**Chapter 1 Introduction**

This dissertation proposes the use of image based rendering with PCA to reconstruct the original image from the training data. This method is intended to save the rendering time and reduce the requirement of devices for complex rendering, such as volume rendering.

???

* 1. **Motivation**

In computer graphics, some of the objects could be rendered by polygons such as buildings, weapons, characters in video game and so on. However, some other requirements are needed to render the inner information for the object or the model is too hard to create as a polygon. Some of them could be rendered by volume rendering with 3D data sets. For example, the CT, MRI and engineering drawing which is used in medical and engineering visualization. Moreover, there might be some transparent objects could also be rendered by volume rendering, such as trees, fuzzy and gas.

With the rapid volume data expansion, the computation became more and more huge quickly. It will cost a lot to render the volume data directly on commodity desktop machine with interaction. Some devices even cannot support the large GPU memory. The lower computing power device would have worse result, such as in mobile devices.

In this case, some image based rendering methods were proposed to overcome the complex computing. Image based rendering has several advantages including speeding up the process and directly showing the result of volume rendering independent of the complexity of 3D datasets.

* 1. **Methodology**

The goal is to use PCA to reconstruct original images and use multithread programming to speed up the process of construction, so that we can test if the image based rendering for volume rendering could be used in real world application.

* 1. **Contribution**

Based on the original paper [1], I re-implement the PCA function in C++ and speed up the reconstruction process with OpenCL module in OpenCV. To reduce edges between cells, blurring was added onto the result images for better performance. Experiments were created to test the quality of reconstruction images, the memory required for reconstruction, the frame rate for real-time application and the relationship between these three variables. By changing the resolution of the same dataset, the cell dimension of each image and the number of components used in PCA, users could choose suitable parameters for this PCA image based rendering on different dataset.

* 1. **Summary of Chapters**

This dissertation is structured as follows:

* Chapter2 gives an overview of previous works done related to imaged based rendering, principle component analysis, volume rendering, and parallel programming. A detailed overview is given to main works related to this project, IBR (image based rendering) with PCA (principle component analysis) on volume rendering data.
* Chapter3 presents the design and implementations of the preprocessing, the CPU reconstruction, and the method of speeding up the reconstruction by using parallel programming.
* Chapter4 details the result gained from experiment. A discussion relating to the results is also presented.
* Chapter5 summarizes the project and provide a discussion of future work.

**Chapter 2 Background**

**2.1 Image based rendering**

In computer graphics, rendering is the process of generating an image from a 2D or 3D scene, which might contain several models. The result of this scene is related to lighting simulation, texture fidelity, shading information and viewpoint direction. However, when the scene is too complex or the subtle light effect in the real world is too difficult to simulate, traditional techniques have some limitation or the cost to simulate is too high. In this case, image-based rendering system was created to improve the realism of the 3D model and used to approximate global illumination effects as well.

The plenoptic function [9] is a function using pencil of rays’ angle of a given point, position for this given point, time dimension and intensity vanes with wavelength, to describe possible environment maps for a given scene from this given point in space. In Plenoptic Modeling [6], McMillan and Bishop discussed finding the disparity of each pixel in stereo pairs of cylindrical images. First, they used a cylindrical projection as the plenoptic sample representation instead of normal six planar projection onto a cube. Then, they reconstructed and resampled image warps that map sample images to arbitrary angle of cylindrical viewpoints by using plenoptic function.

Gortler and Grzeszczuk [7] used a 4D function called Lumigraph. This system is made for sampling the plenoptic function by handheld camera and use these light in Lumigraph. Instead of using 5 parameters to represent position and direction in plenoptic function, they proposed a way to reduce parameters to 4 by using 2D slices of 4D space. Direction is parameterized using two parallel planes (Figure 1). All the point can be represented by 4 coordinates in the Lumigraph. Then, they applied a discrete subdivision onto a 4D space and associate a coefficient and a basis function (reconstruction kernel) with each 4D grid point.

