Scientific Journal of Impact Factor (SJIF): 4.72

e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 4, Issue 6, June -2017

Content based Image Retrieval using Multiview Alignment Hashing

Bhagyodaya Jadhav
Dept. of E&TC,
Siddhant College of Engineering, Sudumbare,
Pune, India

Prof. M. U. Inamdar
Dept.of E&TC,
Siddhant College Of Engineering, Sudumbare
Pune, India

Abstract— *In this paper, content based image retrieval is implemented using multiview hashing techniques. The implementation consists of three steps preprocessing, features extraction , classification. Preprocessing step consists of color to gray conversion and histogram equalization, feature extraction consist of Histogram of Oriented gradients (HOG) for shape detection, Local Binary pattern (LBP) for texture description and color histogram for color similarity. Hashing is a simple , popular and efficient method for nearest neighbor search in large-scale data spaces by high-dimensional feature descriptors into a similarity preserving Hamming space with a low dimension. For most hashing methods,. The proposed method is systematically evaluated on three data sets: Caltech-256; CIFAR-10; and CIFAR-20, and the results show that our method outperform than individual techniques.*

Keywords—*Image Search;LBP; HOG; GIST; Hashing function; Image retrival.*

I. INTRODUCTION

Huge collections of images and video are increasingly available, arising in domains as diverse as community photo collections, scientific image data of varying modalities,news media, consumer catalogs, or surveillance archives. In the last decade in particular, user-generated content is widely stored and shared on the Web. Numbering in to the tens or hundreds of billions, their sheer size poses a challenge for conventional computer vision techniques. For example, every minute more than 20 hours of new video are uploaded to YouTube and 100,000 photos are uploaded to Facebook. Clearly, even real-time methods would be incapable of coping with this deluge of data. Consequently, researchers are exploring new representations and approaches to search large-scale datasets.At the core of visual search is the nearest neighbor problem: given a query, which items in the database are most like it? Despite the simplicity of this problem statement, fast and accurate nearest neighbor search can enable a spectrum of important applications. Efficient algorithms to address the basic similarity search task have received much attention over the years, yielding a variety of tree-based and hashing-based algorithms.

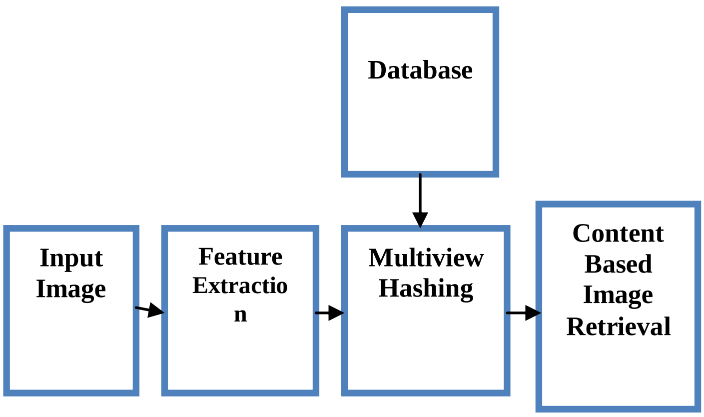

Fig. 1 Generalized Block diagram of system

The paper is organized with four sections: The first section is an introduction including previous research on CBIR. The second section describes the feature extraction used in the method. The third section offers multiview hashing classification. The final section provides experimental results in image retrieval and analysis of corresponding results.

II. FEATURE EXTRACTION

A) Histogram of Oriented Gradients

Dalal and Triggs proposed histogram of oriented gradients in the context of human detection. Their method uses a dense grid of histogram of oriented gradients, computed over blocks of various sizes. Each block consists of a number of cells. These blocks can overlap with each other. For each pixel, I (x, y), the gradient magnitude, m(x, y) and orientation, Θ(x, y) is computed. A local one-dimensional orientation histogram of gradients is formed from the gradient orientations of sample points within a region. Each histogram divides the gradient angle range into a predefined number of bins. The gradient magnitudes vote into the orientation histogram [4].

