Review of Geometric Consistency Preserving Techniques for Content Based Image retrieval

Dr. Vinayak Kottawar
Associate professor, D Y Patil College of Engineering Akurdi , vgkottawar@dypcoeakurdi.ac.in

Dr. Neeta Deshpande
Associate Professor, R. H. Sapat College of Engineering, Nashik

Dr. Vijaykumar S. Jatti
Associate Professor , Symbiosis Institute of Technology, Pune , vijaykumar.jatti@sitpune.edu.in

Abstract :

Today, image data exist with extremely diverse visual and semantic content, and it is rapidly growing in size. This created innumerable possibilities and hence considerations for real-world image search system designers. Manual annotation of images with keywords describing the image content can make it easier to find images of interest, but this takes more time, making this approach very costly [1,2]. Thus, searching in image collection based on visual content is potentially a very powerful technique.

CBIR uses visual contents of an image to search the desired images.It deals with the fundamental problem to mathematically describe the image content (image signature) and then, assessing the similarity between a pair of images based on their signatures. Inspite of the apparent simplicity of this, there are significant obstacles that need to be overcome in order to design an efficient CBIR system.

Keywords: *CBIR, RANSAC,* Dynamic Programming, Image Retrieval

DOI Number: 10.14704/nq.2022.20.9.NQ44510 **Neuro Quantology 2022; 20(9):4461-4468**

Introduction :

It is important to find a good image signature using suitable image features for the design of an efficient CBIR system.Most of the CBIR approaches rely on a preprocessing step of feature extraction, which aims to extract suitable image features (descriptors) such as color, texture, shape and spatial layout, that carries enough information to allow successful retrieval of relevant images from a database containing thousands or millions of images.

It is observed that, the problem of searching images according to these visualsignatures is very challenging since there are many factors affecting the retrieval performance, like resolution, illumination variations and occluded objects. This motivates the need for an intermediate representation such as Visual Word and Vocabulary. Visual Word is a small part of an image which carries some information related to these features.

When the retrieval process considers only the model vectors that represent the visual content of images, it sometimes fails to produce accurate results. This is because the approach based on visual vocabulary, totally ignores the geometry of the extracted visualwords. Two images can contain similar visual words, but in a totally different spatial layout from the other [3]. Thus, the inclusion of a geometry consistence check would be very useful. The RANSAC (Random Sample Consensus) algorithm is popularly used to preserve the geometric consistency during image retrieval [4].

The RANSAC algorithm is a simple, yet powerful, technique that is usually applied to the task of estimating the parameters of a model, using data which is contaminated by outliers. RANSAC estimates a global relation that fits the data, while simultaneously classifying the data into inliers (points consistent with the relation) and outliers (points not consistent with the relation). Due to its ability to tolerate a large fraction of

outliers, the algorithm is a popular choice for a several robust estimation problems.

Visual Information Extraction and Representation

Content-based image retrieval uses the visual contents of an image to represent and index the images. The visual contents of the images in the database are extracted and described by visual content descriptor. The feature vectors of the images in the database forms a feature database. The system then retrieves the desired images based on the similarity of these extracted feature vectors.

A good visual feature should be invariant to the accidental variance introduced by the imaging process. However, there is an inverse relation between the invariance and the discriminative ability of visual features, since a very wide class of invariance loses the ability to discriminate between important differences.

A visual content descriptor can be either global or local. A global descriptor uses the visual features of the whole image, whereas a local descriptor uses the visual features of regions or objects to describe the image content. To obtain the local visual descriptors, an image is usually divided into parts and then features are extracted from each part locally. The easiest way of dividing an image is to partition it, which cuts the image into tiles of equal size and shape. A simple partition does not generate perceptually meaningful regions, butit is a way of representing the global features of the image at a finer resolution. A better method is to divide the image into homogenous regions according to some criterion, using region segmentation algorithms that have been rigorously investigated in computer vision. A more complex way of dividing an image, is an object segmentation that obtain semantically meaningful objects (like ball, car, horse).