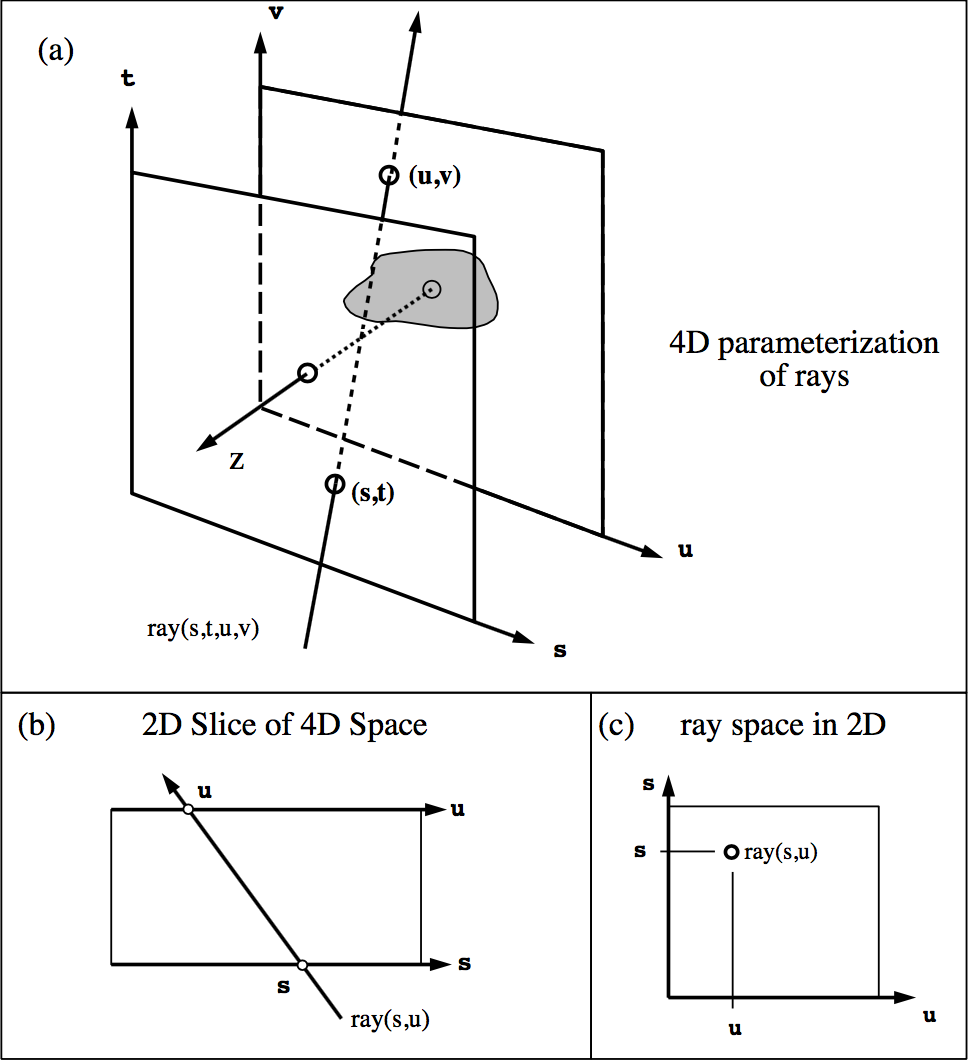


Figure Parameterization of the Lumigraph

Marc and Pat [8] propose a robust way named light field rendering to give more freedom view of image-based rendering. The main idea is to fill in regions of space free of occludes by creating a light field from a set of images corresponding to inserting each 2D slices into the 4D light field representation. Instead of using depth information, they simply combined and resampled available images. A new view could be shown in real time by extracting slices in appropriate direction if a light field is created.

