❏   139

Incremental Tag Suggestion for Landmark Image Collections

Sutasinee Chimlek, Punpiti Piamsa-nga
Department of Computer Engineering, Faculty of Engineering, Kasetsart University,
Chatujak, Bangkok, 10900, Thailand

Article Info	ABSTRACT
	In recent social media applications, descriptive information is collected through user tagging, such as face recognition, and automatic environment sensing, such as GPS. There are many applications that recognize landmarks using information gathered from GPS data. However, GPS is dependent on the location of the camera, not the landmark. In this research, we propose an automatic landmark tagging scheme using secondary regions to distinguish between similar landmarks. We propose two algorithms: 1) landmark tagging by secondary objects and 2) automatic new landmark recognition. The images of 30 famous landmarks from various public databases were used in our experiment. Results show increments of tagged areas and the improvement of landmark tagging accuracy.

Corresponding Author:

Punpiti Piamsa-nga,
Department of Computer Engineering,
Faculty of Engineering, Kasetsart University,
Chatujak, Bangkok, 10900, Thailand.
Email: pp@ku.ac.th

1. INTRODUCTION

The wide availability and low cost of digital photography means that people who travel generally accumulate a very large number of photographs. Furthermore, these images are often uploaded to various social media, such as Facebook. The images not only relate to the people in the photographs but also to the location of where they were taken. For information on identity, social-network applications, such as Facebook and Google Plus, use manual and semi-automatic tagging on shared photographs. For example, face recognition applications will suggest "name tags" of recognised human faces in new photographs and ask users for confirmation. Tags of human faces are text-based tags for use in a text-based search.

Tags for objects in social media must be related to keywords as most queries to the system are still text based [1]-[4]. Keyword tags such as "Tokyo Tower", "Big Ben", and "Taj Mahal" are common queries. The massive number of users on the Internet may help tag and share to improve tag information. However, textual keyword tagging becomes more difficult for non-social applications, such as cataloging home photographs, because it is time consuming and assistance from social networks is rarely available. By observation, most users simply keep a set of photographs in a folder on their hard drive and the folder name is usually based on the associated event, which is not necessarily consistent with the others. Photo album software, such as Google Picasa and Adobe Lightroom, can also be used to manage these images. This software can suggest tags for faces using the same methods as current social networks; however, landmark images are rarely tagged as GPS data may not be available from the photographs. Users regularly tag only a small number of interesting landmarks and leave most other similar images untagged, affecting the search accuracy.

While landmark images can be tagged similarly to faces in social media, the task is more difficult because there are many types of landmarks to be recognized including churches, pagodas, towers, buildings, temples, and houses.

Landmarks in photographs can be tagged automatically [5]-[7]. If the landmark is known and tagged by a user, the features are extracted from these landmark images and used to search for other images by feature index. If a search result is found, an image area at the location of the extracted features is then automatically tagged with the search result. If a landmark in a photograph is new and unavailable in the landmark database, the result should not be found (see Figure 1). However, the classifier must still select the best-ranked answer, which may not always be accurate.
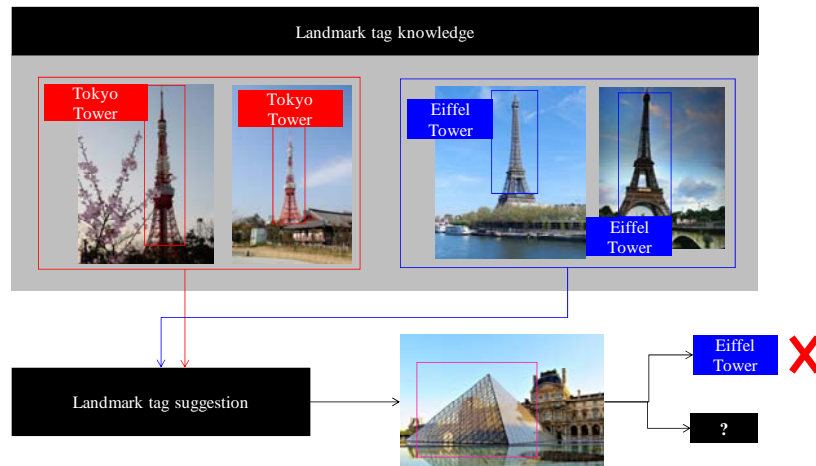
Figure 1. Automatic landmark tagging

One limitation of landmark tagging is that it cannot distinguish between two similar landmarks in different locations. In Figure 2, the Eiffel Tower and the Tokyo Tower have the same shape, but their colors and surroundings are completely different. In most learning machines that generally recognize only the landmark object, both objects would receive the same tag. We believe that some secondary objects in the photograph, such as a Japanese house, may help users to select the correct answer. In other words, the user's prior knowledge may not be sufficient for landmark recognition and considering the secondary surroundings of the landmark may improve search accuracy.
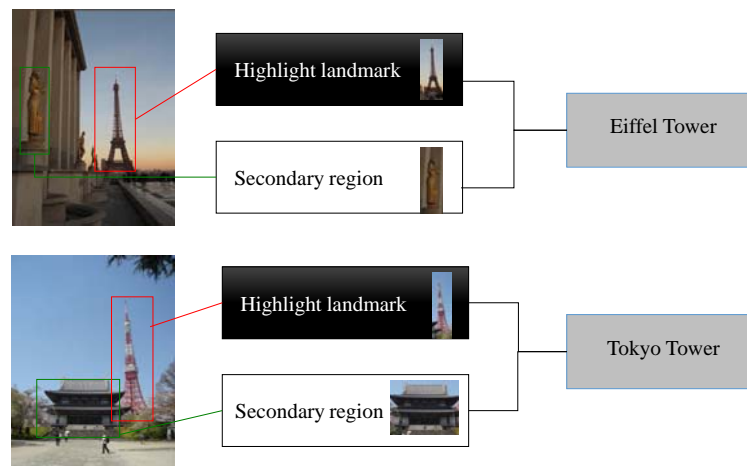
Figure 2. Example of the secondary surroundings of the landmark