Most Content based image retrieval systems perform feature extraction as a preprocessing step, generating global image features like color histogram or local descriptors like shape and texture. A region based color descriptor indexed in 3-D space along with their percentage coverage within the regions is proposed in [14], and observed to be computationallymore efficient in similarity based retrieval than conventional color histograms. The authors state that this compact representation is more efficient in terms of search and retrieval, than the highdimensional histograms, and it has also overcome the limitations associated with earlier schemes such as dimensionality reduction and color moment descriptors. In [15], a multi-resolution histogram approach is proposed which captures the spatial image information efficientlyand observed to be an effective technique to retrieve images. S. Jeong [16] has proposed a technique in which Gaussian mixture vector quantization (GMVQ) is used to extract color histograms that yield better retrieval than uniform quantization and vector quantization with squared error. A set of color and texture descriptors which are thoroughly tested for inclusion in the MPEG-7 standard, and suitable for natural images and video, is described in [17]. These include histogram-based descriptors, dominant color descriptors, spatial color descriptors and texture descriptors suited for browsing and retrieval.

Shape is a key attribute of segmented image regions, and its efficient and robust representation plays an importantrole in the retrieval. A shape similarity measure based on discrete curve evolution to simplify contours is discussed in [18]. Contour simplification helps to remove noise and irrelevant shape features from consideration. C. Schmidhas proposed a new shape descriptor for shape matching, referred to as shape context [19] which is a compact and robust to a number of geometric transformations.

A Dynamic Programming (DP) approach to shape matching has been proposed in [19]. The limitation of this approach is that the computation of Fourier descriptors and moments is slow, although pre-computation may help to produce real-time results.

For characterizing shape within images, reliable segmentation is important, without which the shape estimates will not be meaningful. Even though the general problem of segmentation in the context of human perception is difficult tosolve, there are some interesting new directions, one of the most important approaches is segmentation based on the Normalized Cuts criteria [20]. This approach is based on the theory of spectral clustering and it has been extended to texture image

segmentation by using cues of contour and texture differences [21].

Segmentation based on the mean shift procedure [22], multi-resolution segmentation of low depth of field images [23], a Bayesian framework based segmentation involving the Markov chain Monte Carlo technique [24], and an EM algorithm based segmentation using a Gaussian mixture model [25] are the recently proposed important approaches for segmentation.

Features based on local invariance such as interest points approachis used extensively in image retrieval. Scale and affine invariant interest points that can deal with significant affine transformations and illumination changes is observed to be an effective features for image retrieval [26]. Similarly, wavelet-based salient points have been popularly used for retrieval [27]. The significance of such special points lie in their compact representation of important image regions, leading to efficient indexing and good discriminative power, especially in object-based retrieval. A comparative study of different types of color interest points used in image retrieval is given in [28], while a comparative performance evaluation of the various proposed interest point detectors is reported in [29]. When a huge number of image features are available, one way to improve generalization and efficiency in classification and indexing is to work with a feature subset. An automatic feature subset selection for support vector machines (SVM) is one of the most popularly used algorithms for feature subset selection.

Object Based Image Retrieval

In real world images, the same object often has different appearances, due to the effects like partial occlusion and deformation. Factors, such as illumination or 3D viewpoint change, make the retrieval task even more challenging. Thus, it is useful to retrieve images by producing local feature correspondences between two images based on the assumption that more feature correspondences can be obtained between two images that contain the same object than that of the two irrelevant ones.Feature correspondence is considered as an important property in computer vision. It is widely used as a graph matching task, which tries to optimize a complex objective function based on the feature vectors and spatial constraints [30], [31]. This approach attains better accuracy, but it is computationally expensive, which makes these algorithms inappropriate for image retrieval task.

While the global and local feature extraction techniques presents good results in certain retrieval problems, in the case of ”object”-based applications, they are observed to be less efficient. Thus, most modern algorithms begin with the determination of some interest points within an image. The interest points are the regions that carry some properties such as invariance to various image transformations, illumination, etc. After finding the interest points, theregions are defined in the neighborhood of these points and descriptors are formed.

O. Chuma has proposed an approach in which the representation of local image structure and a matching scheme is presented which is insensitive to many appearance changes [32]. This method is applied to two-view matching problems of images from different modalities. Atechnique to learn and recognize object class models from unlabeled and unsegmented cluttered scenes is presented in [33]. In this work, objects are modeled as flexible constellations of parts. A probabilistic representation is used for all aspects of the object: shape, appearance, occlusion and relative scale. An entropy-based feature detector then applied for region selection within an image.R. Sukthankar has proposed a technique to tackle the problem of near-duplicate image retrieval with a part-based representation of images using distinctive local descriptors extracted from points of interest, which are invariant to several transformations.

