Optimal Feature Matching Method using Bayesian Graph Theory

Wan Hyun Cho[1], In Seop Na[2,*] , Sun Worl Kim[3] and Soo Hyung Kim[2]

[1]*Department of Statistics, Chonnam National University*
[2]*School of Electronic & Computer Engineering, Chonnam National University*
[3]*Radiation Health Research Institute, Korea Hydro & Nuclear Power Co. Ltd.*
whcho@jnu.ac.kr, ypencil@hanmail.net, sunworl@gmail.com, shkim@jnu.ac.kr

Abstract

Local feature matching is an essential component of many image and object retrieval algorithms. Feature similarities between object model and scene graph are complemented with a regularization term that measures differences of the relational structure. In this paper, we present a novel approach to the optimal feature matching using new Bayesian graph theory. First, we will discuss properties of various local invariant feature detectors and descriptors for scale, affine transformation and illumination changes. Second, we propose an efficient features corresponding method using local invariant features and new graph matching algorithm. Main theoretical background of our algorithm is that it can be based on the Bayes theorem and an iterative convex successive projection algorithm used to obtain the global optimum solution for feature matching problem. Finally, we have conducted the comparative experiments between proposed method and existing method on various real images. Experimental results show that our method outperforms clearly rather than the existing algorithms about feature correspondence in two images with rotation or scale transformation and illumination changes.

Keywords: *Local Invariant Feature; Bayesian Graph Theory; Feature Matching Method, Convex SDP*

1. Introduction

Many features matching algorithms have been proposed during last few decades [1-3]. Among these methods, the similarity measure is one of the most powerful tools for feature matching. In order to find the corresponding point for a feature point using the similarity measure, a template window is considered around the feature point and this window is shifted pixel by pixel across a larger search window around an estimated corresponding point, and in each position the similarity between the two regions is measured. The maximum or minimum value of the resultant measurements defines the position of the best match. Normalized cross correlation and SSD (sum of squared differences) are well-known methods for measuring similarity between two regions. In addition to a normalized similarity value, normalized cross correlation has the advantage of being invariant to the linear change between the data sets, which makes the algorithm robust against low varying illumination change the scene. In recently, Torresani, et. al., [4] presented a new approach for establishing correspondences between spare image features. They formulate this matching task as an energy minimization problem by defining a complex objective function of appearance and the spatial arrangement

* Corresponding author. Tel.: +82 62 530 0430; fax: +82 62 530 3439.

E-mail address: ypencil@hanmail.net (I.S. Na).

of the features and used a novel graph matching optimization technique, which they refer to as dual decomposition.

In this paper, we present a novel approach to the optimal feature matching using new Bayesian graph theory. First, we will discuss properties of various local invariant feature detectors and descriptors for scale, affine transformation and illumination changes. Second, we propose an efficient features matching method using local invariant features and new graph matching algorithm. Main theoretical background of our algorithm is that it can be based on the Bayes theorem and an iterative convex successive projection algorithm used to obtain the global optimum solution for feature matching problem. Finally, we have conducted the comparative experiments between proposed method and existing method on various real images.

2. Detectors and Descriptors

A local feature on distinguished region of given image represents an image pattern which differs from its immediate neighborhood. It is usually associated with a change of an image property or several properties simultaneously. Here, we have discussed various detectors and descriptors describing the image characters that are locally invariant with image rotation, scale transformation and illumination changes.

2.1. Local Invariant Feature Detector

The currently most popular distinguished detectors can be roughly divided into three categories; corner based or intensity based, contour based or region based, parametric model based detector or other approaches [5,6,7]. Corner based detectors locate interest points and region which contain a lot of image structure such as edges or intensity changes around points. To measure the change, first and second derivatives of images are used in many different forms and combination. Therefore, these are not suited for uniform regions and region with smooth transitions. Region based detectors regard local blobs of uniform brightness as the most salient aspects of an image and extract contours from image. Parametric model based detectors or other approach take into account the entropy of a region or try to find interest points by matching models or templates to an image. The most popular interest detectors, which give sufficient performance results, are listed: Harris or Hessian point based detectors (Harris, Hessian, Harris-Laplace, Hessian-Laplace), Difference of Gaussian Points (DoG) detector, Entropy Based Salient Region (EBSR) detector, Harris or Hessian affine invariant region detectors (Harris-Affine, Hessian-Affine), Maximally Stable External Regions (MSER) detector, and Edge Based Regions (EBR) detector and Intensity Based Regions (IBR) detector.

