LENSLET IMAGE COMPRESSION USING ADAPTIVE MACROPIXEL PREDICTION

Haixu Han, Xin Jin, Senior Member, IEEE, Qionghai Dai, Senior Member, IEEE

Shenzhen Key Lab of Broadband Network and Multimedia,
Graduate School at Shenzhen, Tsinghua University, Shenzhen 518055, China

ABSTRACT

In this paper, an efficient compression method is proposed for lenslet images captured by plenoptic cameras for recording the spatial and angular light information at a super-high-resolution. After applying a reversible image reshaping method to the lenslet image, a reshaped and regularized image will be generated and compressed by the video codec comprising the proposed adaptive macropixel prediction mode. Based on the analysis of spatial correlations among adjacent macropixels, two spatial prediction modes are proposed as: multi-block weighted prediction mode and co-located single-block prediction mode, to predict the coding unit by minimizing the coding cost. The multi-block weighted prediction is formulated by minimizing the Euclidean distance between the coding unit and co-located blocks in the macropixel structure. Performance evaluations have shown that the proposed method achieves 50.9% of bit-savings on average compared to HEVC. It also outperforms state-of-the-art coding methods drastically.

Index Terms— Lenslet image compression, adaptive macropixel prediction, HEVC, adaptive mode decision.

1. INTRODUCTION

In recent years, the commercial plenoptic cameras like *Lytro* [1] have attracted great interests for their capability of light field acquisition. Based on microlens array architecture as shown in Fig. 1, a lenslet image captured by plenoptic cameras is composed of multiple macropixels that can record both spatial and angular light information. Nevertheless, containing essential 4D light field information results in larger volume of data than traditional images, which desires an efficient compression approach for storage and transmission.

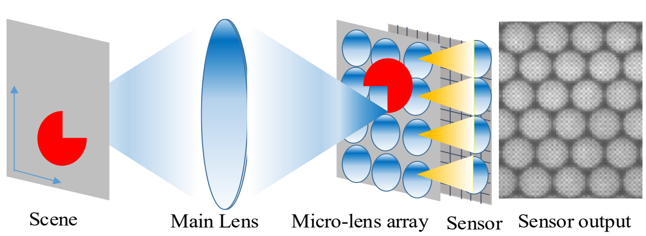


Fig. 1 Image formation process of a standard plenoptic camera.

The existing methods of lenslet image compression can be mainly classified into two categories: compressing the lenslet images directly and compressing a pseudo video sequence generated from lenslet images. The approaches in the first category compressed the lenslet image directly using the existing image/video coder such as JPEG [2]. Some coding tools such as locally linear embedding-based (LLE) prediction [3], disparity compensation [4] and displacement intra prediction [5] were proposed. Also, some coding tools in HEVC screen content coding like Intra-Block Copy (IBC) mode [6] can efficiently improve the compression efficiency. However, the optical imaging correlations among macropixels were not fully exploited yet. For the approaches in the second category, a pseudo video sequence was generated by extracting subaperture images from the lenslet image and afterwards compressed by a video coder [7]. Some approaches like self-similarity compensated prediction [8] were proposed to improve the compression efficiency. However, the computational complexity of such approaches is hundreds of times higher than those in the first category. Furthermore, due to incoherence between macropixels and the coding unit grids in the block-based coding architecture, the spatial correlations among macropixels were not fully exploited by the approaches in these two categories, which greatly affects the compression efficiency.

In this paper, a novel compression method is proposed for lenslet images. Considering the optical characteristics of the lenslet images, an image reshaping scheme proposed by us in [9] is applied to align the macropixel structure with coding unit grids. Based on the regularized lenslet image, an adaptive macropixel prediction including the multi-block weighted prediction mode and the co-located single-block prediction mode is proposed. The multi-block weighted prediction is formulated by minimizing the Euclidean distance between the coding block and co-located blocks in adjacent macropixels. Experimental results demonstrate that the proposed method achieves a maximum of 80.9% bitrate reduction with an average of 50.9% bitrate reduction compared with HEVC [10]. It also provides a significant improvement relative to existing methods like IBC [6] and LLE [3] with an average of 30.3%/25.6% bitrate reduction.

The rest of this paper is organized as follows. Section 2 presents the framework of the proposed method. The proposed adaptive macropixel prediction is described in

Section 3. Experimental results are shown in Section 4 followed by conclusions in Section 5.