Therefore, three challenging problems must be resolved for improved landmark image tagging. First, search accuracy is too low as there is a dearth of manually tagged landmark images. Second, tag suggestions cannot be accurate since it is difficult to distinguish between two similar landmarks. Third, as

new photographs are inserted into the database, new landmarks must be suggested; otherwise, the system outputs the closest result.

In this paper, we propose incremental automatic tagging for landmark images using surrounding regions to distinguish between similar landmarks. This improves landmark tagging accuracy and identifies new tags for new landmark images.

The organization of this paper is as follows: Section 2 is a review related work; Section 3 describes the overview of our proposed framework; Section 4 presents the experimental results. Finally, we conclude this research in Section 5.

2.   RELATED WORK

Typically, landmark image tagging uses learning-based approaches with visual features from image regions corresponding to the landmark. Learning-based approaches generate landmark classifiers for a predefined set of landmark regions using knowledge previously learned from labeled landmark regions [8], [9]. A landmark corresponds to a category from which a classifier is trained. Joshi et al. [10] proposed using image tags and visual content to infer the landmark. Papadopoulos et al. [11] proposed a framework to tag landmarks and events in geo-tagged images. This involves learning from a hybrid image-similarity graph, including visual and tag similarities between images. Raguram et al. [12] proposed an iconic scene graph to build stereo scene models for 3D landmark recognition. In [13], Yunpeng et al. presented a large-scale landmark classification system from geo-tagged Flickr images using multi-class support vector machines. The accuracy of landmark tagging using learning-based approaches is low because the predefined set of landmark images is limited and new landmarks cannot be detected.

For real-world applications, we cannot rely entirely on a predefined set of landmarks and improving the learning algorithm to detect new landmarks is crucial. Several publications suggested methods to support new landmark images by using incremental tagging. For instance, the personal image tagging approach tagged new images with tags mostly used by the query user or their friend [14], [15]. Xirong et al. proposed an algorithm that learns tag relevance by accumulating votes from visual neighbors of the target image on the assumption that different users label visually similar images using the same tags [16], [17]. The feature indexing strategy is further enhances scalability for large-scale applications.

Dong et al. [18] present a scheme that performs incremental tagging until a satisfactory tagging accuracy is reached or the user stops the process. This approach to exemplar selection assumes that exemplar images can represent visually similar image clusters better than the centroids of the obtained clusters, which may not be real samples. The Columbia TAG (Transductive Annotation by Graph) system [19] incorporates graph-based label propagation methods and intuitive graphical user interfaces (GUI) that allow users to quickly browse and annotate a small number of images/videos. The system then refines the labels for the remaining unlabeled data in the collection. The graph-based label propagation methods are constructed from visually similar images and the initial labels of a subset of nodes in the graph are provided by some external filters, classifiers, or ranking systems. This system's performance depends on classifiers for image labeling and the user's choice of positive and negative images/videos for the labels. Thanh-Binh Le et al. [20] proposed an incremental selection strategy which can improved the classification accuracy of semi-supervised learning (SSL) algorithms.

As described above, there are several proposed incremental tagging methods that use graphs for visually similar images and user responses to add new tags for new landmark images. However, these existing methods require complex computations and have limited accuracy because they either depend on visual features from the entire image or highlight only the landmark region and user responses (thereby potentially confusing similar landmarks).

3.   RESEARCH METHOD

In this paper, we propose an incremental landmark tagging framework to support new landmark tagging and address the low accuracy of existing methods resulting from an inability to distinguish between similar landmarks in different locations. We use a hybrid approach combining the learning-based method and the correlation between landmarks and salient surrounding objects using co-occurrence.

In general, the image has two types of regions: the highlighted landmark region and the secondary region. The highlighted landmark region is the prominent or well-known object within a particular landscape that is a common interest area for photographs. Secondary regions are dominant areas related to the landmark that users neglect to tag or that contains objects that cannot be represented by label text.

