TRAINING INVERSE BRDF WITH INCOMPLETE DATA FOR 3D RECONSTRUCTION THROUGH PHOTOMETRIC STEREO

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

SAMET KILECI

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

SEPTEMBER 2014

Approval of the thesis:

TRAINING INVERSE BRDF WITH INCOMPLETE DATA FOR 3D RECONSTRUCTION THROUGH PHOTOMETRIC STEREO

submitted by **SAMET KILECI** in partial fulfillment of the requirements for the degree of **Master of Science in Electrical and Electronics Engineering Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen
Dean, Graduate School of **Natural and Applied Sciences** __________

Prof. Dr. Gönül Turhan Sayan
Head of Department, **Electrical and Electronics Engineering** __________

Prof. Dr. Uğur Halıcı
Supervisor, **Electrical and Electronics Engineering** __________

Examining Committee Members:

Prof. Dr. M. Kemal Leblebicioğlu
Electrical and Electronics Engineering Department, METU __________

Prof. Dr. Uğur Halıcı
Electrical and Electronics Engineering Department, METU __________

Prof. Dr. Yasemin Yardımcı Çetin
Graduate School of Informatics, METU __________

Assist. Prof. Dr. Tolga İnan
Electrical and Electronics Engineering Department, TEDU __________

Dr. Soner Büyükatalay
ASELSAN Inc. __________

Date: __________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: SAMET KILECI

Signature :

ABSTRACT

TRAINING INVERSE BRDF WITH INCOMPLETE DATA FOR 3D RECONSTRUCTION THROUGH PHOTOMETRIC STEREO

Kileci, Samet
M.S., Department of Electrical and Electronics Engineering
Supervisor : Prof. Dr. Uğur Halıcı

September 2014, 47 pages

In this thesis, missing data phenomena seen in a photometric stereo model is dealt with machine learning approaches. Photometric stereo model takes input images acquired with different illuminating conditions and predicts surface properties of an object. Specular regions appear on the images due to reflection for certain angle of light and camera and shadow regions appear because of surface structure of the object and light angle. Since specular and shadow regions degrade the performance of the photometric stereo, in this thesis these regions are handled as regions with missing data by using machine learning approaches. Neural network ensembles are implemented to handle the specular and shadow regions. Networks are trained with full range of BRDF data by omitting the values which have irrelevant intensity information. Once they are trained, test data is assigned to their adequate network by considering the location of missing data. This feature selection and ensemble structure of the networks significantly decrease the effect of missing data. Finally, outputs of each networks are used in the 3D reconstruction, surface structure of the object is successfully obtained with proposed photometric stereo model even in the presence of incomplete data.

Keywords: photometric stereo, missing data, inverse BRDF machine learning

ÖZ

FOTOMETRİK STEREO İLE 3B GERİÇATIM İÇİN EKSİK VERİ İLE TERS BRDF ÖĞRETİLMESİ

Kileci, Samet
Yüksek Lisans, Elektrik ve Elektronik Mühendisliği Bölümü
Tez Yöneticisi : Prof. Dr. Uğur Halıcı

Eylül 2014 , 47 sayfa

Eksik veri örüntü sınıflandırma ve tanıma sistemlerinde zayıflatıcı bir etkiye sahiptir. Bu tezde, fotometrik stereo (FS) modelinde karşılaşılan eksik veri sorunu makine öğrenme yaklaşımları ile çalışılmıştır. FS algoritması farklı aydınlatma koşullarında çekilmiş resimleri kullanıp, nesnenin yüzey yapısı özelliklerini tahmin eder. Görüntülerde belirli kamera ve ışık açılarında malzemenin ayna yansıması özelliği sebebiyle ortaya çıkan parlak bölgeler ve yüzeyin yapısından kaynaklanan gölgeli bölgeler bulunmaktadır. Parlak ve gölgeli bölgeler FS modelinin performansını düşürdüğünden bu bölgeler eksik veri olarak sınıflandırılmış ve bir makine öğrenmesi yöntemi olan yapay sinir ağları kullanılarak aşılmış ve performans artırılmıştır. Ağlar yansıma modelinden elde edilen parlaklık değerleri ile çok parlak ve çok karanlık değerler dışarıda bırakılarak eğitilmiştir. Test fazında, parlaklık vektöründeki anlamlı veriler kombine edilerek en uygun ağa yönlendirilmiş ve her bir ağın çıktısı ile objenin yüzey dikmeleri kestirilmiştir. Yapay sinir ağı gurubu oluşturulması ve parlaklık verilerinin işlenmesi yöntemleri harmanlanarak eksik veri probleminin zayıflatıcı etkisi bertaraf edilmiştir. Son olarak kestirilmiş yüzey dikmeleri kullanılarak nesnenin yüzeyi 3B geri çatımı eksik veriye rağmen başarıyla çalışılmıştır.

Anahtar Kelimeler: fotometrik stereo, eksik veri, ters BRDF, makine öğrenimi

To my wife

ACKNOWLEDGMENTS

I would like to express my greatest gratitude to Prof. Dr. Uğur Halıcı who supported, motivated and guided me throughout this research. I would like to thank ASELSAN Inc. for their support during all levels of this study. This work is also supported by TUBİTAK BİDEB Ms (2228) scholarship. The greatest thanks go to my wife and family for their endless support and encouragement in every part of this period.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

BRDF	Bidirectional Reflectance Distribution Functions
HEOM	Heterogeneous Euclidean Overlap Metric
PS	Photometric Stereo

CHAPTER 1

INTRODUCTION

Photometric stereo is the reconstruction method of surface structure with given images. Shape of an object can be derived from set of images taken in different illumination conditions with special techniques. Thus, quality of images used in such an algorithm directly affects the overall performance. In this thesis, a novel method for photometric stereo is proposed by considering the effect of specular regions placed in the mirror reflection positions. The specular regions in the input images are treated as incomplete data and methods for incomplete (or missing) data applied in photometric stereo.

In general context, missing data is a problem which can be faced in various types of data-driven processes in many areas from medicine to industry. Missing data is the incompleteness of important information and may occur because of the nature of the data itself or appear in acquisition stage of the data. Data – driven mechanisms require complete data in order to work properly. If the data is incomplete or corrupted, it may cause the system to give mistaken results. When the input data into the system is missing, the performance may degrade and lead to inappropriate or inaccurate results. By considering the importance of this case, the following techniques have been developed to cope with the problem to prevent the decrease of the efficiency of the proposed photometric stereo:

1. Deletion of missing features in the data
2. Imputation of missing data by using statistical and machine learning approaches.
3. Estimation of the missing data by using model-based approaches

4. Design of machine learning models which can work with missing data (Neural Network Ensembles)

These approaches have different perspectives. The first one is based on deleting samples with missing features from the data and it may lead to losing useful information as well, this is why it is an inadequate method to use. The second is based on substituting missing variables rather than deletion. In second class approaches, the system replaces the missing features with values obtained from some statistical tools by using the whole data set such as mean or median. In the third approach, a model such as expectation maximization (EM) is designed and most likely values are estimated. In the last approach, the system does not delete or replace the missing data, instead they are neural network ensembles which are designed to handle missing features without manipulating them. It is complicated to utilize in every system but it is the most useful one compared the previous approaches.

The focus of this study is implementation of a novel approach to photometric stereo algorithm in the presence of incomplete data. The inputs of the photometric stereo algorithm are images captured in different light angles provided that camera view is fixed. In this type of configuration, captured images may contain shadows and extra shiny regions because of surface properties of the material and light and view angles. Generally, in most calibrated image acquisition setup, three possible optical cases are encountered; self-shadowing, cast shadowing and specular reflection[1]. The regions where these effects are seen may actually provide important information about the structure of the surface. Hence, pixels to corresponding to these regions should be processed carefully to save the information or photometric stereo model should handle these corrupted areas while giving accurate estimations about the surface structure.

In this thesis, an inverse BRDF model is proposed for photometric stereo algorithm and missing data due to specular reflection is examined. Some machine learning techniques are applied to the problem to deal with the incomplete data.

Bidirectional reflectance distribution functions (BRDF) are used in several shape from reflectance algorithms in the literature. BRDF exhibits the close correlation between the surface orientation and the intensity values of the image. Thus with given lighting and imaging conditions an inverse BRDF can be trained in order to obtain surface orientation from intensities. Proposed method is inverse BRDF which is structured as a set of neural networks that is trained for a specific material . An ensemble of these networks are trained with group of pixels in training phase considering the incomplete features in the input vectors.

Networks are trained with the supervised learning method such as individual entries of input vectors carry the intensity information of a pixel and target vectors are corresponding normal angles given that view angle and the type of the material is fixed.

Before training phase, input data is produced with Bidirectional Reflectance Distribution Function (BRDF). The intensity information for a material with fixed lighting and camera direction is fetched from tabular BRDF data. Therefore, for training, 8 intensity values are obtained with 8 lighting angles and fixed viewing and normal angles. The networks are then trained with these input data and their corresponding normal angles as target vectors. During the tests, objects are synthetically illuminated from eight different angles. Eight synthetic images are then decomposed into the 1 by 8 vectors which each entry of a vector is a pixel illuminated from a certain angle and then fed into the inverse BRDF in order to obtain surface normal angles. Finally, surface normal angles are used in the global integration method in order to construct the surface of the object. The accuracy of the model clearly depends on two factors. One of them is training performance of the inverse BRDF which is related with neural networks. And the other one is the distribution of the specular regions on the images which actually degrades the performance of photometric stereo models in general. The specular regions in these input images can be handled with missing data approaches described above. Some of these approaches are implemented within the inverse BRDF in this work. Inverse BRDF then handlse the shiny regions as well with the described approaches. Performance of the inverse BRDF with the help of missing data handling methods are remarkably increased.

1.1 Organization

The rest of the thesis continues with the following chapters. In the second chapter, detailed background information about classification with missing data and photometric stereo is provided. In the third chapter, the implementation and the results of the study are presented. Finally, in the last chapter, conclusion and possible future works are explained.

CHAPTER 2

BACKGROUND AND RELATED WORK

2.1 Introduction

In this chapter, the definitions and background information about the work presented in this thesis will be explained. Photometric stereo with inverse BRDF with existence of incomplete data problem is divided into two basic sections. At first, the methods that deal with incomplete data in any recognition system will be presented. How incomplete data affects a neural network and which techniques were used in the literature will be explained. Then, photometric stereo methods in the literature will be presented. In that section, reflectance model BRDF will be presented and its close connection with proposed inverse BRDF model will be explained in detail. Definitions and the background information about the incomplete data and the literature review of approaches used throughout this thesis are given in the next section

2.2 Incomplete Data and Recognition Systems

Pattern recognition and classification systems are the mechanisms that evaluate available data to the system and then take actions based on the structure of the data and purpose of where being used[2]. Data structure is the main factor which directs the structure of the classification system. To illustrate, a speech recognition system proposed by Gaikwad et. al. [3] can be viewed as working in four stages; analysis of speech, feature extraction, modeling of the recognition system and testing. In modeling of the recognition system, there are a few approaches which can be categorized as

"Knowledge based approaches", "Statistical based approaches" and "Dynamic time warping" in the literature. The designer should select the appropriate approach to create a successful recognition system considering the structure of the speech signal as input. When considering the speed of the speech signal presented to the system, Dynamic Time Wrapping can be found as must pleasurable approach in order to cope with different speaking speeds. Hence, the structure strictly directs the design of a classification or recognition system.

