Universidade do Minho
Escola de Engenharia

João Carlos Garcia da Cunha Barbosa

RTI-based techniques and tools for digital surrogates

September 2009

João Carlos Garcia da Cunha Barbosa **RTI-based techniques and tools for digital surrogates**

UMinho|2009

Universidade do Minho
Escola de Engenharia

João Carlos Garcia da Cunha Barbosa

RTI-based techniques and tools for digital surrogates

Dissertation for MsC degree in Computer Science

Supervisor:
Alberto José Proença

September 2009

É AUTORIZADA A REPRODUÇÃO PARCIAL DESTA TESE, APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE

Universidade do Minho, ___/___/_____

Assinatura: _____

Acknowledgements

My utmost gratitude goes to my thesis advisor, Dr. Alberto Proença for allowing me to join his team, for his expertise, kindness, and most of all, for his patience. I believe that one of the main gains of this 3 years program was working with Dr. Proença and gaining his trust and friendship. Dr. Proença, it has been an honor to work with you.

My gratitude also goes for Dr. Luís Paulo Santos and Dr. João Luís Sobral, for the numerous hours spent debating RTI, its possible applications and solutions for the problems encountered during this dissertation work. Also to my lab colleagues, specially to Edgar and Rui, for their time spent with me bouncing off ideas.

Thanks to Paulo Bernardes for his enthusiasm, commitment, support and friendship.

Thanks to the D. Diogo de Sousa Regional Archaeological Museum (MDDS) curator Dr. Isabel Silva for opening the doors of the museum and its collection for this research and to Manuel Santos for the endless hours spent in trying new RTI capture approaches.

Thanks to Tom Malzbender from HP Labs for is support, debate of ideas and help thought this desertion work.

Thanks to the Cultural Heritage Foundation for making this study possible by providing its funding. Specially to Mark Mudge, Carla Schroer and Michael Alshey for their continues support, feedback and unquestionable friendship.

I thank my girlfriend, Cristina who stood beside me and encouraged me constantly, and to my sister, Ana for being my best friend and for her continuous support and interest in what I do.

Finally, I would like to express my utmost gratitude to my parents and grand-parents for their faith in me, for teaching me that hard work is always paid in double and that I should never surrender.

Abstract

Digital representations of Cultural Heritage (CH) and Natural Science (NS) artefacts often use collected data from the real world together with computer graphics techniques and skilled human intervention. A scholar will trust these representations when they are extended to digital surrogates, with all empirical provenance information logged in. and use them when they are robust, affordable and easy to access and handle. This was the main context and motivation for this dissertation.

Image-based techniques are a popular way to acquire and model surface materials. Reflectance Transformation Imaging (RTI) provide a powerful and efficient tool to build a model of the surface of an artefact, with a reliable 3D visualization, as analysed in this dissertation.

An automated process pipeline was developed to acquire and build an RTI representation as a digital surrogate, in an open source context, the RTIbuilder. This automated process pipeline uses computer vision algorithms to acquire the information required by the generation stage and offers the user a set of tools for the most commonly used image processing procedures.

The RTIbuilder not only apply software engineering techniques to solve the user's problems, but also addresses most of the concerns posed by the scientific method: it removes the unreliable human factor from the RTI generation process during the workflow automation, by storing and preserving all the empirical provenance that include the unaltered empirical data gathered, the process and parameters history and all human intervention over the gathered empirical information and process parameters.

Workshops and tutorials were given to audiences with CH and NS professionals, and the RTIbuilder was field tested in several locations worldwide. Feedbacks were valuable to tune the application and opened new tracks for future work.

Resumo

Representações digitais de artefactos da herança cultural e das ciências naturais usam dados do mundo real em combinação com técnicas de computação gráfica e intervenção de profissionais especializados. Um investigador apenas poderá confiar nestas representações digitais se elas forem convertidas em substitutos digitais, que deverão incluir toda a informação de proveniência.

Técnicas baseadas em imagem são uma forma tradicional de capturar e modelar materiais. "Reflectance Transformation Imaging (RTI)" disponibiliza uma ferramenta poderosa e eficiente para construir uma representação da superfície de um artefacto, com uma visualização 3D fidedigna, tal como se demonstra ao longo desta dissertação.

Foi desenvolvido um encadeamento de processos automático para capturar e produzir representações RTI enquanto substitutos digitais, num contexto de software aberto, dando origem a uma ferramenta designada RTIbuilder. Esta metodologia usa algoritmos de visão computacional para adquirir e calcular os parâmetros de entrada necessários à geração das representações RTI. Para além disso, disponibiliza um conjunto de ferramentas para as técnicas de processamento mais utilizadas no processo.

O RTIbuilder, para além de aplicar técnicas de engenharia de software na resolução de problemas do utilizador, vai também ao encontro dos principais problemas decorrentes do método científico: remove a falibilidade humana do processo de geração de RTI, uma vez que armazena e preserva a proveniência empírica. A proveniência empírica engloba as observações empíricas recolhidas e inalteradas, a sequência de processos e respectivos parâmetros usados, bem como todo o histórico de manipulação da informação empírica recolhida e parâmetros de processo.

Foram ministrados workshops e tutoriais a profissionais da herança cultural e ciencias naturais e o RTIbuilder foi testado em situações reais em vários locais a nível mundial. Os resultados desses testes foram imprescindíveis para afinar a aplicação e abriu novos caminhos para trabalho futuro.

Contents

Chapter 1

Introduction

1.1 Context

Mankind has always stride to evolve and leave a definitive mark of his existence and achievements for generations to come. Some of the knowledge of ancient cultures transformed society in such a way that their knowledge and culture are still present today, but most of its history has been lost through the centuries. However, there is irrefutable evidence of its presence in the landscape by a collection of sites that must be studied, documented and preserved for future generations.

Natural Sciences (NS) and Cultural Heritage (CH) professionals look into the digital representation as a tool to communicate knowledge and to convey ideas. However, digital representation can be described by three types that distinguish their use: art and entertainment, visualisation and digital surrogates. Computer models and animations of palaeolithic sites 2.5 million years ago are helpful to visualise the way of life and activities of the Neanderthal man. Although, these visualisations are useful to express and efficiently communicate an idea, however the author inevitably introduces speculative data into the model to effectively convey his interpretation of the society behaviour at the time [MMC+08]. In spite of the usefulness of the visualisation model, the speculative contents will lead other scholars to discard those represen-

tations unless a full discloser and description of the speculative factors is explicitly provided[HNP06].

Digital surrogates aim to reliably represent *real world* content in a digital form. A full description of a digital surrogate is given in the next session, but for clarity it must enable scientific study and therefore must be built from inter-subjectively verifiable information. To build scientifically robust and long lasting digital surrogates requires tools that are open (non-proprietary), affordable (in instrumentation and human resources) and easy to use. The literature describes some of these modelling pipelines for archaeological visualisation (e.g., [AFT+04]), but most do not satisfy all the requirements.

The search for scientifically robust digital surrogates is growing for three main reasons: (i) to provide a common framework for scholars and experts to reliably promote credible scientific discussions; (ii) to give remote access of a site data to a wider audience and (iii) to build repositories for long term digital preservation of the human cultural legacy.

Reflectance Transformation Imaging (RTI) developed by Tom Malzbender et. al. 2001 [MGW01] and the related techniques are widely accepted by the NS and CH communities as a techniques with a large potential for generating digital surrogates [MMC+08]. RTI is an image-based re-lighting technique that uses a set of images taken from a single viewpoint, illuminated from different light-source positions, and later processed to allow the generation of re-lighted virtual views of the object surface. RTI techniques are *empirical* by nature, since they rely on photographs to generate the final representation through a mathematical process. RTI has already been used by scholars to document CH artefacts in an efficient way and the tools originally developed were useful enough to help scholars to study the artefacts[HBMG02, FBM+06a].

1.2 Digital Representation vs Digital Surrogate

Creating a digital representation of an object is a common practice in the Natural Sciences and Cultural Heritage communities. Tools and techniques

allow the digitisation of an object, ranging from the simple digital description and 2D images to pure 3D data and models. The creation of digital representation attempts to document *reality* as close as possible, but in most cases the supporting information is sparse and biased by the producer.

This bias is easy to identify when we look carefully into the process of digitisation or to the processing steps usually taken to generate the digital representation. For instance, when photographers need to shoot a rock art engraving, they carefully plan the illumination condition required to capture the details on the surface of the stone. The planning will inevitably require them to make choices of the relevance of some features since they will not be able to set-up an illumination stage able to capture the whole in a single shot. This decision process will introduce bias and the final digital representation will not accurately represent the surface properties. The same problem may arise from the process characteristics when generating, for example, a 3D mesh representation of the surface. The 3D acquisition through 3D range scanners can have sufficient data to accurately represent the surface of the stone but it generally creates highly dense meshes that can be composed of billions of triangles. Working with these highly dense mesh requires a high computational power normally not available, demanding a complexity reduction step to make the mesh less computational intensive. The level of simplification traditionally lies on the judgement of the producer and again requires him to make subjective evaluation of relevant features.

These subjective decisions decrease dramatically the trust level that a scholar may have in a digital representation created by another scholar. The level of trust is directly related with the amount of empirical information preserved and to the provenance data related with the digital representation. A scholar will only be able to trust a digital representation, and therefore use it on his/hers research, if he/she is able to evaluate its quality following a scientific approach.

Scientific methodology when applied to digital representation imposes a set of constrains related with credibility, accessibility and usefulness. Scientific methodology requires the detachment from the observed data (empirical

data) of the observer, and the information needed for scientific inquiry is designated by *provenance*, which includes data acquisition parameters and processing steps. These two types of data, empirical and provenance, are the corner stone of scientific methodology since they allow any scholar to replicate, validate and assert the quality of the method or results and must be fully disclosed. The concept of *empirical provenance* mentioned by Mark Mudge et. al. [MMC+08] enables a digital representation to act as a digital surrogate. Empirical provenance records all the information related to the acquisition and the transformation steps taken to convert the original untreated empirical data that was acquired into the final digital representation.

A digital surrogate is a digital representation by nature, but encapsulates within enough empirical provenance metadata to allow reproducibility and assessment of its quality by any other scholar. The physical detachment of an object from the object itself guaranties the access to all potential data on the object and allows its dissemination through different digital means, specially if the underlined technology is open enough to allow its democratisation. By combining on the digital surrogate the best of digital representation with the scientific methodology, to ensure the quality of the representation, we make sure that it is useful to interdisciplinary research studies.

1.3 Motivation

The RTI techniques use as basis true empirical data, photographs and relies on a well documented mathematical process to generate the final representation. This empirical nature makes the technique a suitable candidate to generate a digital surrogate. However, unless additional data is recorded to establish its provenance it will be difficult for scholars to find the information reliable. The RTI generation process relies only on the prior knowledge of the light source positions, which can be empirically recorded as highlights on spheres carefully placed near the object. For the digital representation to pass through the scrutiny of the scientific methodology information about the light source, data on the camera, lens and pre-generation imaging processing must be fully

recorded and embeded with the RTI representation.

This thesis aims: (i) to produce a proof of concept of the premise that RTI can be used to produce a valid digital surrogate, (ii) to define and built an automated set of tools to generate the digital representation of an object, and (iii) to explore information within the RTI technique to enhance details to help cultural heritage professionals to express their interpretation of the represented artefact.

1.4 Outline

Chapter 2 are presented some concepts related with computer graphics and vision required to understand some of the concepts presented in later chapters. Chapter 3 presents the RTI concept and all related techniques required for data acquisition and processing. Chapter 4 describes the problems related with digital surrogates long term preservation and how they can be addressed by the automation of some of the processing steps within the workflow. It then presents the specification of a software application to address these issues, the RTIbuilde. Chapter 5 describes how the RTIbuilder was implemented and tested, together with a discussion on its deployment. Chapter 6 introduces some final remarks and presents possible future work.

Chapter 2

Surface Materials

The material appearance plays an important role in the human perception of the real world. Through the surface appearance we can determine the materials that a particular object is composed of, roughly calculate age or estimate its value. The surface material also gives important clues to its mechanical properties and how we should handle the object. In art, design and architecture, material properties are carefully studied and placed in order to create specific effects or an atmosphere.

The human perception of an object depends on physical and psychological phenomena. The light that reach our eyes after reflecting on the surface of a object creates the physical stimulus required for the eye physiology to perceive the object; however, the connection to the idea of the object processed in our mind is a consequence of the observer experience. The end result of a computer generated image, which is a flat 2D image, must create the physical stimulus required to generate the same idea of appearance of the object as if directly observed. To achieve this goal we need to build models based on knowledge from light physics, human perception and image formation. For all these reasons computer graphics research has dedicated a special attention to material appearance and material light interaction.

The light physics has been studied since ancient Greece to the most recent quantum electrodynamics and keeps expanding. The fundamental light

physics model today is given by quantum optics which can explain the dual wave-particle behaviour of light, by simplifying the quantum optics model, and the light interaction at the atomic level. However, the model is too detailed to use in image generation. An approach based on the wave behaviour of light given by the Maxwell equation can efficiently map effects such as diffraction, interference and polarization which can be observed in every day scenes. The wave model is still considered to be a highly demanding model in terms of computational resources and the usual computer graphics models tend to ignore these types of phenomena by not considering the wave nature of light. Dropping the wave nature of light, the remaining particle behaviour can be efficiently modelled by the geometric optical rules traditionally used for image generation. The geometric models consider that the light is emitted, reflected and transmitted deprecating the size of the wavelength. Simplification to the geometric optical model can be introduced by neglecting some of the light properties to achieve a balance between accuracy and computational resources.

Although, human perception has been studied for almost the same amount of time as the light physics, the current models from psychology, however numerous, can not combine all the individual characteristics but only a very small fraction of them. These models complexity make it very difficult to incorporate them into the object, or scene, rendering process. Nevertheless, there are some common terms used to describe the look and feel of a surface material such as *shiny blue, matte red or bumpy.* Although crude, these approximation allows us to establish rough relationships between human perception and light material interaction that can be explored to approximate the material appearance. Computer graphics tries to take advantage of the fuzziness of the terms used to describe the surface material to simplify the light, geometry and material models to create a stimulus able to produce a convincing photo-realistic image. For instance, the bump map technique allows the model to use flatter surfaces but still produce an image that will stimulate the observer into perceived shapes that do not exist in the model geometry.

In the next subsection, some of the concepts behind the light physics

modelling will be briefly explored.

2.1 Radiometry

To understand the model involved in measuring light energy we must first explore some of the physical quantities studied by radiometry. In this section the term "light source" makes no distinction between a "real" light source or the reflection of light by a surface onto another.

Radiant power or *Radiant flux* (Φ) defines the amount of light energy flowing (ΔQ) from, to or through a surface per time unit (Δt), expressed in (Watts or Joules/sec).

$$\Phi = \frac{\Delta Q}{\Delta t}, or, \frac{dQ}{dt} \tag{2.1}$$

Radiant power expresses the dependency between the amount of energy flowing and time interval. However, the amount of incident, or departing, energy might be uneven through the surface area. *Irradiance* (E) measures the amount of incident *radiant power* ($d\Phi$) per unit of surface area (dA), while the term *radiante exitance* (M), often called *radiosity* (B) in the global illumination problem domain, measures the amount of departing *radiant power* ($d\Phi$) per unit of surface area (dA).

$$E = \frac{d\Phi}{dA} \tag{2.2}$$

$$M = B = \frac{d\Phi}{dA} \tag{2.3}$$

The *radiant flux* also depends on the solid angle ($d\omega$) and can be expressed by the term *intensity* defined as

$$I = \frac{d\Phi}{d\omega} \tag{2.4}$$

So far, the quantities only express the amount of energy flow per unit of time as dependent on either unit surface area or unit of solid angle. Both dependencies are combined in a single term called *radiance* (L).*Radiance* (L)

specifies the amount of energy flow to or from a surface (dA) in a certain direction, per unit of solid angle ($d\omega$) around the given direction, per unit of projected area perpendicular to the direction of travel:

$$L(\vec{x}, \hat{\omega}) = \frac{dE}{d\omega} = \frac{d^2\Phi}{d\omega A^{\perp}} = \frac{d^2\Phi}{cos\theta d\omega} \tag{2.5}$$

where θ is the angle between the direction of travel and the surface normal.

The concept of radiance can be perceived as *intensity*: an observer looking to two surfaces with an area of 10 m^2 and 1 m^2, respectively, both emitting 1000 Watts, the second will appear brighter, more intense, since the amount of energy per unit of surface area is higher that the first.

Radiance is defined in a point $\vec{x}$ with regard to a given direction, and since radiance can be defined as incident or exitant we will denote this dependency as:

- $L(x \leftarrow \Theta)$ - incident radiance from the direction Θ
- $L(x \rightarrow \Theta)$ - exitant radiance in the direction Θ
- $L(x \leftarrow y)$ - radiance from point y in the direction of point x
- $L(x \rightarrow y)$ - radiance from point x in the direction of point y

Although radiance varies depending on the direction Θ, an important property arises from this dependency since radiance remains constant along straight paths.