Our proposed framework consists of two methods: 1) Landmark tag learning to generate landmark classifiers from labeled landmark images, a method which uses both the highlighted landmark and the

surrounding objects to help improve tag accuracy; and 2) Incremental suggestion to tag existing landmarks from a predefined landmark dataset and add new tags for new landmark images or different views of the landmark image from a predefined landmark dataset. Our framework is shown in Figure 3.
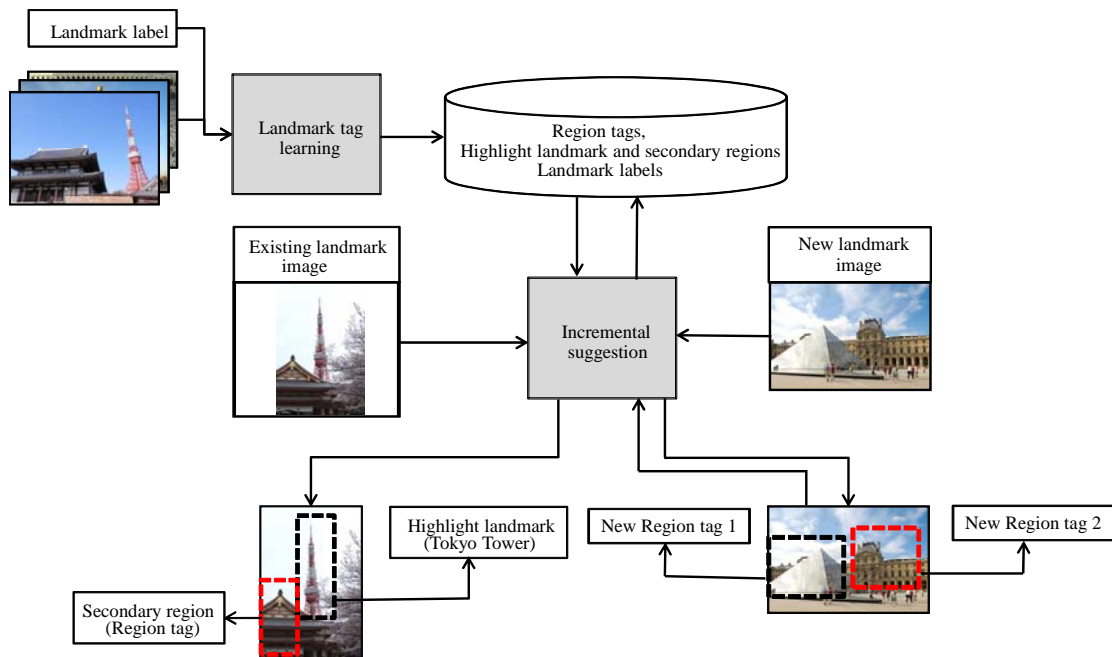

Figure 3. Incremental learning for landmark image tagging

3.1. Landmark Tag Learning

Landmark tag learning consists of two processes: 1) a multiclass classification for tag learning that learns models by distinguishing between the highlighted landmark regions and secondary regions for labeled landmark images. Multiclass learning is required to learn tag models for the landmarks in the dataset; and 2) the construction of a matrix of the highlighted landmark and secondary region co-occurrence to represent their correlation (see Figure 4). We use the landmark classifier model and co-occurrence for the incremental suggestion method.

Highlighted landmark regions correspond to the high probability regions of each landmark image. We determine secondary regions by using significant conditions of probability between the highlighted landmarks and the other regions for each landmark image.

We use a salient detection algorithm to detect the salient regions for feature extraction and K-means grouping. Each cluster is named according to the "region tag" that contains a similar object or visual content. This region tag is used to determine the highlighted landmark region and the secondary region.

Landmark tag learning starts with the highlighted landmark region and secondary regions detected using salient object detection approaches. We assume that salient object detection is able to select the highlighted landmark and secondary regions related to the landmark image while ignoring irrelevant visual information. We use three salient object detection approaches: the method of Itti et al [21], Graph Based Visual Saliency (GBVS) [22], and HouCVPR [23]. Because each of these three models is efficient for specific salient regions in object tagging, using these complementary models increases salient object detection accuracy.

Our salient detection model is shown in Figure 5. Let $M(.)$ represent the detected image area using a salient detection method; d represents the image; r_i represents a region in the image d; and $score(r_i)$ indicates the score corresponding to r_i being part of the salient region detected by $M(.)$, which is given as follows:

$$score(r_i) \equiv score(r_i, M_{Itti} \cup M_{GBVS} \cup M_{HouCVPR}) \tag{1}$$

If $score(r_i)$ is greater than a predefined threshold, each area r_i becomes a salient region. Finally, a set of all salient regions $S(d)=\{S_1, S_2, ..., S_{|s|}\}$ becomes a representation of an image d.