Completeness is another important property of a data and should be considered in most cases while designing a classification system. Real-world applications suffer a common drawback, missing or unknown data[4]. Almost every automatic recognition system can encounter incomplete data. In sonar applications weather conditions affect the sea properties such as temperature and temperature deviation may cause degradation of receiving signal from an object. A video tracking system can encounter clusters such as clouds in front of tracked objects. Clustered video frame is an incomplete data for such a system. This type of missing data comes from the nature of application and can be considered in early design stages of recognition systems. In another case, a single or more sensors of a multiple sensor system can fail during an operation. For example, humidity sensor of a combined sensor group of a weather prediction tool can break down. This failure can be feeding humidity values to the system only once in a day instead of in every hour. In this case, the continuity of humidity data is corrupted and one can see the combined data is particularly missing in some instances of observation. Incomplete weather condition data through sensor set may generate incorrect prediction values upon this failure. In case of getting incomplete data after designing the classification system overall performance can be degraded dramatically. This is why the phenomena of missing data has been an important research area an there are very large studies in the literature for statistical analysis of missing data[4, 5, 6]. It is also examined in pattern recognition literature[4] When two disciplines are compared, statistical analysis requires more effort while pattern recognition gives direct results by using more complex structures.

The way the missing data problem is handled is closely related to how missing data occurs. According to Little and Rubin [7] there are three missing data mechanisms: Missing Completely at Random, Missing at Random and Missing Not At Random.

Missing Completely at Random is a situation in which the possibility of missing data is not associated with any features either missing or observed in the data set. In this condition, it is not probable to differentiate complete data from incomplete. Missing at Random is the condition where the cause of data missing is irrelevant to missing values but relevant to other variables. This ignorable situation is seen when it is possible to predict the pattern of the missing data from other variables which can be an external effect. Missing Not At Random is a non-ignorable case in which the missing data depends on the values that are missing. This case is the most difficult to predict from other variables.

Little and Rubin [7] stated that there are two ways of missing data patterns which are arbitrary and monotone missing patterns. In arbitrary missing pattern, it is probable to observe missing in any place and order of the variables does not have importance. In monotone pattern, the order of variables is important, they have a common order.

In this thesis, some parts of the input data is saturated due to reflectance property of the object. Thus, some numbers of pixels in the input images are treated as missing or incomplete data.

2.2.1 Approaches for Incomplete Data in Pattern Recognition

Pattern recognition with missing data can be seen as two parts of a problem, handling missing values in feature vector and designing an appropriate classification mechanism for this feature vector. In the literature, it has been said that most of the approaches can be grouped into four different types depending on how both problems are solved:

1. Deletion of incomplete cases and classifier design using only the complete data portion.
2. Imputation of missing data and using the edited set learning phase of the classification problem.
3. Use of machine learning procedures, where missing values are incorporated to the classifier.

4. Model-based procedures for treating missingness i.e. data distribution is modeled by expectation-maximization algorithm.

Figure 2.1 illustrates the four groups of approaches.

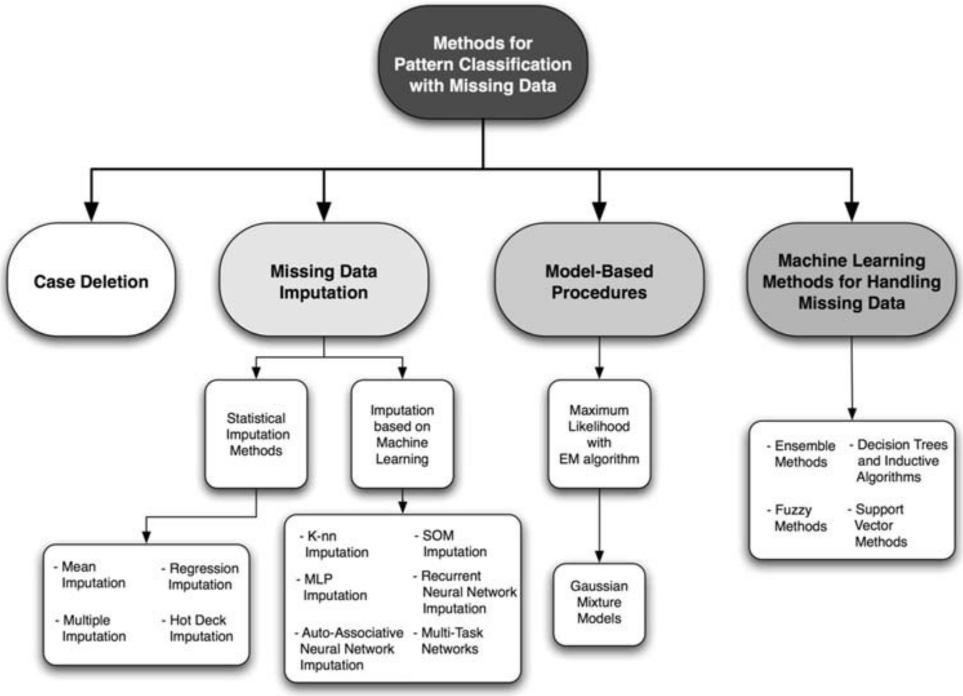


Figure 2.1: Classification with missing data approaches

2.2.1.1 Case Deletion Approach

At first, easiest and mostly common approach is deletion. Deletion based approach handles missing data separately from classification phase. Obtaining a complete dataset by deleting the features missing among all observations is a poor but easy approach. Dataset with missing features in some instances can be reorganized until it contains no missing value in it by just deleting the rows and columns in a simple heuristic. The percentage of missingness in a row or column can be a decision criteria which is to be deleted. With an algorithm which uses this type of criteria, dataset can be cleared from incomplete or missing values without making any assumptions. However, most of the useful data can be lost during deletion operation. Therefore, this approach seems a very poor method[8][9].

2.2.1.2 Imputation Approach

Next approach is named imputation which can be easily applied on the missing data throughout any recognition system. Imputation approach is widely used in the literature while treating the missingness problem as a part of classification problem. This approach can be divided into two subsequent methods considering the tools for implementation of imputation. First one is statistical imputation method based on statistical analysis tools for imputing. There are available options from mean imputation, regression imputation and hot and cold deck imputation[7, 5, 6]. Generally, these methods can be applied to one feature in an observation or multiple missing values within the feature vector.

Mean imputation is the first method as a statistical tool. Missing values are imputed with the mean or mode of rest of the samples among the dataset. Mean is used for continuous values and mode is used for discrete values of feature vectors. The main disadvantage of this method is that it ignores the variability of the data and the correlation between the various components of the data[5]. Farhangfar et al conducted a detailed study over imputation methods and finds mean imputation method least beneficial[10]. Mean imputation improves the classifier performance by at most 1 % for datasets containing significant amount of missing data[11, 12, 10, 13]. However, it can be used as pre-imputation method and combining other imputation methods.

Second imputation method approach is imputation based on machine learning algorithms such as K-nearest neighbor and auto associative neural networks. This approach models imputation mechanism based on available information on dataset [4].

K-nearest neighbor (K-nn) imputation method is actually a common hot deck method. K nearest neighbors are selected from non-missing samples by using a distance function [14]. The most similar samples are used to impute the missing value in selected missing valued feature vector. The existing values of donors are used to calculate the missing value. The calculation depends on data; the mode can be used for discrete and the mean can be used for continuous data type[4]. Closest neighbors' corresponding features can contribute to missing feature with a weight for each. The weights can be adjusted so that closer neighbors in the donors can contribute more. This approach improves the results[4]. The main effective factor in the K-nn imputation is the distance measure. Heterogeneous euclidean overlap metric (HEOM) is used[15, 14] and HEOM is the distance between a pair of vectors named X and Y and dimension of n;

$$D(X,Y) = \sqrt{\sum_{i=1}^{n} d_i(x_i,y_i)^2} \tag{2.1}$$

where $d_i(x_i, y_i)$ is the distance between X and Y on its i'th attribute and defined as:

$$d_i(x_i,y_i) = \begin{cases} 1, & \text{one of the } i\text{'th attributes of } x \text{ and } y \text{ is missing.} \\ d_D(x_i,y_i), & \text{if } i\text{'th attribute of } x \text{ and } y \text{ are discrete.} \\ d_C(x_i,y_i), & \text{if } i\text{'th attribute of } x \text{ and } y \text{ are continuous.} \end{cases} \tag{2.2}$$

The equation above states that distance between an unknown feature with complete vectors feature is 1 so that 1 is the maximum distance. For discrete features d_D assigns 0 if the attributes are the same or assigns 1 if they are different. Finally, for continuous attributes d_C assigns the distance;

$$d_C = \frac{|x_i - y_i|}{\max x_i - \min x_i} \tag{2.3}$$

which is normalized distance and $\max x_i$ and $\min x_i$ are maximum and minimum values of continuous attributes observed in the dataset. Batitsta and Monard [12]reveals that classification accuracy of K-nn imputation is good enough but only when missing features of the sample vectors are not highly correlated to each other. But this study is not performed on large datasets and each dataset has different amount of missing data hence study is not comprehensive in this aspect. Troyanskaya et al [16] compares the K-nn to mean imputation and SVD methods and states that the K-nn method is far better than the other methods.

2.2.1.3 Machine Learning Approaches and Neural Network Ensembles

The approaches described and explained in previous sections deal with missing data problem being apart from the classification problem. In the other words the methods; case deletion, imputation, machine learning and model based approaches handles with only missing data provided that classification task is another problem.

In this section,dealing with incomplete data problem and classification task are examined as a problem not just an imputation.

Neural network ensembles are proposed as methods for classification of incomplete data[17, 18, 19, 20]. Network reduction, a multiple MLP scheme is proposed by Sharpe and Solly[17]. Multiple MLP classifiers are designed so that each of the MLP is responsible to classify each incomplete data combination. Constructing several MLP networks based on combination of missing features deals with classification with incomplete data. However, constructing such a set requires more space and increases the time of the training phase. Krause and Polikar[18] proposed an ensemble of neural network structure which is trained with random subsets of features instead of each combination of missing attrributes. In this structure, each possible combination of input features is not to be guaranteed to be shown. Jian et al proposed another ensemble that uses combinations of complete dataset as inputs in training phase[19]. Unlike the method proposed in[18] this approach uses every possible combinations of complete dataset. Juszczak and Duin[20] developed another ensemble method in which large number of classifiers are trained on each feature so that when incomplete data shows up, classifier output corresponds that missing feature and ensemble makes decision with remaining classifiers.

An ensemble of neural network can deal with the input vectors whose features are randomly missing or incomplete.