In fig. 2.1, assuming that the medium between point $\vec{x}$ and $\vec{y}$ does not participate in the light transport, we can easily deduce that $L(x \leftarrow y) = L(y \leftarrow x)$, by computing the energy transport between them, and as a consequence the radiance does not attenuate with distance. This accounts also for the fact that an human observer, or a camera, which are sensitive to radiance perceive the same colour or brightness independently of the distance of the object. However, this property is not valid if the medium between the points is a participating medium, which will inevitably absorb and scatter light energy.

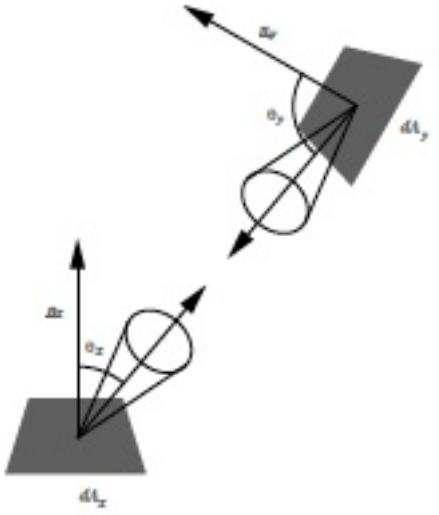


Figure 2.1: Energy transport between two differential surfaces.

In computer graphics, namely in global illumination, once the incident and exitant radiance is known for all surface points, then it is possible to compute the light energy transport through the all three-dimensional scene to generate the final image representation. Since computer graphics also aims a balance between computational resources and reliable physical models if the medium is non-occluding, such as air, it is approximated to vacuum, meaning that the medium will not be taken into consideration, in any computation.

All the above measures and quantities are also dependent on the wavelength of the light energy under consideration. As stated by Nicodemus et. al. 1976 [Nic76] we can incorporate this dependency by considering that the above quantities are integral functions over the full spectrum of the wavelength. For the radiance with wavelength, also know as *spectral radiance* can be computed by simply integrating it over the full wavelength spectrum

$$L(x \leftarrow \Theta) = \int_{spectrum} L(x \leftarrow \Theta, \lambda) d\lambda \tag{2.6}$$

but usually the dependency of radiometry terms is implicitly expressed.

Radiance, due to the presented properties, is the most fundamental quantity used in computer graphics to produce a plausible physically based image to present to the user, and the global illumination algorithm compute its value, for each visible point from the user viewpoint.

A more complete radiometry introduction can be found in the book by McCluney [Ros94] and by Nicodemus in [Nic76].

2.2 Light and Material interaction

The material, and specially the interaction between the material and light energy, plays an important role in modelling the scene. Materials have different behaviours under the same lighting conditions, changing the overall appearance of the object surface. Some materials will exhibit a diffuse behaviour, other will behave as mirrors or even have a transparent behaviour allowing us to "see" through the object. It is also easy to notice that a wall painted with a flat white color, such as "matte white" paint, will not have the same behaviour as a wall covered with perfect mirrors and that it will produce a different illumination result.

To fully model the light interaction with a surface material as presented in fig. 2.2 we need a twelve dimension function to cope with all the spatial, time and wavelength dependencies.

To explain the Figure 2.2 let us assume the light wave, with wavelength λ_i, emitted from a light source placed in the direction $\Theta_i = (\theta_i, \phi_i)$ relative to the tangent plane at point $\vec{x_i}$ on the object surface, as a continuous source of photons flowing in a straight line in the direction of $\vec{x_i}$. Considering the instant t_0 as the moment that the photon leaves the light source, it will take t_i to travel the distance between the light source and the surface point $\vec{x_i}$. The energy of the photon will be scattered through the surface material and eventually leave the surface at point $\vec{x_r}$ in an arbitrary direction $\Theta_r = (\theta_r, \phi_r)$. However, due to the interaction of the photon with the surface material two

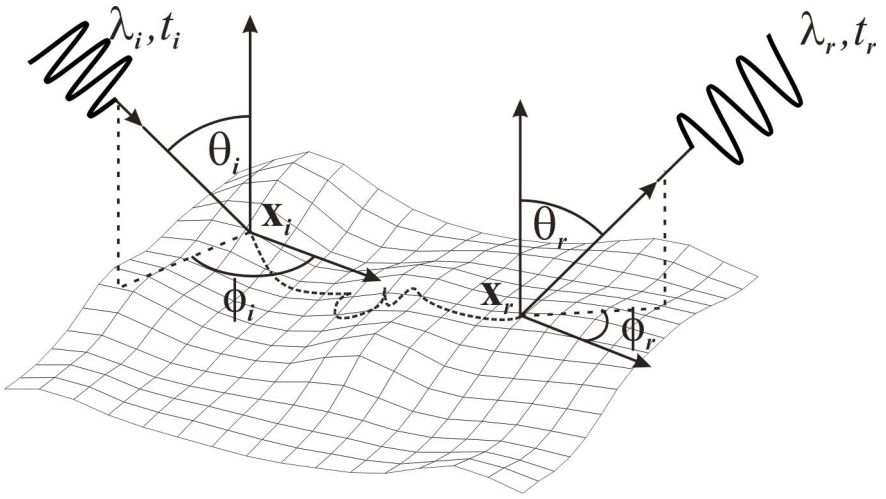


Figure 2.2: Geometrical properties of the energy transport. (from [MMS+04])

phenomena may occur and are known as phosphorescence and fluorescence. When phosphorescence occurs the material absorbs part of the energy of the photon and releases it slowly over time increasing the time gap between the instance t_i and t_r, this phenomena is easily observed in materials that seem to "glow" in the dark. On the other hand, some materials are able to absorb part of the photon wavelength energy converting the absorbed energy to other forms of energy, such as heat or movement, and emitting a new photon with a longer, less energetic, wavelength (λ_r) exhibiting a fluorescence property.

Nevertheless, this generic 12 dimensional function is impractical to model, compute and specially to capture. Computer graphics light models aim to achieve a balance between the accuracy of the light material interaction model and the computation resources by dropping some of the physical phenomena involved. To reduce the complexity of the general 12D reflectance function, some assumption are made:

- light travels at infinite speed;
- the incident and exitant surface point are the same;
- the material does not change the light wavelength;

The simplification reduces the accuracy of the light transport model but simplifies the computation. For instance, to assume that the light travels at infinite speed is not crucial for the human perception, since we are not able to perceive the velocity but it allows the rendering algorithms to use a steady-state approach to the image generation process. These simplification drop 3 dimensional dependencies reducing the function to a 9D function. Additionally, a 8D function known as *bidirectional scattering-surface reflectance distribution function* (BSSRDF) described Nicodemus et. al. 1977 [NRH$^+$77] can be obtained by the discretization of the wavelength into three colour bands (RGB).

BSSRDF is one of the most complete material models used in computer graphics, which models all dependencies, except phosphorescence and fluorescence. Due to the complexity involved in modelling and capture of the BSSRDF it is not often used and is replaced by simpler formulation.

Bi-directional Reflectance Distribution Function (BRDF) is the most general model of reflectance used in computer graphics, at least at the level of detail that we are considering. The BRDF is defined as the ratio between the differential radiance reflected in an exitant direction, and the incident irradiance through a differential solid angle; or, more precisely, the BRDF is defined as the derivative of reflected radiance to incident irradiance

$$f_r(x, \Theta_i \rightarrow \Theta_r) = \frac{dL(x \rightarrow \Theta_r)}{dE(x \leftarrow \Theta_i)} = \frac{dL(x \rightarrow \Theta_r)}{L(x \leftarrow \Theta_i) cos\theta_i d\omega_i} \tag{2.7}$$

and its values as to be positive and may exhibit different values for different wavelengths.

$$f_r(x, \Theta_i \rightarrow \Theta_r.\lambda_i) \geq 0, \forall \Theta \in \Omega \tag{2.8}$$

The basic concept of the BRDF (eq. 2.7) is illustrated in figure 2.3, showing the incoming light ray direction ($\hat{\Theta}_i$), the outgoing light direction ($\hat{\Theta}_r$) towards the observer viewpoint, both defined on the orthogonal axis system formed by the normal $\vec{n}$, tangent $\vec{t}$ and by the binormal $\vec{b}$ at point $\vec{x}$.

BRDF's with different incident light directions are linearly independent

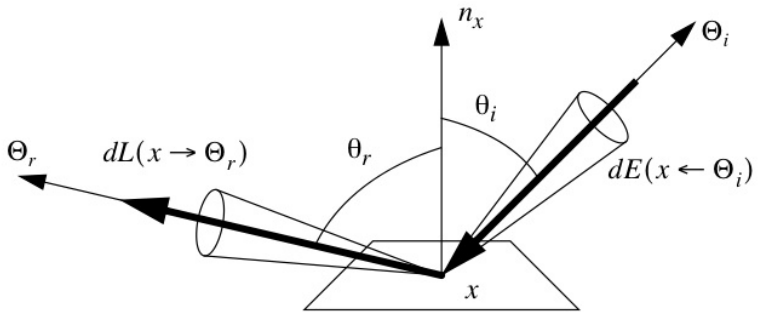


Figure 2.3: BRDF - Light reflection on surface material.

from each other. Therefore, the BRDF as defined in equation 2.7 behaves as a linear function in regard to all incident direction, and to compute the total amount of reflected radiance we must compute the sum of all contributing light sources distributed over the hemisphere above the surface point, assuming that the surface is opaque and non-emissive, meaning that we must integrate equation 2.7 over the surrounding positive hemisphere (Ω^+)

$$dL(x \rightarrow \Theta_r) = f_r(x, \Theta_i \rightarrow \Theta_r) dE(x \leftarrow \Theta_i) \tag{2.9}$$

$$L(x \rightarrow \Theta_r) = \int_{\Omega^+} d\omega_{\Theta_i} f_r(x, \Theta_i \leftrightarrow \Theta_i) L(x \leftarrow \Theta_r) cos(\vec{n_x}, \Theta_i) \tag{2.10}$$

where $cos(\vec{n_x}, \Theta_i)$ is the cosine of the angle formed by the surface normal at point x ($\vec{n_x}$) and vector $\hat{\Theta}$.

For a BRDF to be physically plausible it must observe two fundamental physical principles: *reciprocity* and *energy conservation*. The *reciprocity*, also know as the *Helmholtz reciprocity*, a principle that states that the paths followed by light rays can be reversed; if the incident and exitant direction are interchanged, and the value of the BRDF will remain constant:

$$f_r(x, , \Theta_i \rightarrow \Theta_r) = f_r(x, , \Theta_r \rightarrow \Theta_r) \tag{2.11}$$

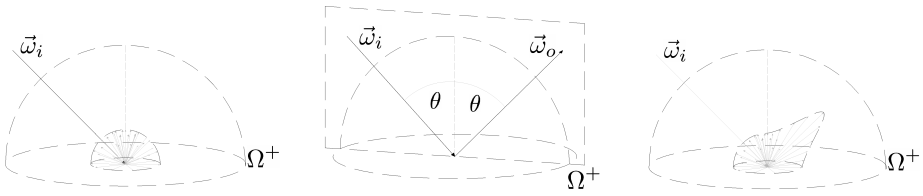
$$f_r(x, , \Theta_r \leftrightarrow \Theta_r) \tag{2.12}$$

where the double arrow notation, in equation 2.12, indicates that the two directions may be freely interchanged. The *energy conservation* principle states that an object material can not reflect more energy than the received energy, but may absorb, or dissipate in another form of energy, a portion of the received energy

$$\int_{\Omega^+} f_r(x, \Theta_i \rightarrow \Theta_r) cos(\vec{n_x}, \Theta_r) d\omega_\Theta \leq 1, \forall \Theta_i \in \Omega^+ \tag{2.13}$$

In general, BRDF's are considered anisotropic, meaning that if the incident and exitant are rotated about the surface normal at point $\vec{x}$ its value will change. Nevertheless, there are many materials where the value of the BRDF does not depend on a specific orientation of the surface, and are considered as isotropic.

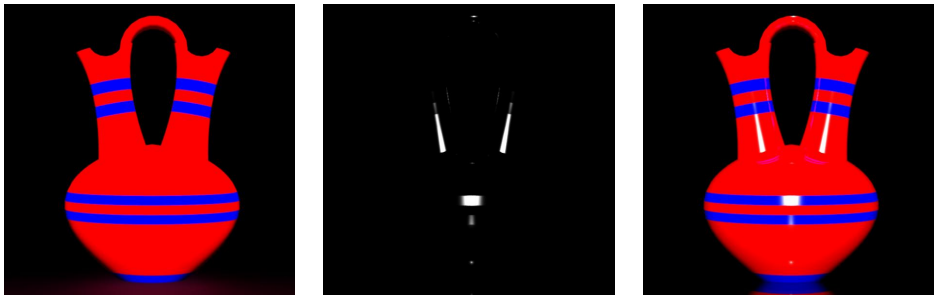
BRDF typically describe opaque materials, and the reflection properties depend on the material the object is composed of (e.g. metal or plastic), surface structure (e.g. rough or polished) and on homogeneity of the material across the surface. Depending on the nature of the BRDF used, the material appearance will be denoted as diffuse (fig. 2.4(a) and 2.5(a)), specular (fig. 2.4(b) and 2.5(b)) or glossy (fig. 2.4(c) and 2.5(c)).



(a) Diffuse (b) Specular (c) Glossy

Figure 2.4: Bidirectional Reflectance Distribution Function albedo

Diffuse reflective materials are characterised for distributing the same portion of incident light uniformly in all directions. Diffuse or *Lambertian* reflection are view independent, i.e., the illuminated surface looks the same to the observer independently of the its viewpoint direction (fig. 2.4(a) and 2.5(a)).



(a) Diffuse (b) Specular (c) Glossy

Figure 2.5: Material appearance

Specular surfaces behave as perfect mirrors and they only reflect light in one specific direction, meaning that, according to Snell's law, the incident and exitant direction make equal angles with the surface's normal and are both coplanar with the surface normal(fig 2.4(b)). To determine the direction of the light reflection we only need to compute:

$$R = 2(n_x.\Theta_i)N - \Theta_i \tag{2.14}$$

However, perfect specular surfaces are an ideal mathematical concept, since the surface has only one exitant direction for which the BRDF is different from 0; therefore the value of the BRDF along that direction is infinite. There are no perfect specular materials in the real world, but some can exhibit this behaviour very closely.

Most surfaces are not perfectly diffuse, and it has been already stated that there are no perfect specular surfaces, so they all are a combination of both reflectance behaviours. In literature these surfaces are often called glossy surfaces, and their BRDF is difficult to model with a single analytical formula.

Although we have limited the scope of the work to opaque materials it is important to mention that some surface materials present a "transparency"[1] behaviour. Strictly speaking, the BRDF is defined over the entire sphere

[1]The correct term would be transmission in opposite to the term reflection

of directions (4π steradians) around a surface point, and the "transparent" side of the BRDF (in this case called *bi-direction transmission distribution function* BTDF) also exhibits properties such as diffusion, specularity and glossyness depending on the characteristics of the materials.

Real materials can have complex BRDF models; nevertheless, various analytical models have been proposed to simulate its behaviour. One of the simplest models proposed for diffuse surfaces is the Lambert's model. Since the amount of exitant radiance is equal in all possible directions over the enclosing hemisphere, the Lambert's model reduces the BRDF to a constant value

$$f_r(x, \Theta_i \rightarrow \Theta_r) = k_d = \rho_r/\pi \tag{2.15}$$

Another, historical, model was presented by B. Phong in [Pho73] that adds to the lambertine model the specular coefficient to create the appearance of glossy surfaces, becoming popular to the point of being directly supported by GPU hardware,

$$f_r(x, \Theta_i \rightarrow \Theta_o) = k_s \frac{R.\Theta_r^n}{N.\Theta_i} + k_d \tag{2.16}$$

since the specular surfaces reflect light along a specific direction R computed from equation 2.14. However, this model does not cope with the two physical constrains of BRDF's, the energy conservation principle and the reciprocity principle. The later proposed modified Blinn-Phong model solves some of these problems, but not all, and is formulated as

$$f_r(x, \Theta_i \rightarrow \Theta_o) = k_s(N.H)^n + k_d \tag{2.17}$$

where H is the halfway vector between Θ_i and Θ_r. However, this model is still not able to capture the realistic BRDF's behaviours.

Some analytical models present suitable behaviours to model real world surfaces and they will be introduced in an informal way since they are not in the scope of this work.

Cook and Torence in [CT82] assumed that a surface is composed of a collection of randomly distributed micro-facets to create their BRDF model

and additionally they also included the Fresnel reflection and refraction terms,

$$f_r(x, \Theta_i \rightarrow \Theta_r) = \frac{F(\beta)}{\pi} \frac{D(\theta_h)G}{(N.\Theta_i)(N.\Theta_r)} + k_d, \tag{2.18}$$

where F is the Fresnel reflection term, the micro-facet distribution is modelled in D and term G models the geometric shadowing.