Imposters or billboards are also a way which can be called image based rendering. The motivation for this technic is to reduce the complexity of rendering a 3D scene such as trees, grass, backgrounds in video game. Without losing resolution, 3D objects could be represented as a 2D image rotate with viewer’s observed direction. There’re two categories of imposters [11], static and dynamically-generated. Static imposter is usually created by artist and showed a right-angle like billboards. Dynamically-generated imposter is regenerated at runtime by rendering an image of a 3D object to a texture.

**2.2 Principle component analysis**

As discussed in 2.1, there are several techniques used in image based rendering. For our purpose, we intend to show the result image of a scene by reconstructing from the dataset of this scene. We intended to reconstruct corresponding image from user’s interaction, rather than rotate one image with different angle according to camera position, like billboards mentioned in 2.1.

In this case, Alakkari [1] proposed to use Principal component analysis (PCA), a statistical method, to compress the dataset and reconstruct images by back projecting process as a method to implement image based rendering. Principal component analysis [12] is a statistical procedure that uses an [orthogonal transformation](https://en.wikipedia.org/wiki/Orthogonal_transformation) to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables. The number of principal components is less than or equal to the smaller of the number of original variables or the number of observations.

Dimensionality reduction using PCA enables pose changes to be visualised as manifolds in low-dimensional subspaces and provides a useful mechanism for investigating face pose [13]. A reconstruction image could be generated by adding mean image of dataset and several number of highest eigentextures (sometimes called eigenfaces) multiplied with projected scores (figure 2).



Figure Projection process (Reconstruction) [5]

Ko Nishino [14] proposed the Eigen-Texture method to create a 3D model of an object from a sequence of range images. By aligning the pixel value from images into the 3D model, it can get a nice control of the mixed reality system. Although image based rendering is simple and handy for a stand-alone object without any background for the virtual reality, there’re some disadvantages, such as cast shadows under real illuminations, when combining these object into a real-world background. Our cell-based PCA process is similar with the generating virtual images process in this approach. The difference is that their result is summation of component virtual images sampled under single illuminations.

**2.3 Volume rendering**

Contrary to surface rendering, volume rendering (figure 3) is a technique to render 3D volume data to show the interior information of the object, usually used in medicine and some transparent objects. There are several ways for the 3D volume data visualization. The most basic volume rendering algorithm is ray-casting which uses straight-forward numerical evaluation of the volume rendering integral [15]. 2D texture [16] and 3D texture [17] mapping were used in these places. However, because of the large data size, it still cost a lot to render the volume data directly on standard hardware in real time. In this case, image-based rendering, as mentioned previously, is an alternative to deal with volume data.

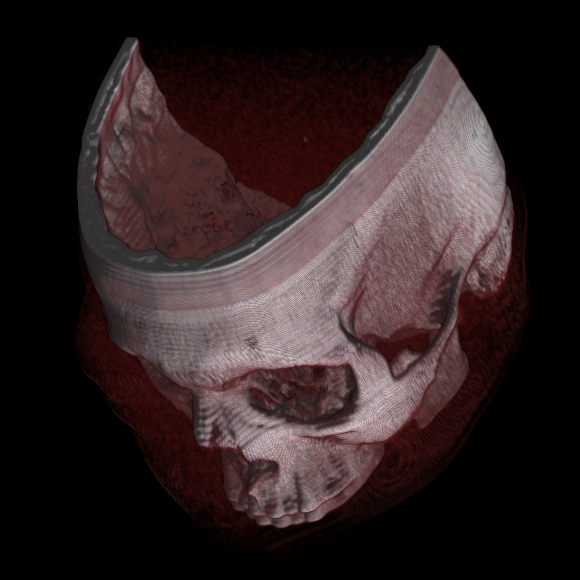


Figure A volume rendered cadaver head using view-aligned texture mapping and diffuse reflection [10]

Chen [18] presented a way to render surfaces from volume data by integrating the fully volume rendered image (the keyview) and the model of the volume. It was made by first constructing the surface model of the volume and texture mapping the keyview onto the geometry. Then use cast-ray to render the rest of new region when changing the view slightly so that they can save several computing processes for visualizing volume data.

