A New Communication System: Invisible Blog
Attaching Blog Messages to Physical Objects

Yuko Kambe and Shinya Nakajima
School of Science and Engineering, Kokushikan University, Setagayaku 154-8515 Tokyo Japan
E-mail: nakajima@kokushikan.ac.jp Tel: +81-3-5481-3336

Abstract **This paper describes the new communication system named "Invisible Blog System (IBS)". IBS enables users to attach their blog messages to physical objects and/or scenic locations. Another user that visits the scenic locations to which one or more IBS messages are attached, can access the messages. The main technological issues in realizing IBS are; how to identify the scenic locations and/or objects and so embed the IBS messages and how to browse them. This paper focuses on the first point and proposes a location identification method. The method utilizes RGB scalar-quantized histograms as a feature vector. Location identification experiments are performed at 20 sites. The identification rate is over 96% on average which confirms that the proposed method is valid and effective.**

I.INTRODUCTION

Communication services such as Blog and SNS (Social Network Service) are becoming very popular. In addition, multimedia content upload/download sites like You-Tube are very well accepted. With these Internet-based services, people can access worldwide information and may receive responses from anywhere in the world, while using a PC at home. However, this may cause new problems including the 'hikikomori' (social withdrawal).

Given this background, our goal is to create a new Internet-based communication media that is strongly dependent on physical (outdoor) places. Our solution is a new communications media named "Invisible Blog System (IBS)" that attaches electronic information like BLOG entries to physical outdoor places and/or objects.

IBS enables users to attach a Blog message to a physical place after taking a picture of the site with a mobile phone that runs IBS. When someone else takes a picture of the place to which one or more IBS messages have been attached, IBS alerts him to the existence of existing IBS messages via the mobile-network. He then can browse the messages and even write a reply to one of the IBS posters. This reply is forwarded to the originator of the message.

The following section gives an overview of IBS. We then describe the scene identification method, which enables IBS messages to be attached to physical places. After reporting the results of a scene identification experiment, we finally give our conclusion and future work.

II.OVERVIEW OF THE INVISIBLE BLOG SYSTEM

IBS enables users to attach a blog message to any place they visit. For instance, if you visit a scenic place and take a picture with your mobile phone, you can attach a blog message to that place. Anyone who visits the same place can find the message and browse it by using an IBS-equipped mobile phone.

Fig. 1 shows how IBS works. As can be seen in the figure, a Blog message can be attached to a physical place and you can find and browse the message using an IBS-equipped mobile phone.

Fig.1 How the Invisible Blog System works.

An overview of the system is given in Fig. 2. When attaching a blog message, the picture of the location and the GPS information are automatically uploaded to the IBS server together with the corresponding blog message. The key to IBS is the scenic point identification engine. GPS provides just rough location information; actually, the accuracy of GPS data is around 40m. The objects we assume here are buildings, houses, monuments, and sometimes entrance doors/exits; the distance accuracy assumed here is around 5m or better.

A number of related works have been reported. MIT developed "Post-It note," which is an RFID tag based system [1]. Post-It note can attach information to physical objects which have RFID tags. IBS can attach blog information to "scenic locations" and does not utilize any local identification system. Object recognition has been studied for a couple of decades [2][3]. However, most "object recognition" methods require a number of pictures of the object to permit recognition; front, back and side pictures for example. In case of IBS, a user does not want to take a number of pictures to attach a blog message. This means that IBS must identify a scenic object from just one picture.