2. PROPOSED SYSTEM ARCHITECTURE

Using HEVC as the video encoder, the proposed encoding system comprising the image reshaping method and the proposed adaptive macropixel prediction mode is shown in Fig. 2.

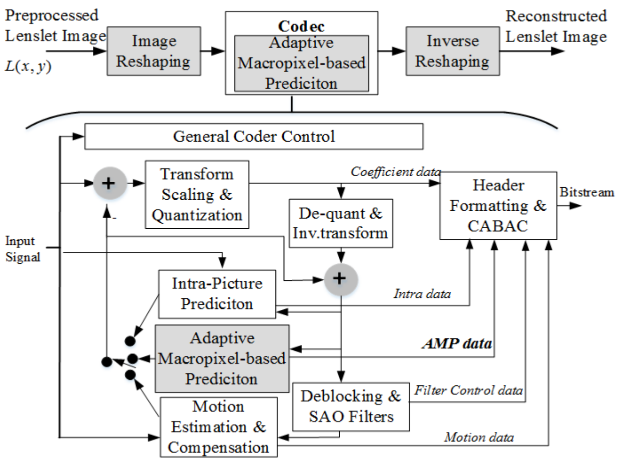
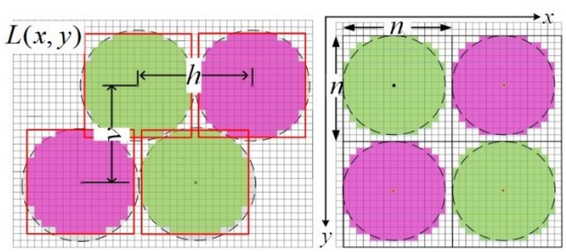


Fig. 2 The proposed encoding system.

First, the preprocessed lenslet image $L(x,y)$ with an orthogonal integer pixel grids shown in Fig. 3(a), which is processed by demosaicing, devignetting, rotation and scaling [11], is fed to the proposed encoding system. Applying the *Image Reshaping* algorithm proposed by us in [9], the macropixels in the preprocessed lenslet image are realigned to the coding unit grid to be block-based coding friendly. *Image Reshaping* mainly consists of two steps: macropixel alignment, which aligns the centers of macropixels and guarantees each macropixel can be fully contained by a non-overlapped 16×16 block as shown in Fig. 3(b) (n=16); and adaptive interpolation, which generates the intensity value for the missing pixels to maximize the intensity continuity among the regularized adjacent macropixels in the reshaped lenslet image. The reshaped image will be 14.2% larger than the preprocessed lenslet image according to the calculation.



(a) (b)

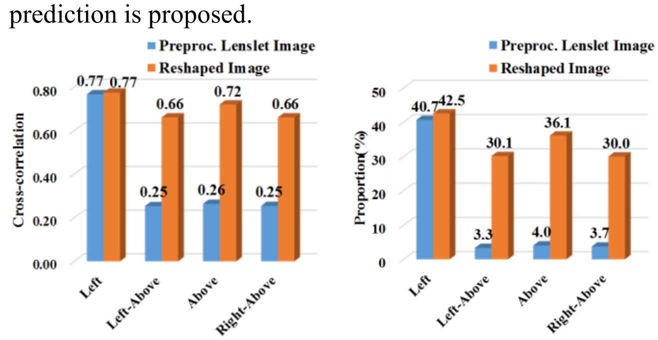
Fig. 3 Macropixel structures: (a) in a preprocessed lenslet image; (b) after image reshaping.

Then, the regularized and reshaped image is compressed by a video codec comprising the proposed adaptive macropixel prediction (AMP). AMP mode predicts the current coding unit using a co-located block in the adjacent macropixel or using the weighted combination of co-located blocks in the adjacent macropixels by solving a Euclidean distance minimization problem. It can fully exploit the distinct correlations among macropixels without requiring any priors of the optical system. The proposed prediction mode is described in detail in Section 3.

Finally, an *Inverse Reshaping* process, which is the inverse process of image reshaping, is performed on the reconstructed image to convert it back to the original lenslet image structure as shown in Fig. 3(a).