2.2 Local Invariant Feature Descriptor

The currently most popular distinguished descriptors can be roughly divided into three categories; distribution based descriptors, filter based descriptors, and other methods [8,9]. Distribution based descriptors use histograms to represent the characteristics of the region. The characteristics could be pixel intensity, distance from the center point, relative ordering of intensity, or gradient. Filter based descriptors use a set of differential operators or different types of filters to describe an interest region. Other approach takes into account the moments or gradient moment with high order degrees to characterize shape and intensity distribution in an interest region. Many different techniques for describing local image regions have been

developed. They are listed: Scale Invariant Feature Transform (SIFT) descriptor, PCA-SIFT, Gradient location and orient histogram (GLOH), shape context, spin images, differential invariants, steerable filters and complex filters, and moment invariants.

3. Efficient Local Invariant Feature using New Graph Matching Method

3.1 Formulation of Probabilistic Feature Matching Algorithm

Here, we will formulate feature matching problem as Bayesian inference framework. The feature matching implies the correspondence between two sets of feature detectors $(P^{(1)}, P^{(2)})$ using their descriptor properties $(D^{(1)}, D^{(2)})$ given from the two input $I^{(1)}$ and $I^{(2)}$. This is represented by the bi-directional detector to detector correspondences mapping $m_1 : P^{(1)} \to P^{(2)}$ and $m_2 : P^{(2)} \to P^{(1)}$. And also it is efficiently represented using an assignment matrix $\mathbf{X} \in \Re^{N_1 \times N_2}$ of nonnegative real numbers, where N_1 and N_2 denote the numbers of feature detectors in $P^{(1)}$ and $P^{(2)}$ respectively. That is, each component $X_{(i,j)}$ of the assignment matrix $\mathbf{X}$ is representing the possibility that a detector $p_i^{(1)} \in P^{(1)}$ matches to some detector $p_j^{(2)} \in P^{(2)}$. Hence, in order to find the optimal assignment matrix, we are going to apply the Bayesian inference principle with the matching problem between two sets of feature detectors $(P^{(1)}, P^{(2)})$.

3.1.1. Prior Distribution of Assignment Matrix: First, we will consider the prior distribution for row and column vectors of the assignment matrix $\mathbf{X}$. Here, we factorize an assignment matrix $\mathbf{X}$ as the N_1 row vectors $\mathbf{x}_1^R, \cdots, \mathbf{x}_i^R, \cdots, \mathbf{x}_{N_1}^R$. Then, each component x_{ij} for one of row vector $(\mathbf{x}_i^R)^T = (x_{i1}, \cdots, x_{ij}, \ldots, x_{iN_2})$ represents the possibility that the detector $p_i^{(1)} \in P^{(1)}$ is mapping to any one of detectors $p_j^{(2)} \in P^{(2)}$. Hence, we can assume that these components satisfy two properties. These are $x_{ij} \geq 0$ for all $j = 1, \cdots, N_2$ and $\sum_{j=1}^{N_2} x_{ij} = 1$. Therefore, we can define a new discrete random variable Y taking with finite number of values $1,2,\ldots,N_2$. And the probability distributions of Y taking y can be parameterized by the component of a vector $\mathbf{x}_i^R$, that is, $p(Y = y) = x_{ij}, y = 1, \cdots, N_2$. Another way to write this is

$$p(y \mid \mathbf{x}_i) = \prod_{j=1}^{N_2} x_{ij}^{I(y=j)}, \tag{1}$$

where $I(y = j)$ denote an indicator function.