Ward et. al. 1992 [War92] presented a BRDF model similar to the Phong model, but replaced the cosine term to a power by an exponential function. This model simulates the surface micro-structure variation, as it appeared in previous optics literature, by averaging the slope of the microscopic surface roughness and using it as a parameter for the exponential term.

Another often used model was presented by Lafortune et. al. [LFTG97] by generalising the original Phong models, but describes specular lobes arround any axis defined with respect to the surface instead of peaks of reflection arround the specular direction; the general formulation can be expressed as

$$f_r(\Theta_i \rightarrow \Theta_r) = f_r(u \rightarrow v) = \rho_s(C_x u_x v_x + C_y u_y x_y + C_z u_z v_z) \tag{2.19}$$

where u and v are the directions Θ_i and Θ_r, in orthogonal axis form by the normal at the incident surface point. The coefficients C_x, C_y and C_z determine the size and the direction of the lobe while ρ_s defines how narrow it is. This model adds also the ability to model retroreflectance property, which is the tendency of some surface materials to reflect the specular component in the direction of the light source.

2.3 Modelling and Acquisition

Generating images that realistically represent materials and object surfaces is one of the most challenging fields of computer graphics. Traditionally, surface geometry models only capture surface details on the triangular meshes up to a certain scale. The microscopic features that influence the reflectance behaviour are simulated through the analytical BRDF models.

This approach is not suitable nor designed to address the modelling of realistic appearance of all real world materials, and due to the lack of correspondence with some of the most interesting physical effects that are difficult to parameterize. However, very limited group of materials, such as metal and shinny plastic, can be efficiently approximated by this approach. The difficulty of parameterization of the analytical BRDF model is even higher when the material requires the modelling of the meso structures present at most surfaces, which play an important role in the appearance of the surface material and are between the macro-scale structures captured in the geometry and the micro structures modelled in the analytical BRDF's.

Some mesoscopic and microscopic structures can be modelled using *normal* or *bump mapping*. Normal maps, or bump maps [Bli77], can be used to transform flat surfaces into curved surfaces by introducing a new per-texel normal to be used in the BRDF evaluation process instead of the surface's geometric normal. Driven by the limitation posed by traditional surface material modelling and the increasing size of memory available in computer graphics cards, image based techniques together with normal mapping are becoming increasingly important, but they fail to accurately represent the surface and material appearance since they are not able to express the light and view dependency of the material.

Some researchers have proposed several distinct approaches to overcome the limitation of the traditional surface material modelling. The most complete model used in computer graphics, presented by Nicodemus [NRH$^+$77], is the 8D BSSRDF, which is able to cope with most of the phenomena present at the 12 parameters of the generic function. BSSRDF require the modelling and estimation of 8 different parameters, which are highly demanding in terms of computer resources, acquisition systems and time requirements. To reduce the complexity of the capture, fitting and rendering processes of real world surface material some authors reduce the complexity of the BSSRDF model by dropping some of the physical dependencies, which limits its use to a particular subset of materials. Fig 2.6 presents a taxonomy adpated from [RM] which represents some of the different levels of simplification possible from the generic 12 function. Opaque homogeneous materials do not require the

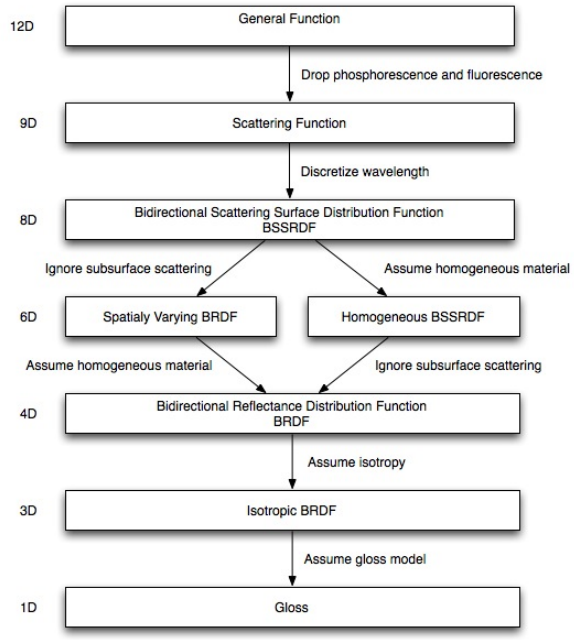


Figure 2.6: A hierarchy of reflectance functions according to taxonomy adapted from [RM]

modelling of a complete BSSRDF. They can be modelled with a simple 4D BRDF, since they do not exhibit a spatial dependency across the surface. Nicodemus [NRH+77] presented a gantry named gonioreflectometer able to place a light sensor, traditionally a spectroreflectometer able to acquire a dense reflectance sample at a single surface point or integrated over a surface area, and a calibrated light source in different angles relatively to the sampled surface material in order to obtain a complete set of samples required to estimate the 4D BRDF parameters.

The gonioreflectometer was used by Murray-Colemann et. al. 1990 [MCS90] (fig. 2.7) in his light studies and later adapted to computer graphics by Sing-Choong Foo in his master thesis [Foo97].

Ward et. al. [War92] proposed a similar device in 1992 (fig. 2.8) using image based technology instead of spectrometer to measure the BRDF. The

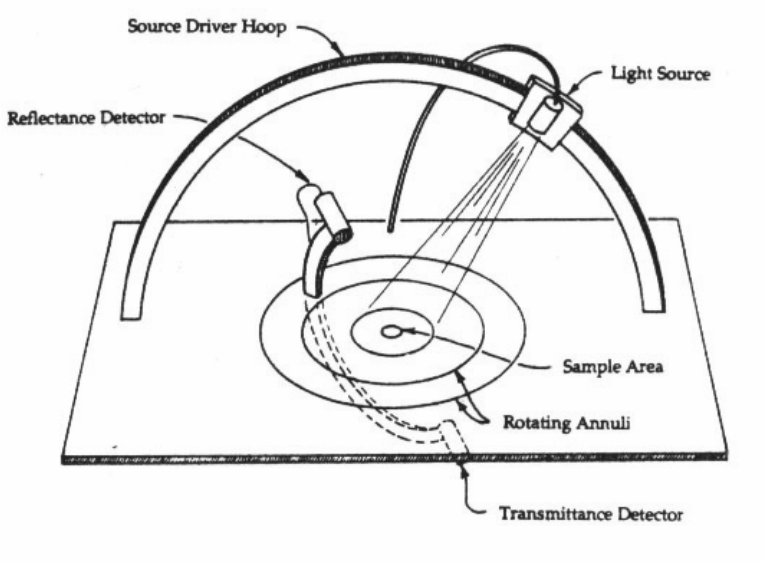


Figure 2.7: A gonioreflectometer designed by Murray-Coleman and Smith (from [MCS90])

key elements of Ward approach are a half-silvered plastic hemisphere and a CCD camera with a fish-eye lens. The combination of the hemispherical mirror and the CCD camera accounts for two of the required degrees of freedom without the mechanical constrains by capturing in a single image a wide number of viewpoints. This approach reduces the time required to capture BRDF samples; however, it requires careful calibration to allow the retrieval of the individual viewpoints. These approaches assume that the material is sampled from a flat surface; however, the approximation can be achieved by using instead a spherical surface, as proposed by Marschner et al.[MWL+99] and later used by Matusik et al.[MPBM03] to measure isotropic BRDFs.

Some authors went further in the acquisition of surface reflection properties by trying to incorporate the spatial dependency into the final representation, using image based approaches, limiting, or dropping, some of the required degrees of freedom of a complete 6D BRDF. *Light fields* and *reflectance fields* are two types of techniques that fall into the category of image-based rendering.

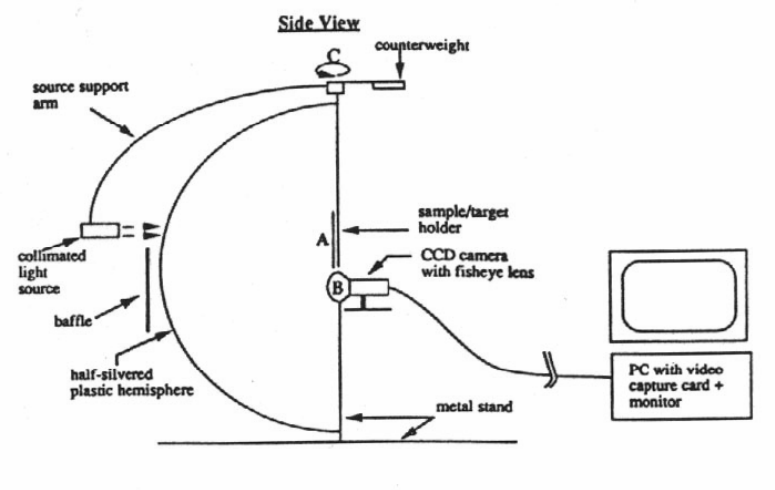


Figure 2.8: Ward's silver hemisphere reflectometer (from [War92])

Light fields became popular with the work of Levoy et. al.[LH96] and Grotler [GGSC96], and describe a set of acquisition techniques that capture the spatial dependency using various viewing angles but neglect the light source direction. The light fields [LH96] and in the lumigraphs [GGSC96] the different samples are acquired using a single light source with a fixed position in space and each image sample, captured with a CCD device, represents a different view angle of the material surface.

Reflectance fields, on the other hand, neglect the viewing direction. A reflectance field is captured using a fixed CCD device and each image register the surface reflectance of the material sample under different illumination angles. Reflectance fields where first explored by Debevec et. al. [DHT$^+$00] to capture the appearance of human faces in a device he called *light stage* where the face of an actor is placed in the centre of a dome while, a light source rotates around him. A reflectance field technique is also the basis of Malzbender et. al. approach [MGW01] presented in the next chapter.

Dana et. al. [DvGNK99] in her studies of computer vision developed a method to capture a complete sample of the 6D BRDF, which is known as *Bidirectional Texture Function* (BTF). BTF where not developed to address

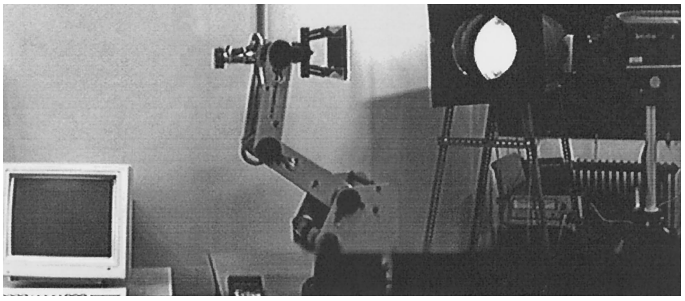


Figure 2.9: BTF capturing device at Columbia University (from [DvGNK99])

model the *Spatial Varying BRDF* (SVBRDF) or *Apparent BRDF* (ABDRF) behaviour of the surface material, but to address problems related to computer vision tasks, such as texture recognition, texture segmentation and shape-from-texture.

Several authors proposed different acquisition system for the BTF approach that can be summarised in three catagories: gonioreflectometer-like, which use some sort of mechanical gantry, a parallel version of the latter using an array of cameras and an approach using mirrors.

The gonioreflectometer-like approach is similar to the classic BRDF measuring setup, but uses instead a CCD camera to enable the sampling of the spatial dependence of the BTF. Dana et. al. [DvGNK99] was the first to present a proposal based on a gonioreflectometer-like setup (fig. 2.9). They used a setup with a fixed and calibrated light source with fresnell filters to parallelize the light, a 640x480 CCD camera, which could be moved around the sample and a robotic arm to rotate the sample in order to account for the remaining required degrees of freedom. With this approach Dana was able to acquire 205 images for each of the 61 samples of real world isotropic materials publicly available on the CUReT database [CU]. While the coarse nature of the sample acquisition is not suitable for high-quality rendering, the main goal of the measuring work done was the acquisition of texture samples for computer vision related tasks, such as texture recognition, segmentation and shape-from-texture. Although the required time for each sample was not

reported, by taking into account the time required to measure a single 4D BRDF with a similar device, it is plausible to assume that it is time consuming. A major disadvantage of the samples that are available is that they are not rectified in the acquisition stage, requiring tedious and time consuming steps to address the problem through computer vision strategies to project each image of the sample onto the same plane in order to rectify it.

McAllister et. al. [MLH02] proposed a similar but improved approach. McAllister also proposed the fitting of the acquired data to an analytical BRDF model per pixel, thus obtaining a SVBRDF. A small number of materials were sampled and are available online for research purposes.

Sattler et. al. [SSK03] replaced the low resolution CCD camera used by Dana with a 13.5 Mpx (4500x3000 pixels) camera. The system can also acquire a denser sample of the angular domain by capturing 6561 images per material sample, instead of the original 205 images. This dense sampling takes 14 hours to acquire, but the authors claim to have acquire a wider variety of material samples, which are online at the Bonn BTF database [oB]. As an improvement to the samples available at the CUReT database, the images in the Bonn BTF database are all registered and rectified.

The main disadvantage of the gonioreflectometer-like setup is the required mechanical movement, which takes time to move each individual element and to post-process each image. The solution allows some flexibility in choosing the element that will be allowed to move (light source, image sensor or material sample) as long as the required degrees of freedom are met. With this premise in mind, a system can be designed to minimize the time required for the acquisition. Sattler et. al. [MMS+04], inspired by the approach followed by Malzbender et. al. [MGW01], proposed a system based on an hemispherical array of 151 digital cameras that were mounted on a rigid hemispherical dome, which due to the parallel nature of the system reduce the acquisition time to about an hour, taking advantage of the fact that the required image rectification step is trivial since there is prior knowledge of the camera positions.

Ward et. al. [War92] used a hemispherical mirror to efficiently speedup

the BRDF acquisition. Following the same basic idea Han and Perlin [HP03] used an array of mirrors to acquire a BTF. Their system uses a kaleidoscope to obtain in the same image the required different view angles, therefore reducing the amount of time and the required image storage space. A major advantage is that the system has no moving parts and consequently the image registration step is simple; on the other hand, each acquired image is of low quality and requires image colour calibration due to the multiple reflections in the mirrors and to the fact that the mirrors are not perfect.

The acquisition of real-world surfaces using a BTF approach will lead to the generation of huge amounts of data. Dana et. al. [DvGNK99] originally used a tabular representation, and developed an interpolation algorithm to render the material from the collected samples. However, this approach is time consuming and does not cope well with real time rendering requirements. Some of the proposed analytical models, as the ones presented on the previous section, are in some extent available on current GPU cards or they are not difficult to implement.

McAllister et. al. [MLH02] proposed a solution to the compression of the BTF by fitting the data onto Lafortune BRDF model. However, to accomplish the estimation of the BRDF model parameters he had to assume that the surface meso-structures are nearly flat, although composed of different materials. The final representation of the BTF is described in a tabular form, where each element represents a point in the 2D space (texel), and for each texel a Lafortune model is fitted using a non-linear optimization method, in this case Levenberg-Marquard. Although the storage of the parameters is done on a similar method to traditional 2D texture, the major drawback is that Lafortune lobes fitting is not straight forward, specially when using Levenberg-Marquard optimization, as reported by the authors. Nevertheless due to the assumption of near flat meso-structures, this approach is restrict to materials that do not exhibit masking and shadowing. To overcome this limitation Daubert et. al. [DLHS01] proposed an additional view dependent scaling term to model the occluding effects. The additional parameter is stored in each pixel as a lookup table for each view direction and interpolated between measured viewing directions, which allows its computation on

graphics card hardware, as successfully used by the authors to render cloth materials.

Meseth et. al. [MMK03] proposed the use of light or reflectance fields to compress the BTF data. They noted that the changes in the light direction are smoother that changes in the view direction, which was already observed by Malzbender mainly for diffuse surfaces [MGW01], and therefore the signal in reflectance fields have a lower frequency than the signal in light fields, which makes them a suitable candidate to approximation by a simple analytical function. Since for each view direction the reflectance field is approximated in each texel by a simple analytical model, the authors proposed the interpolation between the measured view points using a weighted set of closes view directions reflectance fields [MMK03, MMK04].

Until now all the presented approaches to BTF compression use some form of analytical or polynomial model, that in many cases is not suitable for high frequency reflectance signals that arises from shadowing and masking. Some authors proposed a different approach based on the fact that the measured BTF data can be considered a multi-dimensional signal, where the signal processing techniques can be applied to find a more suitable basis function for the data. This approach implies the reorganization of the measured data into 2D arrays to allow the use of linear algebra techniques such as principal component analysis (PCA) and singular value decomposition (SVD). The main different between the proposed strategies differ in the way the data is distributed in the array: McCool et. al. [MAA01] applies a per texel linear decomposition on the per texel ABRDF and tries to find a set of lower dimensional factors to approximate them, while Liu et. al. [LHZ+04] reorders the array so that for each texel in the BTF the discrete ABRDF is a column in the array and then applies a PCA to the complete BTF, keeping only the first k components of their BTF approximation. The latter approach lead to a large array size, which makes the PCA approach highly computation demanding. Hauth et. al. [HEE+02] and Sattler et. al. [SSK03] proposed that the PCA approach be applied only on a view dependent basis only, therefore they only apply the PCA approach to a single slice of the BTF with fixed view direction, and consider each fixed view direction independent from

each other.