The features are extracted from the salient regions and are grouped by Bag of Visual Words (BOVW) [24] and GIST [25] using a K-means clustering algorithm. We assign a specific region tag to each cluster containing similar visual content. Thus, each image is represented by a "Bag of Region tags"—an occurrence vector of salient tags in each image.
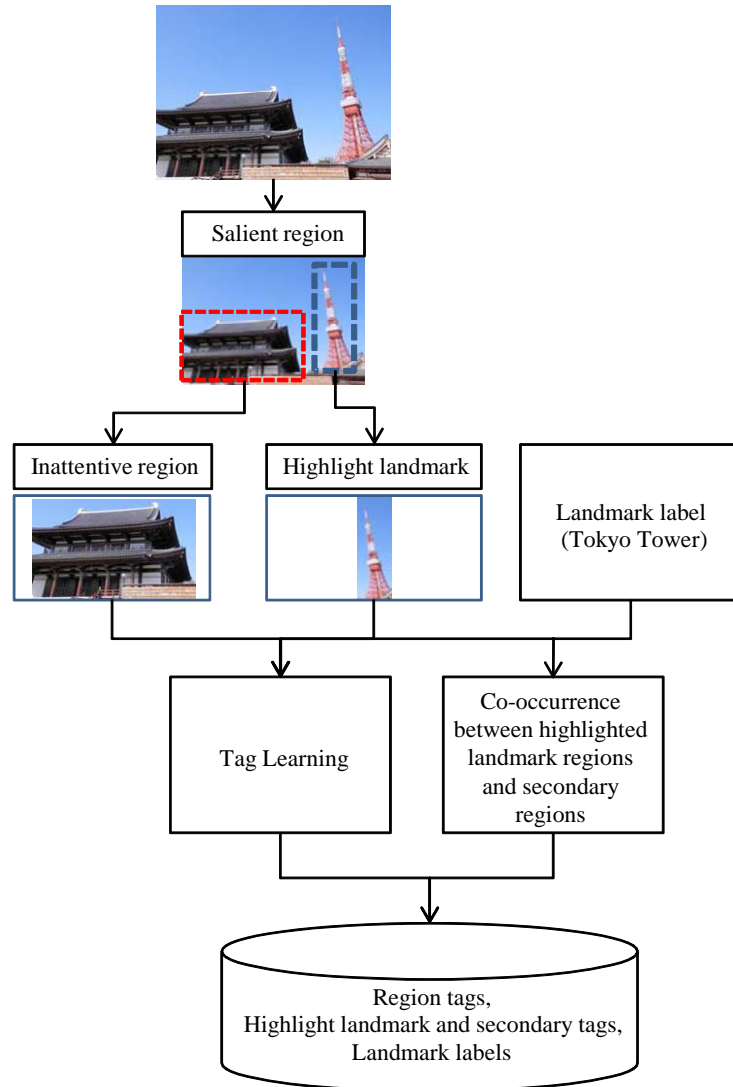

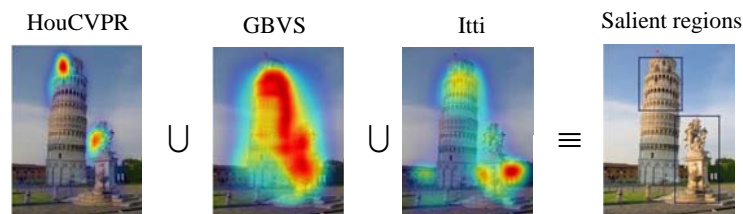Figure 4. Landmark tag learning

Figure 5. Salient detection

Highlighted landmark and secondary tag generation uses an arbitrary threshold that is determined either statistically or by a learning machine. Highlighted landmarks correspond to region tags with a probability greater than the threshold. Secondary regions are determined by region tags with a conditioned

probability with highlighted landmarks of greater than zero. Let L be the set of landmark labels; D be the set of training image data and $D=\{D_1,D_2,...,D_{|L|}\}$ where D_i represents the subsets of images that have label $l \in L$; $C=\{c_1,c_2,...,c_k\}$ be the set of k feature clusters from all salient regions of all images in training set D; region tag $R=\{r_1,r_2,...,r_k\}$ be the set of representatives of C, where r_i is the centroid of c_i; and $P(r_i,D_j)$ be the probability of region tag r_l occurring in the training subset D_j. Highlighted landmark regions and secondary with landmark label j are given as follows:

$$H_j = \{\forall r_i | P(r_i, D_j) > T_j\} \tag{2}$$
$$I_j = \{\forall r_i | P(r_i, H_j) > 0\} \tag{3}$$

All highlighted landmark regions and all secondary regions are defined by $H=\{\forall_j|H_j\}$ and $I=\{\forall_j|I_j\}$, respectively. The threshold T_j is arbitrary.

Each image can be represented by a binary vector of region tags, referred to as the "Bag of Region tags" (BoRtags). If an image d_i is tagged by region tag r_j then $BoRtags(d_i,r_j)=1$, otherwise $BoRtags(d_i,r_j)=0$.

We use multiclass SVM to construct tag classifiers for all $|L|$ landmark labels using BoRtags. Highlighted landmark and secondary region correlation is analyzed using the co-occurrence of the two regions for each landmark label as show in eq.4. We assume that each landmark image has a similar landmark co-occurrence between highlighted landmark and secondary regions. Let $P_{D_l}(H_i,I_j)$ be the co-occurrence between highlighted landmark region Hi and secondary region Ij in the training subset D_l, which is given as follows:

$$P_{D_l}(H_i, I_j) = \frac{P(H_i \cap I_j)}{P(D_l)} \tag{4}$$

All tag classifiers and co-occurrence values between highlighted landmark and secondary regions are used to assign landmark tags for untagged images that of either existing or new landmarks.

3.2. Incremental Suggestion

Typically, automatic tagging defines tags for untagged images using tag classifiers from learning processes. This approach is limited by a restriction on the size of tag models. Hence, the tagging precision and recall measure of automatic tagging are still fairly low. We propose incremental suggestion for improved tagging accuracy for new landmark images or different views of existing landmarks in a dataset, as illustrated in Figure 6.