3. ADAPTIVE MACROPIXEL PREICTION

In the regularized lenslet image, high spatial correlations among adjacent macropixels can be easily detected and exploited. An analysis is conducted by calculating the average of cross-correlation coefficients between the current 16×16 block and its adjacent four 16×16 blocks located left to, left-above, above and right-above the current block. The statistics are retrieved both for the preprocessed lenslet image and the reshaped lenslet image using image "Fountain_&_Vincent" [12] in Fig. 6(c) as an instance. As shown in Fig. 4(a), after applying image reshaping to the lenslet image, the average of cross-correlation between the current block and its four adjacent blocks is much higher than the preprocessed lenslet image. And, the proportion of the blocks with strong correlations, which have the cross-correlation higher than 0.9, in the reshaped image is also much larger than that in the preprocessed lenslet image, as shown in Fig. 4(b). Hence, based on the spatial correlations, an adaptive macropixel prediction mode including the multi-block weighted prediction and co-located single-block prediction is proposed.



(a) (b)

Fig. 4 The statistics of spatial correlations: (a) the average of cross-correlation; (b) the proportion of blocks with cross-correlation higher than 0.9.

A. Multi-block weighted prediction

The multi-block weighted prediction (MWP) is designed to provide a precise prediction using a linear combination of co-located blocks.

1) *Reference blocks*: In order to employ the spatial correlations of adjacent macropixels, four reference blocks co-located in the spatially adjacent macropixels are selected. *The reference block size is consistent with the prediction*

unit size of the current block. As the prediction unit size is the integer multiples of 16, from 64×64 to 16×16, the four reference blocks spatially adjacent to the current unit in the same size are used. The green blocks shown in Fig. 5(a) and (b) are the reference blocks used for prediction unit size 32×32 and 16×16 as instances. As the prediction unit size varies from 8×8 to 4×4, the four blocks co-located in the spatially adjacent reconstructed 16×16 blocks, as the green blocks shown in Fig. 5(c)-(d), are used. If some of them are unavailable, only the available candidates will be used.

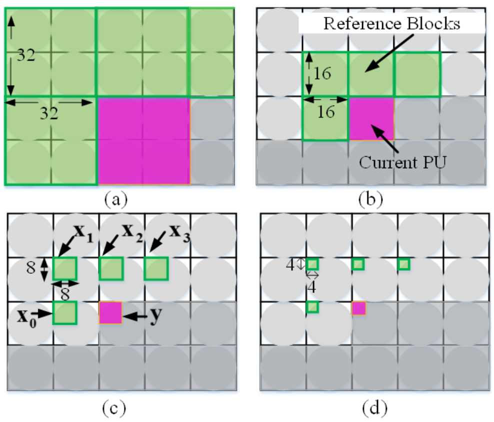


Fig. 5 The current prediction unit, the block in red, and its reference blocks, the blocks in green, as the prediction unit size equals to: (a) 32×32 (similar with that equals to 64×64); (b) 16×16; (c) 8×8; and (d) 4×4.

2)*Weighted prediction*: The current coding unit is predicted by

$$\mathbf{y'} = w_0\mathbf{x_0} + w_1\mathbf{x_1} + w_2\mathbf{x_2} + w_3\mathbf{x_3} \quad , \qquad (1)$$

where $\mathbf{y'}$ represents the prediction of the current unit; $\mathbf{x}_i$ is a reference block (a green block shown in Fig. 5); w_i is the weight value corresponding to $\mathbf{x}_i$, which is derived by minimizing the Euclidean distance between the current block and the reference blocks as

$$\begin{aligned} &\text{minimize} \quad \|\boldsymbol{Xw} - \boldsymbol{y}\|_2^2 \\ &\textit{subject to} \quad \mathbf{1}^T \boldsymbol{w} = 1 \quad , \\ &\qquad\qquad \boldsymbol{w} \geq 0 \end{aligned} \qquad (2)$$

where $\boldsymbol{w}$ is the weight vector arranged by w_i; $\boldsymbol{y}$ is the vectorized current unit; $\boldsymbol{X}$ is a matrix in which each column is a vectorized $\mathbf{x}_i$. The logarithmic barrier method [13] is applied to solve the equation. To encode the weights efficiently, they are scaled and quantized to integer values from 0 to 127. Hence, the coding unit is predicted by

$$\mathbf{y'} = (w_0\mathbf{x_0} + w_1\mathbf{x_1} + w_2\mathbf{x_2} + w_3\mathbf{x_3} + 64) >> 7 \ , \quad (3)$$

where >> denotes a bit shift operation to the right.