Describing the full problematic surrounding the BTF approach would be tedious and out of the focus of this work, therefore we suggest the reading of a STAR report by Muller et. al. [MMS+04] at Eurographics, which includes all the related problems with BTF acquisition, compression, decompression and rendering, as well as comparative analyses of all proposed approaches by other authors.

All of the briefly described approach to model the BSSRDF, the SVBRDF or part of it poses problems when the final goal is to create a digital surrogate that must withstand the scrutiny of the scientific method and address problems related with long term digital archive. Some rely on skilled professional to tune the parameters of the analytical models to described the acquired data, others raise problem of defining the level of information that must be represent, while others create a problem by requiring huge amounts of storage spaces that are not compatible with long-term digital preservation.

The RTI approach purposed by Malzbender et. al. [MGW01], is far from being able to represent the full reflectance properties of a surface. However, it is shown in the following chapters provides a solid mathematical background that can be easily validated, adds value to traditional photographic workflows used by CH and NS professionals and the processing workflow can be easily automated to allow repeatability and the automatic gathering of empirical provenance.

Chapter 3

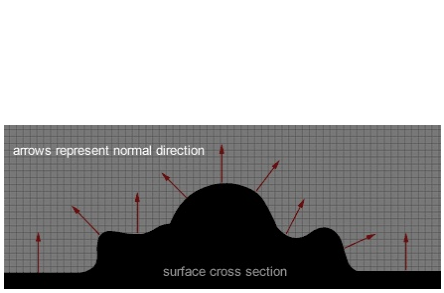
Reflectance Transformation Imaging

The surface reflectance properties of an opaque material can be modelled using a 4D BRDF associated to an infinitesimal point in the object surface. All contributing factors, such as meso-structures and micro-structures, are in some way modelled in the object geometry, in the BRDF model or through normal maps. These approaches require some user handling, which may impair the photorealistic quality of the final rendered image. The user handling do not enable the reproducibility of these type of approaches, which make them incompatible with the digital surrogate philosophy.

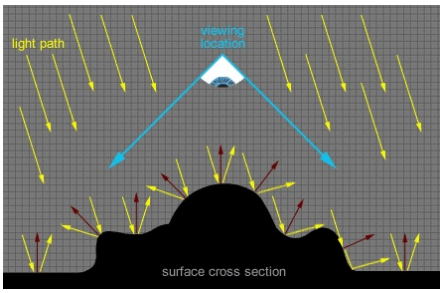
Several methods have been developed in the last decade for better measured approximation to 6D BRDF that differ in complexity and accuracy. Image based re-lighting [AS02, DHT$^+$00, DWT$^+$02, MGW01, WAA$^+$00], also known as reflectance fields, only extracts a slice of the complete 6D BRDF since the capture is limited to a set of images taken from a single viewpoint under different illumination directions. The processing, and compression, of the acquired information distinguishes the different reflectance field approaches, which range from the most simple tabular format to complex analytical models fitting.

Reflectance Transformation Imaging (RTI) presented by Malzbender et.

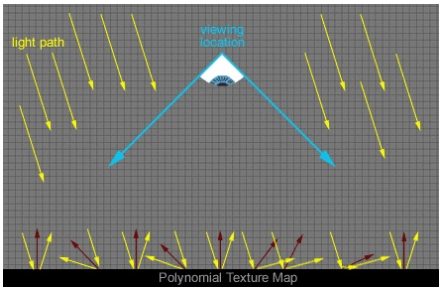
al. at HP Labs [MGW01], is one of the reflectance field techniques. The distinctness of the RTI approach lies in the compression. Malzbender proposed the use of Polynomial Texture Maps (PTM) to handle the compression of the acquired data. The PTM approach stores per pixel the fitted coefficients values of a polynomial function. Although the polynomial approach increases the time required to compress the acquired data, it reduces the decompression stage to a single evaluation of the polynomial function for any given light direction (fig. 3.1) and allows the recovery of surface information such as surface normals, which enables the development of surface details enhancement techniques.



(a) Surface cross section with normal information



(b) Sampled light-surface interaction captured in single image



(c) Final surface re-lighting using PTM's

Figure 3.1: Basic Reflectance Transformation Imaging concepts (kind permission of Mark Mudge)

The Natural Sciences (NS) and Cultural Heritage (CH) professional soon realised the potential of the technique due to the simplicity of the capturing method, the final representation format and to the amount of information that can be recovered from the polynomial function.

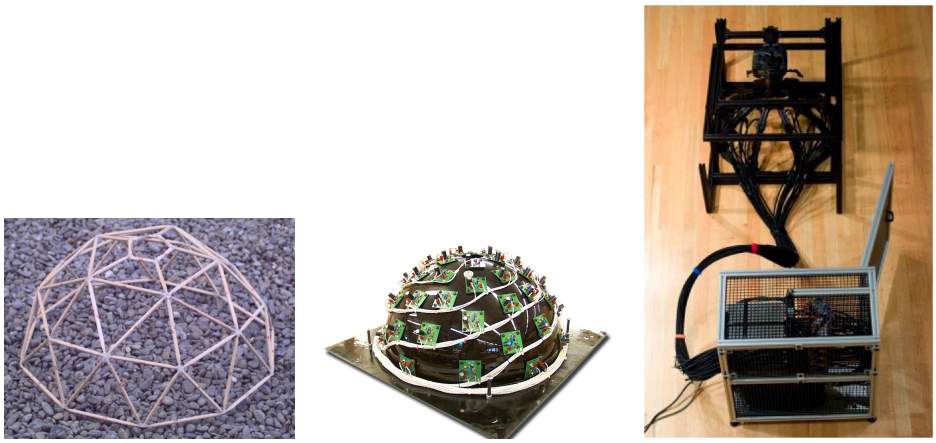
The next section is dedicated to methodology of the acquisition stage of the RTI approach, and the solution proposed for different problems related mainly with the scale of the object surface. Section 3.2 will address the PTM proposed by Malzbender to handle the compression stage of the acquired data and different application of the PTM approach. Section 3.3 targets the surface information recovery from the PTM approach and how it can be handled to enhance selected surface details.

3.1 Image acquisition

The RTI approach is a photometric technique, which requires the acquisition of multiple images of a static object with a fixed camera under varying lighting conditions, similar to the work by other authors such as Debevec in [AS02] and Georghiades in [GBK99]. In the reflectance fields the camera position does not need to be known since the view-dependent phenomena are deprecated in favour of light-dependent phenomena, but the compression stage requires as input the captured image set and the 3D vector that specify the direction of the light source for each image. Several different ways can be used to estimate the light source direction for each image, that vary in complexity, cost and precision.

The original method proposed by Malzbender uses rigid domes where the position of each light source position to the target surface at the centre of the dome, is fixed and known. Depending on the size of the objects the dome approach may have drawbacks that do not allow its use. To cope with these drawbacks, specially with medium to large scale objects, professionals have suggested that the light sources could be manually placed, and their position could be either measured during the shooting session, or could be estimated from highlights in glossy spheres or shadows cast from small sticks near the targeted surface [DHT$^+$00, DCCS06, CDMR04, MMSL06].

The dome style approach establishes a fixed distribution for light sources positions that can be later used by the compression stage. The dome style approach can to decrease the acquisition time and varies in complexity and



(a) Basic icosahedron dome (b) HP labs automated dome (c) CHI automated dome

Figure 3.2: RTI capturing domes ((a) and (b) kind permission of Tom Malzbender, (c) kind permission of Mark Mudge)

cost. The simplest dome is shown in fig. 3.2(a), which requires the photographer to manually reposition the light source to a vertex of the icosahedron after each shot. Fig. 3.2(b) shows a dome, built for the Antikythera Research Project [FBM+06b], with 50 different computer controlled light sources that dramatically reduces the capturing time. The Cultural Heritage Imaging foundation (C.H.I.) took the dome concept even further by introducing controlled wavelength light emitting equipment, allowing their dome to be used with artefacts sensitive to specific wavelength such as ultra-violet. The dome style technique was proven suitable for small objects such as coins [MVSL05], small archeological artefatcs [FBM+06b] or small fossils [HBMG02].

Debevec used a dome style approach to capture reflectance fields of human faces, which required a dome with considerable size [DHT+00]. However, cultural heritage and natural sciences are full of examples of medium to large scale object. The dome style approach in these particular cases is at the most impractical, either due to the size of the dome required or to some physical constraint posed by the artefact or the surrounding environment, such as artefact handling restrictions or the existence of physical barriers.

For such cases Scopigno et. al. [DCCS06] suggested an approach based

on a pre-planning of the light source distribution taking into account the size of the artefact to be captured and the surrounding physical constraints. This approach requires careful measuring to position the pre-computed light sources.

Dana et. al. [CDMR04] used instead a small stick to record the light source direction for each photograph. The projected shadow into a white backgrounds, was later used to recover the light direction for each sampled image. Debevec et. al. [TSE+04] introduced glossy spheres to record the highlights on the surface, for each sampled light direction.

Mudge et. al. [MMSL06] extended the idea into the RTI, recovering the light directions from highlights recorded on a black glossy sphere, placed near the object and captured during the shot session, using simple geometry based on the estimated sphere centre and radius and the relative position of the highlight to the centre of the sphere. The technique is known as *Highlight based RTI* (HRTI).

Mudge also suggested that each image could be analysed to retrieve the sphere centre and radius, along with the relative highlight location using a computer vision approach. This suggestion led to the theme of this thesis and to the development of a software package, the RTIbuilder, that automates the HRTI acquisition pipeline: search the black sphere, find its centre and radius, estimate the centre of the highlights at each sphere image, compute the light directions and generate the final RTI compressed representation.

The combination of the acquisition techniques, dome and highlight based, can address most of the problems posed by the CH and NS communities. The simplicity of both approaches, and specially the freedom to place the light sources in the highlight based approach, eases the acquisition of a fair representation of an artefact. The highlight based approach was able to take the technique from the laboratory to the field and to cope with working restrictions in remote and isolated locations.

Due to this reduced set of requirements, the RTI acquisition approach is also able to offer a low cost alternative to NS and CH professional, as it

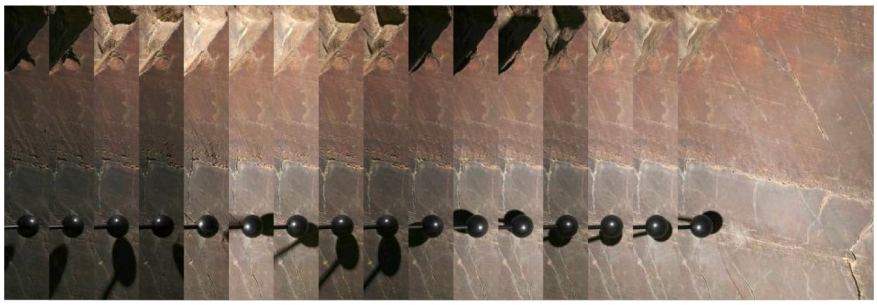


Figure 3.3: Same images of the Piscos Man set at Coa Valey Archeologic Park, images taken using highlight based RTI.

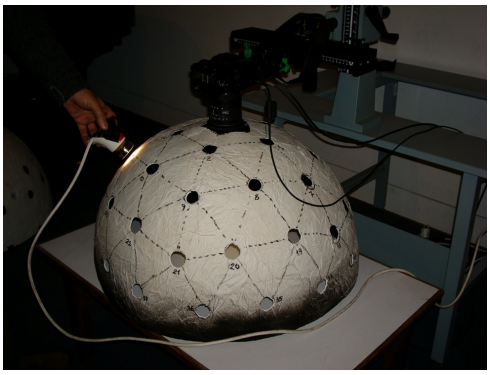


Figure 3.4: Combination of Dome and Highlight based techniques in the D. Diogo de Sousa Regional Archeological Museum.

was demonstrated in the dome built for the Regional Archeological Museum D. Diogo de Sousa. This dome was built using a "papier-mâché" technique (fig. 3.4) and the light source directions were estimated using the highlight based approach. This low cost approach demonstrates that as long as the requirements for the RTI compression stage are met, the acquisition is only limited by the imagination.

3.2 RTI compression stage

The acquisition stage generates a large set of images. With a 12Mpx digital camera these sets will require large amounts of storage space (several GB). Although a range of different compression techniques can be used [MMS+04], most of them will store the information in a tabular format, where an array containing the sampled information is stored for each light direction and pixel. The stored information can the be used to interpolate between the sample light directions to compute the colour values for any unsampled light direction. In this approach the computational intensive task is on the decompression stage.

Malzbender observed that in the samples stored at each pixel the chromaticity remained fairly constant across the sampled light directions and that the luminance accounted for most of the apparent colour change. To express this colour dependence on luminance ($L(u,v)$) at a particular point in the surface (u,v), Malzbender computes the colour of a new light direction using the following equations:

$$R(u,v) = L(u,v)R_n(u,v)$$
$$G(u,v) = L(u,v)G_n(u,v)$$
$$B(u,v) = L(u,v)B_n(u,v)$$

where $(R_n(u,v), G_n(u,v),\ B_n(u,v))$ are a computed unscaled colour in RGB format model. The RGB model was chosen since it is the colour model for most digital cameras, and the conversion to other models, although surveyed, would require additional computation resources.

The second observation that Malzbender made was that for diffuse objects: even with high spatial frequency variation, the dependence of luminance on light directions was very smooth. Combining the two empirical observations, he suggested that the compression of the acquired data could be achieved by approximating per pixel the luminance term to a bi-quadratic

polynomial function:

$$L(u,v,lu,lv) = a_0(u,v)l_u^2 + a_1(u,v)l_v^2 + a_2(u,v)l_u l_v + a_3(u,v)l_u + a_4(u,v)l_v + a_5(u,v) \quad (3.1)$$

where u and v are the spatial coordinates on the texture, and the light direction dependency is express by (l_u, l_v) that represent the light direction projected vector into the tangent plane of the local coordinates system.

The bi-quadratic polynomial function expresses the dependence of the luminance term on light direction, which were registered for each captured image during the acquisition stage. The unscaled colour for each pixel can be easly computed for each individual colour channel in RGB format, then the coefficients $(a_0..a_5)$ from eq. 3.1 can be computed by solving the following linear system:

$$\begin{bmatrix} l_{u0}^2 & l_{v0}^2 & l_{u0}l_{v0} & l_{u0} & l_{v0} & 1 \\ l_{u0}^2 & l_{v0}^2 & l_{u0}l_{v0} & l_{u0} & l_{v0} & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ l_{u0}^2 & l_{v0}^2 & l_{u0}l_{v0} & l_{u0} & l_{v0} & 1 \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_5 \end{bmatrix} = \begin{bmatrix} L_0 \\ L_1 \\ \vdots \\ L_N \end{bmatrix}$$

where N is the number of sampled images. To solve this linear system Malzbender computes the best fit in the L_2 norm using a singular value decomposition (SVD) algorithm. The solution is further simplified by assuming that the light direction variation across the acquired surface is negligible. The simplification, although introducing inaccuracies, has the advantage of only requiring a single SVD computation for a given light direction configuration, which can be applied to each individual texel, reducing dramatically the required computational time. Once the coefficients for an individual texel are determined, they are stored, along with the chromatic information (R_n, G_n, B_n), in a spatial map known as a Polynomial Texture Map (PTM).

The bi-quadratic polynomial functions will be approximated using only a discrete set of values, and consequently will smooth some light effects such as high specular spikes or hard shadowing. The smoothing effect is visible in

the final representation, but only will affect the light space, preserving any spatial dependencies even when the surface exhibits high spatial frequencies. The PTM approach will also convert any point light source into an area light source due to the light direction variation simplification across the surface.

Common sense would expect that by increasing the number of sampled light directions the accuracy of the approximation would increase indefinitely. Scopigno et. al. [DCCS06] did show that computing the difference between normal maps reconstructed from the PTM's generated from the same image set with different number and distribution of light sources, he was able to determine that with a carefully distributed light direction sampling, the threshold was found to be 80 photos. He specifically determined that a set composed of 70 to 80 images would represent the best relation between acquisition effort and accuracy.

The described PTM format stores for each individual texel the value of luminance and the computed chromatic RGB values and is known as LRGB PTM. However, for some applications it is preferable to apply the bi-quadratic polynomial to the individual colour channels and on the RGB PTM. The fitting process is applied directly to each colour channel, storing 6 coefficients per colour channel. The RGB PTM enables the use of the PTM approach on surfaces where the chromatic values varies significantly, due to self-shadowing for instance, or when used to encapsulate unidirectional behaviours, such as time-lapse or depth of focus.

Bump maps are required by traditional texture mapping to introduce small perturbations on the surface normal to achieve a more realistic *look&feel*, and Malzbender provided mathematical tools to convert then into the PTM format [MGW01]. This step is not required for an acquired PTM, since it already encapsulates the behaviour. PTM's were designed to capture the behaviour of mainly diffuse surfaces and, as seen before, the approach will smooth the specular spikes of the surface, but they can be reconstructed, or artificially recreated, using the recovered surface normal's techniques described in the next session. The original work in PTM [MGW01] also includes some transformation to allow its use to encapsulate behaviours such as anisotropic,

fresnel and off-specular effects.

