IMAGE FEATURE EXTRACTION BASED ON SPECTRAL GRAPH INFORMATION

Jieqi Kang, Shan Lu, Weibo Gong, Patrick A. Kelly

University of Massachusetts, Amherst
Electrical and Computer Engineering Department
151 Holdsworth Way, Amherst, MA, 01003

ABSTRACT

Graphs have been widely used in image processing and understanding tasks. We introduce a novel graph generation model which greatly reduces the size of the traditional pixel-based graph. Based on the generated graph, we propose two feature extraction methods which utilize spectral graph information, and apply the features to image. Experiments show that our proposed oscillatory image heat content and weighted heat content spectrum features are more robust to small distortions and changes of viewpoint than the image heat content feature we proposed previously. The features are also capable of capturing important image structural information of the image and perform well alone or in combination with other low-level image features.

***Index Terms*—** Feature extraction, spectral graph information, image classification.

1. INTRODUCTION

Image retrieval and classification is a challenging problem involving tasks that include image pre-processing, feature design and learning algorithms. Among the vast literature exploring these topics, a recent trend is the use of graphs in image representation and understanding tasks. These models have been shown to be effective in a variety of image related applications.

Using graphs for image understanding is natural. Human intelligence and the human vision system are built upon the neurons with their connections in the visual cortex, which forms a complex network[1, 2]. In [3, 4] the author points out that Monte-Carlo simulations of diffusion can be helpful in testing the similarity of complex graphs. Meanwhile, graph based image representation allows the possibility to apply some well-studied techniques in graph theory such as the graph-cuts theory, similarity testing[5, 6] and the spectral graph theory[7]. Successful applications include graph based image segmentation[8, 9, 10, 11], image retrieval, clustering and classification based on graph statistics[12, 13, 14].

In our recent paper[15], based on the mathematical theory of diffusion over manifolds and the asymptotic behavior of the heat content[16, 17, 18], we proposed a new low level heat content image feature for image retrieval. Our preliminary result showed that the heat content feature could improve the image retrieval performance when combined with existing low level features. However, for a mid-size image with more than $10,000$ pixels, the image heat content feature can only be roughly estimated. Moreover, the functions that are summed to generate the feature decay exponentially with rates given by the graph Laplacian eigenvalues; hence, information carried by larger eigenvalues is mostly lost.

In this paper, we first introduce an approach to drastically reduce the size of the traditional pixel-based graph representation. By merging nodes, our model can generate a much smaller graph for the same image, which makes complicated feature extraction methods such as spectral analysis a possible option. We also propose a novel feature extraction model based on the spectral graph information. Experiments show that our re-designed features perform better than the original heat content feature and is a more effective supplement to further improve the performance of traditional feature-based classification.

The rest of the paper is organized as follows. In Section 2, we introduce the graph generation model. In Section 3, we propose two feature extraction models. Section 4 shows the result of three experiments using the new models, and Section 5 is the conclusion.

2. GRAPH GENERATION OF GRAYSCALE AND COLOR IMAGES

Traditional pixel-based graph generation precisely represent the pixel relations in the image. However, it is a challenging task to apply powerful tools such as spectral analysis when the graph is very large. Even when the graph is made more sparse, the size is still too large for a mid-size image input. In this section, we propose a novel graph generation model which has a far smaller number of nodes compared to traditional methods.

2.1. Small-size graph generation by nodes merging

Our new small-size graph generation model is derived from traditional pixel-based models. We reduce the size of the graph by "nodes merging". Specifically, if some pixels form a homogenous area, they can be clustered into an image segment. In our graph generation model, we assign a single node for such an image segment. By doing so, the total number of nodes in the graph is greatly decreased.