Our framework combines tag classifiers with the co-occurrence between highlighted landmark regions and secondary regions to suggest tags. For untagged images, tags are suggested from both tag classifiers and the co-occurrence between highlighted landmark regions and secondary regions. We assume that if the untagged image is a new landmark image, each method would suggest a different tag.

Tags for untagged images are suggested using the following steps. First, the salient regions of the untagged image are detected and their features are extracted. Second, a region tag is assigned to each salient region in the untagged image and BoRtags represent the untagged image. Third, region tags and landmark labels are suggested for the untagged image using a tag classifier model with the suggestion referred to as Tag_{model}. Fourth, a co-occurrence matrix is constructed between the region tags of the untagged image. Let $P_{d_q}(r_i,r_j)$ be the co-occurrence matrix between region tag r_i and region r_j of the untagged image d_q, given as follows:

$$P_{d_q}(r_i,r_j) = P(r_i \cap r_j) \tag{5}$$

Fifth, the co-occurrence of the untagged image is compared with the co-occurrence of each landmark label. Let $Score(d_q,D_l)$ be the distance measure score between the co-occurrence of the untagged image and the co-occurrence of each landmark label as follows:

$$Score(d_q,D_l) = \sum_{i,j=1}^{k} |P_{D_l}(r_i,r_j) - P_{d_q}(r_i,r_j)| \tag{6}$$

Let $P_{D_l}(r_i,r_j)$ be the co-occurrence between region tag r_i and region tag r_j in landmark training subset D_l. The landmark label is suggested for untagged images based on the landmark labels with the lowest

distance measure score between the co-occurrence of the untagged image and the co-occurrence of each landmark label. $Tag_{co\text{-}occurence}$ refers to the landmark label suggested using the co-occurrence of each landmark label. Sixth, the landmark labels suggested using tag classifiers (Tag_{model}) are compared to those suggested using co-occurrence ($Tag_{co\text{-}occurence}$). The untagged image obtains the suggested landmark labels the tag model if $Tag_{model}=Tag_{co\text{-}occurrence}$ or $Score(d_q,D_l)<T_{new}$, otherwise, a new landmark label is suggested. T_{new} is a threshold which is determined statistically from the maximum distance measure score between the co-occurrence of landmark images and the co-occurrence between highlighted landmark regions and secondary regions in each landmark dataset. Finally, if the untagged image is given a new landmark label, we represent the co-occurrence of the new landmark label using the BoRtags of the untagged image. In addition, if the untagged image is assigned a tag model in the database, we update the co-occurrence of the assigned landmark label.
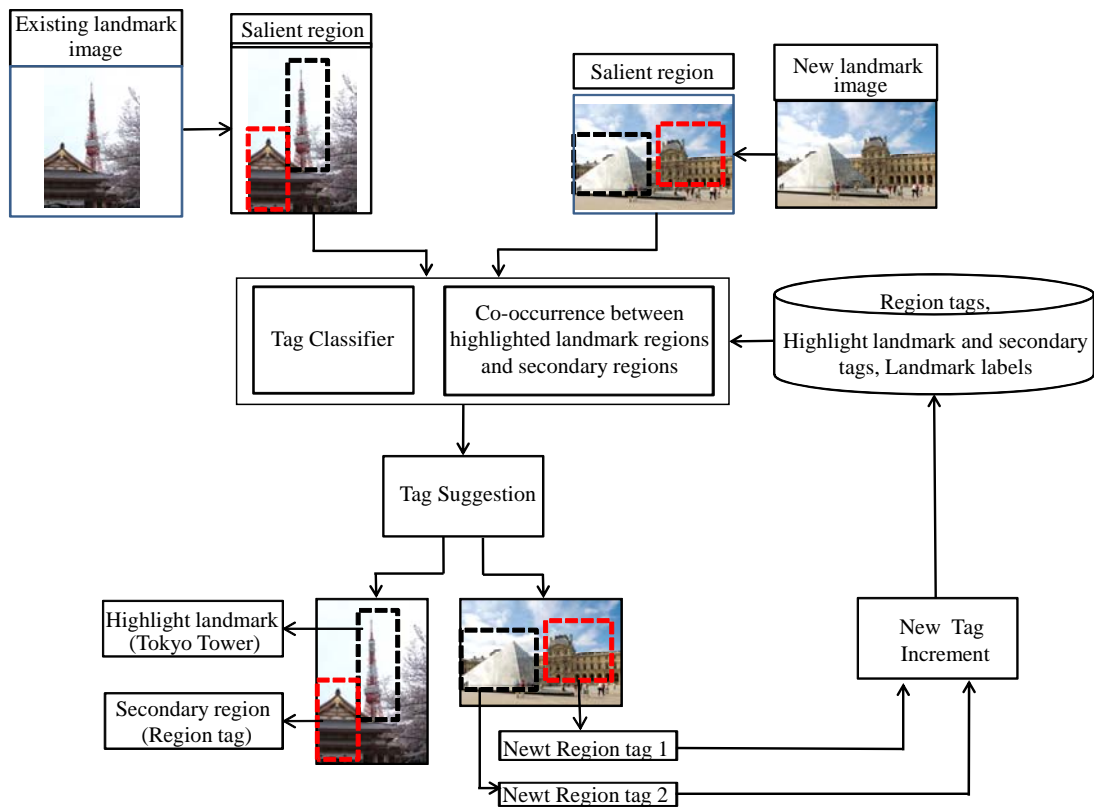