3.3 PTM surface enhancement

The bi-quadratic polynomial function stores the BRDF behaviour at a particular point in the surface. The normal at this point can be computed and using the knowledge from radiometry and light surface material. Since the surface is assumed to be mainly diffuse, the surface normal can be computed from eq. 3.2 by determining the maximum luminance value of the function.

$$\begin{aligned} L(u,v,lu,lv) = \quad & a_0(u,v)l_u^2 + a_1(u,v)l_v^2 + a_2(u,v)l_u l_v + a_3(u,v)l_u \\ & + a_4(u,v)l_v + a_5(u,v) \end{aligned} \tag{3.2}$$

The bi-quadratic polynomial function describes a 2 dimensional parabola and the maximum luminance projected vector can be determined by

$$l_{u0} = \frac{a_2 a_4 - 2a_1 a_3}{4a_0 a_1 - a_2^2}, \forall_{a_0,a_1,a_2} : (4a_0 a_1 - a_2^2) > 0) \wedge a_0 < 0 \tag{3.3}$$

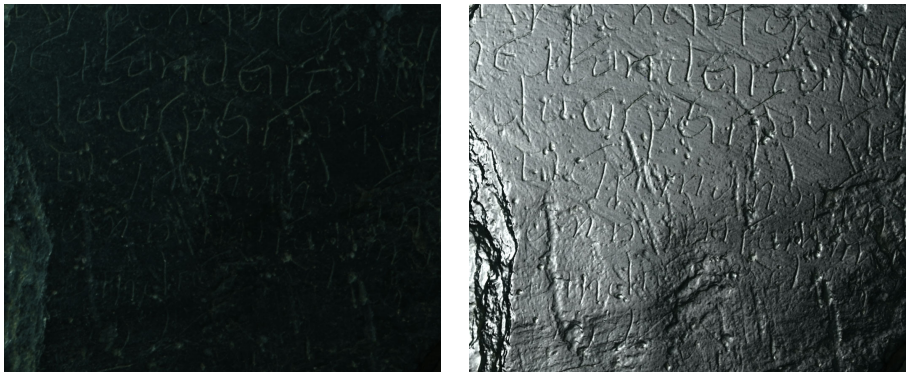
$$l_{v0} = \frac{a_2 a_3 - 2a_0 a_4}{4a_0 a_1 - a_2^2}, \forall_{a_0,a_1,a_2} : (4a_0 a_1 - a_2^2) > 0) \wedge a_0 < 0 \tag{3.4}$$

where l_{u0} and l_{v0} are the projected values of the surface normal, then normal can be recovered as

$$\vec{N} = (l_{u0}, l_{v0}, \sqrt{1 - l_{u0}^2 - l_{v0}^2}), \forall_{l_{u0},l_{vo}} : 0 \leq (-l_{u0}^2 - l_{v0}^2 \leq 1) \tag{3.5}$$

The ability to recover the surface normal field enables the artificial reconstruction of the specular information using for instance the Blinn-Phong analytical model. Specular enhancement enables the observer to discover features that can not be seen with regular light condition, as shown in fig. 3.5 where the inscription on the bottom of the plate can be easily seen and the inscription has a better overall readability when using this enhancement.

The second enhancement method proposed by Malzbender, is to modify the coefficients, by computing new values to increase the directional sensitivity



(a) Normal re-lighting of the PTM (b) PTM specular enhancement

Figure 3.5: Detail on a schist plate (from the e D. Diogo de Sousa Regional Archeological Museum collection)

of the surface to light, while maintaining the normal and the albedo colour unchanged for the maximum luminance. The new luminance coefficients can be computed using:

$$\begin{aligned} a'_0 &= ga_0 \\ a'_1 &= ga_1 \\ a'_2 &= ga_2 \\ a'_3 &= (1-g)(2a_0 l_{u0} + a_2 lv0) + a_3 \\ a'_4 &= (1-g)(2a_1 l_{u0} + a_2 lv0) + a_4 \\ a'_5 &= (1-g)(a_0 l_{u0}^2 + a_1 l_{v0}^2 + a_2 l_{u0} l_{v0} + (a_3 - a'_3) l_{u0} + \\ &\quad (a_4 - a'_4) l_{v0} + a_5 \end{aligned}$$

where g is the gain factor used to increase arbitrarily the Gaussian curvature of the parabola. Fig. 3.6 shows the application of this method on a 3000 year old funerary statuette from ancient Egypt, called an *ushabti*.

Malzbender also realised that the reconstruction from the PTM was not bound by the physical light constrain ($\sqrt{l_u^2 + l_v^2} \leq 1$), which could allow the extrapolation of the reflectance model beyond what is geometrically achiev-

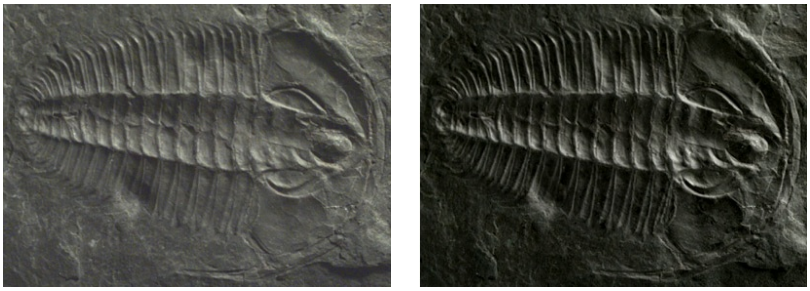
(a) Original photograph (b) Reconstruction from PTM (c) Diffuse gain (g=1.9)

Figure 3.6: Increasing the diffuse gain on a 3000 year old Egyptian funerary statuette, using same light directions (kind permission of Tom Malzbender)

able. Although the enhancement has no physical equivalent, Malzbender realised that we could use light directions that are more oblique that physically possible improving significantly the overall contrast of the image as shown in fig. 3.7.

3.4 Conclusion

Reflectance Transformation Imaging is a technique from the reflectance fields that shares most of the limitations posed by other related image based relighting techniques. Due to its approach for data acquisition, compression and related re-lighting enhancements it provides a set of useful tools to the end user, specially to CH and NS professionals.

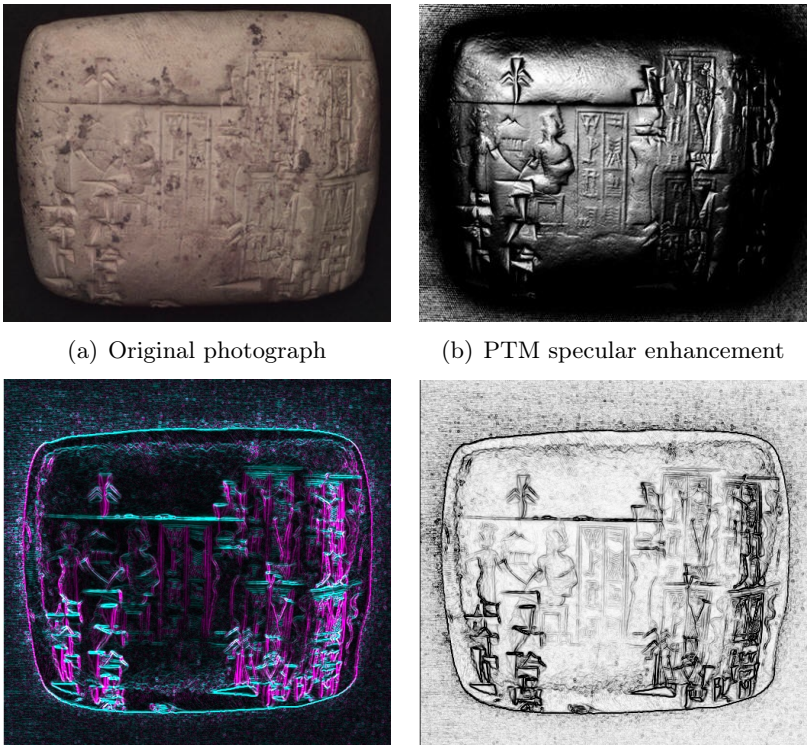


(a) Original photograph (b) Reconstruction from PTM using a extrapolated light source direction

Figure 3.7: Light Direction Extrapolation on a trilobite fossil (kind permission of Tom Malzbender)

The acquisition techniques are able to cope with CH and NS professionals requirements and constraints without imposing significant modifications to their usual workflow. The RTI technique is still based in photographs, and the requirements to use any of the proposed acquisition approaches do not require a long learning curve, as proven during the workshops we gave for the Portuguese CH community. The participants of those workshops demonstrated a particular interest in the highlight based acquisition approach due to its applicability on the field, removing the need for a controlled environment and the ability to cope with the most common restrictions such as cost, physical barriers and time requirements.

An algorithm and applied techniques based in Hemispherical Harmonics (HSH) was developed and presented by Wang and Prabath et. al. [MMC+08] that improves the approximation quality and tries to incorporate the view dependency. However, this approach is not stable yet for deployment and the publication of its results is expected soon. Until then the bi-quadratic polynomial approach use by the PTM is still one of the best compression techniques when compared with other reflectance field approaches. The PTM approach shifts all the computational intensive task to the compression stage, reducing the decompression stage to a simple evaluation of a polynomial function. It also offers a collection of useful enhancement functions that increase and diversify the potential applications of the techniques.



(a) Original photograph (b) PTM specular enhancement

(c) Sobel filter applied to normal field (d) Inverted Gray-scale Sobel filter applied to normal field

Figure 3.8: Exploratory usage of edge detection algorithm on the normal fields of a 4000 year old neo-Sumerian tablet (Original PTM courtesy of Tom Malzbender)

The simplicity on the decompression stage combined with the proposed enhancements offers the end user a collection of exploratory tools that go beyond a computer graphics application. As an exploratory work, computer vision algorithms were applied to enable the enhancement and identification of surface features, which demonstrates the potential of the RTI technique to the application of some traditional image filters to enhance the visualisation capabilities. Fig. 3.8 demonstrates the use of edge detection filters (Sobel filter) for surface feature enhancement, extraction or recognition.

RTI is one of the simplest and easiest image based re-lighting techniques that can be used, and due to the process used to generate the final represen-

tation it has the potential to become a scientifically robust digital surrogate. Specially the highlight based approach is able to record all the required information to allow the validation of the acquisition process: light directions are record on a sphere that is in the image space, the camera information (e.g., camera model, serial number, lens used) is stored by the camera on the raw format in some way (e.g., EXIF, XMP) and the compression process is well documented. However, the process still requires some level of automation to detect and record all the required data, process and empirical provenance information to withstand the scientific method scrutiny.

Chapter 4

RTI Digital Surrogate

The digital age poses new challenges to the NS and CH professionals: to trust and use data collected by other scholars, they must be able to inquiry how the data was collected and how it was processed; only if the given answers are reliable and repeatable they use the data for is own research.

This notion of inquiry, reliability and repeatability are the basic concepts underneath the scientific methodology. The digital age offers a high level of freedom that may act as a double edge sword: on the one side it offers a new insight to the real world and on the other it allows the data to be processed to express, or convey, a single idea or point of view.

Most of the lost information is a consequence of the data processing workflow: in most cases the user is required to participate and to some extend express its personal opinion over some factor. A study [BFMS05] conducted in the most prominent US museums and libraries, which addressed digital imaging workflow, concluded that all workflows added some kind of artistic and personal retouch to the acquired images. The study concluded that this personal touch, for instance the colour calibration, when manually done and not by an automated process, destroys part of the original information. The study strongly suggests that only a complete automated process can originate a reliable digital representation with a reproducible workflow.

The long term digital preservation is another challenging problem. The digital technology creates new and enhanced formats and deprecates others at incredible rates. This rate of change in the digital formats lead software developers to deprecate formats in favour of new and improved ones, which for long time digital archiving meant that a chosen digital format could lead in just a couple of years to a lack of software able to read it . The problem is increased by the fact that most of these formats are proprietary and lack the proper documentation to build a viewer or a converter when the current software is no longer available. The problem could be address by constant refreshment and update of the data to new formats, but the solution would only be practical for very small digital archives.

The challenge posed to the computing science community, and to the scientific community in general, is how to produce a digital representation that can act as a digital surrogate and how the problem of long term digital preservation can be addressed. The common consent is that the solution must incorporate all the methodology behind the scientific process, meaning that the original data must be maintained unchanged, all the processing steps must be logged and any bias information must be fully disclosed. These features create pressure on the research teams to develop tools that are able to give a positive answer to these concerns.

4.1 RTI for a digital surrogate

The RTI technology is in essence a photographic 2D approach that can convey, and retain, true 3D information into its final representation format. If we look closely to the RTI acquisition, processing and compression we will find some of the fundamental steps required for RTI to become a digital surrogate.

The acquisition process requires a set of rules that can be reproducible and can be evaluated in terms of accuracy and quality of the acquisition and adequacy to the subject.

Digital cameras store relevant information on the images that they generate. Current equipments are able to automatically store the camera model,

the type of lens used, the camera parameters at the time of acquisition such as aperture, time of exposure, white balancing, time and date of capture and a wide range of other useful parameters without user intervention.

Together with the camera information the acquisition process should also record for each captured image the light source data, including its position. The storage of light source position enables any scholar to evaluate if the light distribution was adequate to the targeted object surface, and if needed it allows the repetition of the acquisition for comparative evaluation. The original dome style acquisition approach or the acquisition approach proposed by Scopigno et. al. [DCCS06] assumes a constant distance between the object and the light source and only requires the light direction information to be stored in a separate log file, equivalent to a traditional *lab notebook*, and relies on the accuracy of the photographer to register the orientation of the dome or the correct positioning of the light sources. The highlight based acquisition automatically computes and stores the light source direction from the highlights on the surface of a glossy sphere placed next to the object. The light source direction, or position, can be computed using the approach proposed by Mudge et. al. [MMSL06].

The RTI final generation process only uses the acquired information, without destroying it, to generate a new and richer representation, which is computed using a well described and documented mathematical method. These two combined features allow scholars to reproduce the experiment, evaluate the reliability of the representation or use the raw data for new approaches. The fact that the RTI approach is well documented and relies on a mathematical approach makes it also suitable for long term digital archival. Although the currently available format, the PTM, is property of HP Labs and as such not adequate for long term perservation, new formats are been proposed as open source and open standards. The new open-source HSH format being developed at the University of California Santa Cruz by Wang and Prabath et. al. [MMC+08] enable the use of RTI by a wider audience and ensures that the software can be maintained or reproduced at any point in time.

4.2 RTI and Empirical Provenance

The RTI technology is able to address some of the problems posed by the scientific methodology and long term digital preservation. However, the generation process poses some problems that must be addressed in an efficient way.

Traditionally, the scientific inquiry starts by systematically gathering empirical data from the real world and the methodology requires that subsequent processes to be fully documented. The scientific information registered in the lab notebook is known as provenance and explains where the information came from and how it was processed. The provenance must always be recorded with the results, otherwise it would not be possible for another scholar to verify the results and assert the level of the quality of the information.

For a digital surrogate to be widely adopted, it must pass the test of the lab notebook. The concept of Empirical Provenance (EP) presented by Mudge et. al. [MMC+08] is for digital surrogates the equivalent to the lab notebook for non-digital representations, and records the journey of the original untainted empirical data from its initial capture all the way throughout the processing pipeline (fig. 4.1) to its final form as a digital surrogate. As shown in fig. 4.1, the generation process may require image processing priorly to the RTI building stage.

For a given digital surrogate the empirical provenance attributes depend on the tools and methods used to build it. For a single digital photograph the empirical provenance information can be found on the EXIF or XMP data such as: the camera manufacture and model, firmware version, the information required to convert the raw data into image; and in the XMP all the editing operations like cropping, re-sizing or sharpening.

The operations performed during this stage may have a significant impact in the final RTI representation and therefore must be registered. Tools such as Photoshop, from Adobe, are able to store the editing operations in a open standard XML schema for images called XMP, which can be used to gather

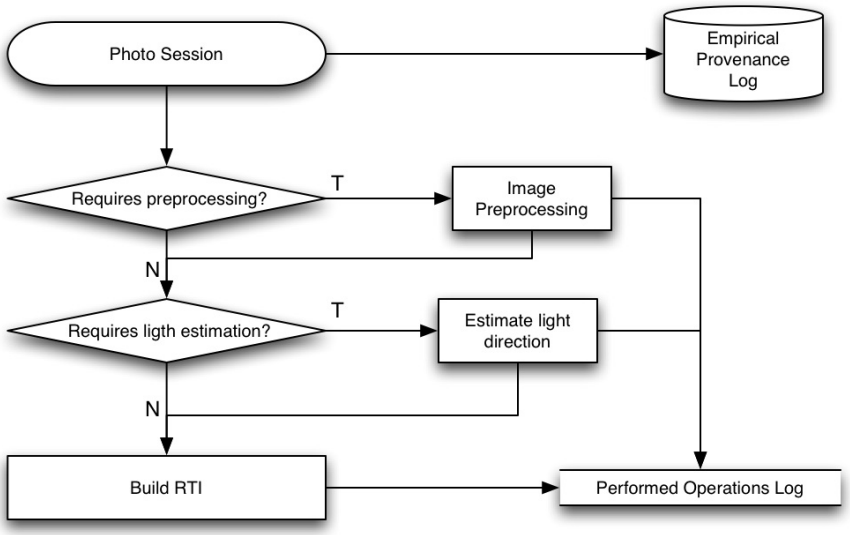


Figure 4.1: Workflow of the pipeline required for generating RTI

the information required for empirical provenance and can incorporate other XML schemes like OASIS [Cov00] or CIDOC-CRM [Doe05] that are already being used in long term digital archival.