In each detection window is divided into cells of size 8×8 pixels and a group of 2×2 cells is integrated into a block as shown in fig.4.1. Block can overlap with each other. The orientation histogram of each cell contains nine bins covering an orientation range of 0 degree–180 degree (unsigned gradients–a gradient vector and its negative vote into the same bin). Each block contains a concatenated vector of all its cells. In other words, each block is represented by a 36-D feature vector (9 bins/cell × 4 cells/block). Each of the HOG descriptor blocks is then normalised based on the energy of the histogram contained within it. Normalisation introduces better invariance to illumination, shadowing and edge contrast. In order to reduce the effect of nonlinear illumination changes owing to camera saturation or environmental illumination changes that affect three-dimensional(3D) surfaces, '2-norm is applied followed by clipping (limiting the maximum values of the gradient magnitudes to 0.2) and renormalizing. The value of 0.2 is determined experimentally using images different illuminations for the same 3D objects. The final step is to combine these normalised block descriptors to form a feature vector.

B) Color Histogram features

In image processing and photography, a color histogram is a representation of the distribution of colors in an image. For digital images, a color histogram represents the number of pixels that have colors in each of a fixed list of color ranges, that span the image's color space, the set of all possible colors. The color histogram can be built for any kind of color space, although the term is more often used for three-dimensional spaces like RGB or HSV. For monochromatic images, the term intensity histogram may be used instead. For multi-spectral images, where each pixel is represented by an arbitrary number of measurements (for example, beyond the three measurements in RGB), the color histogram is N-dimensional, with N being the number of measurements taken. Each measurement has its own wavelength range of the light spectrum, some of which may be outside the visible spectrum. If the set of possible color values is sufficiently small, each of those colors may be placed on a range by itself; then the histogram is merely the count of pixels that have each possible color. Most often, the space is divided into an appropriate number of ranges, often arranged as a regular grid, each containing many similar color values. The color histogram may also be represented and displayed as a smooth function defined over the color space that approximates the pixel counts. Like other kinds of histograms, the color histogram is a statistic that can be viewed as an approximation of an underlying continuous distribution of colors values.

C) Local Binary Pattern

The concept of local binary pattern (LBP) was introduced for texture classification [31], [47], [48]. This approach has many advantages. For example, the LBP texture features have the following characteristics: 1) They are robust against illumination changes; 2) they are very fast to compute; 3) they do not require many parameters to be set; 4) they are local features; 5) they are invariant with respect to monotonic grayscale transformations and scaling; and 6) they have performed very well in many computer vision image retrieval applications. The LBP method has proved to outperform many existing methods, including the linear discriminant analysis and the principal component analysis. In order to deal with textures at different scales, the LBP operator was later extended to use neighborhoods of different sizes. Defining the local neighborhood as a set of sampling points evenly spaced on a circle centered at the pixel to be labeled allows any radius and number of sampling points. When a sampling point does not fall in the center of a pixel, bilinear interpolation was employed In the LBP method where each pixel is replaced by a binary pattern that is derived from the pixel's neighborhood. Each grayscale pixel P of an image is used as a center of a circle with radius R = 1 or 2 (radius R is usually kept very small). M represents the number of samples that determines the number of points that are taken uniformly from the contour of the circle. If needed, these points are interpolated from adjacent pixels. Each grayscale pixel P is compared with these sample points one by one. If the center point P is larger than the current neighborhood sample point I, the result is a binary zero; otherwise, the result is a binary one. When doing this operation, for example, clockwise from a certain starting point, the result will be a binary pattern with lengthM. This operation is illustrated in Fig. 6.

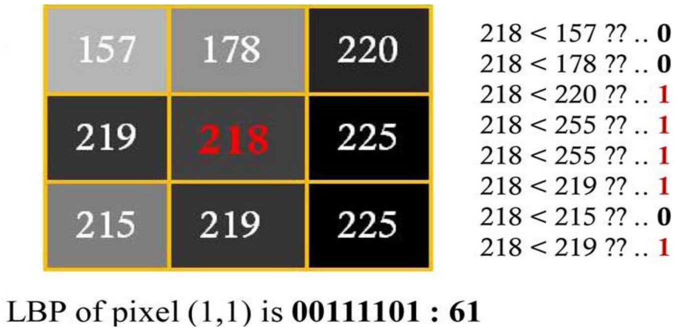

Fig.3 LBP operator example

For our database of images, an (8, 1) circular neighbourhood was used. The segmented images were extracted using k-means clustering, and then the LBP operator was applied on them before calculating the HD (see Fig. 7). Two sets of values were extracted: first, HD of the 80 images without applying LBP, and second, HD of the images after applying LBP. When comparing these two data sets, it was observed that the LBP operator enhanced the overall performance by a very high margin.