Meyer [19] came up with an image-based volume rendering with opacity light fields. He separated the process to three parts, getting ray slices and proxy surface, generating key views and interactively rendering as opacity light fields. This approach can take advantage of the different features of the hardware, leave the heavier rendering work to the good graphics cards and combine the final result in remote application.

Tikhonova [20] used a representation of a multi-layered image, or an explorable image to show interior structure of the object, by simulating opacity changes and recoloring of individual features. From the single-view, a small number of rendered images were generated from 3D data set. Their technique could automatically extract multiple layers. By decompose the data into different layers corresponding to different structures in the data, users could change the opacity and colour of each layer to interact with models without re-rendering whole volume data. This can save GPU and CPU in devices, making it possible to run on low-end hardware.

**2.4 Parallel programming**

**Chapter 3 Design and Implementation**

**3.1 Basic PCA implementation**

In this section, I’ll explain the detail of the PCA implementation which is based on a PCA tutorial [3].

**3.1.1 Put the observed object into the matrix**

The observed object should have same dimension. For example, the observation we handle here are images, the training images should have the same dimension. Every image is put into one vector. If the image is dd dimension, the vector size is dd. If the training images number is n, the matrix with n rows and dd columns is formed as follow,

(1)

where represents the image’s all pixel values (For one channel),

**3.1.2 Subtract the mean**

To get the variance for each pixel, calculate average value in column, get the mean matrix, and use the original matrix to subtract this mean matrix. The adjusted data is calculated as below,

(2)

where is the result from equation (1), is the mean vector with dd elements.

**3.1.3 Calculate the covariance matrix**

Covariance is a similar way like standard deviation and variance to perform “how much the dimensions vary from the mean with respect to each other” [3]. It is always measured between two-dimensional data. For example, two vectors and with 10 elements in each vector, where is the mean of vector , is the mean of vector .

(3)

Covariance matrix is to measure multi-dimensional vector variance from their mean. For a matrix with n rows. The covariance matrix is calculated as below,

(4)

where is calculated by equation (2), is the transpose of the , n is the number of observed vectors. If the is n rows with dd columns, the result of covariance matrix should be dd rows with dd columns.

**3.1.4 Calculate the Eigenvectors and Eigenvalues**

Eigenvectors can be calculated by solving the equation as below

(5)

where  is an eigenvector of the linear transformation , the scale factor  is the eigen-value corresponding to that eigenvector and is the n by n identity matrix. [5]

In this place, the previous result which is calculated as covariance matrix is the matrix in equation (5). The number of eigenvectors and eigenvalues should be dd. Each eigenvector should have dd elements.

**3.1.5 Sort Eigenvalues**

After getting the eigenvectors and eigenvalues, the sorting process with highest to lowest order should be calculated by eigenvalues and sorting the corresponding eigenvectors. The highest eigenvector represents the highest variance for each dimension compared to mean vector. The second eigenvector represents the second highest variance for each dimension compared to mean vector, and so on.

**3.1.6 Choose the number of components and calculate scores**

While using all dd eigenvectors could reconstruct the same original observed vector, choosing some of the highest eigenvectors could also reconstruct the original vector with subtle data missing which means the more eigenvectors we used, the higher quality for the reconstructed vector. For an original observed vector, the projection value (referred to as a score) is calculated by project one eigenvector onto the original vector as below:

(6)

where is original vector, is the transpose of highest eigenvectors, and is the score of this eigenvector projected onto original vector, , .

**3.1.7 Reconstruction of the original image**

The final step is to reconstruct the original image by using the sum of product of some number of highest eigenvectors and corresponding scores projected on one original image. The equation is as follow,

(7)

where is the highest eigenvector of this original vector, is the corresponding score for this highest eigenvector, is the result of the reconstructed vector, , is the number we choose as components, .

