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

SURF Based Image Set Compression

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

[1]Dept. 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: ***The field of image set compression, which refers to the compression of image sets like photo albums and medical images, is becoming increasingly relevant in today's digital world. In recent years, several schemes for image set compression have been proposed. In this paper, we propose a new SURF based image set compression scheme which uses SURF and HEVC. We also provide comparison results of our scheme with SIFT based image compression scheme.***

Keywords: Image Set Compression, Image features, MST, SIFT, SURF, HEVC.

1. Introduction

Use of photos has increased exponentially in recent years. For e.g., Facebook revealed in a white paper that as of 2013, Facebook users had uploaded over 250 billion photos and that it was increasing at a rate of 250 million photos every day [1]. This number is increasing exponentially with features like auto upload from mobile phone. This has prompted researchers to develop efficient methods for storing images and more importantly, set of images. Traditional method of compressing set of images involves compressing each image individually using well established image compression methods like JPEG, JPEG-2000, and PNG. Set of images generally have images which have correlated content. Image compression methods are highly efficient in exploiting correlation within the image but cannot use correlation across images in a set of images. Video compression methods like H.264 and High Efficiency Video Coding (HEVC) exploit inter frame correlation but they assume certain type of correlation between frames which does not exist in set of images. There is a group of techniques for efficient compression of image sets which uses graph based approach. In graph theory based approaches, following steps are involved:

1) Images in the set are first classified into groups of similar images
2) Minimum Spanning Tree (MST) is determined for each group
3) A sequence of images is obtained by traversing the MST for each group
4) The sequence of images are compressed using video compression methods. Some sort of alignment may be used in this step to achieve higher compression

In this paper, we propose a method which uses Speeded Up Robust Feature (SURF) for image grouping as well as MST generation and HEVC for final step of video compression. The rest of the paper is organized as follows: Section 2 provides a review of the graph theory based approach for image set compression, Section 3 reviews keypoint detection in images and use of SIFT and SURF for feature extraction and image matching, Section 4 describes the proposed method of image set compression which uses SURF and HEVC along with experimental results and finally, conclusions are provided in Section 5.

2. Graph Theory Based Image Set Compression

In our earlier paper, we had provided a review of state-of-the-art image set compression techniques [2]. We provide relevant parts of the review in this section of the paper as well for quick reference. Graph theory based compression techniques are the most advanced techniques for image set theory. In [3], Chen et. al. proposed an image set compression scheme which uses MST with all set images as its nodes and prediction cost between the two nodes images of the edge as edge weight. In [3], 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 [4] where root mean squared error between two images is used as the prediction cost.

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 [5]. The method proposed in [5] is called MSTa. All images of the set are used as nodes of the tree in MST methods provided in [3] and [4] but in MSTa, 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, MSTa outperforms all these three schemes. For sets which have similar images, MSTa behaves more like centroid method and for sets with varied images, MSTa behaves more like MST method. In [6], Schmieder et. al. further extend MSTa and provide a hierarchical approach. While MSTa included zero image, average image of the set, and all images of the set to the tree, hierarchical scheme in [5] 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 [7], [8], [9], [10]. Feature based scheme can be invariant to scaling, rotation, and robust to illumination changes and hence better suited for compression of photo albums. In [7], local feature set Scale Invariant Feature Transform (SIFT) is used for clustering images 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 [11]. 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. [8], [9], and [10] also employ similar schemes.

In this paper we propose a method which uses feature set Speeded Up Robust Feature (SURF) for clustering images as well as for generating MST for each cluster.

3. Keypoint (Feature) Detection in Images

Image matching is an important field of research in computer vision and it is generally achieved by either global feature matching or local feature matching. Local feature matching is more stable and involves two stages: interest point detection and their description. Feature detection should have high repeatability and speed and its descriptor should have low number of dimensions.

In 1999, David G. Lowe proposed a local feature descriptor called Scale-Invariant Feature Transform (SIFT). SIFT and its variants have become extremely popular for feature detection. A more detailed description of SIFT is provided by Lowe in 2004 [11]. In 2006, Bay et. al. proposed a local feature descriptor called Speeded Up Robust Feature (SURF) which is partly inspired by SIFT but is much faster [12].

In this section, we shall review SIFT and SURF based local feature descriptors.

3.1 Scale Invariant Feature Transform (SIFT)

Following steps are required for constructing SIFT local feature descriptor:

1. Construct scale space: Scale space is implemented as Difference of Gaussian (DoG) pyramid. Image is consecutively filtered using Gaussian filter followed by sub-sampling. DoG pyramid is formed by evaluating difference between images in consecutive levels of Guassian filtering.
2. Locate potential feature points: Extrema points in the DoG pyramid are obtained by first evaluating local maxima and minima using a neighborhood of 26 values (8 values in same scale and 9 values each in scales above and below) and thereafter improving the localization to subpixel accuracy.
3. Eliminating weak keypoints (Filter low contrast responses and edges): Extrema points whose key value is less than a threshold and extrema points which are poorly localized, i.e. which are edge points, are eliminated. Poorly localized extrema points are determined using Hessian.
4. Orientation assignment: Dominant orientation is assigned to keypoints based on peak in the histogram of orientations in the local neighborhood around the keypoints.
5. Build keypoint descriptors: Final descriptor is computed by taking a neighborhood of size 16x16 around keypoints, dividing them into cells of 4x4, and evaluating histogram of gradient directions within each cell. Histogram uses gradient magnitude and Gaussian function based on distance as weights.