Adobe also developed a new open source image format, the Digital Negative (DNG), which represents a major step forward to empirical provenance gathering. The DNG is able to store the original image in a TIFF format along with the XMP data, as well as the image in the camera manufacturer raw format.

Although some acquisition parameters can be taken from XMP metadata or from the EXIF information stored in the raw data by the camera, they only serve the empirical provenance concerns. The RTI building stage depends on the accuracy of the light directions given for each photograph. For dome style acquisition approach or other similar techniques where the light position is fixed, or pre-computed, the accuracy of the description will depend on the careful alignment of the dome with the surface and the camera. The approach proposed by Scopigno, et. al. [DCCS06] adds extra complexity by requiring

careful measurement of the light position placement. In both cases if the data is carefully described it will meet the requirements of the scientific method allowing the verification of the quality of the capture, but still requires human intervention in parameters that can influence the final RTI result.

Highlight based acquisition approach addresses the problem by automatically recording the light source direction, or position when more than one sphere is used, as an highlight on the surface of the sphere. Since the sphere surface is captured in the same image space as the subject, then the need for careful hand annotation is no longer required and the light direction estimation workflow can be automated. [BSP07], extended on chapter 5 of this thesis, automates the workflow of the post-processing steps required to estimate the light sources directions using computer vision algorithms to: estimate the sphere centre and radius; to detect and estimate the highlight centre; to compute the light directions following the approach in [MMSL06] and to call the final RTI building stage. This automation removes the dependency from the NS and CH professionals and enables the acquisition of empirical provenance information allowing the reproducibility of the process.

The final stage, the RTI creation, relies on a well documented mathematical process, which allows it to be verified by other scholars. Since the process is mathematically based and all the possible operator handling lay on the input parameters to the mathematical formulation, it does not require the additional empirical provenance information other than the full discloser of the input data and the used mathematical method.

4.3 RTIbuilder Specification

The computation of the PTM function requires prior knowledge of the light source position for each of the captured images for a given set. Accurate positioning of a large number of light sources is a hard task and suggests the help of additional tools, such as the use of rigid domes to evenly distribute the position of light sources, or software tool to schematics to help the positing of the light sources during the shooting session.

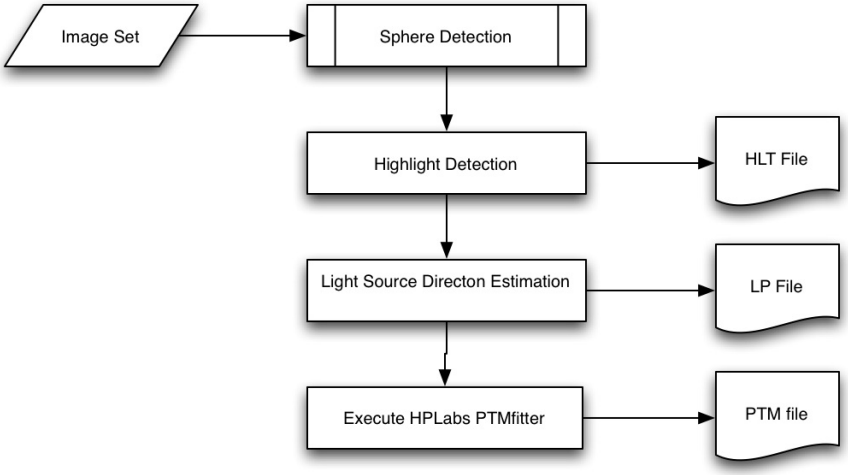


Figure 4.2: Workflow for Highlight based RTI

Figure 4.3: Highligh based capture session in the Idol Fountain in Braga, Portugal. The light source is manually positioned with the help of a string to keep the distance to the centre of the captured inscription constant across all images.

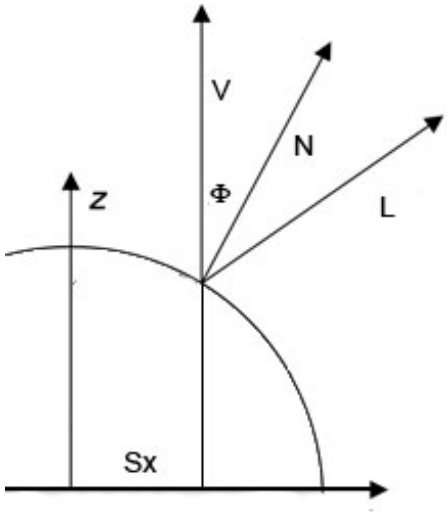


Figure 4.4: Estimation of the light direction from the highlight centre on the sphere surface, where V points to the viewpoint and L to the light source [BSP07].

The highlight based workflow proposed by Mudge et. al. [MMSL06], uses a different approach to the problem (fig. 4.2). Mudge placed a black glossy sphere next to the object and captured the sphere with the object surface. The light source is positioned and since the PTM approach assumes that the distance to the light source is constant, he used a string to keep same the distance across all images (fig. 4.3).

The light source direction is recorded as an highlight in the surface of the sphere. To retrieve the light source direction Mudge uses the geometry of the sphere and assumes that the reflection is well approximated by the Lambert reflection law. Fig. 4.4 suggest a possible approach to compute the light source direction $\vec{L}$ using the highlight coordinates (H_x, H_y) and the sphere centre (C_x, C_y) to estimate the projected surface normal vector at the

highlight centre:

$$S_x = \frac{dx}{r} = \frac{(H_x - C_x)}{r} \tag{4.1}$$

$$S_y = \frac{dy}{r} = \frac{-(H_y - C_y)}{r} \tag{4.2}$$

Now, the unitary surface normal vector can be reconstructed using,

$$(S_x, S_y, \sqrt{1.0 - S_x^2 - S_y^2}) \tag{4.3}$$

The normal vector azimuth can be computed by from cartesian to spherical coordenates by computing

$$\Phi_N = cos^{-1}(\sqrt{1.0 - S_x^2 - S_y^2}) \tag{4.4}$$

which gives the angle between the sphere surface normal ($\vec{N}$) and the view direction ($\vec{V}$). The Lambert reflection law states that the angle between the light source $\vec{L}$ with the surface normal $\vec{N}$ is the same as the angle between the surface normal and view direction $\vec{V}$ assuming that the viewpoint is in the infinite. Therefore, the azimuth of the light source direction is:

$$\Phi_L = 2.cos^{-1}(\sqrt{1.0 - S_x^2 - S_y^2}) \tag{4.5}$$

The Lambert reflection law also states that $\vec{L}$,$\vec{N}$ and $\vec{V}$ are co-planar, which allows the estimation of the remaing spherical angle to a simple computation of

$$\Theta_L = sin^{-1}(y)/sin(\Phi_L) \tag{4.6}$$

The corresponding cartesian 3D vector of the light direction can be obtained by simply computing

$$(sin(\Phi_L) * cos(\Theta_L), sin(\Phi_L) * sin(\Theta_L), cos(\Theta_L)) \tag{4.7}$$

The original process of estimating the light source direction depended on

the user skills to find the centre and radius of the sphere in the image and the centre of the highlight.

However, this estimation can be automated using computer vision algorithms. It has the advantage of transforming the workflow into a non user-dependent and repeatable process, and with the automatic gathering of empirical provenance. This led to the development of a software application to track the light source position, which generates the required files to feed the polynomial fitter application, the PTMfitter (from HPlabs). The integrated platform containing the LPtracker, the PTMfitter and a PTMviewer was later developed and deploy at HP labs website on the PTMbuilder. This platform later evolved into a generic RTIbuilder, containing enhanced versions if the previous LPtracker and supporting more external fitters (PTM, HSH) and viewers.

RTIbuilder aims to fully automate the workflow in fig. 4.2 hiding from the user the configuration parameters required by the vision algorithms or converting them into a language that both NS and CH communities can easily relate to. The application will also be compliant with hardware and multiple operating systems that both NS and CH communities use.

The RTIbuilder automates the following operations:

Sphere detection : to detect one or more spheres in the image and to estimate its centre coordinates and radius, while logging all the images used for the detection algorithm; plus the algorithms, and associated parameters, to detect the sphere and any additional handing performed by the user during this stage.

Highlight detection : to detect the highlight in a sphere, record the highlight absolute centre coordinates with the image file name into a text file for future reference. It must also record all empirical provenance such as algorithms used, algorithms parameters and any user intervention on the process if it occurs.

Light direction/position estimation : as a first simple approach it estimate the light source direction for each image, using the centre and

radius information estimated in the first stage and using the approach proposed by Mudge. The estimated coordinates and respective image file name must be recorded in a text file named LP file to be used by the HP Labs PTMfitter.

To allow the democratisation of the workflow it also must be based on open-source software to allow the contribution of the community and remove the dependency on a single developing team.

As is described in the next chapter, the RTIbuilder began as a simple tool to automated the Highlight RTI approach and evolved in time with the contribution from demonstration, workshops, training and cooperation with the Portuguese CH community and from similar events promoted by C.H.I. foundation.

4.4 Conclusion

As pointed out by Mudge et. al. [MMC+08] a digital surrogate to be widely used by the NS and CH community requires a "great level of simplification, cost reduction, increase easy of use and its compatibility with the current working culture".

The digital photography skills required by the RTI technique are already mastered by the NS and CH professional. To achieve the required level of expertise for RTI acquisition only a small amount of training is required. The RTI acquisition process does not need any special equipment other than the one already in the professional digital photography kit used by the NS and CH professionals. The combination of the two previous factors drastically reduces the cost of implementing a digital workflow for the creation of digital surrogates.

However, the wide adoption of digital workflows is only possible if a new family of image-based empirical data acquisition tools emerges offering automatic post-processing capabilities. The RTIbuilder software, developed during this dissertation work, is one of these tools and is the key element

to convert the RTI based digital representation into digital surrogates. The software, developed using a open-source approach, automates almost all RTI generation pipelines and collect the required empirical provenance information. Other tools developed by the international consortium lead by the C.H.I. foundation, which address both ends of the pipeline, acquisition, final RTI format and retrieval and mapping of the empirical provenance information gather by the application to long term digital archival standards such as OASIS and CIDOC-CRM.

Chapter 5

The RTIbuilder

Reflectance Transformation Imaging (RTI) has been successfully used to document cultural heritage information, from small to medium size objects, recording a significant amount of 3D information that convetional photography is not able to capture. We developed a set of techniques and software tools to significantly speed up the RTI generation and to record the required empirical provenance information.

The original LPtracker tool, developed in collaboration with HP Labs and C.H.I. foundation, to successfully automated the pipeline that was proposed for Highlight based RTI (HRTI). It was the first tool to automate the RTI generation workflow, making it reliable and reproducible. The LPtracker, presented in [BSP07], automatically detects the light source direction from a set of previously captured images using the HRTI approach: it uses image processing techniques coupled with spatial geometry to find the accurate positions of spheres placed next to the object and each highlight on its surface; from these coordinates it computes the direction of the lighting positions, as required by the RTI generation stage.

The application was deployed as an open-source tool and tested worldwide by CH professionals. The application evolved to a more complex and elaborate application by incorporating all the feedback gathered during our collaboration in CH running projects, from workshop we gave to the national

Figure 5.1: The automation pipeline for Highlight RTI

CH community and from similar events promoted by the C.H.I. foundation in Europe and U.S.A.. The software, now named RTIbuilder, incorporates all the features from the original LPtracker and adds more workflow flexibility and image pre-processing capabilities.

Also the original LPtracker was a monolith application, which limited the development of other features since it was required to change the core of the application to include them. The new version, the RTIbuilder, uses a plugin approach to the problem providing mechanisms to achieve *loose coupling* between the different components that can be used in a single workflow. This developing strategy addresses two major issues: the continuos improvement of each component by other teams without the need to rewrite the whole application and the communication schema used to exchange information between the different components, which can be used to record also empirical provenance information.

The next sections describe the methodology used to develop the RTIbuilder and describe some of its key components.

5.1 Automation of the HRTI workflow

The HRTI workflow has a straightforward path, as show in fig. **??**. To extract the light source position using the strategy suggested by Mudge [MMSL06], we first need to estimate the position and size of the sphere placed next to the object and the geometric centre of each highlight as seen at the sphere surface.

(a) Near the Piscos Man engraving

(b) Near Fariseus 1a rock plate

Figure 5.2: Samples of captured spheres with object surfaces

5.1.1 Sphere Detection

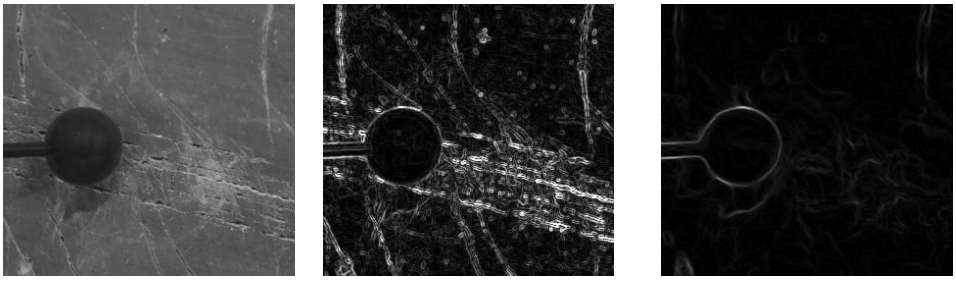
Detection of black spheres in images can be a challenging task in open environments, manely those

The detection process assumes that all images were taken with a fixed camera, e.g., all objects are represented by the same set of pixels, and it is based on three steps:

1. remove the sphere shadows
2. detect the object edges
3. find the sphere area and/or edges, and estimate geometric centre and radius

The complexity required to detect a dark ball of the process is due to the variety of conditions under which the image may be acquired, as these can not be controlled, particularly when operating in open environments.

The HRTI approach as described by Mudge uses a black snooker ball, which due to the light source direction will project a shadow that can be confused with the ball itself during the segmentation phase. The RTI generation relies on a set of images where the light source is uniformly distributed, and consequently the shadow size and direction will change from image to image.



(a) Median image of the complete image set (b) Sobel operator before second median filter (c) Sobel operator after second median filter

Figure 5.3: The first steps for sphere detection in image set from Piscos Man

To take advantage of this phenomena we compute the median for each pixel using all images in the set. This median filter serves two different purposes: helps to remove the projected shadows of the sphere and the noise generated by the highlights in the sphere surface (fig. 5.3(a)).

The resulting image cannot yet be used for sphere detection since most of the targeted surfaces will present significant edges to produce enough noise to interfere with the process (fig. 5.3(b)). To remove part of the noise while preserving the sphere edges an additional 7x7 median filter is applied to the previous image. The additional median filter performs well since the area of the sphere is composed of an homogenous material and presents better results than the traditional approach based on low pass filters (e.g. Gaussian) by not smoothing the sphere edges (fig. 5.3(c)).

A gradient operator (Sobel) is applied to the resulting image, to compute the edge gradient value and simultaneously the direction of the gradient. Both gradient value and direction were later used to speed up the approximation of sphere shape.

In last and final stage, the Hough Transform estimates the sphere center and radius. The traditional approach uses a binarization step prior to the application of the Hough Transform, which to contours thicker than one pixel wide. Although the problem could be solved using a thinning algorithm, a second problem arrises from the binarization technique: the threshold value

will make most of the sphere contour disappear. An accurate detection of the ball centre is crucial to generate accutare RTI representation. To improve the output quality of the sphere centre we compute the Hough transform directly from the gradient image. We use the gradient value to accumulate in the Hough space, as an alternative to the traditional process, which increases by 1 unit the Hough space corresponding to each pixel. Still to create Hough space we need to project each edge pixel in all direction, which makes the process time consuming. Since the Sobel operator is also able to compute the gradient direction, we use it to project the edge point along an arc of $\pi/6$ on that direction decreasing the number of required projections for large radius.

The radius of the sphere is unknown requiring the application of the Hough Transform to the edge contour image with different radius parameters to determine the best centre and radius approach. It is traditional to solve this problem using a 3D array to track position (x,y) and radius. However the computation of the Hough Transform is independent for each given radius, which allows its computation in parallel using a thread to compute for a specific radius and determine the best centre and weight. By gathering the maximum values in an array we can determine the best centre and radius that describes the sphere.