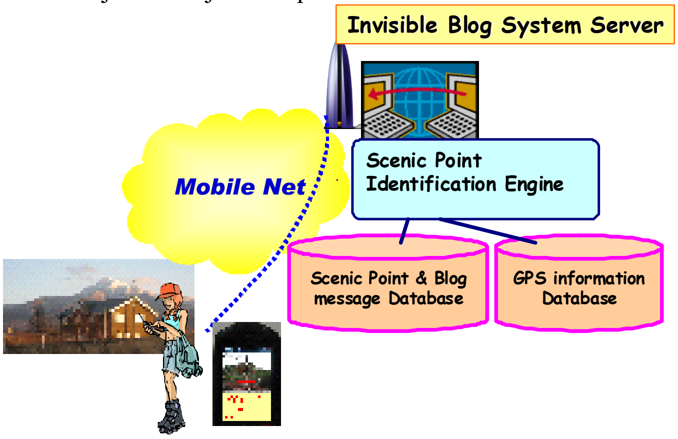

Fig.2 Overview of the Invisible Blog System

III. Scene Identification Method

Scene identification can be broken down into 2 sub processes; color information identification and shape information identification. The former is assumed to be more important than the latter and so is the focus of this paper. We recognize that shape based analysis can play an important role in scene identification; if the weather or time of day changes, the color features may change to some extent.

For color information identification, we scalar quantize the R, G, and B values of each pixel. A histogram for of each set of R, G, and B quantized values is then computed. Finally, the RGB histograms are normalized by the total number of pixels to form the feature vector of the scenic location. Fig.3 shows the feature vector generation procedure.

For scalar quantization, we apply the LBG vector quantization technique [4]; the codebook size is set to 4. The quantization process can be described as follows; i) select a pair of samples (minimum and maximum values), ii) a set of R(or G/B) values can be split into two clusters by comparing the distance from the two minimum and maximum values, iii) compute the centroid of each of the two clusters, iv) re-compute the distance from the two centroids and form revised versions of the two clusters, v) for each pair of clusters, repeat procedures i) to iv) until the number of clusters condition is satisfied.

The feature vector thus has 12-dimensions (4 x R/G/B). Generally speaking, when one takes a picture, the main object tends to be positioned in the center. Accordingly, when computing the histogram, a weight function that prefers the center over the periphery is used. The weight function is simply a trapezoidal curve, see Fig. 4

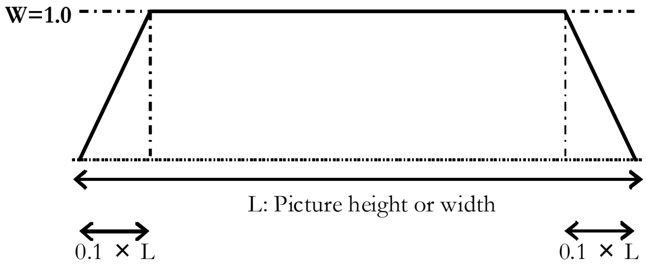

Fig.4 Weight function for computing histogram

The scene identification method is based on discriminant analysis. Given a feature vector of the source picture $\boldsymbol{h}_s$, reference picture feature vector $\boldsymbol{h}_c$, and threshold distance η, scene identification is performed as follows.

If $\mathbf{D}(\boldsymbol{h}_s, \boldsymbol{h}_c) < \eta$, $\boldsymbol{h}_s$ and $\boldsymbol{h}_c$ are taken from the same scenic location, …(1)