III. MULTIVIEW HASHING METHOD

In the section Multiview Alignment Hashing approach is used, referred as MAH. Our goal is to learn a hash embedding function, which fuses the various alignment representations from multiple sources while preserving the high-dimensional joint distribution and obtaining the orthogonal bases simultaneously during the RKNMF. Originally, we need to find the binary solution which, however, is first relaxed to a real-valued range so that a more suitable solution can be gained. After applying the alternate optimization, we convert the real-valued solutions into binary codes..

IV. EXPERIMENTAL RESULTS

The proposed method is systematically evaluated on three data sets: 1) Caltech-256; 2) CIFAR-10; and 3) CIFAR-20 a shown in Fig. 4. Performance of algorithm is evaluated on the basis of percentage cross validation accuracy.

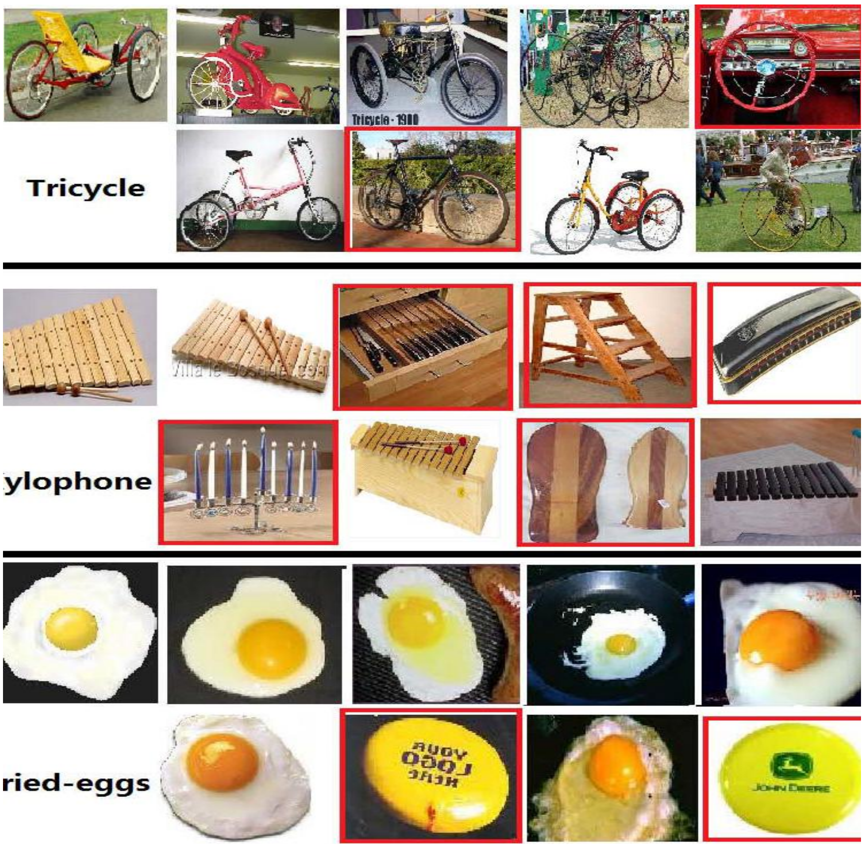

Fig. 4 Sample Caltech 256 Database images

V. CONCLUSION

In this paper, we have presented a novel unsupervised hashing method called Multiview Alignment Hashing (MAH). We incorporate multiple visual features from different views together and an alternate way is introduced to optimize the weights for different views and simultaneously produce the low-dimensional representation. We address this as a nonconvex optimization problem and its alternate procedure will finally converge at the locally optimal solution. For the out-of-sample extension, multivariable logistic regression has been successfully applied to obtain the regression matrix for fast hash encoding.

Fig.5 a) Original Image b) gray Image c) Horizontal Gradient d) Vertical gradients e) Magnitude

References

[1]. L. Shao, L. Liu, and X. Li, "Feature learning for image classification via multiobjective genetic programming," *IEEE Trans. Neural Netw. Learn.Syst.*, vol. 25, no. 7, pp. 1359–1371, Jul. 2014.

[2]. F. Zhu and L. Shao, "Weakly-supervised cross-domain dictionary learning for visual recognition," *Int. J. Comput. Vis.*, vol. 109, nos. 1–2, pp. 42–59, Aug. 2014.

[3]. J. Han *et al.*, "Representing and retrieving video shots in human-centric brain imaging space," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2723–2736, Jul. 2013.