5.1.2 Highlight Detection

The light source direction can be geometrically estimated from the images coordinates of the sphere centre, radius and the centre of the highlight on its surface. Since the ball remains at a fixed position across the image set (only the light position changes across images), the region where we need to search for the highlight can be limited to a circle containing the sphere. The process is then reduced to the detection of potential highlight candidates (more than one when the lights reflect from other surfaces), the selection of the correct highlight candidate and the computation of its centre.

We start the process by removing all image information outside the circle containing the sphere. This step removes all possible interference from information in the sphere background. Now we need to identify all possible

highlight candidates, which is achieved using a threshold algorithm to remove all the information from pixels with a grey colour value lower than a threshold (e.g. 80%) of the brightest pixel value in the image. This step removes the interference from reflections from other surfaces, since the light source will be recorded in the surface with an high brightness value. Several bright regions can appear due to inter-reflections and after threshold they appear isolated with a high brightness value.

A modified version of a labelling algorithm is applied to the image. The modified algorithm uses the traditional approach to identify and isolate each of the remaining bright areas. We use the algorithm also to compute the centre of mass of the highlight shape. We computed centre coordinates by using a weighted average of all pixel in the highlight area, using the grey scale value of the pixel as weight. The final computed value is not a integer value, which is not rounded to provide a more accurate centre position.

Whenever more than one potential highlight region is detected, we assume, based on experimental settings, that the highlight that directly reflects the light source has its centre closer to the sphere centre, since fake highlights due to inter-reflexions are always farther from the centre.

During this stage we also compute an image containing all the highlights with all images in the set. This image quickly gives an overview of the light distribution used during the capture session, allowing a visual evaluation.

5.1.3 Light Position Estimation

The quality of the RTI implementation relies heavily on the accuracy of the estimation of the light source position for each shot, as well as on an uniform distribution of the light sources. The computation of the 3D vector that describes the light source direction or position using the HRTI approach, is computed by a geometric analysis of the relative positions between the centre of the sphere and the centre of each highlight, as proposed in [MMSL06] and described in the previous chapter.

5.2 The RTIbuilder implementation

The original work on the PTMbuilder focused mainly on the LPtracker, which aimed to produce an automated tool to estimate the light position from its highlight on the surface of a glossy sphere, placed next to the object and captured in each image.

The PTMbuilder was designed as a monolith application, where each different stage of the pipeline was hard-coded into the application. This approach produced a stable software package for highlight-based RTI. The PTMbuilder also generates and records some of the required empirical provenance data, especially the data related to the processing pipeline [BSP07].

To add new functionalities to this implementation, they must be hard-coded into the application. By separating each stage of the pipeline into a plug-in component and letting the user specify the required features and the processing pipeline, the application gains more flexibility and code maintenance is reduced, as long as the user respects the pipeline inter-stage dependencies.

A Service Oriented Architecture (SOA) approach was proposed as an architecture style whose main goal is to achieve *loose coupling* between interacting software agents. A SOA service is an independent unit of work, executed by a service provider, to obtain a specific result, to be used by a consumer service. Both provider and consumer roles are played by software agents on behalf of the owner or a user. To achieve *loose coupling*, SOA infrastructure relies on the data exchange between different software agents, by implementing a well-defined message-passing schema. The messaging passage schema will have to ensure two main goals: (i) to coordinate the work among the different agents and (ii) to exchange the required data among them.

The RTIbuilder software package was developed following a SOA like approach, that incorporated all the features developed in the LPtracker application and enabled the definition of new workflows. The RTIbuilder borrows the base principle of *loose coupling* from SOA approach and uses it achieve more flexibility in terms of workflow specification and components development.

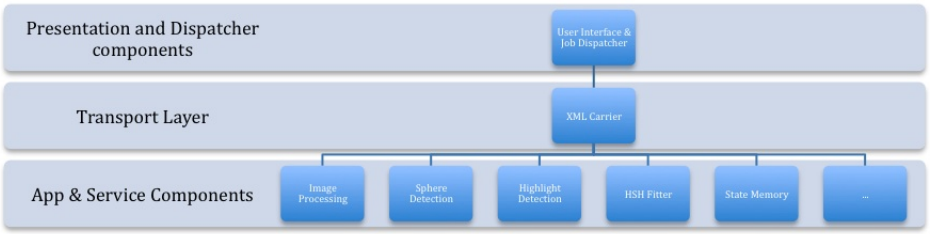


Figure 5.4: The RTIbuilder architecture

The loose coupling in the RTIbuilder is achieved using an underlined communication mechanism between each of the software agent, enabling developers to replace, add or produce new components without having to rewrite the whole application.

Fig. 5.4 gives an overview of the architecture of the RTIbuilder, as described in the specification document [BP09a]. The presentation and dispatcher layer contains the components that interact with the user and dispatch he/she request to the computational modules. The transport layer is responsible to establish and define the communication mechanisms between the application, the user interface and the computation modules. The application and service components layer also contains; the computation components that provideuser interfaces to be used by the components in the presentation layer.

Both, LPtracker and RTIbuilder applications, were developed in Java. Java provides platform independence, which was a crucial requirement for a fast and successful deployment of the applications, and the Java SDK provided most of the required API to develop the applications.

Next subsections describe each of the implemented components for the RTIbuilder and how they interact with each other to produce a valid workflow.

5.2.1 XMLcarrier component

The XMLcarrier is a component of the transport layer of the application, which is responsible for the communication and data exchange between the different components of the system, using a well-defined communication schema.

The XMLcarrier manages a library of methods that are shared by all components to be integrated or designed to work with the RTIbuilder application. The XMLcarrier implements methods to add, extract and validate information, by each component, in the data to be transmitted between them. Although each component may implement its own parser and write for the XMLcarrier schema, the library provides generic methods and ensure compatibility with upcoming updates to the current schema.

The current XMLcarrier schema was developed to describe the final RTI file generation process including the parameters, results and log information for each of the executed components and user defined metadata.

XMLcarrier is organized in five main sections:

Header : contains project information (e.g.: project name, time data, user data, common image metadata, user defined metadata);

File : information about the images used on the generation process;

Process : process pipeline definition and description;

Data : parameter and results description for each defined process;

Log : log information from each component.

A more complete description of the XMLcarrier schema, and how each component stores or retrieves information, is in [BP09b].

The use of XML schemas to handle the inter-layer communications, enables the storage of XML message to a file, which will contain all the required information to ensure that the process is repeatable, reproducible and additional empirical provenance information can be stored for future query.

A major advantage of using XML is that it can be easily parsed by another application allowing the extraction of information for long term digital archiving, using standards such as OASIS [Cov00] and CIDOC-CRM [Doe05].

5.2.2 Sphere detection component

The sphere detection component uses image processing algorithms to detect one or more spheres in an image set.

The component provides a user interface to be used by the presentation component, which wraps all the required user interaction data, inputs and validation, that are required by the computational model. The user interface is similar to the ball detection stage in the LPtracker pipeline.

The user interface contains two panels (fig. 5.5), which with three different areas.

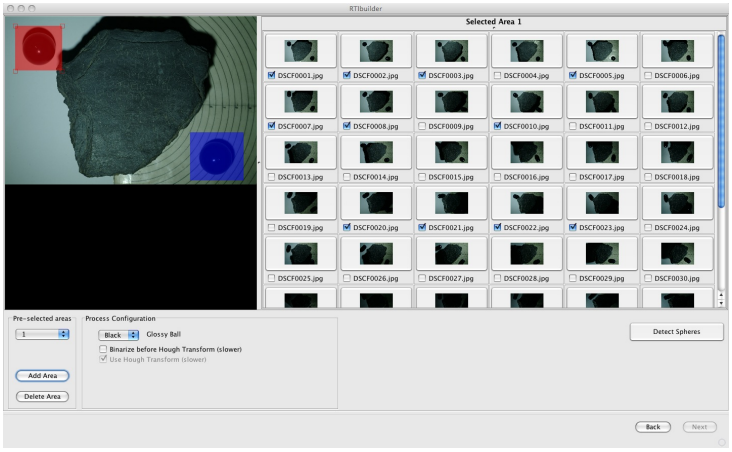
The top left is a preview of the selected image on the right panel, supports drawing of areas containing the spheres to be detected and lets the user perform fine tuning of the detected spheres.

The right panel lets the user select the image to be previewed, and when a specific detection area is available and selected the images that the detection should use for that particular area. After the sphere detection stage it allows the selection of a particular image to preview, for the evaluation of the detection accuracy.

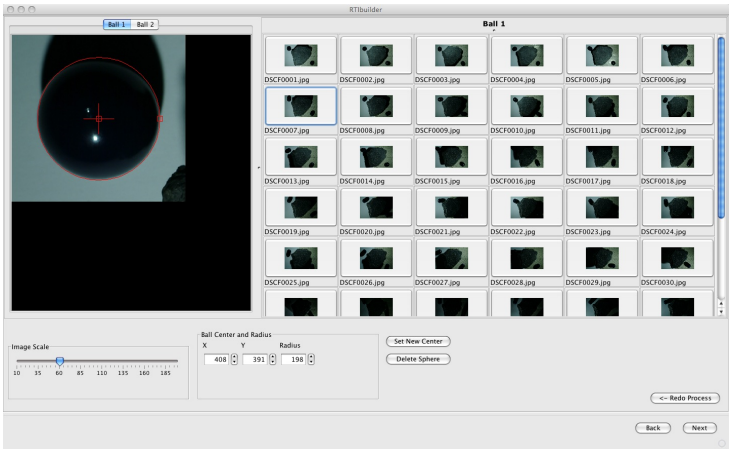
The bottom panel is to configure the stage algorithms. It allows the definition of multiple areas of interest that contain a sphere for detection, the selection of the detection algorithm and the configuration of their parameters. After the sphere detection the user can fine tune or override the automatic detection.

The user interface is also designed to add and extract, to the XMLcarrier message all the required information by the computational algorithms and inform the dispatcher to call the computational module.

When the computational module is called by the application and receives



(a) Sphere detection configuration interface



(b) Sphere detection interface

Figure 5.5: Screen shots of the sphere detection user interface

a valid XML message, it starts to retrieve from it the required parameters. It looks for the areas of interest defined by the user and the related parameters for that particular, area such as images to be used, area location and size, algorithm and associated parameters.

To speed up the process it loads the set of images specified for all areas of interest and crops it using the specified location and size for a particular area. To use all the resources in the system, the sphere detection stage sequence for each area is launched in a different thread. When all threads end their computation, the module gathers the resulting information, such as sphere centre and radius for each area, time spent and error logging, and sends it to the application and user interface together with the original received XML message.

5.2.3 Highlight Detection component

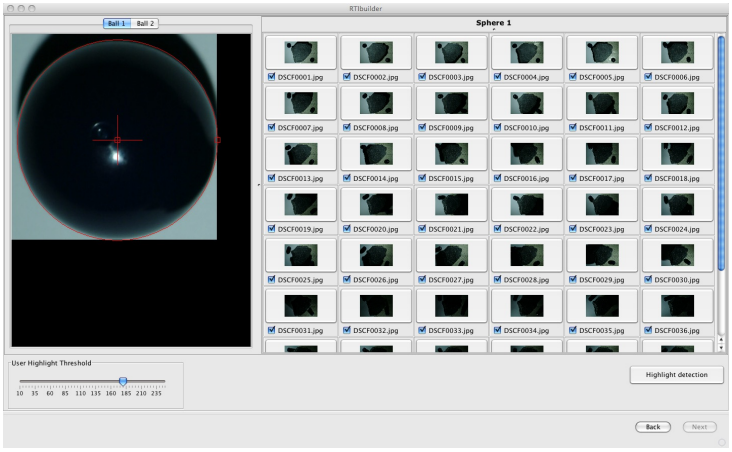
This component estimates the coordinates of the center of a highlight on the surface of the detected sphere. The highlight center coordinates together with the sphere description is used to estimate the light source direction in another module.

The highlight detection uses computer vision labeling algorithms to identify all the highlights recorded on a particular sphere at each image and determine the best candidate as light source direction indicator, computing its center as described in the previous section.

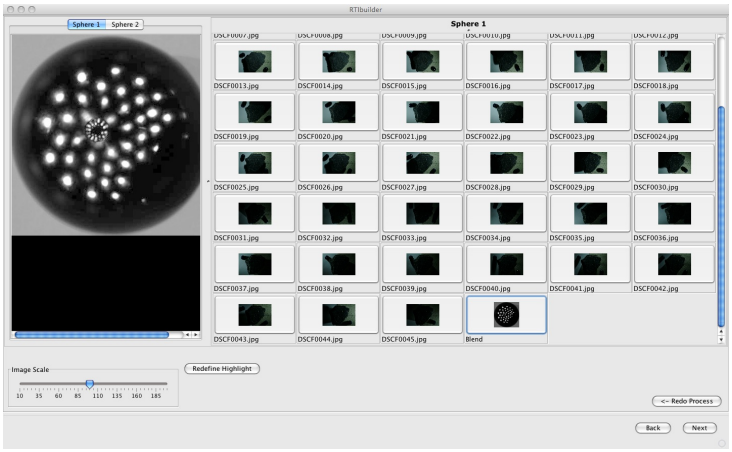
The user interface is similar to the sphere detection interface; the left side panel is used for preview and user hand tuning, the right side panel to select the current image to preview, and the bottom panel to configure the detection algorithms.

The component may accept more than one selected sphere and the preview panel presents each one in a separate tab. The tabs allow the switching between different spheres and have impact on the other panels as well.

The right panel also lets the user to select/dismiss images for the active



(a) Highlight detection configuration interface



(b) Highlight detection tuning interface

Figure 5.6: Screen shots of the highlight detection user interface

sphere, where the application should look for highlights.

The bottom panel allows the definition of the threshold value to be used in the highlight detection on the active sphere. The computational module will discard all pixels bellow the specified threshold value, inside the active sphere.

The interface will package all input data into the current XML message provided by the application and signal the dispatcher to invoke the computational module.

The computational module will query the XML message to obtain the spheres definitions (centre and radius), the highlight parameters for each and all images in the set. For each image in the set it will crop the images using the centre and radius of each defined sphere and store the resulting image in the set of the specified sphere. The original image is discarded from memory as soon as it is not need to save storage space. The computation module launches the highlight detection pipeline described in the previous section in an individual thread for each specified sphere. When all the thread finish their individual work, the highlight centres for each image are stored in the XML message along with the corresponding sphere. In a similar way to the sphere detection module, all empirical provenance information also is stored in the XML message.

The user has now the ability to fine tune the individual detect highlights, or override completely the automatic detection. Any user change to the automatic process is registered and logged in the XML message.

5.2.4 LPcompute component

The LPcompute component will compute for every detected highlight the light source direction used on the capturing session. The details to estimate the light source direction were described in the previous chapter. The final result is the normalized directional vector to the light source, for every image in the file list associated with a specific sphere.

The component does not need any additional user intervention, since it only requires the information from the last two components, which should already be present in the last exchanged XML message. When the dispatcher invokes the component, automatically calls the computation module with the last valid XML message.

The computation module extracts from the XML message the highlight centre coordinates and the respective sphere centre and radius, estimates the light direction and writes back the information to the XML message. If any error occurs logs the error on the XML message as well.

5.2.5 PTMfitter component

This component receives the light positions coordinates or light directions for a set of image files, mapping the information onto a specific format required by the third party application PTMfitter. The component calls the PTMfitter and provides the required input data. The application fitting process is described in chapter 3 and the application is available at http://www.hpl.hp.com/research/ptm/downloads/download.html.

The PTMfitter provided by HP Labs can only handle the JPEG image format, and must be able to convert other images formats to JPEG. This step is required because the RTIbuilder components support a variety of formats that range from bitmaps, to various camera raw formats and the new DNG format proposed by Adobe.

Current component version also handles some basic image processing capabilities, such as cropping and resizing. The image processing methods should be ported to a separate component to be integrated in the workflow. These two image processing functions are often required by end-users, they are currently integrated in the final stage. When these operations are used, they are record in the XML message and executed by the computational part of this component.

The PTMfitter component user interface displays 4 different panels. The panel on the left is the image preview; the middle panel provides configuration

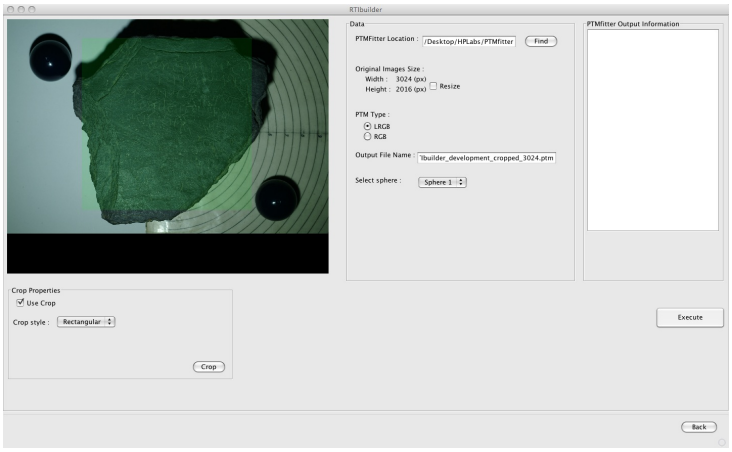


Figure 5.7: Screen shot of the PTMfitter component user interface

options to specify the HP LabsPTM fitter location, to create resized image version of each in the set and to choose which sphere to use, the format type and the name of the PTM file to generate; the panel on the right displays the output of the PTMfitter; while the bottom panel on the left is used to enable/disable the cropping feature and to specify the crop type.