B. Co-located single-block prediction

Although the MWP mode can provide good prediction, it will introduce overhead bits cost by the weights. The overhead bits will affect the compression efficiency especially at low bitrate. Hence, based on statistics of the cross correlations retrieved above, a co-located single-block prediction (CSP) mode is proposed, which only uses a single reference block from the four co-located blocks, as the green blocks in shown in Fig. 5, as the prediction of the current unit. The prediction is formulated as

$$\mathbf{y'} = \mathbf{x}_i \quad . \qquad (4)$$

C. Mode and weight value coding

Applying the proposed prediction modes to luminance component, the mode indexes from 35 to 39 are assigned to MWP and the four modes in CSP. The mode number for luma and chroma is coded same with HEVC intra mode coding. The four weight values in MWP, scaled to integer values between 0 and 127, are coded using the model of intra prediction mode in HEVC for simplicity. Similar with mode coding for luma, the set of three most probable weights selected from two neighboring blocks is established. If the weight value belongs to the most probable weight set, the index in the set is transmitted; otherwise, the weight value is coded using arithmetic model. The entire weight coding process is based on the CABAC model for luma mode coding, which can be further improved.

Finally, as the entire video codec architecture shows in Fig. 2, the encoder will choose the best mode from the proposed adaptive macropixel prediction mode and 35 Intra modes in HEVC based on rate distortion optimization (RDO). Table 1 shows the proportion of intra modes that are selected using image "Fountain_&_Vincent" [12] as an instance. It can be seen that most blocks select the proposed adaptive macropixel intra mode, which demonstrates that five proposed modes can provide a precise prediction for the current block adaptively.

Table 1. Proportion of the proposed prediction modes.

Intra Mode	Mode Index	QP			
		26	32	38	44
Multi-block weighted prediction	35	55.3%	38.4%	26.5%	16.4%
Co-located single-block prediction	36	11.2%	16.4%	16.4%	14.4%
	37	0.7%	1.2%	1.5%	1.5%
	38	8.8%	18.9%	26.3%	33.8%
	39	1.0%	1.7%	1.0%	0.6%
Others	0~34	23.1%	23.4%	28.3%	33.2%

4. EXPERIMENTAL RESULTS

The proposed adaptive macropixel prediction is implemented in HEVC/H.265 Format Range Extension (RExt) reference software (HM-16.9SCM8.0)[14]. Five lenslet images from the database [12] and five lenslet images captured by our own *Lytro Illum* camera, shown in Fig. 6 with spatial resolution 7728×5368 and angular resolution 15×15, are tested. After converting the color space from RGB to YCbCr4:4:4, the lenslet images are compressed using the "All Intra" configuration [15] under QP 26,32,38 and 44.

In order to demonstrate the effectiveness of the proposed method, the eight cases listed in Table 2 are compared.

Using *HEVC* as the benchmark, the performance of the proposed method is demonstrated by comparing with two state-of-the-art algorithms: Intra Block Copy (IBC) mode proposed in [6] and Locally Linear Embedding (LLE) based 3D holoscopic coding method proposed in [3]. BD-Bitrate [16] is calculated, in which PSNR is calculated between the subaperture images rendered from the original lenslet image and those rendered from the reconstructed lenslet image [17], and the bitrate is the number of bits obtained by the compression methods.

Fig. 6 Test lenslet images: (a)-(e) are downloaded from light-field image database [12]; (f)-(j) are captured by our *Lytro Illum* camera.

Table 2. Coding configurations of different methods.