Figure 6. Incremental tag suggestion

We use the co-occurrence matrix $P_{D_l}(H_i,I_j)$ between the highlighted landmark regions and secodary regions in each landmark label to define the landmark dictionary. If $P_{D_l}(H_i,I_j)>0$ then the highlighted landmark H_i and the positive secondary region I_j are defined in the landmark dictionary of landmark L. The landmark dictionary is shown in Figure 7.
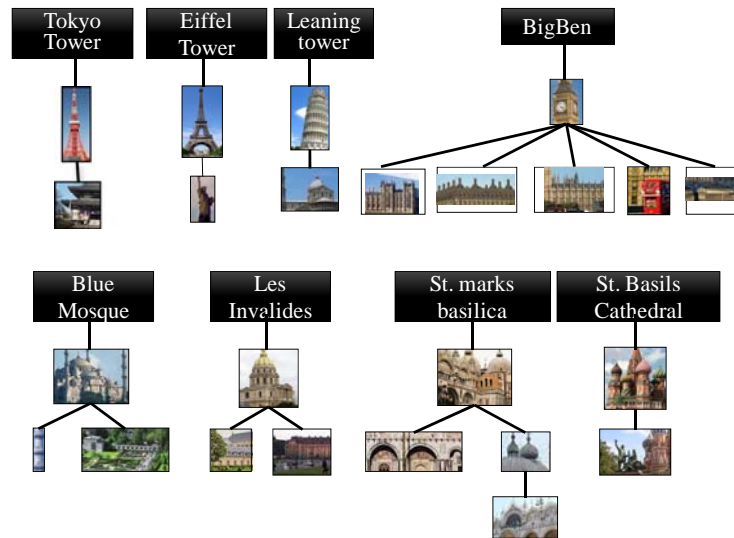
Figure 7. Landmark dictionary

4. RESULTS AND DISCUSSION

We assume that each landmark image contains a highlighted landmark region and other secondary regions and that new landmark images have differing highlighted landmark regions or secondary regions. We aim to evaluate the improvement in the accuracy of landmark tagging when using secondary regions and incremental tag suggestions in comparison to tagging using only the highlighted landmark regions and using only a tag model.

4.1. Data

We conducted experiments using 30 different famous landmarks from around the world. We collected 35,829 images from the Paris Dataset [26], Flickr, and Google by searching each particular landmark name. Each landmark label has about 1100 images and contains around three salient regions.

4.2. Features

We used a 512-dimensional GIST descriptor and a 2500-dimensional BOVW, which are derived from a SIFT descriptor for the feature extraction of salient regions. Each image is represented by 400-dimensional BoRtags.

4.3 Evaluation of Results

We measured the accuracy of the salient selection method and evaluated the performance of landmark tagging in terms of precision, recall, and the F1-measure compared to the ground-truth landmark tags using highlighted landmark regions and positive secondary regions.

The accuracy of the salient selection is evaluated using a ground-truth dataset generated from user-selected highlighted landmark regions. Each detected image has less than three salient regions in the ground-truth images.

Figure 8 shows examples of salient selection results. We grouped landmark images into five categories: temple, tower, church, castle, and art building. The average salient selection accuracy is 82.23% as shown in Table 1.
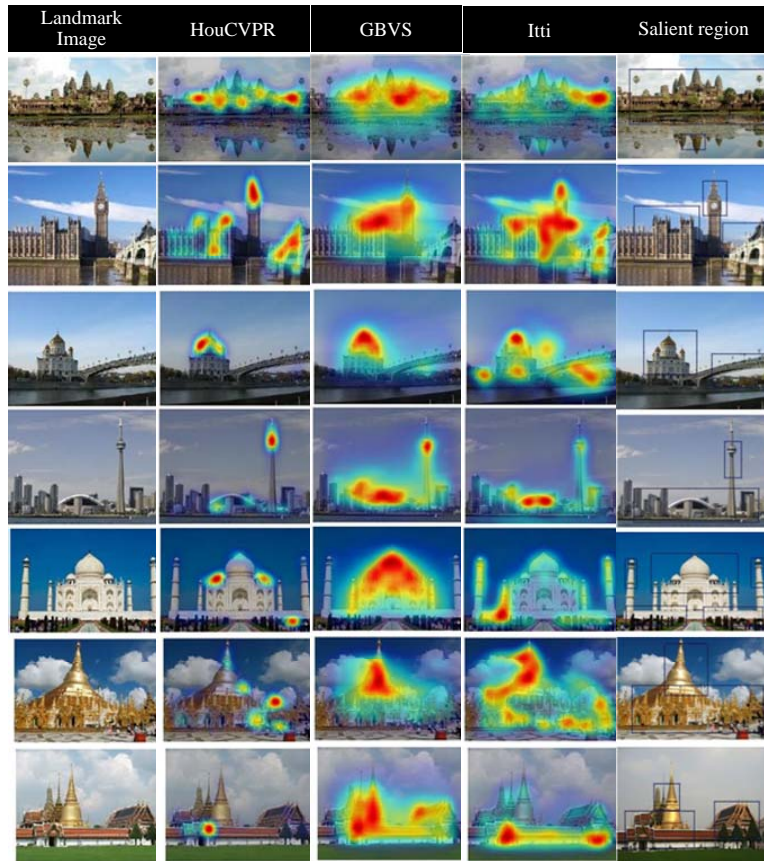
Figure 8. Salient detection