2.2.2 Conclusion and Remarks

Neural network ensembles are suitable solutions for our problem domain. Input vectors with randomly missing or incomplete can be handled in their appropriate neural network in the ensemble. Outputs of each individual network then can be combined in order to create surface information of the object. On the other hand, division and classification of missing data in the input vectors are design problems that should be handled differently in training and testing phase. Thus, the structure of the ensemble is arranged in such a way that complete system will not mis-classify the input vectors. Both in the training and the testing phase, input vectors will be classified to their associated network. In the training phase, all possible input combinations are derived and used in each network. In the test phase, input vectors from the images is going to be assigned to its relevant network. Rest of the work is obtaining surface height (or depth) information from the surface normal angles. In the next sections, some detailed background information about photometric stereo will be presented.

2.3 Photometric Stereo

Photometric stereo, first proposed by Woodham[21] is an approach for the reconstruction of the image surface properties in computer domain. Main purpose of photometric stereo is to recover shape information from a set of gray level or RGB images differing only illumination conditions[22]. Draper and Pridmore states that if it is possible to model the reflectance properties of a surface and the position of the lisght source and camera are known in relation to the object then photometric stereo may recover a dense surface orientation map from gray level images[22].

In Photometric Stereo method, a fixed image is photographed with varying light sources and fixed camera position. The intensity variation in those images depends on the surface normals and the reflectance properties of the object. Photometric stereo uses this dependency in order to obtain surface normals from intensity variations[1]. Depending on the modeling of the reflectance different models can be proposed to solve the surface angle decoding. The keyword *reflectance property* was the main challenging concept in the literature. The surface reflectance of a material is formalized by the notion of the Bidirectional Reflectance Distribution Function (BRDF), which is a 4 dimensional function describing the response of a surface in a certain exitant direction to illumination from a certain incident direction over a hemisphere of directions[23]. In the literature, reflectance models are divided into two groups; analytic and data-driven models. Analytic reflectance models are the approximations of the reflectance characteristics of a surface. Those models are either just empirical formulations that are obtained without analyzing the materials' physical properties or the simplified equations based on physical properties of the material[24]. The weakness of these models is that they are only approximations of reflectance of real materials. Furthermore, most analytic reflectance models are usually limited to describing only particular subclasses of materials – a given reflectance model can represent only the phenomena for which it is designed[24].

Data driven models are constructed with acquired isotropic BRDF and machine learning techniques. Each entry in the BRDF table corresponds to a specific combination of angle of incoming light, view or camera angle and surface normal angle for a surface point. Acquired BRDF entries processed with both linear and non-linear dimensionality reduction techniques and hence tabulated and stored in an efficient way[25]. In the next subsection, the BRDF function and data structure are explained in more detail.

2.3.1 Bidirectional Reflectance Distribution Function

Mathematical representation of BRDF is a function of four variables: two variables specify the incoming light direction, two other variables specify the outgoing light direction. It is defined as a ratio of incoming irradiance $dE_i(\theta_i, \phi_i)$ to the outgoing radiance $dL_r(\theta_r, \phi_r)$;

$$f_r(\theta_i, \phi_i, \theta r, \phi_r) = \frac{dL_r(\theta_r, \phi_r)}{dE_i(\theta_i, \phi_i)} = \frac{dL_r(\theta_r, \phi_r)}{L_i(\theta_i, \phi_i) \cos\theta_i d\omega_r} \tag{2.4}$$

Figure 2.2 represents the coordinate system and the light directions for a unit surface dA .

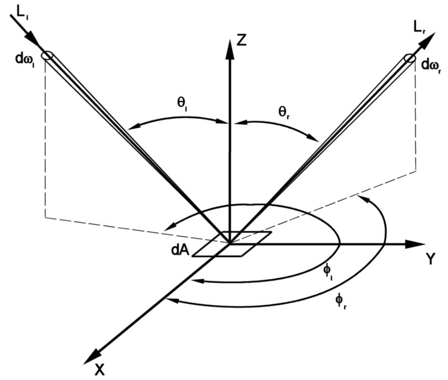


Figure 2.2: BRDF is a function of four variables: two variables specify the incoming light direction, two other variables specify the outgoing light direction. It is defined as a ratio of incoming irradiance to the outgoing radiance.

Isotropic BRDF is an important subclass of BRDFs. The isotropic model is valid for materials for which rotations about the surface normal can be ignored[25]. Isotropic

BRDF function, hence can be written in three variables which θ_r and θ_i are replaced with θ_{diff} . Isotropic BRDF is the function of ϕ_r, ϕ_i and ϕ_{diff} as following;

$$f_r(\theta_i, \theta r, \phi_{diff}) = \frac{dL_r(\theta_r, \phi_{diff})}{d_i(\theta_i, \phi_{diff})} = \frac{dL_r(\theta_r, \phi_{diff})}{L_i(\theta_i, \phi_{diff}) \cos\theta_i d\omega_r} \tag{2.5}$$

Matusik et al used another coordinate system proposed by Syzimon [26] and given in Figure 2.3 The change of variables used because Matusik states that specular peaks

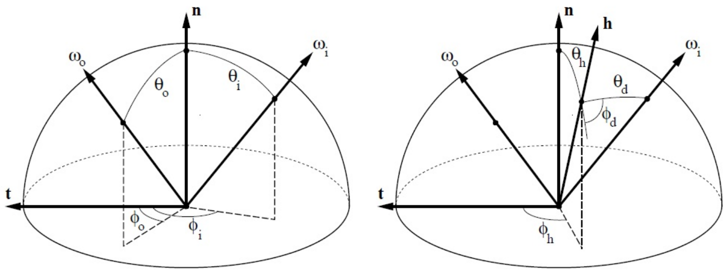


Figure 2.3: The standard coordinate frame is shown on the left. Rusinkiewicz's coordinate system is shown on the right.

were difficult to represent using the natural coordinate system. This new coordinate frame is based on the angles with respect to the half-angle (half-vector between incoming and outgoing directions). With this new coordinate system, sampling density can be varied and and increased near the specular highlight regions. This dataset includes 1,458,000 BRDF entries which includes 90 bins for θ_h and θ_d and 180 bins for ϕ_d. The angle ϕ_d is sampled 180 instead of 360 because of the reciprocity;

$$f(\theta_d, \theta_h, \phi_d) = f(\theta_d, \theta_h, \phi_d + \pi) \tag{2.6}$$

Matusik presents a sample image rendered with this BRDF model. The synthesized version of the acquired image is given in figure2.4.

A simple structure of such an image rendering model can be illustrated with the figure 2.5.

Returning to the photometric stereo case, the specular regions of the input images and shadowed parts are the challenging issues of this kind of systems. Buyukatalay et al proposed[1] an iterative algorithm to albedo and shadowing effects that uses masking of the input images. The performance measurement, total error in calculated surface

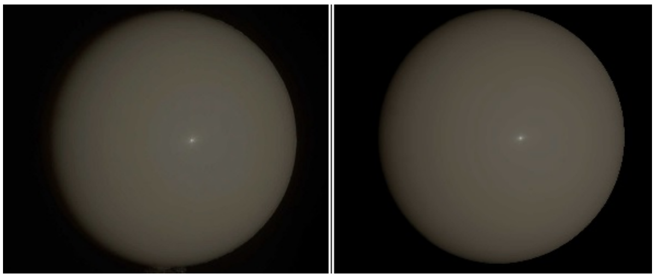


Figure 2.4: Two images of a sphere. Real image is shown on the left,synthesized image using tabulated BRDF data is shown on the right by Matusik et al.

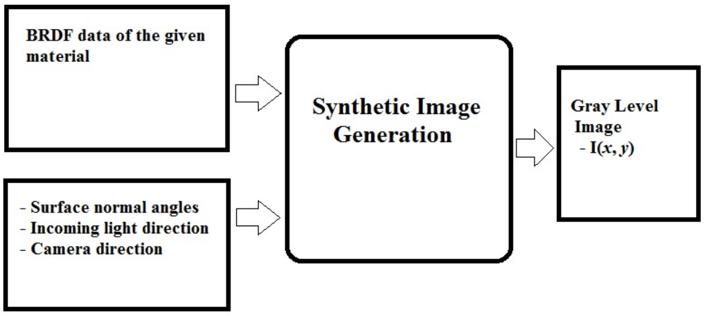


Figure 2.5: This figure illustrates the basic structure of image producing model using the BRDF data, illuminaton and observation information.

normals was reduced to 1.71% from 11.51% when their iterative photometric stereo algorithm was used.

2.4 Summary

There are plenty of methods for reproducing the image surface information from multiple images given that illumination and material type is known. Photometric stereo approaches use multiple images and analytic or data-driven BRDF information of material in order to obtain surface structure accurately. The specular and shadowed regions of the input images are kinds of a missing data to photometric stereo systems. Those areas should be treated specially by using missing data handling methods described in previous sections. We propose a neural network based method for obtaining surface normal angles from given images and illumination information. A set of networks will be trained with BRDF and surface information for a specific type of material and given lighting conditions. Then, those networks called inverse BRDF's are arranged in a way that those can handle incomplete data described in this chapter. In the next chapter, proposed neural network based inverse BRDF model for photometric approach will be explained in detail.

CHAPTER 3

INVERSE BRDF FOR 3D RECONSTRUCTION THROUGH PHOTOMETRIC STEREO

3.1 Introduction

In this chapter, the proposed model, inverse BRDF for 3D reconstruction will be explained. Reconstructing the 3D shapes through photometric stereo requires surface gradients or normal angles as described in the previous chapter.

Firstly, Inverse BRDF model that predicts normal angles of the surface with given multiple images of objects will be presented in detail. Inverse BRDF is a neural network that takes a set of intensity information of a pixel from multiple images and predicts the surface normal of that surface unit. Then, surface normals of the object are used in order to create surface height information as 3D reconstruction.

In the following sections, proposed inverse BRDF with incomplete data and 3D reconstruction through photometric stereo methods are explained.

3.2 Inverse BRDF With Incomplete Data

Inverse BRDF model is proposed in order to obtain the surface normal angles of the surface. In this work, inverse BRDF is designed as a neural network model that produce normal angle properties of surface units as an output. Eight monochrome images of an object is obtained by changing the lighting condition while camera angle is fixed. Those multiple images are fed into the inverse BRDF model as inputs. A

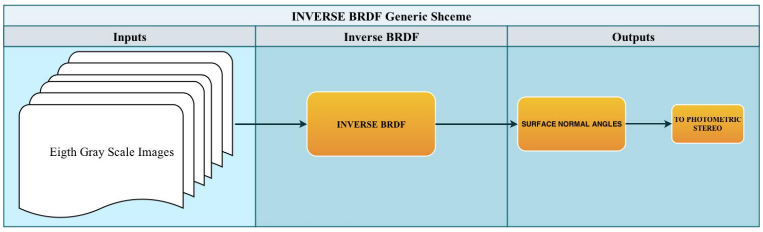


Figure 3.1: Generic Scheme of Inverse BRDF

generic scheme of the inverse BRDF model is given by the figure 3.1. Neural network structure in the inverse BRDF model is selected as feed forward neural networks and radial basis functions.

The networks are trained with input and target vectors by considering the incompleteness of the data.