3.2 Speeded Up Robust Feature (SURF)

Following steps are required for constructing SURF local feature descriptor:

1. Construct scale space: Scale space is constructed using box filters which is an integer approximation of Determinant of Hessian (DoH). This approach implies that scale space can be generated using integral images instead of filtering the image consecutively using Gaussian filter.
2. Locate potential feature points: This step is similar to locating potential feature points in SIFT. Extrema points are obtained using a 3x3x3 neighborhood (26 values in neighborhood) and thereafter improving the localization to subpixel accuracy.
3. Eliminating weak keypoints: Extrema points whose value (approximated DoH) is less than a threshold are eliminated.
4. Orientation assignment: Dominant orientation is assigned based on the sum of all responses within a sliding orientation window covering $\pi/3$ angle. Orientations are obtained using wavelet (Haar) responses weighted with Gaussian function.
5. Build keypoint descriptors: With ‘s’ being the scale of the keypoint, final descriptor is computed by taking a neighborhood of size 20sx20s around the keypoint which is oriented along the keypoint orientation and dividing it into 4x4 sub-regions. For each sub-region, sum of horizontal and vertical gradients and their absolute values are calculated for regularly spaced 5x5 pixels which gives 4 values for each sub-region. Hence, a descriptor of 64 values is obtained which is known as SURF-64. If the sum for positive and negative values is taken separately, 8 values are obtained for each sub-region resulting in a descriptor of 128 values known as SURF-128.

4. SURF and HEVC based Image set compression scheme

We propose an image set compression scheme which uses local feature set SURF-64 instead of SIFT for clustering as well as generating MST of image cluster. We evaluated Prim as well as Kruskal algorithms for MST and both provided same MST output for our test image sets. HEVC video encoding is used for encoding of image set as a sequence of images formed using the

Table 1: Comparison of bits generated and luminance PSNR for various Sequences/Image sets

Sequence/Image Set	Resolution	No. of Frames	Total Bits Generated		Luminance (Y) PSNR	
			SIFT Based	SURF-64 Based	SIFT Based	SURF-64 Based
Castle Dense	1536x1024	19	1924672	1932280	36.22	36.29
Castle Entry	1536x1024	10	1243544	1243544	35.19	35.19
Church Entry	512x768	15	1221680	949200	32.37	31.39
Mall Room	1072x600	7	662424	659528	33.73	33.15
Rock Boat	1296x864	20	2400344	2958296	33.15	33.95
Wadham College	1024x768	5	486376	486376	33.68	33.68

MST. Important HEVC configurations used for encoding are as follows: main profile, IP-only, rate control off, and Qp 32 for I-slice followed by QpOffset of (3,2,3,1) for P-slices.

Table 1 provides comparison of bits generated and luminance Peak Signal to Noise Ratio (PSNR) with clustering/MST generation using SIFT and SURF-64. Figure 1 provides the comparison of bits per pixel generated for the image sequences with schemes based on SIFT and SURF-64. While in some cases SURF-64 performs better than SIFT, the most degradation with SURF-64 as compared to SIFT in terms of bits per pixel is found to be only 0.025 bits per pixel in our test image sets. Hence, it can be concluded SIFT and SURF-64 give similar performances in terms of output. Since SURF is almost 3x faster that SIFT [12], our scheme is preferred to the SIFT based scheme. It is also observed that in small and simple image sets, SIFT and SURF-64 give same output but their output differs when image set become complex.

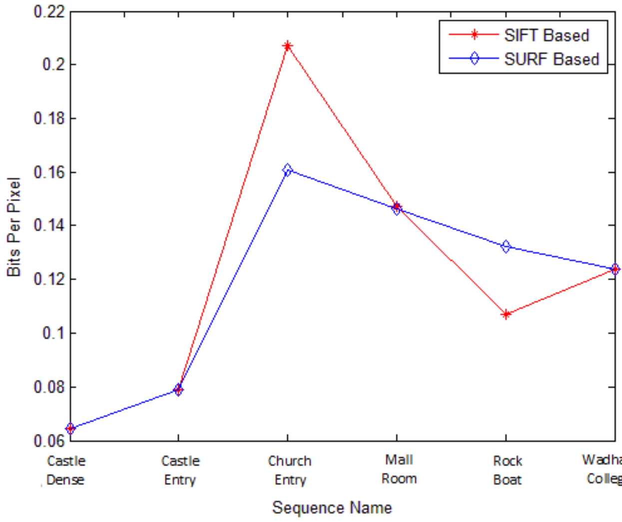

Figure 1: Comparison of bits per pixel generated for image sequences with schemes based on SIFT and SURF-64

5. Conclusions

In this paper, we proposed a SURF based image set compression scheme which uses SURF for clustering as well as MST generation and HEVC for encoding the images in MST as a sequence of video frames. Results show that bits generated and PSNR obtained with SURF are similar and within acceptable limits to that obtained using SIFT and as SURF is much faster that SIFT, our SURF based scheme is preferred to SIFT based scheme.

References

[1] "A focus on efficiency", www.internet.org/efficiencypaper [Accessed: Jul. 14, 2015].
[2] Z. Bohari and M. S. Raeen, "Overview of state-of-the-art image set compression techniques," International Journal of Science and Research, vol. 4, issue 4, Apr 2015.
[3] 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.
[4] C. Nielson and X. Li, "MST for lossy compression on image sets," Proceedings of the Data Compression Conference, pp. 463, 2006.
[5] 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.
[6] 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.
[7] Z. Shi, X. Sun and F. Wu, "Feature based Image Set Compression," IEEE Conference on Multimedia and Expo (ICME), 2013.
[8] 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.
[9] 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.
[10] 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.
[11] David G. Lowe, "Distinctive image features from scale invariant keypoints," International Journal of Computer Vision, vol. 60, pp. 91-110, 2004.
[12] H. Bay, T. Tuytelaars, and L. V. Gool, "SURF: Speeded Up Robust Features," European Conference on Computer Vision (ECCV), pp. 404-417, 2008.