Moreover, we assume that a conjugate prior for each row probabilistic vector $\mathbf{x}_i^R$ is the Dirichlet distribution:

$$p(\mathbf{x}_i^R \mid \alpha) \sim D(\alpha_1, \cdots, \alpha_{N_2}) = \frac{\Gamma(\sum_j \alpha_j)}{\prod_j \Gamma(\alpha_j)} \prod_j x_{ij}^{\alpha_k - 1} \tag{2}$$

where $\Gamma(\cdot)$ denotes the gamma function and $\alpha_1, \cdots, \alpha_{N_2}$ are positive numbers. The hyper-parameter α_j can be interpreted as a virtual occurrence for value x_{ij}. Large α_j corresponds to strong prior knowledge about the distribution and small α_j corresponds to ignorance. Then,

using properties of the Gamma distribution and Dirichlet distribution, we can generate the pseudo random variable for the $(i, j)^{th}$ component x_{ij}^R of each row probabilistic vector $\mathbf{x}_i^R$ as follows:

$$x_{ij}^R \mid (\alpha_1, \cdots, \alpha_{N_2}) \sim \frac{\Gamma(\alpha_j, \beta)}{\sum_{k=1}^{N_2} \Gamma(\alpha_k, \beta)}, \ j = 1, \cdots, N_2 \tag{3}$$

where $\Gamma(\alpha, \beta)$ denote the Gamma distribution with parameter α and β.

Similarly, we factorize an assignment matrix $\mathbf{X}$ as the N_2 column vectors $\mathbf{x}_1^C, \cdots, \mathbf{x}_j^C, \cdots, \mathbf{x}_{N_2}^C$. We assume that a conjugate prior for each column vector $\mathbf{x}_j^C$ is the Dirichlet distribution:

$$p(\mathbf{x}_j^C \mid \alpha) \sim D(\alpha_1, \cdots, \alpha_{N_1}) = \frac{\Gamma(\sum_i \alpha_i)}{\prod_i \Gamma(\alpha_i)} \prod_i x_{ij}^{\alpha_i - 1} \tag{4}$$

where $\Gamma(\cdot)$ denotes the gamma function and $\alpha_1, \cdots, \alpha_{N_1}$ are positive numbers. Then, using properties of the Gamma distribution and Dirichlet distribution, we can generate the pseudo random variable for the $(i, j)^{th}$ component x_{ij}^C of column vector $\mathbf{x}_j^C$ as follows:

$$x_{ij}^C \mid (\alpha_1, \cdots, \alpha_{N_1}) \sim \frac{\Gamma(\alpha_i, \beta)}{\sum_{k=1}^{N_1} \Gamma(\alpha_k, \beta)}, \ i = 1, \cdots, N_1 \tag{5}$$

where $\Gamma(\alpha, \beta)$ denote the Gamma distribution with parameter α and β.

3.1.2. Likelihood Function for Assignment Matrix: Second, in order to solve the feature matching in a Bayesian inference principle, we need to construct the $(N_1 \times N_2)$ likelihood function matrix $\mathbf{L}$ with likelihood components $l_{(l,m)}$ that represents the possibility of some detector $p_l^{(1)} \in P^{(1)}$ matching to any detector $p_m^{(2)} \in P^{(2)}$ for the assignment matrix $\mathbf{X}$. It can be derived from the weighted combination of the matching matrix $\mathbf{Y}$ that is induced from the similarity for a pair of descriptors defined at two detectors and the distance matrix $\mathbf{D}$ that represents the spatial distance between locations of detectors.

First, we consider the matching matrix $\mathbf{Y}$ that is induced from the similarity for a pair of descriptors corresponding two detectors. We define edges $e_{ij}^{(1)} \in E^{(1)}$ and $e_{ab}^{(2)} \in E^{(2)}$ as a pair of detectors $(p_l^{(1)}, p_j^{(1)})$ and $(p_a^{(2)}, p_b^{(2)})$. For each pair of $e_{ij}^{(1)} \in E^{(1)}$ and $e_{ab}^{(2)} \in E^{(2)}$, the similarity matrix S of size $N_1 N_1 \times N_2 N_2$ is defined by having its elements as similarities:

$$\mathrm{S}\left(\mathrm{e}_{\mathrm{ij}}^{(1)}, \mathrm{e}_{\mathrm{ab}}^{(2)}\right) = \exp(-\mid \parallel \mathrm{e}_{\mathrm{ij}}^{(1)} \parallel - \parallel \mathrm{e}_{\mathrm{ab}}^{(2)} \parallel) \tag{6}$$

