International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

Overview of State-of-the-Art Image Set Compression Techniques

Z. Bohari[1], M. S. Raeen[2]

[1]Depatment of Electronics and Communication Engineering, All Saints' College of Technology, Bhopal, India

[2]Head of Department, Department of Electronics and Communication Engineering, All Saints' College of Technology, Bhopal, India

Abstract: ***Image set compression refers to compression of image sets like photo album, medical image sets, and satellite image sets. The use of digital images has increased drastically in the last decade. Hence, the field of image set compression has become far more important. Several methods for compressing image sets have been proposed in the literature. These range from simple approaches which only use minimum and maximum value of pixels to complex approaches which use graph theory and feature extraction algorithms. In this paper, we provide an overview of various image set compression techniques including state-of-the-art methods.***

Keywords: Image Set Compression, Graph Theory, Representative Signals, Similar Images, SIFT, HEVC

1. Introduction

In the last decade, number of images stored on personal computers, cloud, and web servers has increased exponentially. Due to this increase in the use and storage of images, a lot of research is being done towards efficient methods for storing sets of images. Traditional method of compressing a set of images uses image compression techniques like JPEG, JPEG2000, and PNG. While these image compression techniques are very effective in compressing individual images, they can only exploit redundancies within an image and cannot make use of the redundancies across different images in a set of images. Video is generally thought of as a sequence of frames or images. These sequences of frames are assumed to have certain type of correlation which is exploited by video compression techniques like HEVC and H.264. Unfortunately, images in sets of images stored on media devices do not have any defined correlation as in the case video. This difference is primarily because images in a set a generally captured from different angles, with random zooming, and even of totally different objects. Hence, new techniques are being developed for compression of set of images. Apart from photo albums, image set compression is also used for compressing medical image sets and satellite image sets.

The image set compression techniques can be broadly classified as following:

- Representative Signal (RS) based approaches
- Graph theory based approaches
- Other approaches

In this paper, an overview of various methods for image set compression is provided. The rest of the paper is organized as follows: Section 2 reviews methods based on representative signals. Section 3 reviews methods based on graph theory. Section 4 provides an overview of other image set compression methods like quadtree based methods and Karhunen-Loeve method. Finally, conclusions are provided in Section 5.

2. Representative Signal Based Approaches

Techniques which use RS based approach try to represent a set of images by some RS and then compress the RS as well as differences between images in the set and the RS. In case of huge set of images, smaller set of correlated images are formed using clustering algorithms and then this approach is applied independently on each smaller set. This approach is very effectively for highly correlated set of data like medical images but the problem with this approach is its inability to deal with outliers effectively.

2.1 Min-Max Differential Method

The use of Min-Max image pair, i.e. a maximum value image and a minimum value image, for image set compression was proposed by Karadimitriou and Tyler in 1996 [1]. Each pixel in Min-image is computed as the minimum value of the corresponding pixel across all images in the set. Similarly, each pixel in Max-image is computed as the maximum value of the corresponding pixel across all images in the set. The method provided in [1] is called Min-Max Differential (MMD) method. In MMD method, pixels in the difference image are generated by subtracting each pixel from the corresponding pixel in the minimum image or the maximum image. The pixel in the minimum image or the maximum image is chosen such that it results in a smaller difference.

For any compression scheme, encoders and decoders must have synchronized states. In [1], encoder uses the minimum image or maximum image for the difference as long as the difference value is not larger than $(max\_pixel - min\_pixel)/2$. Thereafter, it encodes this difference value and switches to the other image. This difference value serves as an indication to the decoder to switch between minimum and maximum images and thereby remain synchronized without additional information to be transferred between encoder and decoder for switching between minimum and maximum images.