Input data fed into the network in the training phase are one by eight intensity vectors.

$$\mathbf{I} = \begin{bmatrix} i_1 & i_2 & \cdots & i_8 \end{bmatrix} \tag{3.1}$$

where i_k is a intensity value obtained by a data driven BRDF model for a specific material. For any i_k, camera angle is fixed while light angle is changing in between eight directions. For example an intensity vector for a fixed material, **I**, is produced by fixing the camera angle, $\theta_c = 0, \phi_c = 0$, and the lighting angle is changing from, $[\theta_l = \frac{pi}{6}, \phi_l = 0]$ to $[\theta_l = \frac{pi}{6}, \phi_l = \frac{7pi}{4}]$, with incrementing ϕ_l by $\frac{pi}{4}$. Target data fed into the network with this input vector is,

$$\mathbf{n} = \begin{bmatrix} n_x & n_y & n_z \end{bmatrix} \tag{3.2}$$

where n_x , n_y and n_z are scalar components of the Cartesian representation of normal vector. Training data is constructed by scanning all possible surface normal angles and their corresponding intensity vectors.

In the training and test stages, networks are either trained by considering the incompleteness of train data or tested with pre-processed input data. Imputation and neural network ensemble models are designed and tested with some type of synthetic surfaces. Following subsections describes those designs.

3.2.1 Inverse BRDF with Imputation Methods

Imputation methods detect and remove the incomplete data from the input vectors in the test phase and replace those entries with relevant values in order to improve network performance for the pixels that is marked as incomplete.

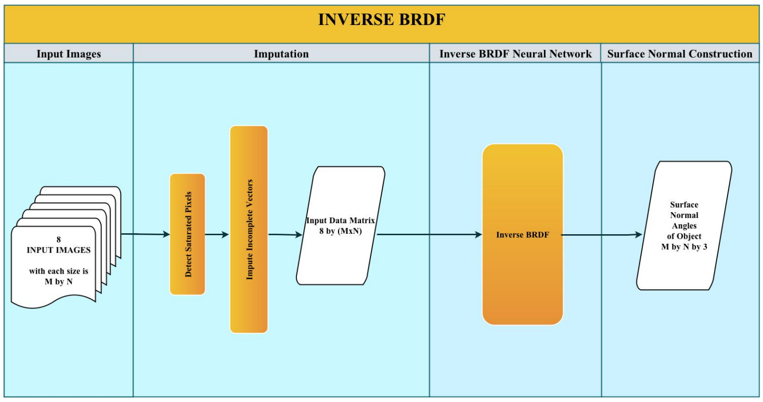


Figure 3.2: Inverse BRDF Model with imputation approaches integrated

Block diagram in the figure 3.2 represents the Inverse BRDF model with imputation methods described in previous chapter.

Imputation is applied to the train and test data for the entries in input vectors which are the specular regions in the images. Those entries are detected so that the intensity level of a pixel increases a threshold T_s then it is marked as missing. The threshold T_s is calculated for each image and depending on the standard deviation of the intensity profile. Figure 3.3 illustrates the intensity variation from a slice of given image. The

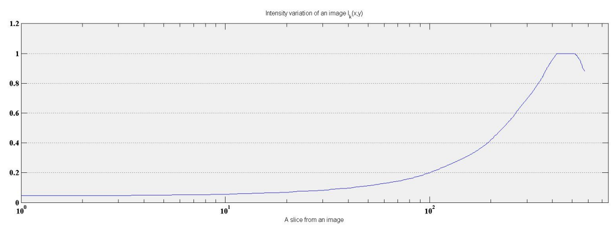


Figure 3.3: Intensity variation of a slice from semisphere image

pixels which exceed the threshold are marked as missing. From the figure 3.3 it is seen that a significant amount of pixels are actually saturated thus will be marked as missing. After pixels that corresponding intensity levels are upper than a threshold is found they are imputed with following methods:

- Mean Imputation
- K-nn Imputation

Following sections explains each imputation method.

3.2.1.1 Mean Imputation

Mean imputation is a very basic statistical tool for imputing the missing features in the data. Marked pixels are imputed with either their sample average corresponds to the average of intensity of 8 images or average of instance that corresponding to the mean of the intensity of that image. For an image sequence $I_k(x, y)$ $k = 1, 2, 3, ...8$; test data is constructed as following matrix **I**;

$$\mathbf{I} = \begin{matrix} I_1(1,1) & I_2(1,1) & \cdots & I_8(1,1) \\ I_1(2,1) & I_2(2,1) & \cdots & I_8(2,1) \\ \vdots & \vdots & \ddots & \vdots \\ I_1(m,n) & I_2(m,n) & \cdots & I_8(m,n) \end{matrix} \tag{3.3}$$

where m and n are the image dimensions. And one sample S for the Inverse BRDF model is a row of **I**, which is actually a pixel's intensity value, calculated with respect to the light directions and objects normal. For example if the pixel of the $I_j(u, v)$ marked as missing; $m_j(u, v)$, the missing value is imputed with its sample average;

$$m_j(u,v) = \frac{\sum\limits_{k=1 \& k \neq j}^{8} I_k(u,v)}{7} \tag{3.4}$$

or its instance average which corresponds to mean of the image;

$$m_j(u,v) = \frac{\sum\limits_{x,y=1,1}^{x,y=m,n} I_j(x,y)}{(m-1) \times (n-1)} \tag{3.5}$$

provided that summation term in both equations 3.5 and 3.4 does not contain missing value, if it does then can be taken as 0.

3.2.1.2 K-nn Imputation

K-nn imputation is another tool based on machine learning approaches. The missing value $m_j(u, v)$ is imputed with the contribution of its neighbors so that K nearest neighbors are found with HEOM distance by using equation 2.1. Particularly, 5 neighbors contributes to the missing feature k of the sample S defined as $S(k)$

$$S(k) = S(n_1)(k)\times w_1 + S(n_2)(k)\times w_2 + S(n_3)(k)\times w_3 + S(n_4)(k)\times w_4 + S(n_1)(k)\times w_5 \tag{3.6}$$

where $S(n_1)(k)$ is the closest neighbor of the incomplete sample S and it is 1 by 8 vector;

$$S(n_1)(k) = \begin{bmatrix} I_1(x,y) & I_2(x,y) & \cdots & I_8(x,y) \end{bmatrix} \tag{3.7}$$

and weight vector **w** is chosen as sum of all weights are 1 and w_1, the closest neighbors weight is higher than the others.

$$\mathbf{w} = \begin{bmatrix} w_1 & w_2 & \cdots & w_5 \end{bmatrix} \tag{3.8}$$

3.2.1.3 Training and Testing Inverse BRDF With Imputation Methods

Inverse BRDF is modeled as neural network scheme which consists of multi layer perceptrons (MLP). MLP, as feed forward neural network, is designed as 8 input layer neurons, 40 hidden layer neurons and 3 output layer neurons. Figure 3.4 illustrates the structure of the network. In the training phase 8 Input layer neurons take input of one by eight sized intensity vector from the input side train data matrix , **TD**. In each step, at the output side, corresponding one by three sized surface normal vectors from the target data matrix, **N** is shown to the output layer neurons. Network is trained until a satisfactory mean square error is obtained. There are approximately from 100 to 200 number of epochs done in the training phase.

In the training dataset the vector entries are actually the BRDF entries of Matusik's data set. Each feature in a sample of this set corresponds to an intensity value of a surface normal when 8 different light angles are considered. Namely,

$$\mathbf{TD} = \begin{bmatrix} TD_1(1,1) & TD_2(1,1) & \cdots & TD_8(1,1) \\ TD_1(2,1) & TD_2(2,1) & \cdots & TD_8(2,1) \\ \vdots & \vdots & \ddots & \vdots \\ TD_1(m,n) & TD_2(m,n) & \cdots & TD_8(m,n) \end{bmatrix} \tag{3.9}$$

where TD is the train data corresponding to the surface normal angle w in spherical coordinates. m is 1 deg bins from 0 deg to 90 deg as zenith angle and n is 1 deg

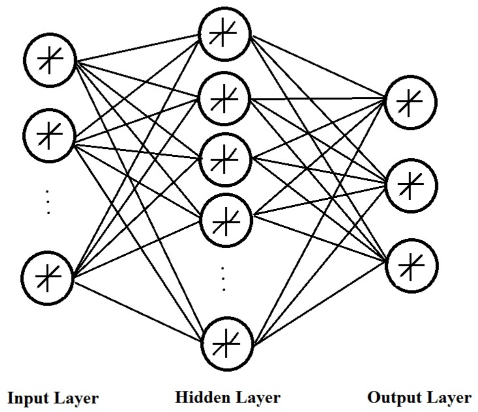


Figure 3.4: MLP consists of 8 input, 40 hidden layer and 3 output neurons.

bins from 0 deg to 360 deg as zenith angle. The target data is the surface normals in Cartesian coordinates has following structure;

$$\mathbf{N} = \begin{matrix} n_x(1,1) & n_y(1,1) & n_z(1,1) \\ n_x(2,1) & n_y(2,1) & n_z(2,1) \\ \vdots & \vdots & \vdots \\ n_x(m,n) & n_y(m,n) & n_z(m,n) \end{matrix} \tag{3.10}$$

with the same convention for **TD** .

In the test phase input data is imputed with the methods given at the previous sections. The output is predicted normal angles $\tilde{\mathbf{N}}$ with following convention;

$$\tilde{\mathbf{N}} = \begin{matrix} n_x(1,1) & n_y(1,1) & n_z(1,1) \\ n_x(2,1) & n_y(2,1) & n_z(2,1) \\ \vdots & \vdots & \vdots \\ n_x(m,n) & n_y(m,n) & n_z(m,n) \end{matrix} \tag{3.11}$$

Predicted normal angles are then used in the 3D Reconstruction phase in order to create height information of the object. In addition, these estimated normal angles are used in error calculations in order to measure the performance of the Inverse BRDF system.

3.2.2 Inverse BRDF With Neural Network Ensembles

The neural network ensembles are structures of interconnected multiple neural networks. Those ensembles are designed in order to handle special cases in recognition process such as incomplete data.

Considering the incompleteness of the input data, multiple networks are arranged so that each of the network should efficiently recognize each incomplete data combination. Moreover, the network ensemble should handle all possible combinations of incomplete data in order to manage a good performance without any loss of input data. Thus, an inverse BRDF with multiple networks should be designed and trained with such a special method in order to meet performance criteria. Following subsections exhibits the design of the input data handling methods and neural network structures of the inverse BRDF.

3.2.2.1 Training Inverse BRDF With Neural Network Ensembles

When the input train data is analyzed it is seen that many of the samples **I** contain saturated pixels due to specular regions. For instance let **I** is a vector, randomly chosen in the train data **TD**, and let $[i_1, i_2, ..., i_8]$ are the intensities of this pixel. As an example let i_6 have the local maximum or it is saturated, i.e. its gray level is 255. This intensity level in the input vector does not carry any useful information for network. Hence it should not be used in training of a network in the structure. This feature is missing because the surface normal for that pixel is too close to the half vector of camera angle θ_c, ϕ_c and light angle θ_{l6} and ϕ_{l6} and light reflected from that pixel is too much when compared its neighborhoods. Figure 3.5 illustrates that input vector.

