applications (DEXA'07), LNCS Vol.4653, pp.790–800 (2007).

[11] Etzioni, O., Reiter, K., Soderland, S., and Sammer, M.: "Lexical Translation with Application to Image Search on the Web," Proc. of Machine Translation Summit XI (2007).

[12] Hou, J., Zhang, D., Chen Z., Jiang, L., Zhang, H., and Qin, X.: "Web Image Search by Automatic Image Annotation and Translation," Proceedings of the 17th International Conference on Systems, Signals and Image Processing (IWSSIP'10), pp.105–108 (2010).

[13] Google Web Search, http://www.google.com/ (2010).

[14] Wikipedia - List of colors, http://en.wikipedia.org/wiki/List_of_colors (2010).

[15] Japanese Industrial Standards Committee: "Names of Non-Luminous Object Colours," JIS Z 8102:2001 (2001).

[16] Smith, J. R. and Chang, S.-F.: VisualSEEk: A Fully Automated Content-Based Image Query System, Proceedings of the 4th ACM International Conference on Multimedia (ACM Multimedia'96), pp.87–98 (1996).

[1] http://www.teu.ac.jp/tangible/