2.2 Min-Max Prediction Method

Karadimitriou and Tyler extended the MMD method to develop Min-Max Prediction (MMP) method in [2]. MMP method is more complex than MMD but also provide better results. The Min-Max image pair in MMP is same as in MMD. Value of a pixel in minimum image and maximum image provides the limit of the range of value which this pixel can take in the image set. MMP exploits the assumption that values of neighboring pixels are near-by and fall in the same area within the range. In MMP, the range for every pixel is divided into N equi-distant levels and it predicts the value of a pixel using the level information of previous pixels. The prediction is calculated as follows:

$$Pred(i) = min(i) + L(i-1)*(max(i) - min(i))/N$$

where Pred(i) is the predicted value for the ith pixel, min(i) is the value of the ith pixel in the minimum image, max(i) is the value of the ith pixel in the maximum image, L(i-1) is the level of the (i-1)th pixel, and N is the number of level. The level of the ith pixel is calculated using the actual pixel value and it is used for calculating the prediction value of the next pixel. Instead of predicting the level using the previous pixel, i.e. (i-1)th pixel, using (L(left) + L(upper))/2 and (L(left) + L(upper) – L(upperleft)) has also been proposed where L(left) is the level of the left pixel, L(upper) is the level of the upper pixel, and L(upperleft) is the level of the upper-left pixel.

2.3 Centroid Method

Karadimitriou and Tyler also proposed a scheme using an average image (RS) instead of Min-Max image pair in [3]. In this scheme, an average image of all images in the set is computed and the difference image is generated by subtracting the image with the average image. A model for pixel prediction in natural images has been provided in [3].

El-Sonbaty et. al. extended the Centroid method in [4] by recursively using median image instead of average image. The median image of the input image set is calculated and a set of difference images is generated using the input image set and the median image in the first stage. In the second stage, median image of this set of difference images is calculated and a second set of difference images is generated using the first set of difference images and the new median image. This step is repeated a pre-determined number of times. The results in [4] show that the scheme provides improvements only till two stages.

2.4 Template Method

For cases where the similarity in the set of images is known apriori, template method is useful. Wang and Yan proposed a template based scheme compression for images of forms in [5]. An image of a blank form was using as a template in [5]. Similar to the centroid method, the difference images in template method are generated by subtracting input set images with the template image. This method is useful only for pre-defined image set.

3. Graph Theory Based Approaches

RS based approaches are effective for tight image set but performance of RS based schemes is affected badly by outliers. Graph theory based schemes overcome this problem. In graph theory, a Minimum Spanning Tree (MST) is a tree which connects all nodes and whose total weight is less than or equal to any other tree which connects all nodes. In [6], Chen et. al. proposed an image set compression scheme which used MST with all set images as its nodes and prediction cost between the two nodes images of the edge as edge weight. In [6], the mean absolute residual obtained from motion estimation between two images is taken as the prediction cost (edge weights) and MST is obtained using this prediction cost. The MST is thereafter used to create a pseudo-sequence of images and video encoding scheme is used for generating compressed bitstream. A similar scheme is also provided by Nielsen and Li in [7] where root mean squared error between two images is used as the prediction cost.

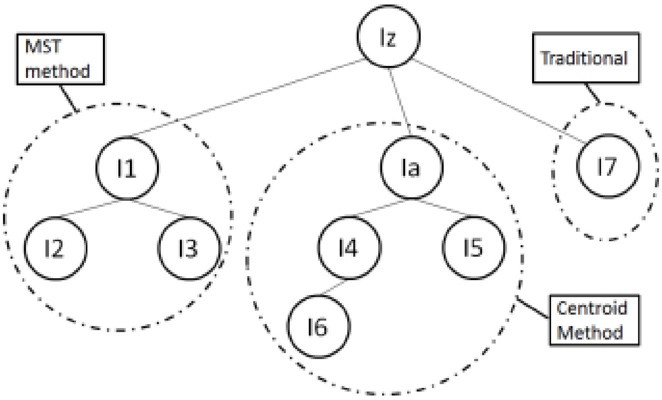

Figure 1: A Minimum spanning tree for MST_a which uses zero image (Iz), average image (Ia), and all images (I1–I7). Parts of the tree which behave like traditional scheme, centroid method, and MST method are marked.