[4]. J. Han, K. N. Ngan, M. Li, and H.-J. Zhang, "A memory learning framework for effective image retrieval," *IEEE Trans. Image Process.*, vol. 14, no. 4, pp. 511–524, Apr. 2005.

[5]. J. Han, S. He, X. Qian, D. Wang, L. Guo, and T. Liu, "An objectoriented visual saliency detection framework based on sparse coding representations," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 12, pp. 2009–2021, Dec. 2013.

[6]. V. Gaede and O. Günther, "Multidimensional access methods," *ACM Comput. Surv.*, vol. 30, no. 2, pp. 170–231, Jun. 1998.

[7]. A. Gionis, P. Indyk, and R. Motwani, “Similarity search in high dimensions via hashing,” in *Proc. Int. Conf. Very Large Data Bases*, vol. 99. 1999, pp. 518–529.
[8]. Y. Weiss, A. Torralba, and R. Fergus, “Spectral hashing,” in *Proc. Neural Inf. Process. Syst.*, 2008, pp. 1753–1760.
[9]. J.Wang, S. Kumar, and S.-F. Chang, “Semi-supervised hashing for largescale search,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 12, pp. 2393–2406, Dec. 2012.
[10]. W. Liu, J. Wang, S. Kumar, and S.-F. Chang, “Hashing with graphs,” in *Proc. Int. Conf. Mach. Learn.*, 2011, pp. 1–8.
[11]. K. Grauman and R. Fergus, “Learning binary hash codes for largescale image search,” in *Machine Learning for Computer Vision*. Berlin, Germany: Springer-Verlag, 2013, pp. 49–87.
[12]. J. Cheng, C. Leng, P. Li, M. Wang, and H. Lu, “Semi-supervised multi-graph hashing for scalable similarity search,” *Comput. Vis. Image Understand.*, vol. 124, pp. 12–21, Jul. 2014.
[13]. P. Li, M. Wang, J. Cheng, C. Xu, and H. Lu, “Spectral hashing with semantically consistent graph for image indexing,” *IEEE Trans.Multimedia*, vol. 15, no. 1, pp. 141–152, Jan. 2013.
[14]. C. Xu, D. Tao, and C. Xu, “Large-margin multi-view information bottleneck,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 8, pp. 1559–1572, Aug. 2014.
[15]. T. Xia, D. Tao, T. Mei, and Y. Zhang, “Multiview spectral embedding,” *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 40, no. 6, pp. 1438–1446, Dec. 2010.
[16]. C. Xu, D. Tao, and C. Xu. (2013). “A survey on multi-view learning.” [Online]. Available: http://arxiv.org/abs/1304.5634
[17]. B. Xie, Y. Mu, D. Tao, and K. Huang, “m-SNE: Multiview stochastic neighbor embedding,” *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 41, no. 4, pp. 1088–1096, Aug. 2011.
[18]. M. Wang, X.-S. Hua, R. Hong, J. Tang, G.-J. Qi, and Y. Song, “Unified video annotation via multigraph learning,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 5, pp. 733–746, May 2009.
[19]. M. Wang, H. Li, D. Tao, K. Lu, and X. Wu, “Multimodal graph-based reranking for web image search,” *IEEE Trans. Image Process.*, vol. 21, no. 11, pp. 4649–4661, Nov. 2012.
[20]. J. Yu, M. Wang, and D. Tao, “Semisupervised multiview distance metric learning for cartoon synthesis,” *IEEE Trans. Image Process.*, vol. 21,no. 11, pp. 4636–4648, Nov. 2012.
[21]. S. Kim and S. Choi, “Multi-view anchor graph hashing,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2013, pp. 3123–3127.
[22]. S. Kim, Y. Kang, and S. Choi, “Sequential spectral learning to hash with multiple representations,” in *Proc. 12th Eur. Conf. Comput. Vis.*, 2012,pp. 538–551.
[23]. S. Kumar and R. Udupa, “Learning hash functions for cross-view similarity search,” in *Proc. Int. Joint Conf. Artif. Intell.*, 2011, p. 1360.
[24]. D. Zhang, F. Wang, and L. Si, “Composite hashing with multiple information sources,” in *Proc. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, 2011, pp. 225–234.
[25]. Y. Kang, S. Kim, and S. Choi, “Deep learning to hash with multiple representations,” in *Proc. IEEE 12th Int. Conf. Data Mining*, Dec. 2012, pp. 930–935.
[26]. D. D. Lee and H. S. Seung, “Learning the parts of objects by nonnegative matrix factorization,” *Nature*, vol. 401, no. 6755, pp. 788–791, Oct. 1999.
[27]. S. Z. Li, X. Hou, H. Zhang, and Q. Cheng, “Learning spatially localized, parts-based representation,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Dec. 2001, pp. I-207–I-212.
[28]. P. O. Hoyer, “Non-negative matrix factorization with sparseness constraints,” *J. Mach. Learn. Res.*, vol. 5, pp. 1457–1469, Dec. 2004.
[29]. Q. Gu and J. Zhou, “Neighborhood preserving nonnegative matrix factorization,” in *Proc. Brit. Mach. Vis. Conf.*, 2009, pp. 1–10.
[30]. D. Cai, X. He, J. Han, and T. S. Huang, “Graph regularized nonnegative matrix factorization for data representation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 8, pp. 1548–1560, Aug. 2011.
[31]. D. Zhang, Z.-H. Zhou, and S. Chen, “Non-negative matrix factorization on kernels,” in *Proc. 9th Pacific Rim Int. Conf. Artif. Intell.*, 2006, pp. 404–412.
[32]. S. An, J.-M. Yun, and S. Choi, “Multiple kernel nonnegative matrix factorization,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2011, pp. 1976–1979.
[33]. L. van der Maaten and G. Hinton, “Visualizing data using t-SNE,” *J. Mach. Learn. Res.*, vol. 9, no. 11, pp. 2579–2605, Nov. 2008.
[34]. J. C. Bezdek and R. J. Hathaway, “Some notes on alternating optimization,” in *Proc. Adv. Soft Comput. (AFSS)*, 2002, pp. 288–300.
[35]. S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[36]. W. Zheng, Y. Qian, and H. Tang, “Dimensionality reduction with category information fusion and non-negative matrix factorization for text categorization,” in *Proc. 3rd Int. Conf. Artif. Intell. Comput. Intell.*, 2011, pp. 505–512.
[37]. D. D. Lee and H. S. Seung, “Algorithms for non-negative matrix factorization,” in *Proc. Neural Inf. Process. Syst.*, 2000, pp. 556–562.
[38]. S. Baluja and M. Covell, “Learning to hash: Forgiving hash functions and applications,” *Data Mining Knowl. Discovery*, vol. 17, no. 3, pp. 402–430, Dec. 2008.