Here $\parallel \mathrm{e}^{(\mathrm{k})} \parallel$ is denoted with the norm of edge. Then, we can compute the $(N_1 \times N_2)$ matching matrix $\mathbf{Y}$ with components $y_{(l,m)}$ that represents the possibility of some detector $\mathrm{p}_{\mathrm{l}}^{(1)} \in \mathrm{p}^{(1)}$ matching to any detector $\mathrm{p}_{\mathrm{m}}^{(2)} \in \mathrm{p}^{(2)}$ from the similarity matrix S as following form. That is, it can be obtained by summing the similarities for all pairs of edges containing detectors $p_l^{(1)} \in P^{(1)}$ and $p_m^{(2)} \in P^{(2)}$:

$$\mathbf{Y} = (y_{(l,m)}), \quad y_{(l,m)} = \sum_{\substack{e_{ij}^{(1)} | p_i^{(1)} = p_l^{(1)} \\ e_{ab}^{(2)} | p_b^{(2)} = p_m^{(2)}}} S(e_{ij}^{(1)}, e_{ab}^{(2)}) \tag{7}$$

Second, we think about the distance matrix $\mathbf{D}$ that represents the spatial distance between positions for a pair of detectors contained two images. In matching problem for two detectors, we have to consider the local information which represents the spatial relationship of neighbors for detectors. The possibility of some detector $p_i^{(1)} \in P^{(1)}$ in image $I^{(1)}$ matching to a near detector $p_m^{(2)} \in I^{(2)}$ is higher than far away detector $p_m^{(2)} \in I^{(2)}$ in image $I^{(2)}$. The local information is formulated into the distance matrix $\mathbf{D}$ as follows:

$$\mathbf{D} = (d_{(l,m)}), \quad d_{(l,m)} = \exp(-\frac{|| p_l^{(1)} - p_m^{(2)} ||}{\max_{p_m^{(2)} \in I^{(2)}} || p_l^{(1)} - p_m^{(2)} ||}) \tag{8}$$

Here, the large value of $d_{(l,m)}$ amplifies the matching potentiality but the small value of $d_{(l,m)}$ attenuates this property. Finally, we combine the matching matrix and the distance matrix to construct the likelihood function matrix $\mathbf{L}$. First, in order to agree with scales for elements of the matching matrix $\mathbf{Y}$ and the distance matrix $\mathbf{D}$, we apply the bi-stochastic normalization scheme with two matrices. Next, adopting the proper weighting parameter π, the likelihood function matrix $\mathbf{L}$ is obtained by using the following equation:

$$\mathbf{L} = \pi \mathbf{Y} + (1-\pi)\mathbf{D}, \ 0 \le \pi \le 1 \tag{9}$$

Hence, we use the i^{th} row and j^{th} column element $l_{(i,j)}$ of the likelihood matrix $\mathbf{L}$ as the likelihood function that represents the possibility of some detector $p_i^{(1)} \in I^{(1)}$ matching to any detector $p_j^{(2)} \in I^{(2)}$. That is, we have that

$$\mathrm{l}_{(\mathrm{i,j})} = \mathrm{Likelihood}\left(\mathrm{p}_{\mathrm{i}}^{(1)} \Leftrightarrow \mathrm{p}_{\mathrm{j}}^{(2)}\right), \mathrm{i} = 1, \cdots, \mathrm{N}_1, \mathrm{j} = 1, \cdots, \mathrm{N}_2 \tag{10}$$

3.1.3. Posterior Distribution of Assignment Matrix: Finally, by combining the prior distribution $p(x_{ij} | \alpha)$ and the likelihood function $L((p_i^{(1)}, p_j^{(2)}) | x_{ij})$ of $(i,j)^{th}$ component x_{ij} of row vector $\mathbf{x}_i^R$ and column vector $\mathbf{x}_j^C$ using Bayes formula, we have obtained the posterior distribution $p(x_{ij}^R | \alpha, L_{ij})$ of $(i,j)^{th}$ component x_{ij}^R of row vector $\mathbf{x}_i^R$ and the posterior distribution $p(x_{ij}^C | \alpha, L_{ij})$ of $(i,j)^{th}$ component x_{ij}^C of column vector $\mathbf{x}_j^C$ respectively as follows:

$$(x_{ij}^R | \alpha, L_{ij}) = \frac{(x_{ij}^R | \alpha) L((p_i^{(1)}, p_j^{(1)}) | x_{ij}^R)}{\sum_{k=1}^{N_2} (x_{ik}^R | \alpha) L((p_i^{(1)}, p_k^{(1)}) | x_{ik}^R)}, \ i = 1, \cdots, N_1, \ j = 1, \cdots, N_2 \tag{11}$$

$$(x_{ij}^C | \alpha, L_{ij}) = \frac{(x_{ij}^C | \alpha) L((p_i^{(1)}, p_j^{(1)}) | x_{ij}^C)}{\sum_{k=1}^{N_1} (x_{kj}^C | \alpha) L((p_k^{(1)}, p_j^{(1)}) | x_{kj}^C)}, \ i = 1, \cdots, N_1, \ j = 1, \cdots, N_2. \tag{12}$$

Here, we have repeated this algorithm iteratively by taking the posterior probabilities $(x_{ij}^R | \alpha, L_{ij})$ and $(x_{ij}^C | \alpha, L_{ij})$ as initial value of the hyper-parameter α_j again. If the difference between values of one step before and after step can be ignored, we will stop the iteration.

Finally, we have obtained that the row and column matching matrices $\mathbf{X}_{\mathrm{R}}$ and $\mathbf{X}_{\mathrm{C}}$ between two features $(P^{(1)}, P^{(2)})$ are defined by

$$\mathbf{X}^{\mathrm{R}} = \begin{pmatrix} \mathbf{x}_1^{R\,T} \\ \vdots \\ \mathbf{x}_{N_1}^{R\,T} \end{pmatrix} \text{ and } \mathbf{X}^{\mathrm{C}} = \begin{pmatrix} \mathbf{x}_1^{C} & \cdots & \mathbf{x}_{N_2}^{C} \end{pmatrix} \tag{13}$$

3.1.4 Computation of Optimal Feature Matching Solution by Convex Problem Relaxation: Next, we consider an optimization problem for recovering the optimal solution $\mathbf{X}^*$ from probabilistic matching matrices $\mathbf{X}^{\mathrm{R}}$ and $\mathbf{X}^{\mathrm{C}}$ by minimizing the distance between $\mathbf{X}$ and $\mathbf{X}^{\mathrm{R}}$ or between $\mathbf{X}$ and $\mathbf{X}^{\mathrm{C}}$ defined by the following form:

$$\mathbf{X}^* = \arg\min_{\mathbf{X}\geq 0} \left(D_{KL}(\mathbf{X}^{\mathrm{R}} \,||\, \mathbf{X}) \right) \text{ or } \mathbf{X}^* = \arg\min_{\mathbf{X}\geq 0} \left(D_{Kl}(\mathbf{X}^{\mathrm{C}} \,||\, \mathbf{X}) \right)$$
$$\text{Such that } \mathbf{X}\mathbf{1} \leq \mathbf{1},\ \mathbf{X}^T\mathbf{1} \leq \mathbf{1},\ \mathbf{1}^T\mathbf{X}\mathbf{1} = k \tag{14}$$

In this case, we are going to use the distance function as the Kullback-Leibler divergence that is a non-symmetric measure of the difference between two probability distribution $\mathbf{P}$ and $\mathbf{Q}$. It is defined to be

$$D_{KL}(\mathbf{P}\,||\,\mathbf{Q}) = \sum_{(i,j)} P(i,j) \log \frac{P(i,j)}{Q(i,j)} \tag{15}$$

Moreover, we assume that k detectors are matching at two images.

This is a special case of the general convex semi-definite program (SDP) which has the following standard form:

$$\min \ \mathrm{Tr}[\,\widetilde{Q}\,\mathbf{X}\,] \text{ s.t. } \mathrm{Tr}[\,A_i\,\mathbf{X}\,] = C_i \ \text{ for } i = 1,\cdots,m,\ \ \mathbf{X} \geq 0 \tag{16}$$