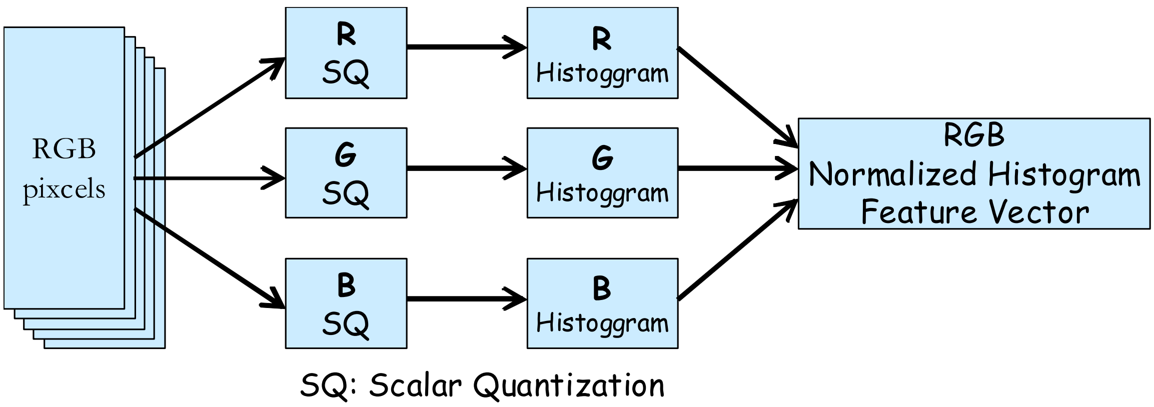

Fig.3 RGB feature vector generation

where D(x,y) stands for Euclidian distance normalized by the variance of each dimension;

$$D(h_s, h_c) = \sum_{i=1}^{12} \frac{(h_s(i) - h_c(i))^2}{\text{var}_i} \quad , \ \ldots(2)$$

where $\mathbf{var_i}$ stands for the variance of i-th dimension values. We assume that the distance variable follows a normal distribution. This is likely to be true when the invisible blog system server holds huge database of scenic pictures. Given the distance x between two pictures, the probability that x was taken from the same place, distance distribution $f_s(x)$, is computed by;

$$f_s(x) = p(x \mid g_s) = \frac{1}{\sqrt{2\pi}\sigma_s} \exp\left[-\frac{1}{2}\left(\frac{x-\mu_s}{\sigma_s} \right)^2 \right] , \ \ldots(3)$$

where g_s means "taken from the same place", μ_s, σ_s are the average distance and standard deviation of the same place distribution. The probability that x is a different place, distance distribution $f_d(x)$, is computed by;

$$f_d(x) = p(x \mid g_d) = \frac{1}{\sqrt{2\pi}\sigma_d} \exp\left[-\frac{1}{2}\left(\frac{x-\mu_d}{\sigma_d} \right)^2 \right] , \ \ldots(4)$$

where g_d means "taken from a different place", μ_d, σ_d are average distance and standard deviation of the different place distribution. Discrimination function L is computed as follows.

$$L = \log f_s(x) - \log f_d(x) \ldots(5)$$

If L>0, distance x can be taken as indicating that the images are of the same place. Solving the equation for L=0 yields the optimum boundary η.

IV. Scene Identification Experiment

To evaluate the proposed scenic point identification method, we performed an identification experiment.

We prepared a couple of angled photographs of 20 scenic locations. Figure 5 shows two pictures taken from basically the same place.

Fig.5 Two pictures taken from the same place.

As mentioned in the previous section, the Euclidian distance of normalized histogram of scalar quantized RGB was used as the identification metric.

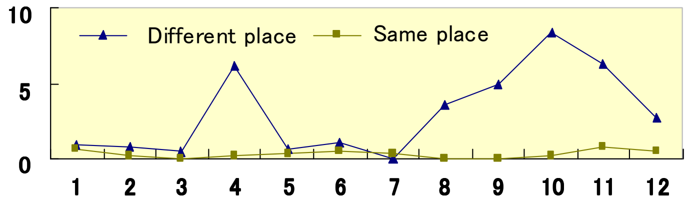

Fig. 6a Dimension-wise distance curves of feature vectors: Place#1.

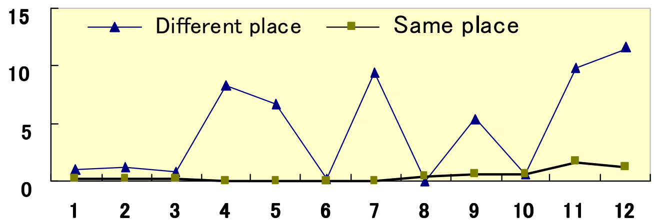

Fig. 6b Dimension-wise distance curves of feature vectors: Place#2.