Geometric Consistency Preserving Techniques for CBIR

When the retrieval process considers only the model vectors that represent the visual content of images, sometimes it fails to retrieve accurate results. This is because the object based approachescompletely ignores the geometry of the extracted interest points. Two images can contain similar visual words, but in a totally different spatial layout from each other. Thus, the inclusion of a geometry consistence check would be very useful. RANSAC is the most popular approach used for preservation of

geometric consistency. It is a simple, yet powerful, technique that is usually applied to the task of estimating the parameters of a model, using data that may be contaminated by outliers. RANSAC estimates a global relation that fits the data, while simultaneously classifying the data into inliers (points consistent with the relation) and outliers (points not consistent with the relation). Due to its ability to tolerate a large fraction of outliers, the algorithm is a popular choice forseveral robust estimation problems.

The working of RANSAC is based on the 'hypothesis and verify' framework. A minimal subset of the input data points is randomly chosen and model parameters are estimated from this subset. The model is then evaluated on the entire dataset and its support (the number of data points consistent with the model) is determined. This 'hypothesis and verify' loop is repeated until the probability of finding a model with better support than the current model falls below a certain predefined threshold. Usually RANSAC find the correct solution even with high degrees of contamination; however, the number of samples required to do so increase exponentially, and the associated computational cost is very large. There have been a number of efforts to increase the efficiency of the fundamental RANSAC algorithm. Some of these strategies [34] aim to optimize the process of model verification, while others [35] seek to modify the sampling process in order to generate more useful hypotheses. While these efforts have shown considerable potential, none of them are directly applicable in situations where real-time performance is important. Relatively fewer efforts have been taken towards the aim of formulating RANSAC in a manner that is suitable for real-time implementations.

Nister describes the preemptive RANSAC framework, where a fixed number of hypotheses are evaluated in a parallel, multi-stage setting [36]. In this case the aim is to find, the best possible solution from a restricted set of hypotheses within a fixed time constraint.

The preemptive RANSAC framework is useful for real-time implementation but there are few limitations of this scheme. One of the crucial limitations of preemptive RANSAC is its inherent non-adaptiveness to the data. The selection of a fixed number of hypotheses implies that a good prior estimate of the inlier ratio is available; in practice, this is often not the case. For low contamination problems, preemptive RANSAC is often slower than standard RANSAC, since it evaluates many more hypotheses than necessary. On the other hand, when the inlier ratio is too low, preemptive RANSAC is unlikely to find a good solution, since it does not test enough hypotheses.Fundamental RASNAC and its evolution is studied and described in the following subsections.

Standard RANSAC

RANSAC is widely applied to assess the parameters of a model using data which are contaminated by outliers. The data that are consistent with the desired model are called inliers and the rest are outliers.

The input of the RANSAC algorithm is a set of data points. The objective is to find the optimal parameters * of a model that maximize the number of inliers. RANSAC performs two steps, generating a hypothesis of the parameters and verifying it with the data.

(i) In the hypothesis generation step, a sample $S_k$ofdata points are randomly chosen from and ahypothesis ^{k} of the parameters is computed from S_k .

(ii) In the hypothesis verification step, the support set I_k , i.e., the set of inliers consistent with ^{k} is calculated and the quality of ^{k} is specified by the cardinality of I_k, | I_k |. The two steps are repeated until the probability $\eta$of finding a better model falls under a threshold η_0. The termination criteria can be expressed as: after k samples, the probability η of all these k samples being contaminated by at least one outlier is

$$\eta = (1- \varepsilon^m)^k \leq \eta \quad (2.1)$$

where ε is the fraction of inliers in . Since ε is unknown earlier, it is updated once a new maximum is reached in the iteration. Algorithm for standard RANSAC is given in Fig. 2.1

Input: ,$\eta_0(= 0.01),\Delta(=3)$

Repeat until$\eta \leq \eta_0$

1. Hypothesis generation

a. Randomly choose a sample S_k of $m=7$ from Ω

b. Compute the fundamental matrix $F_k$using the 7-point algorithm [14].

There will be one or three real solutions.

2. Hypothesis verification

a. Calculate the support set I_k of F_k. If there are three solutions, retain the

solution with the largest support set.

b. If I_k is larger than any previous set, i.e., a new maximum is reached, store

$I_k.I^*= I_k$, and update the fraction of inliers $\epsilon= |I^*|/n$.