Table 1. Salient selection accuracy

Landmark Type	True Positive Rate (%)
Temple	85.50
Tower	78.31
Church	87.82
Castle	81.00
Art Building	78.51
Average	82.23

We evaluated tag suggestion using highlighted landmark and secondary regions with multiclass SVM classifiers with 10-fold cross validation (executed in WEKA). This served as a prediction model for landmark tags on 21 landmark images. We experimented with three sets of regions: 1) highlighted landmark only (E1), 2) entire image (E2), 3) highlighted landmark and secondary regions (E3). E1 and E2 use GIST and BOVW features as input. E3 is our proposed method and uses BoRtags and E4 is the ground-truth method. Table 2 shows the performance of the three methods.

Table 2. Performance of landmark tagging using tag classifier model of 21 landmark images

Method	Set of regions	F1-measure
E1	Highlight landmark only	0.730
E2	Whole image	0.831
E3	Highlight landmark and secondary salient	0.885

The results demonstrate a 15.5% improvement in the F1-measure when using our proposed method (E3) over the method using the highlighted landmark region only (E1). Our method (E3) is also superior to the using the entire image (E2).

Incremental tag suggestion for new landmark images is evaluated in comparison to four methods: 1) tag co-occurrence (P1), 2) tag model with 21 landmarks (P2), 3) tag model with 30 landmarks (P3, baseline

method), 4) our proposed method of incremental tag suggestion (P4). We experimented with 21 existing landmarks in the tag codebook and seven new landmarks. Our proposed method (P4) achieves an accuracy of 0.855, outperforming the other methods, as shown in Table 3.

Table 3. Accuracy of incremental tag suggestion for new landmark images

Method	F1-measure
Tag co-occurrence (P1)	0.803
Tag model with 21 landmarks (P2)	0.641
Tag model with 30 landmarks (P3)	0.804
Incremental tag suggestion (P4)	0.855

Figure 9 shows accuracy of tag suggestion for seven new landmarks. The results of our proposed method (P4) are very similar to those of the baseline method (P3) for tags such as "Louvre", and "Pantheon Rome". The results of our proposed method (P4) are better than those of the baseline method (P3) for tags such as "Moulin Rouge", "Potala Palace", and "St. Marks Basilica".

5.   CONCLUSION

We propose an incremental automatic landmark tagging framework that can tag the highlighted landmark regions and secondary regions of landmark images and videos to achieve improved tag accuracy and search performance. Our proposed method discriminates new landmark tags from existing landmark tags by using a combination of tag models and tag co-occurrence matrices. The tag models are constructed using multiclass SVM classifiers and salient detection extracts highlighted landmark regions and secondary regions for each landmark label. We use SIFT and GIST as features in the salient regions. Our experimental results show that combining the information from highlighted landmark regions and secondary regions can improve F1 performance by 15.5% compared with using only highlighted landmark regions and 5% compared with using salient regions from entire images. Our incremental tag suggestion can tag new landmark images with an accuracy of 0.855.

We are currently improving our technique to increase the accuracy of new landmark image tagging to support large landmark image databases.
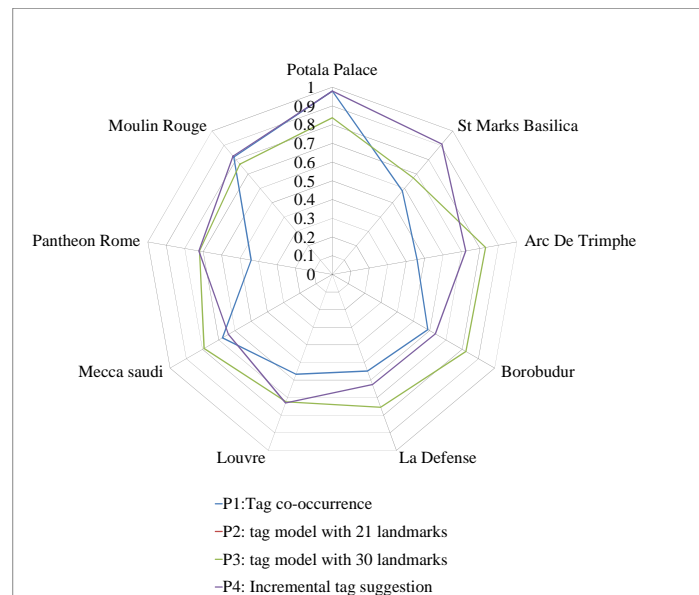