**3.2 Small training data with CPU, image based rendering**

To build up a prototype of the reconstruction and save the preprocessing time, I used 36 training images, at a resolution of 300\*300 pixels. These images were saved screenshots by rotate the camera around a soldier [2] model’s head horizontally for 360 degrees (10 spacing for the azimuthal angle). At first, I directly used cell-based method to crop the whole image to cells instead of using the whole image as the observation. Like the previous experiment [1], I used the same cell-dimension 2020 to create an observed matrix with 36 rows and 400 columns. 2020 pixels values were put into one row (vector) which means each column represented one pixel value. In this case, there were 225 observed matrixes to save the whole images information for each channel.

**3.2.1 Crop to cell and put the image pixel value into the matrix**

The raw rbg image has three channels and the specific location’s pixel value is obtained by the corresponding width and height location. For OpenGL, this can be made by using SOIL library. Then, the image pixel value can be accessed by a char array. Eigen library [4] was used here for matrix calculation. 255 observed matrix is represented as below,

. . . (8)

**3.2.2 Calculate eigenvectors**

According to the 3.1.2, we can get the adjusted data for each matrix by subtract the mean vector in each row. Then, for the soldier head example I used previously, the covariance matrix is calculated as below:

(9)

where is adjusted data of , is the transpose of , is corresponding covariance matrix of , .

In Eigen library, the function could be used to calculate eigenvectors and eigenvalues by offered covariance matrix. Then, the bubble sorting method was used to sort eigenvalues and the corresponding eigenvectors were sorted as well.

**3.2.3 Reconstruct the original cell image and put cells together to a whole image**

After getting the ordered eigenvectors, we can project them onto every original image we want to reconstruct to obtain the corresponding scores. The result of these eigenvectors and scores were saved in files. Every time, instead of spending long time for preprocessing, we can read these eigenvectors and scores from files to reconstruct any original cell image. The final step is to put these reconstructed cells together back to the whole image. Three channels pixel values were put back to the char array for texture loading in OpenGL.

**3.3 Small training data with Cuda (GPU)**

**3.3.1 Parallel programming design for reconstruction**

GPU is powerful for multithread computing. However, there are several ways to complete a complex computing. For here, the input we got are eigenvectors and scores read from files. The output is a whole image pixel values. The most significant procedure inside was reconstruction for each cell image. And this procedure would be in a loop for number of cells time for reconstructing the whole image. CPU computing scudo-code is as below,

1: for all do

2: for all do

3:

4: end for

5:

6:

7: end for

is the number of cells, is the number of components we chosen, is the mean vector for each cell, is the function that resize reconstructed cell data from row to corresponding dimension and put the specific cell to corresponding position in a whole image.

Considering the same reconstruction process for one cell and the delay for transporting time between host memory and device memory, the best optimization is to use multithread programming in the process of reconstructing one cell. Cuda computing scudo-code is as below,

1: for all do

2:

3:

4:

5: end for

Everything was the same as CPU computing except that the reconstruction loop was replaced by a function.

**3.3.2 Detail of Cuda speeding up**

In this function, ???

**3.4 Large training data with OpenCV**

**3.4.1 Preprocess with OpenCV**

**3.4.2 Speed up reconstruction by OpenCL (GPU)**

**3.5 Experiment for real-world application**

**Chapter 4 Results and Evaluation**

The following chapter details the result gained from the implementation. Some experiment based on image quality, frame rate (reconstruction speed), and memory required are discussed and evaluated.