Output: If the non-randomness requirement [17] is satisfied, compute F^* from I^* using the normalized 8-point algorithm [14], [18] and output the cardinality of the support set of F^*. Otherwise, output zero.

Fig. 1 Algorithm for Standard RANSAC

As discussed in [37], using 7-point algorithm in the iteration has an advantage that using 7 correspondences instead of 8 can decrease the probability of selecting contaminated samples in the hypothesis generation step. Normalized 8point algorithm is applied in the final step since it performs better than the 7-point algorithm when nine or more correspondences are available, which is shown in [38]. The non-randomness requirement [39] prevents the RANSAC algorithm for choosing an incorrect model which by chance gains large support.

Experimental Results

Various experiments were carried out for a comprehensive empirical study of our object based retrieval approach. We have used the COREL photo database of ten thousand images. For getting tractable computation, we have selected different folders to have a database of almost six thousand images. The database is divided into number of concepts of different complexities for quality assessment.

The approaches, SRANSAC, LORANSAC, PRO-SAC, Fast RANSAC were evaluated. 5 kd-trees are used for nearest neighbor search. The four tables, from Table 4.2 to Table 4.5, present the average result of 5 runs, with the maximal number of samples allowed being limited respectively to be 200, 100 and 50.

Table.1 shows the performance of four approaches when 25 (Recall-25) and 50 (Recall-50) top ranked images are displayed to the user. when no constraint is imposed on the number of samples, SRANSAC performs best with an advantage in accuracy and the average number of samples generated per image pair is over 1000.

	Recall (25)	Recall (50)	RANSAC Time
SRANSAC	0.68	0.72	37.5
LORANSAC	0.67	0.72	34.5
PROSAC	0.68	0.71	34.1
Fast RANSAC	0.66	0.71	33.2

Table 1 Performance of Four Approaches with no Limitations on Number of Samples Drawn

	Recall (25)	Recall (50)	RANSAC Time
SRANSAC	0.64	0.70	5.8
LORANSAC	0.65	0.70	5.5
PROSAC	0.66	0.71	5.6
Fast RASAC	0.67	0.71	6.2

Table 2 Performance of Four Approaches with Number of Samples Allowed are 200

	Recall (25)	Recall(50)	RANSAC Time
SRANSAC	0.64	0.70	3.8
LORANSAC	0.65	0.70	3.9
PROSAC	0.66	0.71	4.2
Fast RANSAC	0.67	0.70	4.1

Table 3 Performance of Four Approaches with Number of Samples 4 Allowed are 100

	Recall (25)	Recall (50)	RANSAC Time (ms)
SRANSAC	0.62	0.68	1.7
LORANSAC	0.62	0.68	1.8
PROSAC	0.63	0.68	1.6
Fast RANSAC	0.64	0.69	1.5

Table.4 Performance of Four Approaches with Number of Samples Allowed are 50

By decreasing the number of samples allowed, a RANSAC time is observed, e.g., from 37.57ms to 1.73ms for SRANSAC. Fig. 4.11 shows the decrease in measure R(25). FAST RANSAC has the smallest decrease in accuracy due to its excellent property of early termination on relevant images. But even for FAST RANSAC, a big drop is observed when the number of samples is less than 50.

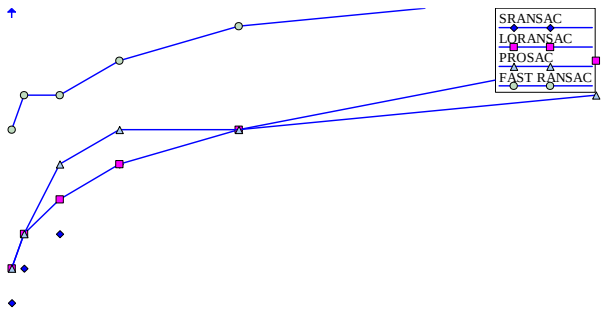

Fig. 2 Effect on RANSAC Time When Number of Samples Allowed has Increased from 10 to 500

Conclusion :

We have explored the problem of object based image retrieval using local feature correspondences. Particularly, we have studied the distinctive properties of several RANSAC variantsand utilized them to speed up the image retrieval procedure.

References:

[1] A.W. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content Based Image Retrieval at the End of the Early Years," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no.12, pp. 1349-1380, Dec.2000.