Those neighbors actually come from the images illuminated from the angles:

- θ_{l4} and ϕ_{l4}
- θ_{l5} and ϕ_{l5}
- θ_{l7} and ϕ_{l7}
- θ_{l8} and ϕ_{l8}

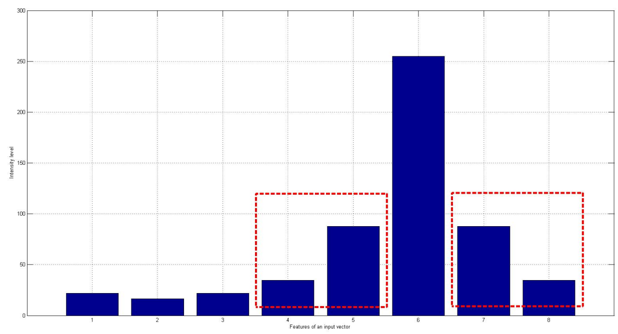


Figure 3.5: An example input vector of inverse BRDF. Note that feature i6 is saturated. Then, it is omitted and its left and right 2 neighborhoods are extracted to create new more useful input data.

A new vector with this extracted features, called **I'**, which consist of;

$$\mathbf{I'} = \begin{bmatrix} i_4 & i_5 & i_7 & i_8 \end{bmatrix} \tag{3.12}$$

This vector is then classified as the element of the "Network 2". The vector **I'** is then fed into the "Network 2" with its target vector **N'**.

Once all train data **TD** is scanned with this method, all input and target vector pairs are assigned to their network. The neural network ensemble is designed in a way that if an intensity is saturated on any vector, then it is not being used and its left two and right two neighbors are extracted from the vector to create a new input vector. The intensity information of those entries have certainly more useful information for the inverse BRDF by considering the incoming light and camera angles.

In the training stage, the networks in the ensemble are trained by using these pre-processed input-target pairs Following table 3.1 illustrates this idea in more general;

In the test phase, similar classification for individual vectors are applied and their outputs are combined at the end of the line. Feed forward multi layer perceptrons and radial basis functions are trained with this input vectors. In the next section, testing phase of the inverse BRDF is described.

Table 3.1: Table shows the most frequently selected features among train data vectors during input vector classification stage.

Img/Net	Net. 1	Net. 2	Net. 3	Net. 4	Net. 5	Net. 6	Net. 7	Net. 8
$I_1(x, y)$		✓	✓				✓	✓
$I_2(x, y)$	✓		✓	✓				✓
$I_3(x, y)$	✓	✓		✓	✓			
$I_4(x, y)$		✓	✓		✓	✓		
$I_5(x, y)$			✓	✓		✓	✓	
$I_6(x, y)$				✓	✓		✓	✓
$I_7(x, y)$	✓				✓	✓		✓
$I_8(x, y)$	✓	✓				✓	✓	

3.2.2.2 Testing Inverse BRDF With Neural Network Ensembles

Figure 3.6 illustrates the proposed inverse BRDF model in block diagram.

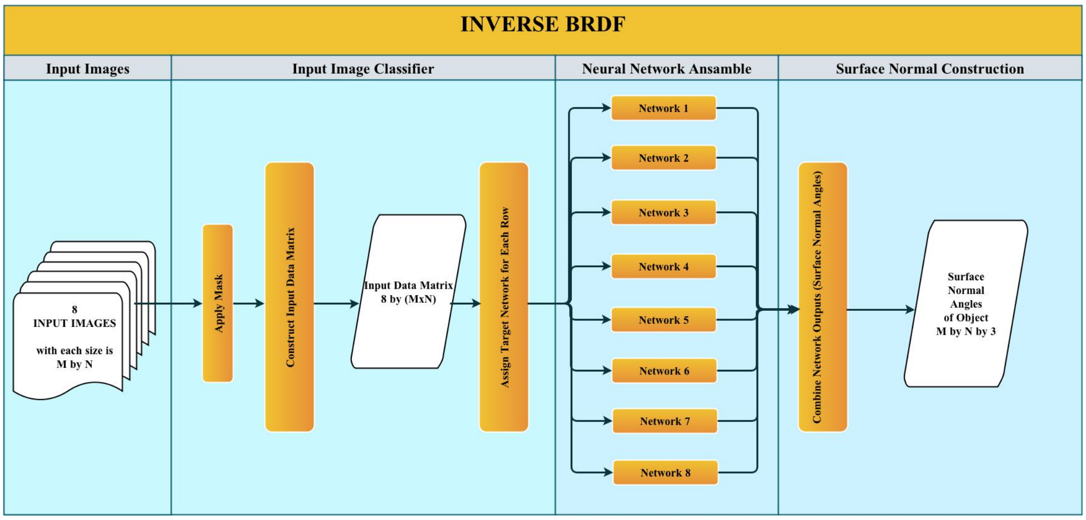


Figure 3.6: Inverse BRDF Model with Neural Network Ensemble

The proposed model has three internal mechanisms; first block for classifying the input data on line in testing mode after applying masks for objects in the images. Second one shows independent neural network models that take their inputs and produce predicted surface normals with respect to the given set of pixels. Lastly estimated surface normals obtained from each independent networks are combined together in order to create single surface normal map of the object.

First internal mechanism of compact structure arranges the input data into the groups so that the each group matches with its dedicated network. In the test stage, eight images of an object are fed into the first block as a matrix of intensity data. This matrix has eight colons, each one is the successive intensity values of a surface unit. As described before each feature comes from an image illuminated from specific angle of light. Thus a row in this matrix carries the intensity information from eight directions.

Firstly, data matrix is marked with a mask to extract the pixels that belong to the image. Then, pixels of interest are fed into the input vector classifier. After the classification every pixel in the object is assigned to its relevant network. Figure 3.7 shows the location of the pixels assigned to the networks.

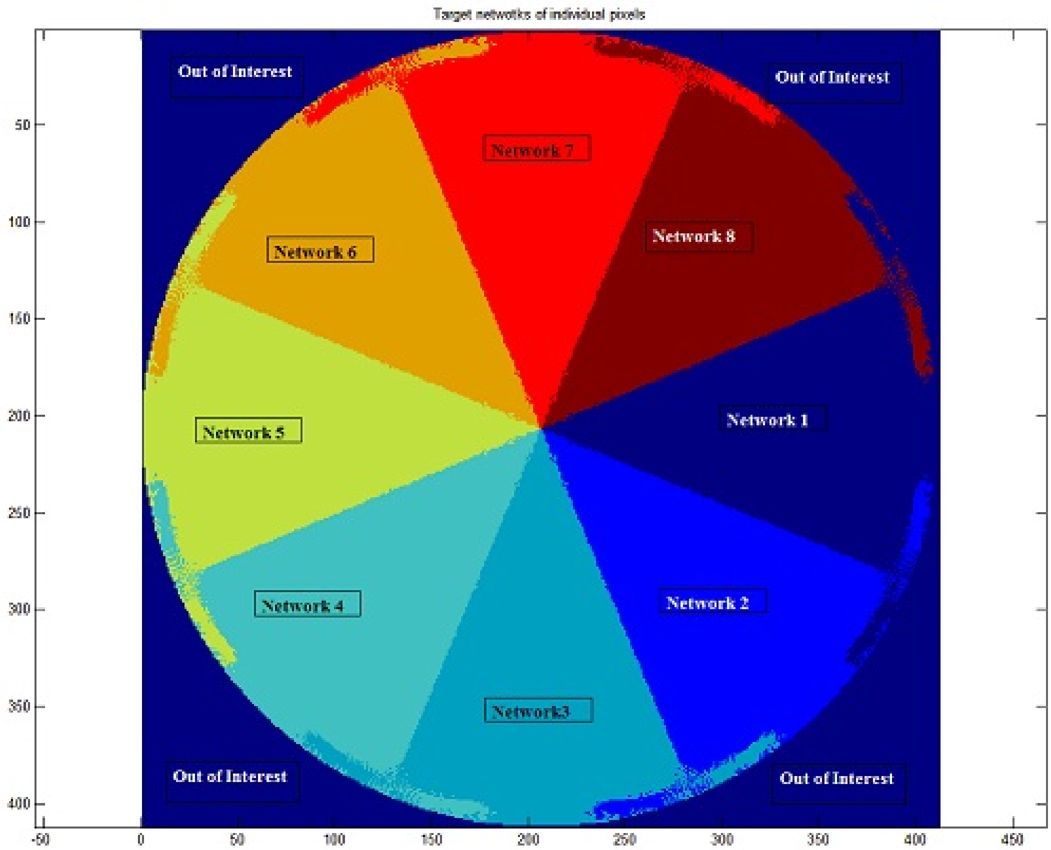


Figure 3.7: Pixels of an object assigned to the networks.

Lastly, network outputs, estimated surface normals are combined in a single matrix as an output of Inverse BRDF. The estimated surface normals from inverse BRDF are then compared with the original normal map to justify the proposed models performance. In the next section, 3D reconstruction with the estimated surface normals will be explained.

3.3 3D Reconstruction With Estimated Surface Normals

3D shape reconstruction can be defined as determining the depth or height information from the gradient field of an object. Proposed photometric stereo model has two internal steps. First step is estimating the surface normal at each surface units. Estimated surface normals are then used in 3D reconstruction phase in order to obtain height information of the object. In this work, the method of convolution of gradient field over the surface is used. Convolution is made in the frequency domain. In the first step, gradient field of the object is calculated from surface normal components. Then the Fourier transforms are obtained by using Fast Fourier Transform (FFT) algorithm. After that, multiplication in Fourier domain is calculated, that is actually yields the height information in Fourier domain. Also Lastly, inverse Fourier transform over the image is applied. Resulting data is normalized and hence surface depth (or height) construction is completed. The results are compared with the ground truth height data and performance of the overall design is measured.

In the following chapter, experimental results with the synthetic images are given.

CHAPTER 4

EXPERIMENTAL RESULTS

The proposed inverse BRDF models were tested by using synthetic images. A semi-sphere, a sombrero and the Mozart's face were used as shapes for synthetic images. Multiple images rendered with eight different light conditions were given into the inverse BRDF and their shape was reconstructed. In each steps outputs were compared with ground truth and performance matrices were obtained. The performance of inverse BRDF has been measured with the following criterions;

- Normal Error
- Intensity Error
- Height Error

Following sections defines these three criterions.

4.1 Normal Error

Normal error between the ground truth normal matrix and the estimated normal matrix were calculated with dot product of two normal vector matrix;

$$E_n = < n_{estimated} \cdot n_{groundtruth} > \tag{4.1}$$

where $n_{estimated}$ is the estimated normal matrix, output of the model and $n_{groundtruth}$ is the matrix contains original surface normals.

