International Journal of Intelligent Information Systems
2015; 4(2-2): 1-4
Published online February 11, 2015 (http://www.sciencepublishinggroup.com/j/ijiis)
doi: 10.11648/j.ijiis.s.2015040202.11
ISSN: 2328-7675 (Print); ISSN: 2328-7683 (Online)

Sparse spectral hashing for content-based image retrieval

Li Jun-yi, Li Jian-hua

School of Electronic Information and Electrical engineering, Shanghai JiaoTong University, Shanghai 200240, China

Email address:
leejy2006@163.com (Li Jun-yi), ljh888@sjtu.edu.cn (Li Jian-hua)

To cite this article:
Li Jun-yi, Li Jian-hua. Sparse Spectral Hashing for Content-Based Image Retrieval. *International Journal of Intelligent Information Systems.* Special Issue: Content-Based Image Retrieval and Machine Learning. Vol. 4, No. 2-2, 2015, pp. 1-4. doi: 10.11648/j.ijiis.s.2015040202.11

Abstract: In allusion to similarity calculation difficulty caused by high maintenance of image data, this paper introduces sparse principal component algorithm to figure out embedded subspace after dimensionality reduction of image visual words on the basis of traditional spectral hashing image index method so that image high-dimension index results can be explained overall. This method is called sparse spectral hashing index. The experiments demonstrate the method proposed in this paper superior to LSH, RBM and spectral hashing index methods.

Keywords: Hashing Index, Sparse Dimensionality Reduction, Laplacian Image

1. Introduction

There are often hundreds of visual features extracted from images. These high-dimension features give rise to huge difficulties for machine learning algorithms such as image similarity study and semantic analysis. To solve this problem, index technology of image high-dimension features becomes a research hotspot in recent years.

Although multi-dimension technology represented by R Tree and KD Tree have gained certain progress, the researches show that time expenditure of most multi-dimension index structures is exponential order, unsuitable for high-dimension situation (such as dozens of dimensions). Besides, the query efficiency is even lower than that of sequential scanning of original data. Meanwhile, how to guarantee data Semantic Hashing [8] (i.e. similarity calculated in index space keeps consistent with original high-dimension space) becomes a hot issue.

In this aspect, LSH(Locality Sensitive Hash) index method [5,6] is proposed. LSH maps high-dimension features into embedded subspace through a group of hash functions to reach high-dimension index purpose. In LSH, hash functions must meet the following conditions: after harsh function mapping, conflict probability of any two high-dimension data is in direct proportion to the distance of data points among original high-dimension space. Since LSH generates index coding based on probability model, it is hard to gain stable results in actual applications. In addition, with the rise in coding digits, LSH accuracy rate improves slowly. Different from random index of LSH, some index technologies based on machine learning are put forward, such as RBM (restricted Boltzmann machine RBM) [8] and stump Boosting SSC[9]. RBM utilizes two-layer unoriented graphics model to generate RBM random index and present exponentially distributed data. Researches show RBM will gain better index properties than LSH[11]. But, due to complexity of RBM, accuracy and efficiency cannot be ensured at the same time, “Boosting” is a technique to enhance generalization ability of machine learning method. It repeatedly constructs weak classifiers through giving training data different distribution weight, and then weak learning devices are combined to generate strong classifiers to gain machine learning results. Researches show Boosting-based index method is also more effective than LSH index coding, but slightly weaker than RBM[11]. But, Boosting is still faced with the problems of high complexity and low high-dimension index efficiency.

To overcome the above problems, Spectral Hashing (SH) index technology based on spectral analysis is proposed [11]. SH introduces eigenfunction for high-dimension data sample. Binary coding is directly conducted for high-dimension data dimension reduction through Principle Component Analysis (PCA). SH can not just improve index efficiency, but also can keep consistent between sample distance calculated in index space and original high-dimension space. But, SH method applies PAC to reduce dimension for original space in index coding process so that all high-dimension features (or visual words) participate in coding. In practice, generally semantics implied in an image is represented only with several distinctive features, rather than introducing other unrelated features in image expression

Based on such consideration, this paper introduces Sparse Principle Component Analysis (SPCA) in SH index coding process and puts forward corresponding global optimization solution to establish explainable binary coding for large-scale image data and fulfill image index. This paper calls such method Sparse Spectral Hashing (SSH) index.