The first step of our model is similar to our original pixel-based graph model[15]. We generate a vertex for every pixel in the image and the weight of any edge is determined by a distance measurement between the corresponding pixels. For color images, the original RGB representation is transformed to the CIELAB format with every pixel represented by a five-dimensional vector $[L, a, b, x, y]$. L is an intensity measurement, a, b describes the visual color and x, y represents the location of the pixel on the image. We compute two distances to form the definition of the edge weight. For any two pixels p_i, p_j, the color distance is defined as $d_c = \sqrt{(L_i - L_j)^2 + (a_i - a_j)^2 + (b_i - b_j)^2}$ (for grayscale image we just use the intensity part) and the geometrical distance is defined as $d_s = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$. The edge weight between any two pixels p_i and p_j is defined as

$$w_{ij} = e^{-d_c^2/\sigma_I} \cdot e^{-ds^2/\sigma_X}, \tag{1}$$

where σ_I and σ_X are controlling parameters. We can also use our original graph generation model [15] in this step to produce pixel-based graph for the image.

The pixel-based graph describes similarities between pixels and captures important structural information on the image. Our goal is to merge similar nodes together while keeping the major structure of the graph intact. A MinMax K-Means clustering algorithm[19] is applied to acquire a fast segmentation for the image.

Suppose we generate k different-size segments S_1, S_2, …, S_k after clustering. We merge all the nodes in each segment together and form k new nodes s_1 to s_k to be the vertices in our new graph. All the edges connected to any node in set S_i from outside in the original graph will then become links connected to the single node s_i in the new graph. At the same time, all the inside edges in each segment are summed up to be a self-loop edge for the corresponding new node s_i. Precisely, the weight between nodes s_m and s_n in the adjacency matrix of our "after-merging" graph is defined as

$$A(m, n) = \sum_{p_i \in cc_m} \sum_{p_j \in cc_n} w_{ij}. \tag{2}$$

We can easily prove that the total degree of the graph does not change. The structure of the graph also remains intact. Considering a random walk on our new graph, the steady-state distribution of the random walkers can be easily calculated from the original distribution by adding entries from the same cluster together. Although there is some information loss inside every image segment, the general heat diffusion pattern is largely equivalent to the original one due to the fact that these image segments are mostly homogenous patches.

A natural image usually has large image segments. Depending on the parameters set for the K-Means clustering, for a standard configuration, the size of the graph normally reduces to less than one percent of the original size, yet most of the structural information of the image still remains. For example, the generated graph only contains hundreds of nodes for a typical 128×128 as shown input in figure 4. The model not only generate a much smaller graph which makes eigen-decomposition no longer a big challenge, but also naturally captures important image information through node merging of homogenous image segments.

3. IMAGE FEATURE BASED ON GRAPH SPECTRAL INFORMATION

Spectral analysis is a powerful tool to extract useful structural information on the graph. For a large-size graph, fast eigen-decomposition becomes very difficult. The image heat content feature [15] is designed to contain all the spectral information which can also be estimated by a fast Monte-Carlo algorithm. However, by drastically reducing the size of the graph, we can directly acquire the precise result of the eigen-decomposition, which gives us more possibilities for designing features based on the spectral information.

3.1. The image heat content feature

We briefly introduce the original heat content feature first to lay the foundation of our new approaches. For a symmetric image graph representation G, let $D = diag[d_u]$ be the diagonal degree matrix. The normalized graph Laplacian [7] of the graph G is defined as $\mathcal{L} = D^{-1/2} L D^{-1/2}$, where $L = D - A$. Suppose that vertex set V is divided into $V = iD \cup \partial D$, where iD is the interior nodes and ∂D is the boundary. Then the following heat equation describes the heat-flow dynamics on the graph:

$$\begin{cases} \frac{\partial h_t}{\partial t} = -\mathcal{L} h_t \\ h_t(u, v) = 0 \text{ for } u \in \partial D, \end{cases} \tag{3}$$