This error function yields 1 for an entry if two of the normal vector are exactly same and stands for zero error between two vectors. Similarly the output will be -1 if two of the normal vectors are just opposite to each other that represents maximum error. So the E_n between any two vector will be in the range of $[-1, 1]$. In order to obtain an abstract result, 1 is substracted from the entries E_n which changes the boundaries to $[-2, 0]$ which -2 represents maximum error and 0 represents minimum error. Additionally if absolute value of the error is taken, error boundaries will be $[0, 2]$ which will be more meaningful. Then, averaging the result over the number of pixels was used as error metrics. Finally, percentage of the result was presented for each result. The error was defined in equations as followings;

$$\mathbf{E(N)} = \frac{\sum |E_n - 1|}{No.Of Pixels} \tag{4.2}$$

recalling E_n as

$$E_n =< n_{estimated} \cdot n_{groundtruth} > \tag{4.3}$$

4.2 Intensity Error

Intensity error was calculated by the difference at the intensities between the ground truth images and the rendered images with the estimated normals. The intensity error metric was calculated with the following equation:

$$\mathbf{E(I)} = \frac{\sum |I_e - I_g|}{No.Of Pixels} \tag{4.4}$$

where I_e is the estimated monochrome image and I_g is the ground truth image.

4.3 Height Error

Height error was calculated by the difference at the depth or height between the ground truth surface and the reconstructed surface. The equation for the height error metric is the following:

$$\mathbf{E(H)} = \frac{\sum |H_e - H_g|}{No.Of Pixels} \tag{4.5}$$

where H_e is the estimated height data and H_g is the ground truth height data.

Experimental results are given in the next section with supported figures and images.

4.4 Performance Of Inverse BRDF

Inverse BRDF with Neural network ensemble provided very good results. Three shapes with three different materials BRDF were used in synthetic image generation for testing the network. Surface normals of each surface were calculated from ground truth height data. The synthetic images were rendered using the "sphere", "sombrero" and "Mozart's" face with using ground truth surface normals. In addition, at the rendering phase, three different materials BRDF data were used. First one was called "Red Specular Plastic" which exhibited very specular property. Second one was the "Black Soft Plastic" whose reflection property was in the middle of very specular and very diffuse. Last material was the "Red Plastic" that was very diffuse material in reflection.

The image sets rendered with given BRDFs were then fed into the inverse BRDF with feed forward neural network and radial basis function network. At the first step, their surface normals were estimated with inverse BRDFs. Estimated surface normals were used in the normal error calculation. Then, images for each estimated surface normals were rendered again and intensity difference between the ground truth and the rendered images was calculated. After that, height information was obtained with estimated surface normals by using the presented algorithm as 3D reconstruction. Finally, height error was calculated for each image set that represented different shapes and materials.

4.4.1 Error Metrics for Inverse BRDF

Following tables show the numerical performance measurements for three material types and for each network type.

Recalling from previous sections, there are three error measurements for specific outputs of the model. First one is normal error, **E(N)**, defined as error between the ground truth surface normals and the estimated surface normals. Second one is average intensity error, **E(I)**, for reproduced images from normal angles and it is the difference between those and original input images. Last one is height error, **E(H)** which is the height difference between the original object and 3D recunstructed shape.

In this thesis three surface types, semi-sphere, sombrero and Mozart's face were worked. Two types of neural network structures were implemented and their performance was measured. First one is feed forward neural networks, the second one is neural networks with radial basis functions.

There are three tables given in this section, first table shows the results for the most specular reflective material, Red Specular Plastic. The second table represents the similar results for the moderate reflective material, Black Soft Plastic. At last, the results were exhibited for very diffusive surface Red Plastic.

For all tables, first row indicates the material type. Second row stands for surface types. At the third row, abbreviated network types are placed. FF is the feed forward neural networks, RB stands for neural network with radial basis functions. At the left colon error metrics are shown for every combination of surface types and network types. Table 4.1 represents the results for the material red specular plastic, which represents very specular property.

Table 4.1: Performance Of The Inverse BRDF for a specular material, Red Specular Plastic

Error Metric	Red Specular Plastic					
	Semi Sphere		Sombrero		Mozart's Face	
	FF	RB	FF	RB	FF	RB
E(N), %	2.59×10^{-1}	3.22×10^{-1}	4.81×10^{-3}	1.51×10^{-3}	1.58	1.58
E(I), %	2.53×10^{-2}	3.36×10^{-2}	1.02	0.99	0.014	0.147
E(H), %	4.73	4.93	1.46×10^{-2}	2.51×10^{-1}	5.79	5.78

Table 4.2 represents the results for the material black soft plastic, which represents moderate specular property of reflection.

Table 4.2: Performance Of The Inverse BRDF for, Black Soft Plastic

Error Metric	Black Soft Plastic					
	Semi Sphere		Sombrero		Mozart's Face	
	FF	RB	FF	RB	FF	RB
E(N), %	0.27	0.33	6.12×10^{-3}	1.13×10^{-2}	1.58	1.58
E(I), %	0.22	0.31	0.17	0.29	0.66	0.70
E(H), %	3.2	3.56	1.34	2.61	5.79	5.76

Table 4.3 represents the results for the red plastic, which represents very diffuse property of reflection.

Table 4.3: Performance Of The Inverse BRDF for a diffuse material, Red Plastic

Error Metric	Red Plastic					
	Semi Sphere		Sombrero		Mozart's Face	
	FF	RB	FF	RB	FF	RB
E(N), %	0.26	0.32	0.44	1.19	1.7	1.86
E(I), %	2.53	3.11	4.49	7.72	3.84	4.98
E(H), %	4.55	5.22	2.98	3.84	5.92	6.12

4.4.2 Performance of Imputation Methods

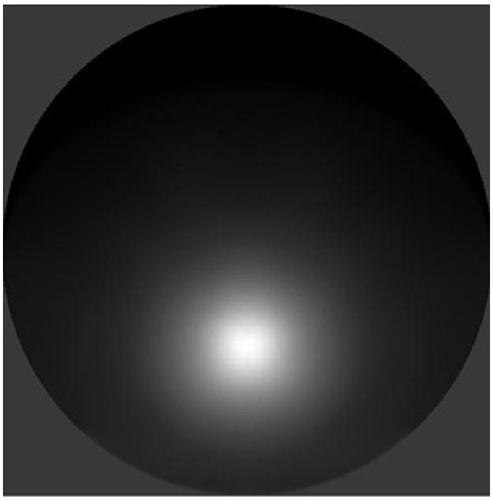
Imputation methods worked in this thesis did not yield a good performance comparing with the neural network ensemble method. Results were computed by similar metrics with the previous sections. Semisphere with the green plastic material were worked for mean and K-nn imputation methods. Table 4.4 represents the results for the sake of completeness.

Table 4.4: Performance of the Imputation Methods

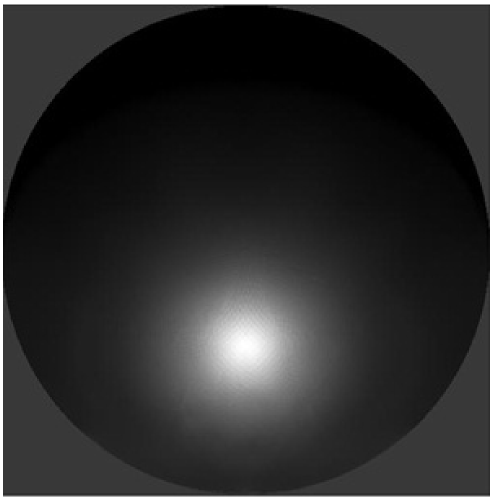
Error Metric	Mean Imputation	K-nn Imputation
E(N), %	24,72	20,23
E(I), %	0.85	0.77
E(H), %	19.97	18.57

4.4.3 Graphical Performance and Visual Results for Inverse BRDF

Following figures shows the input images, output images from estimated surface normals and output of the 3D surface reconstruction phases.

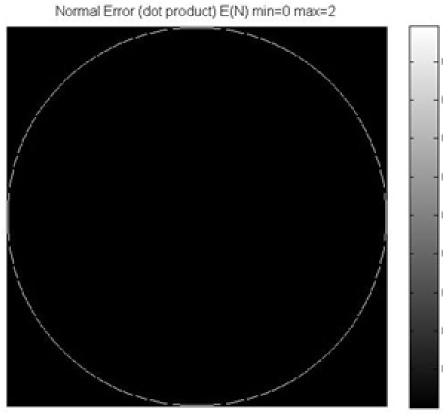


(a) Semisphere input image

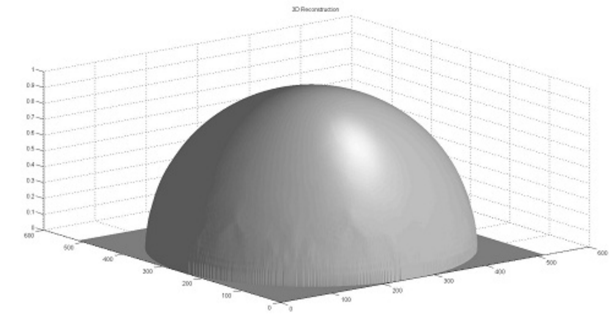


(b) Output image obtained after 3D reconstruction for semisphere

Figure 4.1: Semisphere shape, One of the input and output image



(a) Semisphere Normal Error



(b) Reconstructed 3D shape for semisphere

Figure 4.2: Semisphere shape, Normal Error on the surface and 3D Reconstructed Shape

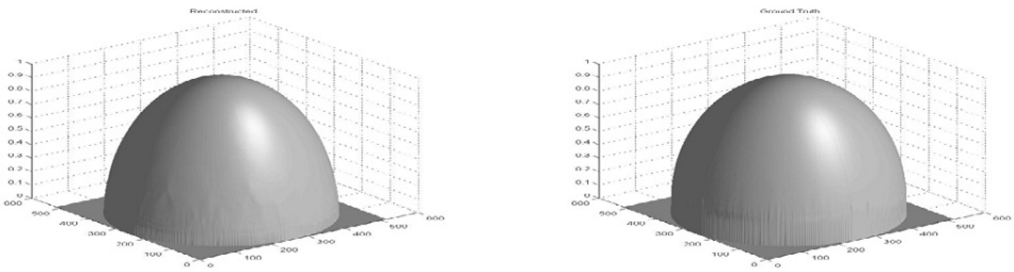
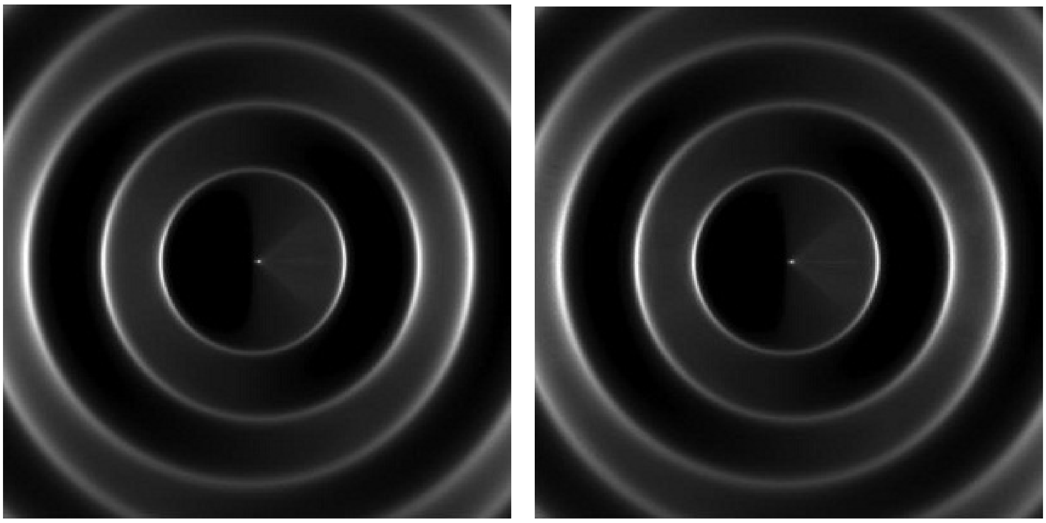