2. SSH

2.1. Relevant Definitions and Hypotheses

A training set composed of N images $\{(x_i): \mathrm{i}=1,2,\ldots \mathrm{N})\}$ is given, where x_i means d-dimension eigenvector of the ith image, and d means the number of visual words in the training set. Θ is the index function of d-dimension vector x_i mapped to m-dimension Hamming space vector y_i from Euclidean space. Θ can be defined as follows:

$$\Theta : x_i \in R^d \rightarrow y_i \in \{-1,-1\}^m \tag{1}$$

A good index function Θ must have the following characteristics: 1) Θ is semantic hash function. In other words, if Euclidean distance between vector x_i representing the ith image and vector x_j representing the jth image is very close, corresponding result after they pass Θ index is also very close to Hamming distance; 2) the index result gained by Θ is efficient. In other words, original data of the whole image data set are mapped by Θ, relatively few coding digits are needed to express original high-dimension image data; 3) the mapping process of Θ index coding can be explained. For every image, just a few visual words used to distinguish semantics are needed for expression.

Favorable index coding should be efficient and keep similarity of the data indexed in original space [11]. In other words, the probability that the result of a bit in index coding is 1 and -1, and each bit is not correlated. SH coding defines the following objective function and constraint conditions to gain the index results:

$$\min imize : trace(Y^T LY)$$

$$\begin{aligned} & subject : \mathrm{Y(i,j)} \in \{-1,-1\} \\ & Y^T 1 = 0 \\ & Y^T Y = I \end{aligned} \tag{2}$$

Where, $L = D - W$ is Laplacian matrix; $W \in R^{N*N}$ is similarity matrix, $W(\mathrm{i,j}) = \exp(-\| x_i - x_j \| / \varepsilon^2)$; D is diagonal matrix, with diagonal element of $D(i,i) = \sum_j^N W(i,j)$. $Y(i,j) \in \{-1,1\}$ makes sure index coding is binary coding; $Y^T 1 = 0$ makes sure the probability that index coding is 1 and -1 is 50%, while $Y^T Y = I$ make sure every bit of index coding is not correlated.

Solving Equation (2) is a NP problem. SH relaxing index coding result is binary condition so that Equation (2) is solvable. That is, SH converts solving Equation (2) to solving the minimum eigenvalue of Laplacian matrix L. After solving Equation (2) is converted to dimensionality reduction problem of Laplacian eigenmap, PCA is directly introduced in SH to carry out dimensionality reduction for original data.

However, in PCA dimensionality reduction process, every dimension of original data participates in dimensionality reduction in the form of linear combination. It is hard to gain physical interpretation of this process. For given image training set, over-completed visual words can be usually gained. An image can be fully expressed only with several visual words, i.e. an image is usually related to a limited number of visual words. For example, visual words related to colors may be more suitable for expressing rainbow, while visual words related to shapes are more suitable for expressing automobile.

This paper uses SPCA[12] in SH index to replace PCA, transforms traditional PCA to non-convex regression form to gain SPC so that index coding is more interpretable. This algorithm in this paper is called SSH index. Since SPCA is a non-convex algorithm, convex optimization algorithm is thus adopted to gain globally optimal solution of SPC.[4]

Assuming SPC p of Laplacian matrix L is a d-dimension vector, the following optimization problem can be gained through giving a constraint to cardinality of p and removing unrelated limiting conditions [12]:

$$\begin{aligned} & \min imize : p^T Lp + \rho Card^2(p) \\ & subjectto : p(i) \in \{-1,1\} \\ & p^T 1 = 0 \end{aligned} \tag{3}$$

Where, $Card(p)$ means cardinality of p; parameter ρcontrols sparse degree. Solving Equation (3) is still a NP problem. However, we can find out corresponding positive semidefinite convex optimization problem [4]:

$$\begin{aligned} & \min imize : trace(LP) + \rho 1^T |P| 1) \\ & subjectto : p(i) \in \{-1,1\} \\ & p^T 1 = 0 \end{aligned} \tag{4}$$

Where, $P = pp^T$, every element of $|P|$ is the absolute value of corresponding elements in matrix P. Equation (4) can be solved through recursion [4].

The above paper gives SSH solving process. The vector after Euclidean space dimensionality reduction can be transformed to vector of Hamming space through directly taking threshold value. But, a problem is still not solved, i.e. how does the images outside training set index and code? In recent years, there have been some methods to solve this problem [2]. Main thought is to transform the eigenvector to eigenequation. Through assuming every original eigenvector belongs to a manifold subspace and obeys multi-dimension even distribution, this problem can be solved through eigenequation of weighted Laplace-Beltrami operators [11].