Test cases	**Configurations**
HEVC	HEVC RExt Profile, bypassing the gray blocks in Fig. 2.
IBC	HEVC RExt Profile + IBC mode proposed in [6]
LLE	HEVC RExt Profile + LLE method proposed in [3]
IR	HEVC RExt Profile + our previously proposed IR in [9]
AMP	HEVC RExt Profile + proposed adaptive macropixel prediction
IR+IBC	HEVC RExt Profile + IR+ IBC
IR+LLE	HEVC RExt Profile + IR+ LLE
IR+AMP	HEVC RExt Profile +IR+AMP

Compression efficiency comparison results are summarized in Table 3. It can be seen that the proposed method *Image Reshaping* (*IR*) and *Adaptive Macropixel-based prediction* (*AMP*) always outperforms *HEVC* with 14.7%/23.7% bitrate reduction on average, respectively. The overall performance of the combination of the two tools is much higher than the summation of performance improvement by applying the two tools individually, i.e. average bitrate reduction achieved by *IR+AMP* is 50.9% which is larger than the summation of 14.7% and 23.7%. Compared with other two approaches *IBC* [6] and *LLE* [3], the proposed method (*IR+AMP*) can also achieve 30.3%/25.6% bitrate reduction on average. After integrating the image reshaping method with LLE and IBC as *IR+LLE* and *IR+IBC*, the proposed methods can also outperform them by an average of 18.0%/16.5% bitrate reduction. Fig. 7 shows the rate-distortion curve for two randomly selected lenslet images *Vase* and *Magic Cube*. It further demonstrates the proposed method outperforms HEVC and other two compression methods at all testing bitrates.

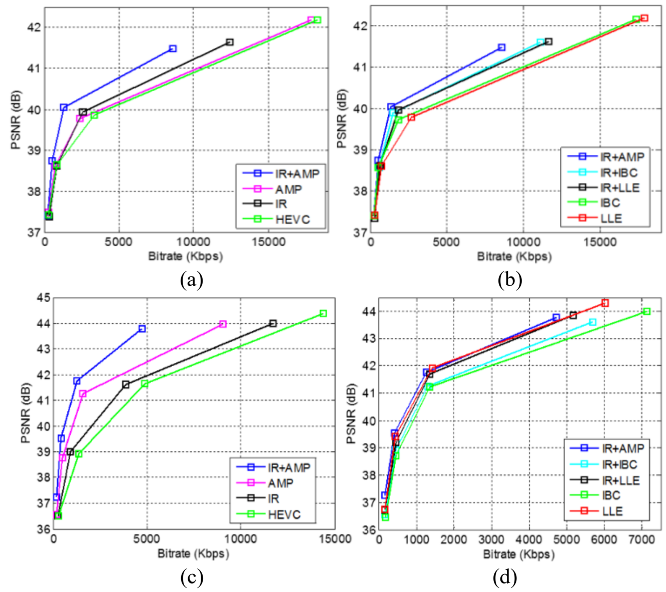


Fig. 7 Rate-distortion results for the test image: (a)(b): Magic Cube; (c)(d): Houses&Lake.

5. CONCLUSIONS

In this paper, we propose a novel lenslet image compression method based on adaptive macropixel prediction. The proposed method can efficiently exploit the optical imaging characteristics of lenslet images and achieve obvious bitrate savings. A peak bitrate reduction of 80.9% and an average bitrate reduction of 50.9% can be achieved compared with HEVC. It also outperforms the state-of-the-art coding tools like IBC and LLE. Extending the method to the light field video is under investigating as the future work.

6. ACKNOWLEDGEMENTS

This work was supported in part by NSFC project of Guangdong 2014A030313733 and the Guangdong Special Support Plan for the scientific and technological innovation young talents 2016TQ03X998, China.

Table 3. BD-Bitrate performance of the proposed scheme.

Image Name	IBC vs.HEVC	LLE vs.HEVC	**IR** vs.HEVC	**AMP** vs.HEVC	**IR+AMP** vs. IBC	**IR+AMP** vs. LLE	**IR+AMP** vs. IR+IBC	**IR+AMP** vs. IR+LLE	**IR+AMP** vs. HEVC
Ankylosaurus_&_Diplodocus	-45.58%	-43.6%	-30.0%	-43.5%	-44.6%	-48.1%	-23.9%	-24.49%	-70.8%
Color_Chart_1	-66.16%	-74.5%	-15.6%	-47.6%	-41.5%	-19.3%	-27.0%	-13.42%	-80.9%
Foutain_&_Vincent	-32.2%	-41.2%	-1.5%	-23.2%	-14.1%	-0.1%	-12.1%	-4.6%	-42.1%
Vespa	-23.5%	-21.4%	-1.6%	-19.8%	-23.0%	-25.1%	-17.7%	-18.1%	-41.3%
Houses&Lake	-53.6%	-61.4%	-34.0%	-51.4%	-38.7%	-23.5%	-19.4%	-13.1%	-72.7%
Cards	-19.8%	-30.0%	-8.6%	-12.0%	-30.6%	-19.4%	-20.9%	-24.9%	-45.3%
Ferriara_Opendoor	-13.4%	-20.9%	-12.5%	-10.0%	-24.7%	-16.7%	-12.6%	-8.3%	-35.6%
Vase	-13.4%	-11.8%	-14.4%	-13.8%	-33.7%	-35.1%	-24.0%	-23.9%	-43.6%
Lamp&Book	-4.8%	-1.2%	-15.2%	-1.5%	-22.4%	-25.6%	-6.0%	-9.4%	-26.6%
MagicCubic	-27.2%	-10.7%	-13.9%	-14.4%	-29.5%	-43.2%	-15.9%	-24.8%	-49.8%
Average	-30.0%	-31.7%	-14.7%	-23.7%	-30.3%	-25.6%	-18.0%	-16.5%	-50.9%

