Learning Based Compression for Real-Time Rendering of Surface Light Fields

Ehsan Miandji   Joel Kronander   Jonas Unger
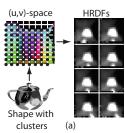
C-Research, Linköping University

Figure 1: *An illustration of surface light fields along with a visualization of clusters is shown in (a). A scene rendererd using our approach is shown in (b). Close ups of the teapot rendered using CPCA and our method, both with the same storage cost, are shown in (c) and (d), respectively. The scene was rendered at 52 fps using our method and 35 fps using CPCA.*

1 Introduction

Photo-realistic rendering in real-time is a key challenge in computer graphics. A number of techniques where the light transport in a scene is pre-computed, compressed and used for real-time image synthesis have been proposed, e.g. [Ramamoorthi 2009]. We extend on this idea and present a technique where the radiance distribution in a scene, including arbitrarily complex materials and light sources, is pre-computed and stored as surface light fields (SLF) at each surface. An SLF describes the full appearance of each surface in a scene as a 4D function over the spatial and angular domains.

An SLF is a complex data set with a large memory footprint often in the order of several GB per object in the scene. The key contribution in this work is a novel approach for compression of SLFs enabling real-time rendering of complex scenes. Our learning-based compression technique is based on exemplar orthogonal bases (EOB) [Gurumoorthy et al. 2010], and trains a compact dictionary of full-rank orthogonal basis pairs with sparse coefficients. Our results outperform the widely used CPCA method [Miandji et al. 2011] in terms of storage cost, visual quality and rendering speed. Compared to PRT techniques for real-time global illumination, our approach is limited to static scenes but can represent high frequency materials and any type of light source in a unified framework.

2 Learning based SLF compression

As described in Figure 1a, an SLF can be thought of as a set hemispherical radiance distribution functions (HRDF) regularly distributed over the surface. This can be described as a 3rd-order tensor $F \in R^{N \times m_1 \times m_2}$, where N is the number of spatial samples in (u,v)-space, and $m_1 \times m_2$ is the (angular) resolution of the HRDFs. EOB is based on training a set of $K \ll N$ full-rank orthogonal basis pairs (exemplars) such that projecting each data point onto *one* basis pair would lead to the most sparse coefficient matrix while minimising the $L2$-error. This can be expressed by minimizing the following energy function:

$$E(\{U_a, V_a, S_{ia}, M_{ia}\}) = \sum_{i=1}^{N} \sum_{a=1}^{K} M_{ia} \|H_i - U_a S_{ia} V_a^T\|^2$$

subject to

$$U_a^T U_a = V_a^T V_a = I, \ \forall a, \ \|S_{ia}\|_0 \leq T \text{ and } \sum_a M_{ia} = 1, \ \forall a,$$

where matrices $H_i \in R^{m_1 \times m_2}$ correspond to HRDFs; T represents the sparsity of the coefficient matrix $S_{ia} \in R^{m_1 \times m_2}$ and $M \in R^{N \times K}$ is a binary matrix associating each HRDF to its corresponding exemplar pair ($U_a \in R^{m_1 \times m_1}$ and $V_a \in R^{m_2 \times m_2}$). Using an iterative algorithm this can be efficiently minimized [Gurumoorthy et al. 2010], resulting in a set of exemplars (basis pairs)

U_a and V_a for $a = 1 \ldots K$. In order to best exploit the coherence in data, HRDFs are represented as matrices instead of vectors. To assist the convergence of EOB, we do a pre-clustering using K-Means to group similar HRDFs. Hence a different set of exemplars is trained for each cluster. Although this approach adds to memory footprint, it will improve accuracy and convergence of EOB.

For training the exemplars, we randomly select $20 - 70\%$ of the HRDFs with a probability distribution proportional to $\|H_i\|_2$. The value of T is fixed during training. Given trained exemplars, nullifying $m1m2 - T$ coefficients from the optimal projection matrix $S_{ia} = U_a^T H_i V_a$ results in a sparse coefficient matrix. This is done in a greedy manner by selecting an exemplar pair and incrementally adding the most significant coefficients until the reconstruction error falls below a threshold. Therefore the sparsity is different for each HRDF during the testing phase. The non-zero elements of S_i are stored as a vector of triplets with elements $S_{it}, t = 1 \ldots T_i$ containing indices and corresponding values. Hence our method produces a very compact dictionary with sparse coefficients adapted to the given SLF function.

3 Rendering and results

To reconstruct outgoing radiance at a point and along a view ray, we first find the corresponding cluster c and exemplar pair a for that point (denoted U_{ca} and V_{ca}). Due to uniform spatial sampling this can be done in $O(1)$. Our reconstruction method can be formulated as follows:

$$H_i(\xi_1, \xi_2) = \sum_{t=1}^{T_i} U_{ca}(S_{it}(1), \xi_1) \times S_{it}(3) \times V_{ca}(S_{it}(2), \xi_2)$$

where $H_i(\xi_1, \xi_2)$ is an element in the HRDF to be reconstructed. Figure 1b shows a scene rendered in real-time (52 fps using a Geforce 460) using our method including global illumination effects not possible to render with other real-time methods.

References

GURUMOORTHY, K., RAJWADE, A., BANERJEE, A., AND RANGARAJAN, A. 2010. A method for compact image representation using sparse matrix and tensor projections onto exemplar orthonormal bases. *Image Processing, IEEE Transactions on 19*, 2 (feb.), 322 –334.

MIANDJI, E., KRONANDER, J., AND UNGER, J. 2011. Geometry independent surface light fields for real time rendering of precomputed global illumination. In *SIGRAD 2011*, Linkping University Electronic Press, 27–34.

RAMAMOORTHI, R. 2009. Precomputation-based rendering. *Found. Trends. Comput. Graph. Vis. 3*, 4 (Apr.), 281–369.