**4.1 Reconstruction of Images**

**4.1.1 Small training images results, reconstructed with CPU and GPU**

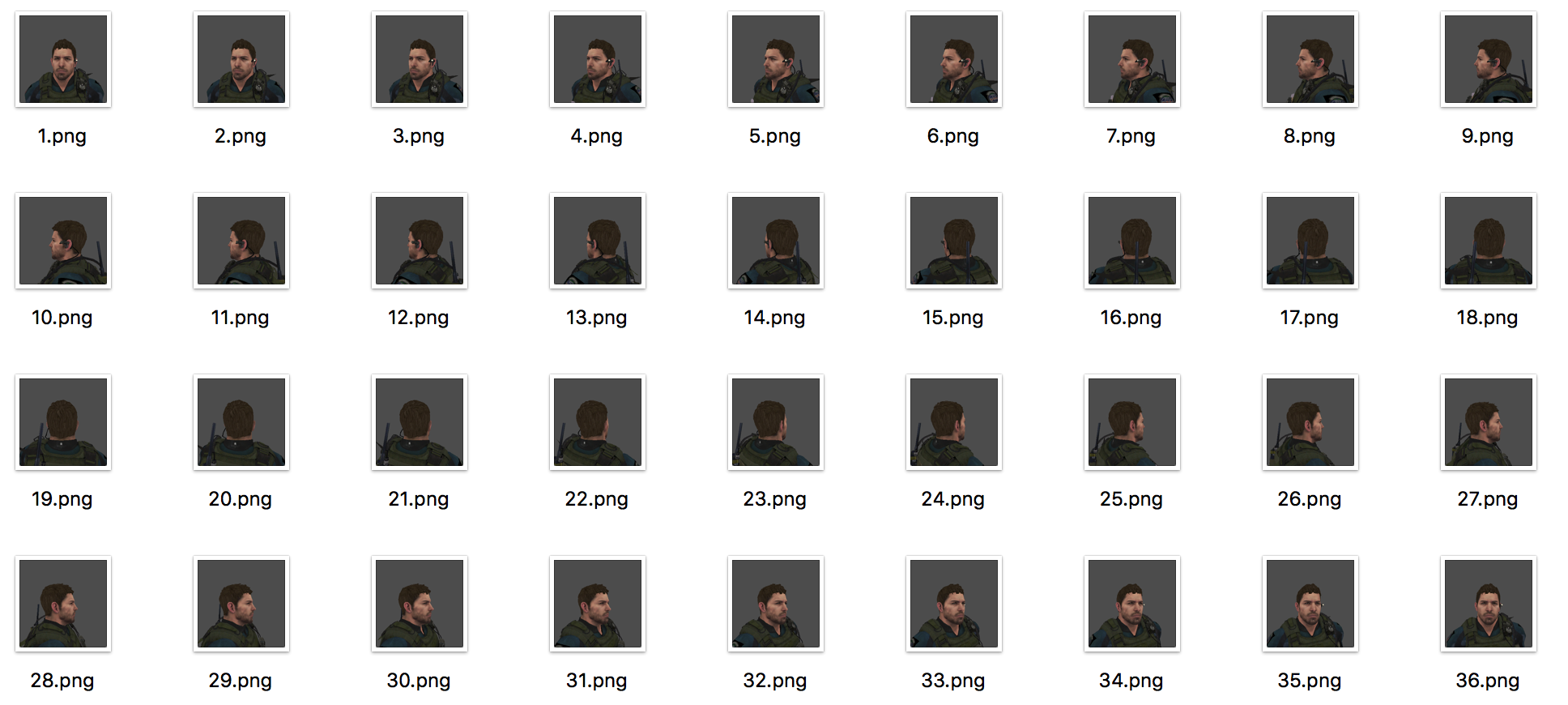


Figure Head training data with 36 images

Figure Left: reconstructed by 15 eigenvectors. Right: reconstructed by 20 eigenvectors.

Figure Left: reconstructed by 30 eigenvectors. Right: original image.

First implementation of PCA with CPU is based on solider head sample. From human eyes, the higher the eigenvectors used, the better the quality of the reconstruction compared to original image.