Fig 6a and 6b show the dimension-wise distance curves of the feature vectors from the same and different places. As can be seen from the figures, the separation between the same and different place feature vectors is significant.

From a pair of pictures from 20 locations, we computed average distances and their standard deviations. Fig. 7 shows the average distances and standard deviations (STD) of the same and different places. As can be seen in the figure, the difference between the same and different places is remarkable and in fact, is statistically significant. By solving (5) for L=0, the optimum discrimination boundary η is found to be 0.44. This boundary value was used in a scene identification experiment; for 20 places (2 pictures for each place) the identification rate was 96%. The statistical parameters were computed from the same set of pictures. The identification result thus reflects a "closed test".

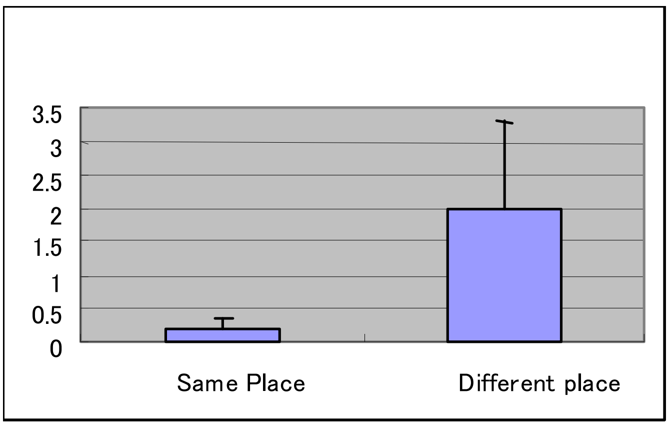

Fig.7 Average distances and standard deviations of the same and different places.

V. CONCLUSIONS

This paper proposes the 'real-world' based communication media named Invisible Blog System (IBS). IBS enables users to attach blog messages to physical scenic locations and the blog messages can be browsed by others who visit the same place. The main technical issue is how to identify the scenic locations at which the blog messages were attached. We proposed an identification method which includes the use of scalar quantization based RGB histograms. A location identification experiment was conducted using two different images for 20 places. The closed test results indicated the proposed method can identify the similarity of pictures with an accuracy of at least 96%. A statistical analysis yielded the optimum discrimination threshold of 0.44.

With regard to realizing the IBS service, a number of technical issues remain unsolved. There are at least two major issues. One is how to handle two pictures of the same scenic view that have different color conditions. The other concerns the mobile phone network and the configurations of the IBS server.

The current scenic point identification method strongly depends on the color features of the images and shape information may be needed to ensure good identification when the color changes due to weather and/or time of day. The scale-invariant features (SIFT) based approach is effective [5] for shape and texture matching. We plan to apply SIFT based analysis to IBS. Moreover, to overcome the weakness of color information analysis, more structural aspects should be taken into account. To improve the identification accuracy, we plan to use RGB vector quantization features of multiple key points of the scenic location picture.

REFERENCES

[1] www.engadget.com/2008/05/01/mit-reinvents-the-post-it-note-with-post-it-notes an Engadget web article

[2] T.M. Strat and M.A. Fischler, "Context-Based Vision: Recognizing objects using information from both 2-D and 3-D imagery," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.13, No.10, pp. 1050-1065, 1991.

[3] H. Murase and S.K. Nayar, "Visual learning and recognition of 3-D objects from appearance," *International Journal of Computer Vision*, Vol.14, No.9, pp. 5-24, 1995.

[4] Y. Linde, A. Buzo, and R. M. Gray, "An Algorithm for Vector Quantizer Design," *IEEE Transactions on Communications,* Vol. Com-28. pp. 84-95, 1980.

[5] D.G.Lowe, "Object recognition from local scale-invariant features", *Proceedings of International Conference on Computer Vision*, pp.1150-1157, 1999.