7. REFERENCES

[1] Lytro, https://www.lytro.com/.

[2] Higa R S, Chavez R F L, Leite R B, et al. “Plenoptic image compression comparison between JPEG, JPEG2000 and SPITH”. Cyber Journals: JSAT, 2013, 3(6).

[3] Lucas L F R, Conti C, Nunes P, et al. “Locally linear embedding-based prediction for 3D holoscopic image coding using HEVC[C]”, Signal Processing Conference (EUSIPCO), 2014 Proceedings of the 22nd European. IEEE, 2014:11-15.

[4] Liu D, An P, Ma R, et al. “Disparity compensation based 3D holoscopic image coding using HEVC,” Signal and Information Processing (ChinaSIP), 2015 IEEE China Summit and International Conference on. IEEE, 2015: 201-205.

[5] Y. Li, M. Sjostrom, R. Olsson, U. Jennehag, “Coding of focused plenoptic contents by displacement intra prediction,” in IEEE Transactions on Circuits and Systems for Video Technology, vol.PP, no.99,2015: pp.1-1.

[6] J. Xu, R. Joshi, and R. A. Cohen, "Overview of the Emerging HEVC Screen Content Coding Extension", 2015: IEEE Transactions on Circuits and Systems for Video Technology, Vol. 26, Issue 1, pp. 50-62.

[7] F. Dai, J. Zhang, Y. Ma, Y. Zhang, “Lenselet image compression scheme based on subaperture images streaming,” IEEE International Conference on Image Processing IEEE, 2015: 4733 – 4737.

[8] Conti C, Soares L D, Nunes P. “HEVC-based 3D holoscopic video coding using self-similarity compensated prediction[J],” Signal Processing: Image Communication, 2016, 42: 59-78.

[9] H. Han, X. Jin and Q. Dai, “Plenoptic image compression based on linear transformation and interpolation,” 2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA), Korea, 2016, pp. 1-6.

[10] Sullivan G J, Ohm J, Han W J, et al. “Overview of the High Efficiency Video Coding (HEVC) Standard,” IEEE Transactions on Circuits & Systems for Video Technology, 2012, 22(12):1649-1668.

[11] D. G. Dansereau, O. Pizarro and S. B. Williams, “Decoding, Calibration and Rectification for Lenselet-Based Plenoptic Cameras,” Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on, Portland, OR, 2013, pp. 1027-1034.

[12] Plenoptic images database, http://mmspg.epfl.ch/EPFL-light-field-image-dataset.

[13] Boyd, Stephen, and Lieven Vandenberghe. “Convex optimization.” Cambridge university press, 2004.

[14] Downloaded from: https://hevc.hhi.fraunhofer.de-/svn/svn_HEVCSoftware/tags/HM-16.9+SCM-8.0/.

[15] O. C. Au, X. Zhang, C. Pang, and X. Wen, “Suggested Common test conditions and software reference configurations for Screen Content Coding,” Joint Collaborative Team on Video Coding (JCT-VC),Torino, JCTVC-F696 , July, 2011.

[16] Bjontegaard, Gisle, “Calculation of average PSNR difference between RD-curves,” ITU-T VCEG-M33, 2001.

[17] Evaluation:http://mmspg.epfl.ch/files/content/sites/mmspl/files/shared/LF-GC/CFP.pdf.