I didn’t focus on the relationship about result quality, number of components and frame rate in this place. Instead, the reconstruction process was focused at first. Calculation for reconstruction based on CPU was used as normal. The reconstruction time for the whole image was 10 seconds in devices with i7 core CPU, 2.3Hz, 2.4G memory? By using Cuda kernel function, the reconstruction time for the whole image was 2 seconds, 5 times faster than CPU computing.

Reference:

[1] Alakkari, S., & Dingliana, J. (2016, September). Volume visualization using principal component analysis. In Proceedings of the Eurographics Workshop on Visual Computing for Biology and Medicine (pp. 53-57). Eurographics Association.

[2] Soldier model link: <https://free3d.com/3d-model/chris-15987.html>

[3] Smith, L. I. (2002). A tutorial on principal components analysis. *Cornell University, USA*, *51*(52), 65.

[4] Eigen library: <http://eigen.tuxfamily.org/index.php?title=Main_Page>

[5] Eigenvector and eigenvalue: <https://en.wikipedia.org/wiki/Eigenvalues_and_eigenvectors>

[6] MCMILLAN, L., ANDBISHOP, G. Plenoptic modeling: An image-based rendering system. In Computer Graphics, Annual Conference Series, 1995, pp. 39–46.

[7] Gortler, S. J., Grzeszczuk, R., Szeliski, R., & Cohen, M. F. (1996, August). The lumigraph. In Proceedings of the 23rd annual conference on Computer graphics and interactive techniques (pp. 43-54). ACM.

[8] Levoy, M., & Hanrahan, P. (1996, August). Light field rendering. In Proceedings of the 23rd annual conference on Computer graphics and interactive techniques (pp. 31-42). ACM.

[9] Adelson, E. H., & Bergen, J. R. (1991). The plenoptic function and the elements of early vision.

[10] Volume rendering: <https://en.wikipedia.org/wiki/Volume_rendering>

[11] Dynamic 2D Imposters: <http://www.gamasutra.com/view/feature/130911/dynamic_2d_imposters_a_simple_.php>

[12] Principle components analysis: <https://en.wikipedia.org/wiki/Principal_component_analysis>

[13] GONG S., MCKENNA S., COLLINS J. J.: An investigation into face pose distributions. In *Automatic Face and Gesture Recognition, 1996., Proceedings of the Second International Conference on* (1996), IEEE, pp. 265–270

[14] NISHINO K., SATO Y., IKEUCHI K.: Eigen-texture method: Appearance compression based on 3D model. In *Computer Vision and Pat- tern Recognition, 1999. IEEE Computer Society Conference on.* (1999), vol. 1, IEEE.

[15] Engel, K., Hadwiger, M., Kniss, J., Rezk-Salama, C., & Weiskopf, D. (2006). *Real-time volume graphics*. CRC Press.

[16] Rezk-Salama, C., Engel, K., Bauer, M., Greiner, G., & Ertl, T. (2000, August). Interactive volume on standard PC graphics hardware using multi-textures and multi-stage rasterization. In *Proceedings of the ACM SIGGRAPH/EUROGRAPHICS workshop on Graphics hardware* (pp. 109-118). ACM.

[17] Wilson, O., Van Gelder, A., & Wilhelms, J. (1994). *Direct Volume Rendering Via 3D Terxtures*. Computer Research Laboratory [University of California, Santa Cruz].

[18] Chen, B., Kaufman, A., & Tang, Q. (2001). Image-based rendering of surfaces from volume data. In *Volume Graphics 2001* (pp. 279-295). Springer, Vienna.

[19] Meyer, M., Pfister, H., Hansen, C., Johnson, C., Meyer, M., Pfister, H., ... & Johnson, C. (2005). Image-based volume rendering with opacity light fields. *no. UUSCI-2005-002. Tech Report*.

[20] Tikhonova, A., Correa, C. D., & Ma, K. L. (2010, March). Explorable images for visualizing volume data. In *PacificVis* (pp. 177-184).

[21]