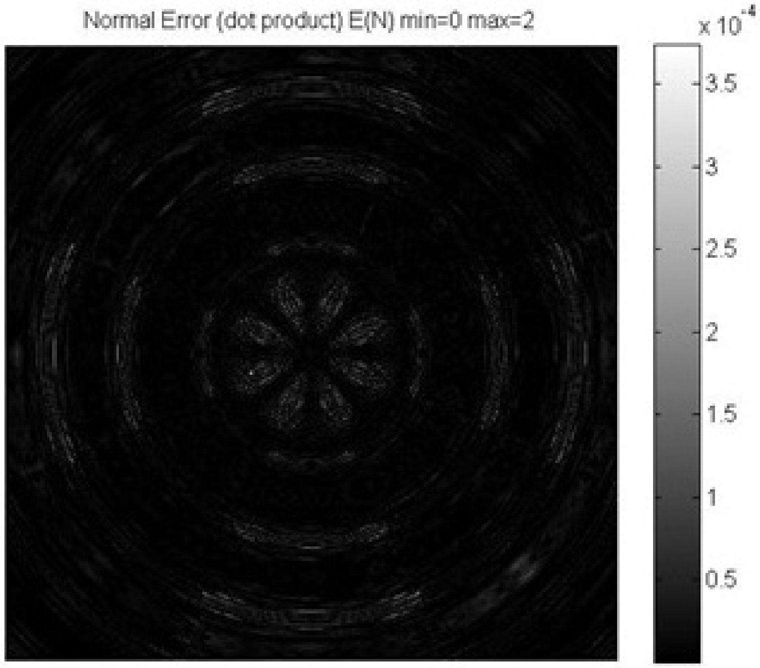
Figure 4.3: Side by side view of Reconstructed 3D shape and Ground Truth



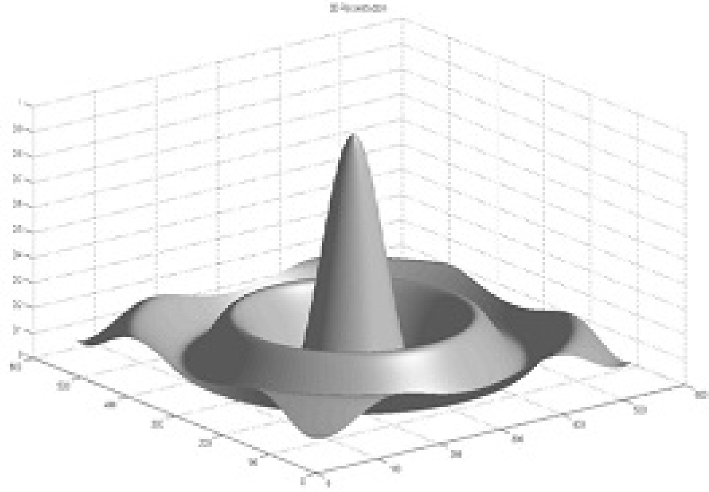
(a) Sombrero input image

(b) Output image obtained after 3D reconstruction for somberro

Figure 4.4: Sombrero shape, One of the input and output image



(a) Sombrero Normal Error



(b) Reconstructed 3D shape for sombrero

Figure 4.5: Sombrero shape, Normal Error on the surface and 3D Reconstructed Shape

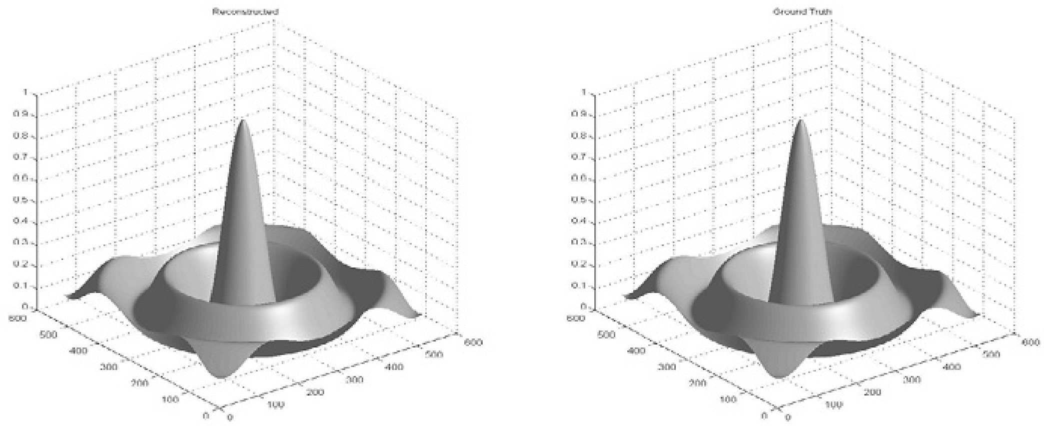
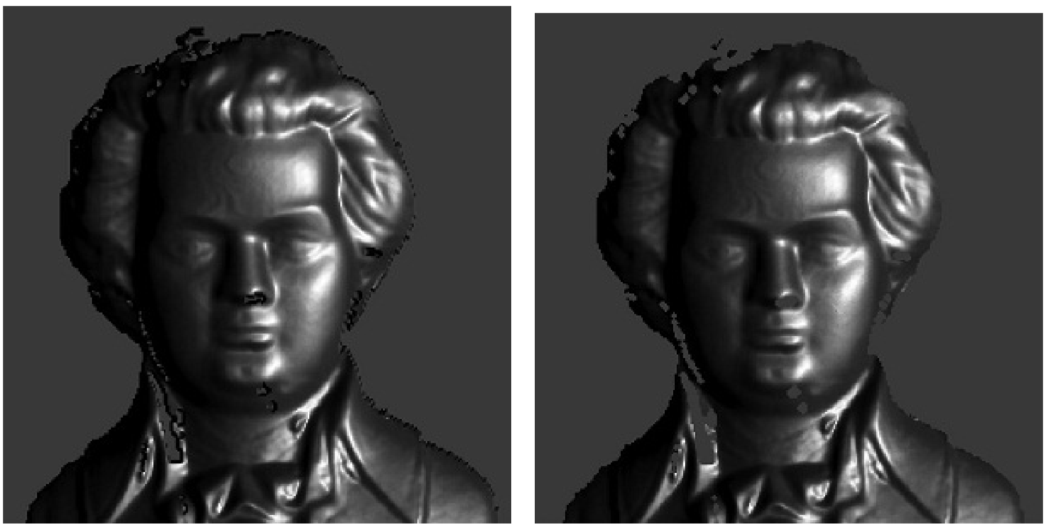


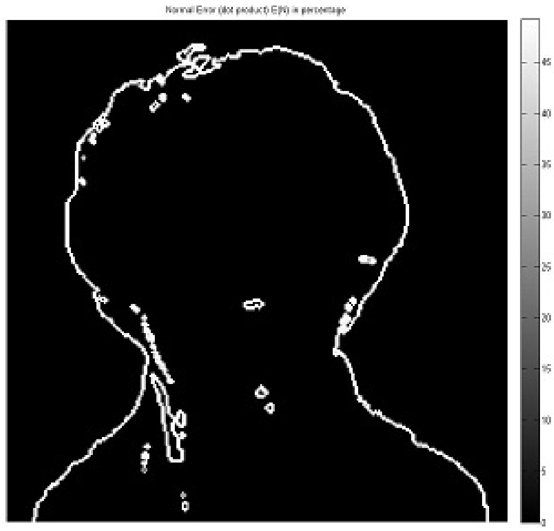
Figure 4.6: Side by side view of Reconstructed 3D shape and Ground Truth



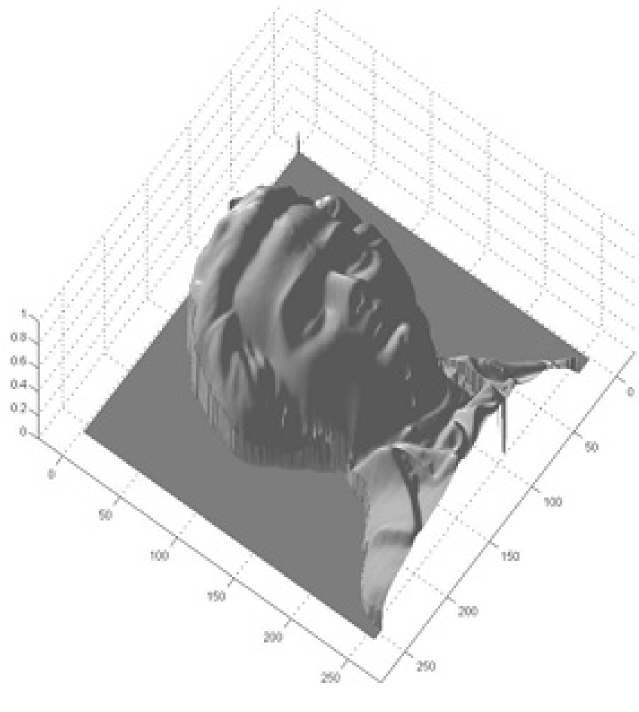
(a) Mozart's Face Input Image

(b) Output image obtained after 3D reconstruction for Mozart's face

Figure 4.7: Mozart's Face shape, One of the input and output image



(a) Mozart Normal Error



(b) Reconstructed 3D shape for Mozart's Face

Figure 4.8: Mozart's Face shape, Normal Error on the surface and 3D Reconstructed Shape

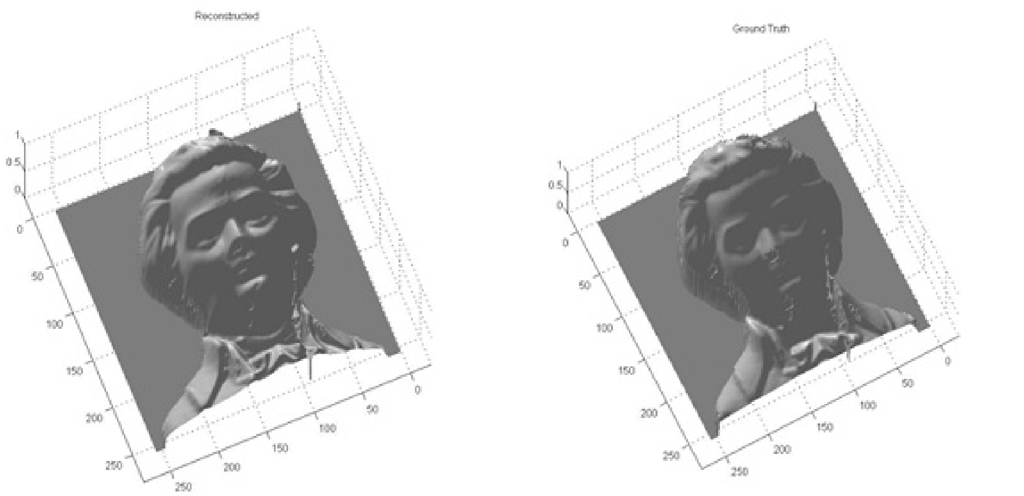


Figure 4.9: Side by side view of Reconstructed 3D shape and Ground Truth

CHAPTER 5

CONCLUSION AND FUTURE WORKS

5.1 Conclusion

Deductions made from available data for a system depend on the completeness and quality of the data in the system. Thus, when the data is incomplete, some problems may arise. Hence methods to deal with missing data have been a great interest for research. The way used to deal with missing data is related to how data becomes missing in the system. Thus, designing a photometric stereo model that handles the incomplete data is very specific problem.