with the initial condition $h_0(u, u) = 1$ if $u \in iD$. Suppose $\Lambda = diag[\lambda_i]$ is the diagonal eigenvalue matrix and Φ is the eigenvector matrix of the interior part of $\mathcal{L}$. The solution to the heat equation is $H_t = e^{-\mathcal{L}t} = \Phi e^{-\Lambda t} \Phi^T$ and for each entry of H_t, we have $H_t(u, v) = \sum_{i=1}^{|iD|} e^{-\lambda_i t} \phi_i(u) \phi_i(v)$, where ϕ_i is the ith column vector in Φ. The heat content $Q(t)$ is defined as

$$Q(t) = \sum H_t(u, v) = \sum_{uv} \sum_{i=1}^{|iD|} e^{-\lambda_i t} \phi_i(u) \phi_i(v). \tag{4}$$

The heat content feature can be improved in different aspects. We will first propose a modified oscillatory heat content feature based on an asymmetric graph in the following subsection. Second, instead of using a summation of exponentially decaying functions, which is largely determined by the component functions with small eigenvalues, we propose a novel feature extraction model directly based on the spectral information of the graph.

3.2. Oscillatory image heat content

We first modify the symmetric graph into an asymmetric one by simply giving very small weights to edges from low average-intensity segments to high average-intensity segments. While this may seem arbitrary, the directional asymmetry in the resulting graph actually carries more information about the image structure.

The eigenvalues and eigenvectors of the normalized graph Laplacian $\mathcal{L}$ of the asymmetric graph can be complex valued, and exist as complex conjugates. The solution to the heat equation of this asymmetric graph Laplacian $\mathcal{L}$ is $H_t = \Phi e^{-\Lambda t}\Psi$. $\Lambda = diag[\lambda_i]$ is the diagonal eigenvalue matrix. Φ and Ψ are the right and left eigenvector matrices ($\Psi = \Phi^{-1}$). For any complex conjugates pairs of eigenvalues $\lambda = a + bi$ and $\overline{\lambda} = a - bi$ with corresponding eigenvectors ϕ, ψ, $\overline{\phi}$ and $\overline{\psi}$, suppose that $\sum_{uv}\phi(u)\psi(v) = \alpha + \beta i$, then $\alpha - \beta i = \overline{\sum_{uv}\phi(u)\psi(v)} = \overline{(\sum\phi)(\sum\psi)} = (\sum\overline{\phi})(\sum\overline{\psi}) = \sum_{uv}\overline{\phi}(u)\overline{\psi}(v)$. The summation $e^{-\lambda t}\sum_{uv}\phi(u)\psi(v) + e^{-\overline{\lambda}t}\sum_{uv}\overline{\phi}(u)\overline{\psi}(v)$ is equal to $e^{-(a+bi)t}(\alpha+\beta i) + e^{-(a-bi)t}(\alpha-\beta i) = 2e^{-at}(\alpha\cos(bt)) + \beta\sin(bt))$, which is a real-valued function. Therefore, the total summation is still real valued and can be written as

$$Q(t) = \sum_{i=1}^{|iD|} e^{-a_i t}\left(\sqrt{\alpha_i^2+\beta_i^2}\sin(b_i t + \arctan\frac{\beta_i}{\alpha_i})\right). \quad (5)$$

Compared to the original heat content, the new oscillatory heat content (OHC) is no longer a summation of purely exponentially decaying functions. It becomes an oscillatory function which contains components of different frequencies, amplitudes and phases. These variations can be helpful in classification tasks. We can also amplify the asymmetric part of $\mathcal{L}$ to generate more pairs of complex eigens by defining a new matrix $\mathcal{L}_k = (\mathcal{L}+\mathcal{L})/2 + k(\mathcal{L}-\mathcal{L})/2$ where $k > 1$.