[2] R. Datta, D. Joshi, J. Li, and J.Z. Wang, "Image Retrieval: Ideas, Influences, and Trends of the New Age," ACM Comput. Surv., vol. 2, no. 2, pp. 1-60,May 2008.

[3] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," Int. J. Comput. Vision, vol. 60, no. 2, pp. 91–110, 2004

[4] Smith R. and Chang S. F. "An Image and Video Search Engine for the World-Wide Web", In Proceedings of SPIE. Vol. 3022, pp. 84–95, 2011.

[5] Ma W. and Manjuath B., "Natra: A Toolbox for Navigating Large Image Databases", In Proceedings of the IEEE International Conference on Image Processing. pp. 568–571, 2002

[6] Pentland, A. Picard, "Photobook: Tools for Content-based Manipulation of Image Databases", In Proceedings of SPIE. Vol. 2185, pp. 34–47, 2008.

[7] Sripanikulchai, K. Maggs, "Efficient Content Location Using Interest Based Locality in Peer-to-Peer Systems". In Proceedings of IEEE INFOCOM, pp. 222-228, 2003.

[8] Sia, K. C. NG, "Bridging the P2P and WWW Divide with DISCOVIR—DIStributedCOntent-based Visual Information Retrieval", In Poster Proceedings of The Twelfth International World Wide Web Conference, Poster ID S172. Hungary.

[9] Smith R. and Chang S. F. "An Image and Video Search Engine for the World-Wide Web", In Proceedings of SPIE. Vol. 3022, pp. 84–95, 2011.

[10] J. Li, D. Joshi, and J. Z. Wang, "Stochastic Modeling of Volume Images with a 3-D Hidden Markov Model," Proc. IEEE International Conference on Image Processing, pp. 188-192, 2004.

[11] J. Li and J. Z. Wang, "Automatic Linguistic Indexing of Pictures by a Statistical Modeling Approach," IEEE Trans. Pattern Analysis and Machine Intelligence, 25(9), pp. 1075–1088, 2003.

[12] J. Malik, S. Belongie, T. K. Leung, and J. Shi,

"Contour and Texture Analysis for Image Segmentation," International Journal of Computer Vision, 43(1), pp. 7-27, 2001.

[13] Y. Deng, B. S. Manjunath, C. Kenney, M. S. Moore, and H. Shin, "An Efficient Color Representation for Image Retrieval," IEEE Trans. Image Processing, 10(1):140–147, 2001.

[14] E. Hadjidemetriou, M. D. Grossberg, and S. K. Nayar, "Multiresolution Histograms and Their Use for Recognition," IEEE Trans. Pattern Analysis and Machine Intelligence, 26(7), pp.831–847, 2004

[15] J. Li, R. M. Gray, and R. A. Olshen, "Multiresolution Image Classification by Hierarchical Modeling with Two Dimensional Hidden Markov Models," IEEE Trans. Information Theory, 46(5) pp. 1826–1841, 2000.

[16] K. Mikolajczk and C. Schmid, "A Performance Evaluation of Local Descriptors," Proc. IEEE Conference on Computer Vision and Pattern Recognition, pp. 192-196, 2003.

[17] E. G. M. Petrakis, A. Diplaros, and E. Milios, "Matching and Retrieval of Distorted and Occluded Shapes Using Dynamic Programming," IEEE Trans. Pattern Analysis and Machine Intelligence, 24(4), pp. 509–522, 2002.

[18] K. Tieu and P. Viola, "Boosting Image Retrieval," International Journal of Computer Vision, 56(1/2), pp. 17-36, 2004.

[19] J. Malik, S. Belongie, T. K. Leung, and J. Shi, "Contour and Texture Analysis for Image Segmentation," International Journal of Computer Vision, 43(1), pp. 7-27, 2001.

[20] D. Comaniciu and P. Meer, "Mean Shift: A Robust Approach Toward Feature Space Analysis," IEEE Trans. Pattern Analysis and Machine Intelligence, 24(5), pp. 603–619, 2002.

[21] J. Z. Wang, J. Li, R. M. Gray and G. Wiederhold, "Unsupervised Multiresolution Segmentation for Images with Low Depth of Field," IEEE Trans. Pattern Analysis and Machine Intelligence, 23(1), pp. 85–90, 2001.