In our case, the objective function for the semi-definite relaxation is given by the following form:

$$f(\mathbf{X}) = \sum_{(i,j)} x^R_{(i,j)} \log \frac{x^R_{(i,j)}}{x_{(i,j)}} \quad \text{or} \quad f(\mathbf{X}) = \sum_{(i,j)} x^C_{(i,j)} \log \frac{x^C_{(i,j)}}{x_{(i,j)}} \tag{17}$$

where $x_{(i,j)}$, $x^R_{(i,j)}$ and $x^C_{(i,j)}$ are respectively the $(i,j)^{th}$ component of matrix $\mathbf{X}$, $\mathbf{X}^{\mathrm{R}}$ and $\mathbf{X}^{\mathrm{C}}$ and the convex SDP constraints sets are defined by:

$$C_1 = \{\mathbf{X} | \mathbf{X}\mathbf{1} \leq \mathbf{1},\ \mathbf{X} \geq 0\},\ C_2 = \{\mathbf{X} | \mathbf{X}^T\mathbf{1} \leq \mathbf{1},\ \mathbf{X} \geq 0\} \text{ and } C_3 = \{\mathbf{X} |\ \mathbf{1}^T\mathbf{X}\mathbf{1} = k,\ \mathbf{X} \geq 0\} \tag{18}$$

To find the globally optimal solution $\mathbf{X}$ of object function $f(\mathbf{X})$, we want to use a generalization of successive Bregman projections as proposed in [10]. Before we consider the algorithm, we need to define a sub-problem operator $P_j(H)$ for $j = 1,2,3$ as follows:

$$P_j(H) = \arg\min_{X \in C_j} f(X) - < X, H > \tag{19}$$

Where $< X, H >= \sum_{(i,j)} A(i,j)B(i,j)$ is the dot product between two matrices. Then, by repeated applications of $P_j(H)$ in a cycle manner, we obtain a primal-dual block update algorithm which is defined as follows.

Algorithm:

Initial Step: Define $\lambda_j^{(0)}, X_j^{(0)}$, $j = 1,2,3$ and set $\lambda_j^{(0)} = 0$ and $X_j^{(0)} = \mathbf{X}^{\mathrm{R}}$, $j = 1,2,3$

Use the convention $X_0^{(t)} = X_3^{(t-1)}$

Iteration Step: Iterate on $t = 1,2,\ldots$ until convergence:

For $j = 1,2,3$:

$$\mathbf{X}_j^{(t)} = P_j(\lambda_j^{(t-1)} + \nabla f(\mathbf{X}_{j-1}^{(t)})) \ , \ \lambda_j^{(t)} = \lambda_j^{(t-1)} + \nabla f(\mathbf{X}_{j-1}^{(t)}) - \nabla f(\mathbf{X}_j^{(t)})$$

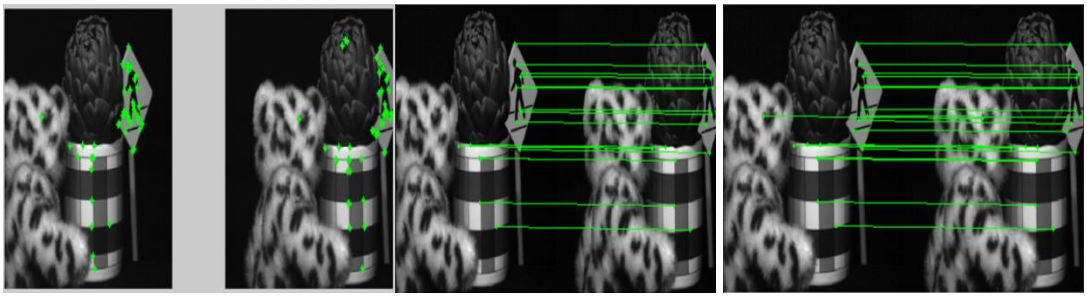
Final Step: At convergence with T iterations, the optimal solution is given by

$$\mathbf{X}^* = \frac{1}{3}(\mathbf{X}_1^{(T)} + \mathbf{X}_2^{(T)} + \mathbf{X}_3^{(T)})$$

The algorithm employs successive Bregman projections and is derived using the framework of Frenchel Duality.

4. Experimental Results

In this section, we evaluate the comparative performances of our method (BGMA) with several existing feature matching approach such as Lowe's shift matching (SM) [11] by conducting experiments on various real image data.



(a) (b) (c)

Figure1. Experiments on images from Toole with rotation and translation transformation. (a) 47 and 48 Hessian detectors extracted from left and right image, (b) SM: 16 correct matches out of 47(left image) and 48(right image) detected points, (c) BGMA: 19 correct matches out of 47(left image) and 48(right image) detected points

We test on two image pairs taken from http://vasc.ri.cmu.edu/idb/html/motion/. The candidate local features are generated using the Hessian detector and shift descriptor. Figure 1 shows that our method is better matching than SM method in rotated and translated images.