Gergel et. al. extended MST based image compression scheme and provided a unified framework for adaptively selecting between traditional method of compressing individual images, centroid method, and MST method in [8]. The method proposed in [8] is called MST_a. All images of the set are used as nodes to the tree in MST method provided in [6] and [7]. But in MST_a, zero image, average image of the set, and all images of the set are used as nodes of the tree. By incorporating the use of traditional method, centroid method, as well as MST method, MST_a outperforms all these three schemes. For sets which have similar images, MST_a behaves more like centroid method and for sets with varied images, MST_a behaves more like MST method. Figure 1 show an example of a minimum spanning tree which illustrates this point.

In [8], Schmieder et. al. further extended MST_a and provided a hierarchical approach. While MST_a included zero image, average image of the set, and all images of the set to the tree, hierarchical scheme in [8] forms clusters in the image set and then includes zero image, average images of each cluster, and all images of the set to the tree.

Recently, feature based prediction cost for building MST has been proposed [10], [11], [12], [13]. Feature based scheme can be invariant to scaling, rotation, and robust to illumination changes and hence better suited for compression of photo albums. In [10], local feature set Scale Invariant Feature Transform (SIFT) is used for clustering image in the image set as well as for generating MST of each cluster. Clustering is achieved using a modified k-mean algorithm wherein the distance between two images is defined as the mean absolute distance between matched 128D SIFT descriptors [14]. This SIFT based distance is also used as edge cost for generating MST within each cluster. A global alignment which involves SIFT based transformation and brightness adjustment is applied on predictor image so as to achieve lower residual image. Block based motion estimation is applied to obtain the residual image and the residual image is entropy encoded using HEVC compatible entropy encoder. [11], [12], and [13] also employ similar schemes.

4. Other Approaches

Apart from RS based methods and graph theory based methods, methods like quadtree based method and karhunen-loeve method have also been proposed for image set compression. Medical image sets like MRI and CT scan, which are inherently 3D datasets, have been compressed using 3D dataset compression scheme in [15]. 3D dataset compression cannot be used for general images sets as their properties are very different from 3D datasets.

4.1 Quadtree Based Methods

Quadtree is a tree datastructure in which each node either has exactly four children or no children (leaf node). A region quadtree recursively partitions the space into four equal quadrants and is commonly used to represent images, especially binary images. For representing images using region quadtree, image region is recursively split into four quadrants unless the node image region become homogenous i.e. all pixels in the node image region are same. Vassilakopoulos et. al. proposed a quadtree based set compression method using an overlapping structure in [16] and evaluated the scheme for binary images. In [16], overlapping portions of quadtrees of two consecutive images is determined and the quadtree of the second image is modified to refer the common parts directly from quadtree of the first image. The results in [16] show that compression is as high as 90% for slowly changing binary images. In [17], Jomier et. al. have proposed generic quadtree structure which combine all images into a single quadtree with descriptive nodes. Generic quadtree helps in efficiently working on one image or simultaneously on any set of images.

4.2 Karhunen-Loeve Method

Kathunen-Loeve Transform (KLT), also known as Hotelling transform, is a technique for principal component analysis. Musatenko and Kurashov proposed an image set compression method using KLT in [18]. In [18], Gram-Schmidt orthogonalization is used to generate a set of basis vectors which decorrelates images in the set. The basis vectors are arranged in decreasing order of importance and they are encoded using embedded zero-tree wavelet coding such that more bits are given to basis vectors of high importance and less bits are given to basis vectors of low importance. Images in the set are represented as linear combination of basis vectors and the coefficients corresponding to less important basis vector can be quantized more than those corresponding to more important basis vectors. KLT is computationally very expensive and hence not preferred for practical applications. Moreover, images in the set should be similar so as to have a small set of basis vectors. Another problem with using KLT for image set compression is that a single (first) image decoding requires recoding of all basis vectors.

5. Conclusions

The field of image set compression has become very important for compressing photo albums as the use of images has increased exponentially. Compression of medical image sets is another important application of image set compression. Medical image sets have high correlation among its images whereas photo albums have clusters of correlated images. RS based methods seem to be most suitable for medical image set compression and MST methods seem to be most suitable for image sets like photo albums. Recent researches in MST methods have made use of advanced feature detection algorithms like SIFT and even incorporated modules from state-of-the-art video compression codecs like HEVC.