Photometric stereo is an algorithm which uses images taken in different light conditions and BRDF information of material to construct the surface properties of an object. In this work, inverse BRDFs are trained in order to obtain surface normal angles and then surface normals are used to generate surface structure of object. Input images contain specular and shadow regions which cause loss of information about surface normals. Predicted surface normals with the effect of incomplete input data may cause degradation on surface depth information.

In this thesis, inverse BRDF, a neural network based photometric stereo method with the presence of incomplete data was proposed and implemented.

As described in early chapters, incomplete data in any prediction system can be handled by several methods including imputation, deletion and some special machine learning techniques. To eliminate the degrading effect of incomplete data due to specular regions of the images, imputation and one of the machine learning methods, neural network ensembles techniques are applied to the problem.

Imputation is actually replacing the incomplete features from the vectors with new values by using two techniques. One is the mean imputation and the second one is imputation with K-nn technique. Both of the two methods actually increase the general performance of the system a bit but not enough to get meaningful results. Imputation by K-nn performs better than naive mean imputation. They both presented better results for the specular and shadow regions however, the results for these regions decrease visual quality and did not preserve the continuity of the whole surface.

On the other hand, the most effective solution, neural network ensembles performed better than imputation methods. Neural network ensemble model with inverse BRDFs has been found a very robust and novel technique to recover the loss of information in the input data. The inverse BRDF ensemble was found as a case specific well combined method of deletion and neural network based solution.

A set of networks together were used in the model to eliminate the effect of incomplete input data. In the training phase, each network was trained with the whole BRDF data with the selected features. In feature selection intensity values in the input vectors were used as decision. Whenever a maximum in a vector was found, its left and right neighbors were cropped and fed into network in the training. Specular regions brought saturated intensity values to the data and shadowed regions made the intensity levels very close to the zero in the input vectors. This selection removed data coming from specular and shadow regions.

Feed forward neural networks and also radial basis functions were used for inverse BRDFs in order to compare their performance.

Once they were trained, test data was assigned to their adequate network by considering the location of incomplete portions of data like in the training data assigning phase. When an input data vector was assigned to its most suitable network, the irrelevant features of the input data were then deleted from vector. This deletion removed the disturbing effect of highly saturated pixels and the black pixels due to shadow. Then each network performed better and more robust because the values in the input features were always in the range of training data. Also, this classify and delete technique assigned every single input data to a specific network without leaving out any of them. Then, outputs of each networks were combined in a single matrix called surface normal angles. Finally, these surface normal angles were used in surface reconstruction phase in order to obtain a surface depth information.

The proposed inverse BRDF model for photometric stereo was tested with synthetic images and the following results were reached:

- Specular reflectance in the input images decreased the performance of inverse BRDF without considering incomplete data.
- Handling incomplete data with imputation methods led to better results.
- Implementation of neural network ensembles with inverse BRDFs provided very good results.

At the end of this study, three major contributions were identified. Firstly, training an inverse BRDF as a neural network can be used as a powerful photometric stereo algorithm when combined with the surface gradient integration technique. However, specular regions due to mirror reflection property of the material can decrease the performance of inverse BRDFs. Secondly, the results demonstrated that ensembles of inverse BRDFs can eliminate the effect of this specular regions significantly. Combining the deletion method with neural network ensembles to treat the incomplete data results in an increase at the performance of the photometric stereo. Lastly, ensembles of inverse BRDFs can perform even better in very specular materials. The results given in the tables 4.1, 4.2 and 4.3 show that height error is smaller in the red specular plastic which carries the most specular reflectance property.

Considering the network type of the inverse BRDF, results show that feed forward neural networks and radial basis functions perform almost the same but feed forward neural network is slightly better. This results recall the fact that feed forward neural networks are more suitable to the generic regression problems like inverse BRDF.

On the other hand, two limitations of the study were found. First, although using BRDF and neural network structure may solve the specularity problem, it cannot present a solution to shadowing problem. The other limitation is that the general performance of the system is highly dependent on the performance of the neural network which maps intensity to surface normal. If the performance of the classification system itself is low, missing data handling methods are not usable.

5.2 Future Works

Lots of additional work can be introduced by using the inverse BRDF proposed in this study and models similar to it. This can improve the efficiency of photometric stereo more and expand this tool to another problem domains. The followinsg are the list of possible works which might be started in the future:

- In this study, only grey level intensity images were worked. In the future, RGB images may be used in photometric stereo
- The performance for the real objects may be worked if a calibrated imaging setup build in order to take images.
- For this work, only one illumination combination was used, it might be possible to use more than one light or environmental illumination models in the future.
- It is possible extend this model in order to generate a surface reconstruction of multi texture objects.

These future works have a great potential on the improvement of proposed photometric stereo model.

REFERENCES

[1] S. Buyukatalay, O. Birgul, and U. Halici. Iterative photometric stereo with shadow and specular region detection for 3d reconstruction. In *Signal Processing and Communications Applications Conference, 2009. SIU 2009. IEEE 17th*, pages 305–308, April 2009.

[2] Richard O. Duda, Peter E. Hart, and David G. Stork. *Pattern Classification*. Wiley, 2000.

[3] Santosh K. Gaikwad, Bharti W. Gawali, and Pravin Yannawar. A review on speech recognition technique. In *International Journal of Computer Applications*, 2010.

[4] Pedro J. Garcıa-Laencina and Anıbal R. Figueiras-Vidal Jose-Luis Sancho-Gomez. Pattern classification with missing data: a review. In *Neural Computing and Applications*, pages 263–282. Springer-Verlag London Limited, 2009.

[5] J.L. Schafer. *Analysis of Incomplete Multivariate Data*. Chapman and Hall, 1997.

[6] Allison P.D. Missing data. In *Sage University Papers Series on Quantitative Applications in the Social Sciences*, 2001.

[7] Roderick J. A. Little and Donald B. Rubin. *Statistical Analysis with Missing Data 2nd edn.* Wiley, 2002.

[8] John K. Dixon. Pattern recognition with partly missing data. In *IEEE Transactions On Systems, Man, and Cybernetics*, pages 617–621, 1979.

[9] James N Morgan and Frank M Andrews. *Multiple Classification Analysis,2nd edition*. Ann Arbor : Institute for Social Research, 1975.

[10] Alireza Farhangfar, Lukasz Kurgan, and Jennifer Dy. Impact of imputation of missing values on classification error for discrete data. *Pattern Recognition*, 41:3692 – 3705, 2008.

[11] E. Acuna and C. Rodriguez. The treatment of missing values and its effect in the classifier accuracy. In *D. Banks, L. House, F.R. McMorris, P. Arabie, W. Gaul (Eds.), Classification, Clustering and Data Mining Applications,*, pages 639–648. Springer, Berlin, Hedielberg, 2004.

[12] Gustavo E.A.P.A., Batista, and M.C. Monard. An analysis of four missing data treatment methods for supervised learning. *Applied Artificial Intelligence*, 17:519–533, 2003.

[13] D.J. Mundfrom and A. Whitcomb. Imputing missing values: the effect on the accuracy of classification. *Multiple Linear Regression Viewpoints*, 25, 1998.

[14] Batista G and Monard MC. Experimental comparison of k-nearest neighbour and mean or mode imputation methods with the internal strategies used by c4.5 and cn2 to treat missing data. Technical report, University of Sap Paulo, 2003.

[15] Jerez JM, Molina I, Subirats JL, and Franco L. Missing data imputation in breast cancer prognosis. In *BioMed'06*, pages 323–328. ACTA Press Anaheim, CA, 2006.

[16] Troyanskaya O, Cantor M, Alter O, Sherlock G, Brown P, et al. Missing value estimation methods for dna microarrays. In *Bioinformatics*, volume 6, pages 520–525, 2001.

[17] P.K. Sharpe and R.J. Solly. Dealing with missing values in neural network-based diagnostic systems. *Neural Computing and Applications*, 3(2):73–77, 1995.

[18] S. Krause and R. Polikar. An ensemble of classifiers approach for the missing feature problem. In *Neural Networks, 2003. Proceedings of the International Joint Conference on*, volume 1, pages 553–558 vol.1, July 2003.

[19] Kai Jiang, Haixia Chen, and Senmiao Yuan. Classification for incomplete data using classifier ensembles. In *Neural Networks and Brain, 2005. ICNN B '05. International Conference on*, volume 1, pages 559–563, Oct 2005.

[20] Piotr Juszczak and RobertP.W. Duin. Combining one-class classifiers to classify missing data. In Fabio Roli, Josef Kittler, and Terry Windeatt, editors, *Multiple Classifier Systems*, volume 3077 of *Lecture Notes in Computer Science*, pages 92–101. Springer Berlin Heidelberg, 2004.

[21] Robert J. Woodham. Shape from shading. chapter Photometric Method for Determining Surface Orientation from Multiple Images, pages 513–531. MIT Press, Cambridge, MA, USA, 1989.

[22] R. J. Draper and T.P. Pridmore. Reflective photometric stereo. In *Image Processing and Its Applications, 1997., Sixth International Conference on*, volume 1, pages 409–413 vol.1, Jul 1997.

[23] Abhijeet Ghosh, Shruthi Achutha, Wolfgang Heidrich, and Matthew O. Toole. Brdf acquisition with basis illumination. In *Computer Vision*.

[24] Wojciech Matusik. *A Data-Driven Reflectance Model*. PhD thesis, Massachusetts Institute of Technology, 2003.

[25] Wojciech Matusik, Hanspeter Pfister, Matt Brand, and Leonard McMillian. A data-driven reflectance model. Technical report, MIT, 2003.

[26] Szymon M. Rusinkiewicz. A new change of variables for efficient brd representation. Stanford University.