2.2. Binary Index Coding of SSH

For given training set including N images $X \in R^{Nxd}$, mapping function Θ maps d-dimension X of Euclidean space to m-dimension Y of Hamming space. The process of solvingΘis divided into two steps:

1) Solve m sparse principle vectors through Formula (4), and map $X \in R^{Nxd}$ to $B \in R^{Nxm}$.

Calculate covariance matrixΣ of X. its SPC p can be solved through convex optimization stated previously. Update Σaccording to Formula (5).

$$\Sigma = \Sigma - (p^T \Sigma p) pp^T \quad (5)$$

Repeat this process for m times and gain m SPC $\{p_1, \cdots, p_m\}$. These principal component vectors serve as column vectors of the matrix and gain matrix M. eigenmatrix B after dimensionality reduction of $N \times m$ is thus gained through $B = X \times M$.

2) Map Euclidean space matrix B to Hamming space matrix Y.

Define the jth vector of matrix B as $B_{(:,j)}$, and δ_j^k can be defined as follows:

$$\delta_j^k = 1 - e^{-\frac{\varepsilon^2}{2}\left|\frac{k\pi}{B_{(:,j)}^{\max} - B_{(:,j)}^{\min}}\right|^2} \quad (6)$$

Where, $k = 1, \ldots N$; $B_{(:,j)}^{\max}$ and $B_{(:,j)}^{\min}$ refer to the maximum value and the minimum value of $B_{(:,j)}$;ε is a constant. For each column vector $B_{(:,j)}$, N δ_j^k can be solved. Thus, $N \times m$ $\delta_j^k (k = 1, \ldots N; j = 1, \ldots m)$ can be gained. Sort δ_j^k, take the first m δ_j^k and express them as $\{\delta_1^{\min}, \ldots, \delta_m^{\min}\}$.

Assuming binary coding corresponding to x_i is $y_i \in \{-1, 1\}^m$, the jth mapping value $y(i, j)$ can be solved according to the following function:

$$y(i, j) = \Theta(\delta_j^{\min}, B(i, t)) - \sin\left(\frac{\pi}{2} + \frac{k\pi}{B_{(:,t)}^{\max} - B_{(:,t)}^{\min}} B(i, t)\right) \quad (7)$$

Where, $\delta_j^{\min}$ is the jth minimum value of $\{\delta_1^{\min}, \ldots, \delta_m^{\min}\}$, which is solved through the t column of k and B; $B_{(:,j)}^{\max}$ and $B_{(:,j)}^{\min}$ refer to the maximum value and the minimum value of $B_{(:,t)}$, $i = (i, \ldots N), j = (1, \ldots m)$. Transform it to binary coding through regarding 0 as the threshold value.

3. Experiment

Before you begin to format your paper, first write and save the content as a separate text file. Keep your text and graphic files separate until after the text has been formatted and styled. Do not use hard tabs, and limit use of hard returns to only one return at the end of a paragraph. Do not add any kind of pagination anywhere in the paper. Do not number text heads-the template will do that for you.

Finally, complete content and organizational editing before formatting. Please take note of the following items when proofreading spelling and grammar:

3.1. Experimental Data set and Feature Expression

This paper compares properties of SSH index algorithm on two image data sets (Oxford5k and MCG-WEBV) as well as E2LSH, RBM and SH.)

Oxford5k: including 5062 11 landmark images of University of Oxford. This data set provides the standard answer of artificial labeling. In this experiment, after SIFT local features are extracted from each image, K-means clustering algorithm is used to gain 300 visual words to express original image data.

MCG-WEBV: this data set contains 80031 videos of YouTube website with the highest click rate from December 2008 to February 2009. This data set provides 828-dimension vectors extracted from key frames of videos. 3814 images are drawn randomly in this experiment.

This paper takes 1.5% of original mean Euclidean distance as neighbor threshold value which serves as the standard [11]. F1 and AUC serve as measurement standards.

3.2. Experimental Results

Table 1 and Table 2 show index results of two data sets. m means digits of index coding. The boldface means the best result under each index coding digit. It can be seen that as a random mapping index algorithm, index property of E2LSH changes little as the rise in the number of index digits. SSH obtains the best results on F1 and AUC measurement standards.