References

[1] K. Karadimitriou and J. M. Tyler, “The min-max differential method for large-scale storage and compression of medical images,” Proceedings of Annual Molecular Biology and Biotechnology Conference, Baton Range, LA, USA, 1996.
[2] K. Karadimitriou and J. M. Tyler, “Min-max compression methods for medical image databases,” ACM SIGMOD Record, vol. 26, no. 1, pp. 47–52, Mar. 1997.
[3] K. Karadimitriou and J. M. Tyler, “The Centroid method for compressing sets of similar images,” Pattern Recognition Letters, vol. 19, no. 7, pp. 585-593, May 1998.
[4] Y. El-Sonbaty, M. Hamza, and G. Basily, “Compressing sets of similar medical images using multilevel centroid technique,” Proceedings of the 7th Conference on Digital Image Computing Techniques and Applications, C. Sun, H. Talbot, S. Ourselin, and T. Adriaansen, Eds., Sydney, Australia, Dec. 2003.
[5] J. Wang and H. Yan, “Form image compression using template extraction and matching,” Proceedings of Visual Information, Dec. 2000.
[6] C. P. Chen, C. S. Chen, K. L. Chung, H. L. Lu, and G. Tang, “Image set compression through minimal-cost prediction structures,” IEEE International Conference on Image Processing, pp. 1289–1292, 2004.

[7] C. Nielson and X. Li, “MST for lossy compression on image sets,” Proceedings of the Data Compression Conference, pp. 463, 2006.
[8] B. Gergel, H. Cheng, C. Nielsen, and X. Li, “A unified framework for image set compression,” In Proceedings of the 2006 International Conference on Image Processing, Computer Vision, & Pattern Recognition, pp. 417–423, 2006.
[9] A. Schmieder, B. Gergel, H. Cheng, and X. Li, “Hierarchical Minimum Spanning Trees for Lossy Image Set Compression,” Proceedings of the 2008 International Conference on Image Processing, Computer Vision, & Pattern Recognition, pp. 57–63, 2008.
[10] Z. Shi, X. Sun and F. Wu, “Feature based Image Set Compression,” IEEE Conference on Multimedia and Expo (ICME), 2013.
[11] H. Yue, X. Sun, and F. Wu, “Cloud based Image Coding for Mobile Devices –Toward Thousands to One Compression,” IEEE Transaction on Multimedia, Issue. 99, 2012.
[12] Z. Shi, X. Sun and F. Wu “Multi-model prediction for image set compression,” Visual Communications and Image Processing (VCIP), pp. 1-6, Nov. 2013.
[13] Z. Shi, X. Sun, and F. Wu, “Photo Album Compression for Cloud Storage Using Local Features,” IEEE Journal on Emerging and Selected Topics in Circuits and Systems, vol. 4, no. 1, pp. 17-28, March 2014.
[14] David G. Lowe, "Distinctive image features from scale invariant keypoints," International Journal of Computer Vision, vol. 60, pp. 91-110, 2004.
[15] X. Qi and J. M. Tyler. A progressive transmission capable diagnostically lossless compression scheme for 3D medical image sets. Information Sciences, vol. 175, no. 3, pp. 217–243, 2005.
[16] M. Vassilakopoulos, Y. Manolopoulos, and N. Economou. Overlapping quadtrees for the representation of similar images. Image and Vision Computing, vol. 11, no. 5, pp. 257–262, 1993.
[17] G. Jomier, M. Manouvrier, and M. Rukoz. Storage and management of similar images. Journal of the Brazilian Computer Society, vol. 6, no. 3, pp. 13–25, 2000.
[18] Y. S. Musatenko and V. N. Kurashov. Correlated image set compression system based on new fast efficient algorithm of Karhunen-Loeve transform. In C. C. J. Kuo, S.-F. Chang, and S. Panchanathan, editors, Proceedings of The International Society for Optical Engineering: Multimedia Storage and Archiving Systems III, pp. 518–529, Oct. 1998.