3.3. Weighted heat content spectrum

Starting from the original heat content in equation 4, we define the heat content spectrum as $[\alpha_1, \ldots, \alpha_{|iD|}]'$, in which

$$\alpha_i = \sum_{uv}\phi_i(u)\phi_i(v) = \left[\sum_u \phi_i(u)\right]^2. \quad (6)$$

The original heat content can then be expressed as $Q(t) = \sum_{i=1}^{|iD|}\alpha_i e^{-\lambda_i t}$. The heat content spectrum has some interesting properties. First, the summation of all the αs equals the number of interior nodes. Because $\lambda_1 \approx 0$, we have the first eigenvector $\phi_1 \approx \left[\sqrt{\frac{d_i}{m|iD|}}, \ldots, \sqrt{\frac{d_n}{m|iD|}}\right]^T$ (m is the mean degree of the graph) and $\alpha_1 \approx (\sum_i\sqrt{d_i})^2/(m|iD|)$, which is directly related to the image intensity distribution. α_1 reaches its maximum $\max(\alpha_1) \approx |iD|$ only when all the vertices have the same degree in the graph. For the rest of the spectrum, the general shape of the weight spectrum captures some major structural information of the graph.

Our new feature utilizes the information of both the eigenvalue distribution and the heat content spectrum. Every eigenvalue λ_i of the normalized graph Laplacian satisfies $\lambda_i \in [0, 2]$. We first generate k overlapping intervals I_1 to I_k on $[0, 2]$. The middle point of each interval I_i is located on $(2i-1)/k$ and the length of the interval is $2/k + \varepsilon(1 - |(k - 2i + 1)/k|)$. The weighted heat content spectrum (WHCS) $f = [f_1, f_2, \ldots, f_k]$ is then defined as $f_x = \sum_{i:\lambda_i \in I_x}\alpha_i$.

4. EXPERIMENTAL RESULTS

We first illustrate the proposed new features (OHC and WHCS) by two experiments that show the ability of the OHC feature to differentiate images in different classes and the robustness of the WHCS feature to view point changes. We then show the improvement of our new features compared to the heat content feature combined with traditional low-level features in a handwritten digits recognition experiment.

4.1. Oscillatory heat content illustration

We select thirty-five images from seven categories in the COREL dataset [20]. The images are shown in figure 1. We calculate both the heat content based on the original graph in [15] and the oscillatory heat content. Figures 2 and 3 shows the results. The same color curves represent images in the same category. We can see that although the original heat content can somewhat differentiate these images, the oscillatory heat content performs much better in either clustering or differentiating. The richer frequency behavior of the new feature better represent more diverse images and textures, which will eventually be helpful in a large scale image retrieval task.

4.2. Weighted heat content spectrum illustration

We test the weighted heat content spectrum with the COIL100 [21] dataset. K is set to be 6 in the K-Means clustering algorithm. $\sigma_I = 0.01$ and $\sigma_X = 4$ in the graph generation procedure. Figure 4 shows some examples of the images and their corresponding WHCS feature vectors. We can see that the WHCS is robust to small viewpoint changes yet still captures the differences between images. The feature is also effective at differentiating between different objects.

COIL100 contains 100 different objects in 72 different view angles. We select 8 different views for each object to form the reference set and the rest to be the testing set. Table 1 shows the resulting classification rate using a 1-nearest-

Fig. 1. Seven categories of similar images from COREL dataset

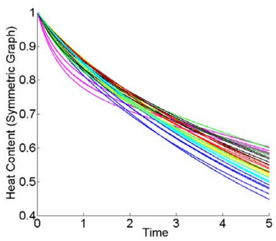

Fig. 2. Heat Content

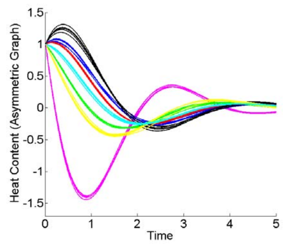

Fig. 3. Oscillatory HC

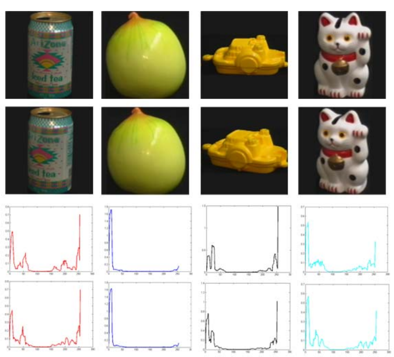

Fig. 4. Weighted heat content spectrum

neighbor classifier. The new WHCS is much better than the original HC because it is directly generated by the full-size image. When the feature is used with the multi-scale color histogram, the classification rate is only a little less than the state of the art hand-designed method.