Table 1. *Experimental results of Oxford5k data set*

m	F1				AUC			
	SSH	*SH*	*E2LSH*	*RBM*	*SSH*	*SH*	*E2LSH*	*RBM*
2	0.2135	0.2134	0.1055	0.2088	0.5088	0.5085	0.5085	0.5085
4	0.2135	0.2135	0.1055	0.1791	0.5088	0.5088	0.5085	0.5043
8	0.2136	0.2136	0.1055	0.1595	0.5097	0.5097	0.5085	0.4933
16	0.2493	0.2286	0.1055	0.1649	0.5980	0.5520	0.5085	0.4991
32	0.3579	0.3273	0.1046	0.1054	0.7246	0.6928	0.5085	0.4812

Table 2. *Experimental results of MCG-WEBV data set*

M	F1				AUC			
	SSH	**SH**	**E2LSH**	**RBM**	**SSH**	**SH**	**E2LSH**	**RBM**
2	0.4003	0.3826	0.0664	0.3255	0.6312	0.6117	0.6073	0.5971
4	0.4960	0.4688	0.0664	0.3227	0.7299	0.7197	0.6073	0.5482
8	0.5652	0.4481	0.0664	0.3161	0.7671	0.6998	0.6073	0.5475
16	0.5489	0.0611	0.0664	0.3112	0.7432	0.6075	0.6073	0.5486
32	0.3706	0.0027	0.0664	0.3040	0.6792	0.5975	0.6073	0.5503

4. Conclusions

This paper introduces SPCA in traditional SH and designs global optimal solution so that high-dimension image index coding become more effective and interpretable. Besides, this paper also discusses image index coding mode outside the training set. Experimental results show SSH is superior to other similar algorithms.

References

[1] M. Belkin and P. Niyogi. Towards a theoretical foundation for Laplacian-based manifold methods. Journal of Computer and System Sciences, 74(8):1289 - 1308, 2008.

[2] Y. Bengio, J. Paiement, P. Vincent, O. Delalleau, N. Le Roux, and M. Ouimet. Out-of-sample extensions for Ile, isomap, mds, eigenmaps, and spectral clustering. In Proceedings of Advances in Neural Information Processing Systems (NIPS), page 177, 2004.

[3] S. Berchtold, C. Bohm, H. Jagadish, H. Kriegel, and J. Sander. Independent quantization: an index compression technique forhigh-dimensional data spaces. In Proceedings of International Conference on Data Engineering (ICDE), pages 577 - 588, 2000.

[4] A. D, a' rAspremont, L. E. Ghaoui, M. I. Jordan, and G. G. Lanckriet. A direct formulation for sparse PCA using semidefinite programming. Proceedings of Advances in Neural Information Processing Systems (NIPS), 2004.

[5] M. Datar, N. Immorlica, P. Indyk, and V. Mirrokni. Locality-sensitive hashing scheme based on p-stable distributions. In Proceedings of Annual Symposium on Computational Geometry, pages 253 - 262. ACM, 2004.

[6] Q. Lv, W. Josephson, Z. Wang, M. Charikar, and K. Li. Multi-probe LSH: efficient indexing for high-dimensional similarity search. In Proceedings of International Conference on Very Large Data Bases (VLDB), pages 950 - 961, 2007.

[7] R. Salakhutdinov and G. Hinton. Learning a nonlinear embedding by preserving class neighbourhood structure. In AI and Statistics, 2007.

[8] R. Salakhutdinov and G. Hinton. Semantic hashing. International Journal of Approximate Reasoning, 50(7):969 - 978, 2009.

[9] G. Shakhnarovich, P. Viola, and T. Darrell. Fast pose estimation with parameter-sensitive hashing. In Proceedings of IEEE International Conference on Computer Vision (ICCV), page 750, 2003.

[10] Y. Tao, K. Yi, C. Sheng, and P. Kalnis. Quality and efficiency in high dimensional nearest neighbor search. In Proceedings of International Conference on Management of Data (SIGMOD), pages 563 - 576. ACM, 2009.

[11] Y. Weiss, A. Torralba, and R. Fergus. Spectral hashing. In Proceedings of Advances in Neural Information Processing Systems (NIPS), pages 1753 - 1760, 2009.

[12] H. Zou, T. Hastie, and R. Tibshirani. Sparse principal component analysis. Journal of Computational and Graphical Statistics, 15(2):265 - 286, 2006.