Figure 9. Accuracy of tag suggestion for seven new landmarks

REFERENCES
[1]　M. Wang, et al, "Assistive tagging: A survey of multimedia tagging with human-computer joint exploration", *ACM Comput. Surv.*, vol. 44, no. 4, pp. 1-24, 2012.
[2]　G. Li, "Improving Relevance Feedback in Image Retrieval by Incorporating Unlabelled Images", *TELKOMNIKA Indonesian Journal of Electrical Engineering*, vol. 11, no. 7, pp. 3634-3640, 2013.
[3]　Z. Wei, et al, "An Efficient Content Based Image Retrieval Scheme", *TELKOMNIKA Indonesian Journal of Electrical Engineering*, vol. 11, no. 11, pp. 6986-6991, 2013.
[4]　S.V. Canneyt, et al, "Discovering and Characterizing Places of Interest Using Flickr and Twitter", *Int. J. Semant. Web Inf. Syst.*, vol. 9, no. 3, pp. 77-104, 2013.
[5]　I. Memon, et al, "Travel Recommendation Using Geo-tagged Photos in Social Media for Tourist", *Wirel. Pers. Commun.*, vol. 80, no. 4, pp. 1347-1362, 2015.
[6]　H. Deljooi, et al, "A Novel Semantic Statistical Model for Automatic Image Annotation Using the Relationship between the Regions Based on Multi-Criteria Decision Making", *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 4, no. 1, pp.37-51, 2014.
[7]　A. Popescu, et al, "*MonuAnno: automatic annotation of georeferenced landmarks images*", in Proceedings of the ACM International Conference on Image and Video Retrieval, 2009, pp. 1-8.
[8]　Y. Zhang, et al, "*An Image Recognition Method Based on Scene Semantics*", in Proceedings of the International Conference on Information Engineering and Applications (IEA), 2012, pp. 461-468.
[9]　H. Liu, et al, "Robust and accurate mobile visual localization and its applications", *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 9, pp. 1-22, 2013.
[10]　X. Li, et al, "*Modeling and Recognition of Landmark Image Collections Using Iconic Scene Graphs*", in Proceedings of the 10th European Conference on Computer Vision: Part I, 2008, pp. 427-440.
[11]　L. Yunpeng, et al, "Landmark classification in large-scale image collections", in *Computer Vision, 2009 IEEE 12th International Conference on*, 2009, pp. 1957-1964.
[12]　D. Joshi, et al, "Inferring photographic location using geotagged web images", *Multimedia Tools Appl.*, vol. 56, no. 1, pp. 131-153, 2012.
[13]　S. Papadopoulos, et al, "Cluster-Based Landmark and Event Detection for Tagged Photo Collections", *MultiMedia, IEEE*, vol. 18, no. 1, pp. 52-63, 2011.
[14]　I. Ivanov, et al, "Geotag propagation in social networks based on user trust model", *Multimedia Tools and Applications*, vol. 56, no. 1, pp. 155-177, 2012.
[15]　Q. Fang, et al, "Discovering Geo-Informative Attributes for Location Recognition and Exploration", *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 11, pp. 1-23, 2014.
[16]　X. Li, et al, "*Learning tag relevance by neighbor voting for social image retrieval*", in Proceedings of the 1st ACM international conference on Multimedia information retrieval, 2008, pp. 180-187.
[17]　L. Xirong, et al, "Learning Social Tag Relevance by Neighbor Voting", *Multimedia, IEEE Transactions on*, vol. 11,no. 7, pp. 1310-1322, 2009.
[18]　D. Liu, et al, "Semi-Automatic Tagging of Photo Albums via Exemplar Selection and Tag Inference", *Trans. Multi.*, vol. 13, no. 1, pp. 82–91, 2011.
[19]　J. Wang, et al, "Columbia TAG System - Transductive Annotation by Graph Version 1.0.",*ADVENT Technical Report #225-2008-3 Columbia University*, 2008.
[20]　T.B. Le, et al, "On incrementally using a small portion of strong unlabeled data for semi-supervised learning algorithms", *Pattern Recogn. Lett.*, vol. 41, pp. 53–64, 2014.
[21]　L. Itti, et al, "A model of saliency-based visual attention for rapid scene analysis", *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 20, no. 11, pp. 1254–1259, 1998.
[22]　J. Harel, et al, "Graph-based visual saliency", in *the Advances in Neural Information Processing Systems 19*, 2007, pp. 545–552.
[23]　Xiaodi Hou, et al, "Saliency Detection: A Spectral Residual Approach", in *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, 2007, pp. 1–8.
[24]　D.G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints", *Int. J. Comput. Vision*, vol. 60, no. 2, pp. 91–110, 2004.
[25]　A. Oliva, et al, "Modeling the Shape of the Scene: A Holistic Representation of the Spatial Envelope", *Int. J. Comput. Vision*, vol. 42, no. 3, pp. 145–175, 2001.
[26]　J. Philbin, et al, "The Paris Dataset", Department of Engineering Science, University of Oxford, http://www.robots.ox.ac.uk/~vgg/data/parisbuildings/, accessed Aug. 27. 2014.

BIOGRAPHIES OF AUTHORS

Sutasinee Chimlek, she received her B.Sc. in computer science from Chiang-Mai University in 1996 and in her M.Sc. in information technology from King Mongkut's Institute of Technology Ladkrabang in 2001. She is currently a Ph.D. candidate in computer engineering at Kasetsart University (KU). Her research interests include pattern analysis and multimedia processing.

Punpiti Piamsa-nga, he received his B.Eng. and M.Eng. in electrical engineering from Kasetsart University (KU) in 1989 and 1993, respectively and his D.Sc. in Computer Engineering from George Washington University in 1999. He is currently an associate professor in computer engineering at KU. His research interest is pattern recognition and multimedia processing.