Feature	Classification rate
HC + Color Hist	85.3%
WHCS + Color Hist	92.1%
Typical state of the art[22]	95%

Table 1. Classification rates of COIL100

4.3. MNIST image classification experiment

The MNIST database [23] is a widely used hand written digits benchmark containing a training set of $60,000$ images and a testing set of $10,000$ images. We simulate the image classification task using combinations of different basic features including intensity histogram (Hist), intensity moments (Mo), Gabor coefficients (Gabor), gray-level co-occurrence matrix (GLCM), edge directions histogram (Edge), heat content feature (HC) and the two proposed features oscillatory heat content (OHC) and weighted heat content spectrum (WHCS). The length of the HC, OHC and WHCS is set to be the 10. The combined feature contains all traditional features. All the features are computed for forty-nine 10×10 average-positioned blocks. Logistic and linear kernel supported vector machine (SVM) [24] classifiers are used in the experiment.

Logistic Classifier	Self	w/ HC	w/ OHC	w/ WHCS
Hist + Gabor + Edge	2.45%	2.34%	**2.18%**	2.20%
Hist + GLCM + Edge	2.77%	2.47%	2.41%	**2.39%**
Mo + Gabor + Edge	2.41%	2.37%	2.28%	**2.12%**
Mo + GLCM + Edge	2.54%	2.48%	**2.23%**	2.28%
Combined feature	2.29%	2.18%	2.08%	**1.98%**
SVM Classifier	Self	w/ HC	w/ OHC	w/ WHCS
Hist + Gabor + Edge	1.47%	1.45%	1.42%	**1.28%**
Hist + GLCM + Edge	1.69%	1.62%	**1.50%**	1.52%
Mo + Gabor + Edge	1.54%	1.48%	1.42%	**1.37%**
Mo + GLCM + Edge	1.60%	1.57%	1.45%	**1.42%**
Combined feature	1.39%	1.36%	1.31%	**1.23%**

Table 2. Classification error rate

Table 2 shows that the performance of the combined features with the OHC and the WHCS are better than the ones with the heat content feature in all situations. In most cases feature sets that include WHCS perform slightly better than those using OHC. This result further illustrates that the proposed two features based on graph spectral information contain unique and useful image information, and could be useful supplement to traditional feature extraction for image classification tasks.

5. CONCLUSION

In this paper, we propose a general approach to reduce the size of pixel-based graph generation and to use graph spectral information for image feature extraction. The proposed image features improve on the original heat content feature and demonstrate an ability to serve as a useful supplement to traditional image feature extraction. Although there are still open questions on the theoretical side, in particular on details of the relationship between the heat content spectrum and graph structure, the experiment results support that the proposed method effectively captures useful image information.

6. ACKNOWLEDGEMENTS

This work is supported in part by the United States National Science Foundation grant CNS-1239102 and Army Research Office Contract W911NF-12-1-0385.

7. REFERENCES

[1] E. Bullmore and O. Sporns, "Complex brain network: graph theoretical analysis of structural and functional systems," *Nature Reviews Neuroscience*, vol. 10, pp. 186–198, 2009.

[2] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: Uses and interpretations," *NeuroImage*, vol. 52, pp. 1059–1069, 2010.

[3] W. Gong, "Can one hear the shape of a concept?," in *Proceedings of the 31st Chinese Control Conference (Plenary Lecture)*, Hefei, China, July 2012, pp. 22–26.