[22] Z. Tu and S.-C.Zhu, "Image Segmentation by Data-Driven Markov Chain Monte Carlo," IEEE Trans. Pattern Analysis and Machine Intelligence, 24(5), pp. 657–673, 2002.

[23] C. Carson, S. Belongie, H. Greenspan, and J. Malik, "Blobworld: Image Segmentation Using Expectation-maximization and Its Application to Image Querying," IEEE Trans. Pattern Analysis and Machine Intelligence, 24(8), pp. 1026-1038, 2002.

[24] Y. Zhang, M. Brady, and S. Smith, "Segmentation of Brain MR Images Through a Hidden Markov Random Field Model and the Expectation-Maximization Algorithm," IEEE Trans. Medical Imaging, 20(1), pp.45–57, 2001.

[25] K. Mikolajczk and C. Schmid, "A Performance Evaluation of Local Descriptors," Proc. IEEE Conference on Computer Vision and Pattern Recognition, pp. 192-196, 2003.

[26] Q. Tian, N. Sebe, M. S. Lew, E. Loupias, and T. S. Huang, "Image Retrieval Using Wavelet-Based Salient Points," Journal of Electronic Imaging, 10(4), pp. 835–849, 2001.

[27] O. Chum and J. Matas, "Geometric Hashing with Local Affine Frames," in Computer Vision and Pattern Recognition, vol. 1, 2006, pp. 879–884.

[28] R. Fergus, P. Perona, and A. Zisserman, "Object Class Recognition by Unsupervised Scale-Invariant Learning," in Computer Vision and Pattern recognition, vol. 2. IEEE Computer Society; 1999, 2003, pp. 1575–1589.

[29] V. Ferrari, T. Tuytelaars, and L. Van Gool, "Simultaneous Object Recognition and Segmentation from Single or Multiple Model Views," International Journal of Computer Vision, vol. 67, no. 2, pp. 159–188, 2006.

[30] J. Li, R. M. Gray, and R. A. Olshen, "Multiresolution Image Classification by Hierarchical Modeling with Two Dimensional Hidden Markov Models," IEEE Trans. Information Theory, 46(5) pp. 1826–1841, 2000.

[31] J. Li, D. Joshi, and J. Z. Wang, "Stochastic Modeling of Volume Images with a 3-D Hidden Markov Model," Proc. IEEE International Conference on Image Processing, pp. 188-192, 2004.

[32] J. Li and J. Z. Wang, "Automatic Linguistic Indexing of Pictures by a Statistical Modeling Approach," IEEE Trans. Pattern Analysis and Machine Intelligence, 25(9), pp. 1075–1088, 2003.

[33] J. Malik, S. Belongie, T. K. Leung, and J. Shi, "Contour and Texture Analysis for Image Segmentation," International Journal of Computer Vision, 43(1), pp. 7-27, 2001.

[34] Q. Tian, N. Sebe, M. S. Lew, E. Loupias, and T. S. Huang, "Image Retrieval Using Wavelet-Based Salient Points," Journal of Electronic Imaging, 10(4), pp. 835–849, 2001.

[35] J. Z. Wang, J. Li, R. M. Gray and G. Wiederhold, "Unsupervised Multiresolution Segmentation for Images with Low Depth of Field," IEEE Trans. Pattern Analysis and Machine Intelligence, 23(1), pp. 85–90, 2001.

[36] R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision, second ed. Cambridge Univ., pp.321-326, 2003.

[37] O. Chum, J. Matas, and J. Kittler, "Locally Optimized RANSAC," Proc. Ann. Symp. German Assoc. for Pattern Recognition (DAGM '03),pp. 236-243, 2003.

[38] R. Hartley, "Projective Reconstruction and Invariants from Multiple Images," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 16, no. 10, pp. 1036-1041, Oct. 1994.

[39] O. Chum and J. Matas, "Matching with PROSAC - Progressive Sample Consensus," IEEE Conf. Computer Vision and Pattern Recognition, vol. 1, pp. 220-226, Jun. 2005.

[40] R. Hartley, "In Defence of the 8-point Algorithm," Proc. Fifth Int'l Conf. Computer Vision (ICCV '95), pp. 1064-1070, Jun. 1995.

[41] O. Chum and J. Matas, "Randomized RANSAC with ,T_{dd} Test," Proc. British Machine Vision Conf. (BMVC '02), vol. 2, pp. 448-457, 2002.