5. Conclusions

We have presented the mathematical formula for the feature matching problem using both local invariant features and Bayesian graph matching algorithm. Our method considers the properties of detectors and descriptors extracted from given image. And it also induces the bi-directional detector to detector probabilistic matching algorithm. Main advantage of our algorithm is that it can be achieved by Bayes theorem and an iterative convex successive projection algorithm in order to the global optimum solution for feature matching problem. Experimental evaluations demonstrate that our method clearly outperforms the previous matching approaches on various real images.

Acknowledgements

This work was supported in part by the Korea Research Foundation Grant by the Korean government (KRF-2010-0013757)) and this research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2011-0029429).

References

[1] E. Trucco and A. Verri, "Introductory Techniques for 3-D Computer Vision", Prentice Hall**, (1998)**.

[2] R. Szeliski, "Computer Vision: Algorithms and Applications", Springer, **(2010)**.

[3] D. T. Kien, "A Review of 3D Reconstruction from Video Sequences", ISIS MediaMill3D technical report series, University of Amsterdam, The Netherlands, **(2005)** [Technical report].

[4] L. Torresani, V. Komogorov and C. Rother, "Feature correspondence via graph matching: models and global optimization", Proceeding of the 10th ECCV: Part 2, Springer-Verlag, **(2008)**.

[5] P. M. Roth and M. Winter, "Survey of Appearance-Based Methods for Object Recognition", Technical Report ICG-TR-01, Inst. For Computer Graphics and Vision, Graz University of Technology, Austria, **(2008)** [Technical report].

[6] K. Mikolajczyk and C. Schmid, International Journal of Computer Vision, vol. 60, no. 63, **(2004)**.

[7] T. Tuytelaars and K. Mikolajczyk, "Local Invariant Feature Detectors: A survey", Computer Graphics and Vision, vol. 3, **(2007)**, pp.177-280.

[8] K. Mikolajczyk and C. Schmid, IEEE Trans. on PAMI, vol. 27, no. 10, **(2005)**, pp. 1615.

[9] E. Balmashnova and L. M. J. Florack, Journal of Math Imaging Vision, vol. 31, no. 121, **(2008)**.

[10] C. Schelleward and C. Schnor, "Probabilistic Subgraph Matching based on Convex Relaxation", In Proc. Int. Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR'05), Springer, vol. 3757 of LNCS, **(2005)**, pp. 171-186.

[11] D. G. Lowe, International Journal of Computer Vision, vol. 60, no. 91, **(2004)**.

Authors

Wan Hyun Cho received his B.S.and M.S. degree in Mathematics fro m Chonnam National University, Korea in 1977 and 1981 respectively. And his Ph.D degree in Statistics from Korea University, Korea in 1988, He is a professor in the Department of Statistics, Chonnam National Univ ersity, Korea. His research interests are statistical modeling, pattern recog nition, and image processing.

In Seop Na received his B.S., M.S. and Ph.D. degree in Computer Science from Chonnam National University, Korea in 1997, 1999 and 2008, respectively. Since 2012, he has been a research professor in Department of Computer Science, Chonnam National University, Korea. His research interests are image processing, pattern recognition, character recognition and digital library.

Sun Worl Kim received the B.S. degree in Information Statistics from Korea National Open University, Korea in 2004, and the M.S. and Ph.D degrees in Statistics from Chonnam National University in 2006 and 2011. Currently, she is working in the radiation health research institute. Her research interests are statistical modeling, pattern recognition, and image processing.

Soo Hyung Kim received his B.S. degree in Computer Engineering from Seoul National University in 1986, and his M.S. and Ph.D degrees in Computer Science from Korea Advanced Institute of Science and Technology in 1988 and 1993, respectively. From 1990 to 1996, he was a senior member of research staff in Multimedia Research Center of Samsung Electronics Co., Korea. Since 1997, he has been a professor in the Department of Computer Science, Chonnam National University, Korea. His research interests are pattern recognition, document image processing, medical image processing, and ubiquitous computing.