[4] W. Gong, "Transient response functions for graph structure addressable memory," in *Proceedings of the 52th IEEE Conference on Decision and Control*, Florence, Italy, December 2013, pp. 6978 – 6985.

[5] D. Koutra, A. Parikh, A. Ramdas, and J. Xiang, "Algorithms for graph similarity and subgraph matching," 2011, Available at `https://www.cs.cmu.edu/~jingx/docs/DBreport.pdf`.

[6] P. Papadimitriou, A. Dasdan, and H. Carcia-Molina, "Web graph similarity for anomaly detection," *Journal of International Service and Applications*, vol. 1(1), pp. 19–30, 2010.

[7] F. Chung, *Spectral Graph Theory*, American Mathematical Society, 1997.

[8] Jianbo Shi and Jitendra Malik, "Normalized cuts and image segmentation," *IEEE Transaction on Pattern Anylysis and Machine Intelligence*, vol. 22, pp. 888–905, 2000.

[9] PedroF. Felzenszwalb and DanielP. Huttenlocher, "Efficient graph-based image segmentation," *International Journal of Computer Vision*, vol. 59, no. 2, pp. 167–181, 2004.

[10] L. Grady, "Random walks for image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28(11), pp. 1768–1783, 2006.

[11] Sara Vicente, Vladimir Kolmogorov, and Carsten Rother, "Graph cut based image segmentation with connectivity priors.," in *IEEE Conference on CVPR*, 2008, pp. 1–8.

[12] I. K. Park, II D. Yun, and S. U. Lee, "Color image retrieval using hybrid graph representation," *Image and Vision Computing*, vol. 17, pp. 465–474, May 1999.

[13] A.R. Backes, D. Casanova, and O.M. Bruno, "A complex network-based approach for boundary shape analysis," *Pattern Recognition*, vol. 42, pp. 54–67, 2009.

[14] Z. Harchaoui and F. Bach, "Image classification with segmentation graph kernels," in *IEEE Conference on Computer Vision and Pattern Recognition*, Minneapolis, USA, June 2007, pp. 1–8.

[15] Jieqi Kang, Shan Lu, Weibo Gong, and Patrick A. Kelly, "A complex network based feature extraction for image retrieval.," in *IEEE International conference on image processing*, 2014.

[16] P. Mcdonald and R. Meyers, "Diffusion on graphs, poisson problems and spectral geometry," *Transaction of the American Mathematical Society*, vol. 354(12), pp. 5111–5136, 2002.

[17] P. Mcdonald and R. Meyers, "Isospectral polygons, planar graphs and heat content," *Proceedings of the American Mathematical Society*, vol. 131(11), pp. 3589–3599, 2003.

[18] S. Butler, "Interlacing for weighted graphs using the normalized laplacian," *Electronic Journal of Linear Algebra*, vol. 16, pp. 90–98, 2007.

[19] Grigorios Tzortzis and Aristidis Likas, "The minmax k-means clustering algorithm," *Pattern Recognition*, vol. 47, pp. 2505–2516, 2014.

[20] James Z. Wang, Jia Li, and Gio Wiederhold, "Simplicity: Semantics-sensitive integrated matching for picture libraries," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23(9), pp. 947–963, 2001.

[21] S. A. Nene, S. K. Nayar, and H. Murase, "Columbia object image library (coil-100)," *Technical Report CUCS-006-96*, 1996.

[22] Allen Y. Yang, Subhransu Maji, C. Mario Christoudias, Trevor Darrell, Jitendra Malik, and S. Shankar Sastry, "Multiple-view object recognition in band-limited distributed camera networks," in *Third ACM/IEEE International Conference on Distributed Smart Cameras*, Como, Italy, 2009, pp. 1–8.

[23] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86(11), pp. 2278–2324, 1998.

[24] C. C. Chang and C. J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011, Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`.