[39]. C. E. Shannon, "A mathematical theory of communication," *ACM SIGMOBILE Mobile Comput. Commun. Rev.*, vol. 5, no. 1, pp. 3–55,
Jan. 2001.
[40]. D. Zhang, J. Wang, D. Cai, and J. Lu, "Self-taught hashing for fast similarity search," in *Proc. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, 2010, pp. 18–25.
[41]. Y. Lin, R. Jin, D. Cai, S. Yan, and X. Li, "Compressed hashing," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 446–451.
[42]. D. Cai, X. He, and J. Han, "Spectral regression for efficient regularized subspace learning," in *Proc. IEEE 11th Int. Conf. Comput. Vis.*, Oct. 2007, pp. 1–8.
[43]. D. W. Hosmer, Jr., and S. Lemeshow, *Applied Logistic Regression*. New York, NY, USA: Wiley, 2004.
[44]. G. Griffin, A. Holub, and P. Perona, "Caltech-256 object category dataset," California Inst. Technol., Pasadena, CA, USA, Tech. Rep. CaltechAUTHORS:CNS-TR-2007-001, 2007.
[45]. A. Krizhevsky and G. Hinton, "Learning multiple layers of features from tiny images," Dept. Comput. Sci., Univ. Toronto, Toronto, ON, Canada, Tech. Rep. TR-2009, 2009.
[46]. A. Torralba, R. Fergus, and W. T. Freeman, "80 million tiny images: A large data set for nonparametric object and scene recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 11, pp. 1958–1970, Nov. 2008.
[47]. A. Oliva and A. Torralba, "Modeling the shape of the scene: A holistic representation of the spatial envelope," *Int. J. Comput. Vis.*, vol. 42,no. 3, pp. 145–175, May 2001.
[48]. N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*,Jun. 2005, pp. 886–893.
[49]. T. Ahonen, A. Hadid, and M. Pietikäinen, "Face recognition with local binary patterns," in *Proc. 8th Eur. Conf. Comput. Vis.*, 2004, pp. 469–481.