5.2.6 HRTI preview component

The HRTI preview component all the task related with the beginning of the automation process: (i) requests the user for the location of the data to be used; (ii) enables the addition of user defined information to be stored along with the process empirical provenance information; (iii) creates the original XML message with the image location, image unique identifiers, process component sequence and user and machine information; handles the creation the preliminary project configuration and the initial XML message to be used in the HRTI workflow sequence.

The user interface is similar to the previous components, on the top left a preview panel for a quick overlook of the loaded image set, on the right the icon buttons to select the active image.

In the bottom panel the user can specify a directory containing the cap-

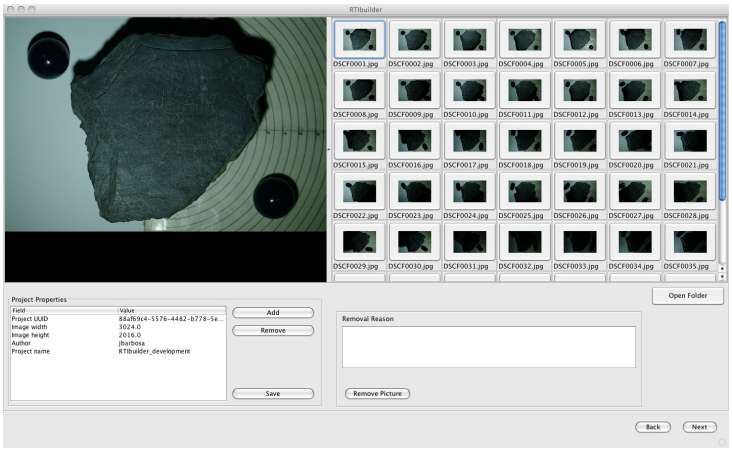


Figure 5.8: Screen shot of the HRTI preview and project configuration component user interface

tured images to be used in the RTI generation process. Following RTIbuilder specification, this component will search the provided directory for specific directories. The first directory the application will probe is the *raw-images* where the raw images download from the camera should be store. If the application is able to recognise the raw image format, it will use the raw format for all components in the workflow. Some raw formats are still proprietary and only readable by the manufacture software, which will required their conversion to a readable format such as JPEG or DNG. The DNG is readable through Adobe DNG SDK, it stores different scales of the image in a TIFF format, the metadata associated with the image (time, date, camera information, lens infortmation, etc), which can be used for empirical provenance purposes, and the raw image itself. The JPEG is read directly using the Java SDK and the application will search for a directory called *jpeg-exports* where the converted images should be.

Also in the bottom panel, the user has the opportunity to add additional metadata to be stored in the XML message exchanged between the components and stored along with the final RTI representation. The user can for instance add information about the people involved, geographic location, descriptions, observations or any other information that might be useful in

the future. The panel allows also the removal of the current selected image, forcing a statement by the user why it is being removed.

5.3 Testing and Deployment

A tool like RTIbuilder is only useful if it is able to cope with the best practices of the CH and NS communities, if is affordable and allows its wide use.

To be able to test the application and determine the accuracy of the computer vision algorithms a set of images was generated using global-illumination rendering techniques. These images provided the precise location of the light source for each individual image and enabled the fine tuning of the detection and computing algorithms. The algorithms were all tuned using real image sets to improve, mainly, the detection of sphere in noisy backgrounds.

The decision to make the application open-source was easily taken from the beginning of the project. From the very beginning the code was always released under GPLv2 Licence to allow its usage by CH and NS without any associated cost and specially to allow the contribution by other developers to the application. This posed a major strain in the development to document, register and log all major changes in the specification and produced code.

Since only the wide use of the application ensures its validity and usability, the decision to promote *hands-on* workshops and collaboration projects with the national CH and NS communities was taken. Two different workshops were given at the D. Diogo de Sousa Regional Archeological Museum for the Portuguese CH and NS communities. These workshops also included other members of the community that addressed the importance of a valid digital surrogate and alerted to the need of best practices of long term digital preservation. In the last of these workshops, the C.H.I foundation team, Mark Mudge and Carla Schroer, joined the speakers panel to shared with the audience their experience with RTI and their vision for the future of digital surrogates and long term digital preservation. The workshops the were designed to promote the technique and give first hand experience on the techniques and software. The C.H.I. foundation has been also promoting

workshop in the U.S. and abroad about photographic best practices for CH professionals were the RTIbuilder plays an important role and Portuguese team has also joined some events outside Portugal to promote the usage of RTI as digital surrogate.

C.H.I foundation leads an international consortium for the development of RTI technology, which the Portuguese team is a part of. Each members of the team is addressing a small fraction of the research required to convert RTI representation into digital surrogates able to address long term digital preservation constraints. The research state-of-the-art presented and debated in a tutorial during the Eurographics 2008 conference[MMC+08].

The participation in the workshops enabled the gathering of user feedback about the software usability and features to enhance its application. The suggestions gathered ranged from simple user interface usability features, such as the strategy to select the sphere or reposition the detected highlight, to more complex features such as image free cropping. The request by C.H.I foundation and the feedback from all members of the international consortium also were incorporated into the software, which lead mainly to the transformation of the software from a monolithic application, to a more flexible application able to incorporate work developed by other teams.

Along with the promoted workshops, the team also joined CH professionals in field research to understand the major constrains and difficulties of working outside a controlled environment. This is the case of the exploratory work done at the Lapedo Child archeological site near Leiria and the Idol Fountain in Braga [BBA+08]. By request of the scientific supervisor of the Lapedo Child archeological site, Francisco Almeida, the RTI technique was applied to document part of the excavation site. The data gathered during this collaboration allowed the adjustment of the software to cope with the strain of field work, specially in remote and isolated places. The experience lead to realisation that the software would have to be more lightweight, fast and that features not considered to be relevant such as the image with all the highlights play an important role when accessing the quality of the light distribution.

The work developed in collaboration with the Archeology Unit of the University of Minho, the D. Diogo de Sousa Regional Archaeological Museum and the Braga City Council Archeology Unit, raised the challenge even higher since the Idol Fountain poses a major problem related with its size. The results of the collaboration were presented in [BBA+08] and point out a possible solution for objects of a considerable size that cannot the captured in a single image, althought future work is still required to make the proposed techniques work together.

Currently the Portuguese team is collaborating with the archeology research team at Salzedas and S. João da Tarouca Monasteries, under the scientific supervision of Luis Sebastian. This research team is using the RTI approach to document some of the artefacts at both locations for documentation purposes and to allow its sharing with other scholars. The project, although in a early stage, has already been proven to be valid and an useful asset by the research team and raised some interesting features request that will, in due time, incorporated into RTIbuilder.

The workshops and the established collaborations provided an environment to test the application under controlled condition and provided valuable insight into the CH community practices, their constrains and needs. This approach provided solid grounds for the development of the application and the improvement of its overall usability. The relations established provide a continuous follow on information in both direction and ensures the continuous improvement of the application by incorporating knowledge from the CH and CS communities.

Chapter 6

Conclusion and Future Work

6.1 Concluding Remarks

The RTI approach is being used by the CH and NS communities to generate digital representation of artefacts and to explore the information within. Although the RTI approach is not able to store the complete 3D information on its final representation, it is able to take advantage of the human perception to convey true 3D information. The illusion of 3D information combined with the texture information give scholars a fair digital representation of the artefact. Several methods have been proposed to take advantage of the partial 3D information within the RTI representation to enhance the surface and texture details [MGW01, MWGA06], which allow scholars to visualise details that can not be seen by the human eye, nor captured by any other technique.

Although the RTI approach is based on empirical data, collected by a fixed camera position under varying light directions, the subsequent processing steps may lead scholars to discard the information if they are not carefully logged, registered and disclosed. The ability to question the provenance of the final digital representation is the corner stone of the scientific methodology. No scholar will trust final RTI representation if he/she is not able to reproduce all the steps taken to generate it. This is one of the major drawbacks of using any digital representation and also a distinguishing factor between digital

representations and digital surrogates.

Digital surrogates use the concept of *empirical provenance* to track all the provenance information associated with the process workflow required to transform the unaltered empirical information gathered in the final digital representation. However, for the workflow to withstand the scrutiny of the scientific method the use of human skill must be reduce to the minimum through the automation of each individual step. This ensures the independence of the workflow from user bias and enables the process reproducibility by describing a set of well defined and a documented automated procedures, which can be assessed in terms of validity and quality.

The RTIbuilder was designed to convert digital representations into digital surrogates using the RTI technique. An earlier version of the application, the LPtracker, was developed to automate the workflow purposed in [MMSL06] for Highlight based RTI, using computer vision algorithms. The LPtracker, presented in [BSP07], employed several computer vision approaches to automate the first steps of the HRTI approach: (i) to detect the sphere in the image set and estimate its centre and radius; (ii) to find the highlight in the sphere surface and estimate its centre. With the information gathered in the two previous steps and the mathematical formulation described [MMSL06], the LPtracker estimated the light direction at each image, as required by the RTI generation step.

The original work with LPtracker provided a proof of concept that the HRTI workflow could be automated. Subsequent trials on the field with CH professionals revealed constraints on the application and on some methods proposed in [BSP07]. The original work of HRTI [MMSL06] used black snooker balls to record the highlight for each image, but in same cases, specially where the background is all black, the sphere is difficult to detect. In [BSP07] it is proposed the use of red snooker balls instead of black spheres, since they are easier to find in the market and due to its red colour it would produce an high value on the red channel of the image. The later property enabled the use of simple computer vision algorithms (region labelling) to easily track the position of the sphere in the image space. However, the strongest

point of using the red glossy spheres is also their major drawback. The trials at Lapedo Child archeological site shown that if the artefact surface is not carefully shield from the sphere reflection on to the surface, it will change its reflectance properties.

The field studies at the Lapedo Child archeological sites and at the D. Diogo de Sousa Regional Archeological Museum also show that the application must be able to cope with the available computational resources. The major constraint revealed in this field test is related to the required memory space, which lead to the development of different strategies to reduce the use of memory. In same cases, as for instance the sphere detection steps, studies where conducted to fine tune the used algorithm by reducing the amount of required image samples to efficiently and accurately estimate the sphere centre and radius. On other cases, specially between stages, all relevant intermediate information is stored to disk while the remaining data is discarded. Although the time required to load the information from disk increases at each individual stage, field studies revealed that due to the reduced amount of memory used and to the fine tuning of the algorithms the application performs faster.

The gathered experience of the collaborations with the national CH community and with the accumulated experience of the international C.H.I foundation, through the workshops C.H.I. promoted, and the knowledge provided by other members of the team was debated in [MMC+08].

The discussion and subsequent meetings with the C.H.I consortium members changed the development of the LPtracker. Although the original LPtracker version was able to address the HRTI process workflow, the meetings revealed that the application should be more flexible and allow the incorporation of more image processing stages into the application. From this meetings it also became clear that the next logical step would be to address long-term digital archive problems by mapping the empirical provenance information into a standard such as the CIDOC-CRM [Doe05], which involved changing the LPtracker empirical provenance log into a more suitable format.

The final result of the discussion was a new specification of the application that became know as RTIbuilder and was documented in [BP09a]. The

RTIbuilder specification used most of the approach previously tested in LPtracker and addressed the problem of flexibility by using a modular approach to its components. The specification also describes specific programming interfaces that must be used when developing a new component for the application. These interfaces enable the application to dynamically load new components to be used in the processing workflow, while enabling developers to produce new components independently of the main application code. Since the application is able to dynamically load components, the user is also able to specify his/her own workflow.

With this approach the application became more flexible and offers the opportunity to other developers to enrich the application with added functionalities. However, to coordinate the operation between different components a well defined communication schema, named XMLcarrier, was defined [BP09b]. The XMLcarrier schema serves two purposes: (i) to coordinate the communication between the different components in the application and (ii) to store the required empirical provenance information. It makes sense to bundle the two items together, since the empirical provenance information records the processing history, which is no more than the sequence of components used and its respective parameters along with the user intervention over those parameters. To enable the future mapping into a long-term digital archive using standards such as the CIDOC-CRM, the XMLcarrier specification enables the storage of additional user defined information.

The RTIbuilder and the LPtracker are able to show that the generation of a digital surrogate using the RTI techniques is possible. However, for a digital surrogate to be widely accepted the technique must also be affordable, cope with the current CH and NS professional best practices and provide added value to the currently used workflows. Currently, the only format publicly available for RTI final representation is the PTM format. Although the PTM is freely available to CH and NS communities, it is owned and patented by HP Labs creating a major drawback to the adoption of the RTI technique for digital surrogates. Wang and Prabath presented in [MMC+08] part of their work in Multi-view RTI. In their work they also introduce a new RTI basis function using hemispherical harmonics (HSH), which will be publicly

available under open-source license. HSH format remove all barriers posed by the current PTM format and enable a wider use of the RTI representation.

To validate that both the LPtracker, bundled together with the PTM tools from HP Labs in a package call PTMbuilder, and the RTIbuilder would meet these criteria, several workshops, collaboration projects and talks were promoted for the CH and NS national communities, along with participation in international workshops such as [SP09]. The workshops and talks promoted for the Portuguese community lead to the improvement of the application and to the development of new strategies for RTI acquisition. For instance, due to the lack of economic resources reported by the workshop participants, it was proposed and tested during one of the workshops the use of low cost *papier-mache* dome in combination with the HRTI approach to speed up the time required to acquire small object.

From the workshops some interesting collaboration projects where started, as for instance the collaboration with the archeological research team at S. João da Tarouca Monastery where the traditional documentation technique is been replaced by the RTI approach, and is revealing some additional features that are being considered for the next RTIbuilder release. The Idol Fountain project presented in [BBA+08], in collaboration with the Braga City Council Archaeology Unit, the University of Minho Archaeology Unit and the D. Diogo the Sousa Reginal Archeological Museum also started from one of the promoted workshops and provided new research challenges that are still far from solved.

Although far from complete, the work developed during this thesis demonstrates that the RTI approach is a valid candidate for digital surrogate representation and the feedback received from the participants in the different workshops, collaboration projects and members of the C.H.I consortium encourages to pursuit the track of research and development taken so far.

6.2 Future Work

These workshops were the seed for some interesting collaboration projects, such as the collaboration collaboration with the research team at S. João da Tarouca Monastery as proven to be a fruitful relationship since they are using the full capabilities of the RTI technique, including sharing the artefacts with remote researchers. Although the final study is not yet concluded, the preliminary feedback pointed out some concerns already reported by other projects and from members of the C.H.I. international consortium.

The major drawbacks pointed out to the RTI approach is the lack of accuracy introduced by the parallax effect of the camera lens, the lack of more points of view to enable the disclosure of some surface details and the need of additional applications to perform some of the required pre-processing steps.

To improve the accuracy of the light position estimation, when using the HRTI technique, a second sphere was introduced in the scene. A more accurate light source position can be obtained using the triangulation of estimated light directions from each sphere. The original work proposed by Mudge [MMSL06] assumes that the error is negligible when considering the view point at the infinite. However, this error can be significantly reduced by estimating the light source accurate position. A prototype already computes the light source position from the highlights recorded in two spheres, but it still lacks sufficient accuracy, specially in grazing angles, to be useful.

Wang and Prabath proposed an approach to Multi-view RTI, but their work was not yet fully published and in the work presented at [MMC+08] they limited the scope of the work to a single rotation axis. The studies done by Meseth et. al. [MMK03], suggest a possible solution to combine the view dependency and light dependency into the same final RTI representation. A similar approach combining the works perviously referred may result in a complete representation of a SVBRDF, and it should be worth to explore.

The combination of complete 3D information and surface reflectance properties would increase significantly the added value of the RTI technique to cre-

ate digital surrogates, with the advantage of including in the same workflow the photogrammetry workflow currently used by CH and NS professionals.

The development of the RTIbuilder is yet far from over. The application has not yet reached a maturity level to generate a digital surrogate, although by collecting data from different sources such as the information stored in XMP, we are able to reconstruct the complete journey of the acquired empirical data to its final RTI format.

The RTIbuilder should cover all the processes in the workflow, from the moment the image is transmitted from the digital camera to the computer, processed by the different stages until it reaches its final digital representation with all the empirical information attached. This implies the application must be able to replace some commercial 3^{rd} party tools, such as Adobe Photoshop, used to perform some basic image processing (white balance calibration, colour calibration, raw image conversion, image cropping, etc...) which has a significant impact in the RTI final representation.

Bibliography

[AFT+04] Peter Allen, Steven Feiner, Alejandro Troccoli, Hrvoje Benko, Edward Ishak, and Benjamin Smith. Seeing into the past: Creating a 3d modeling pipeline for archaeological visualization. In *3D Data Processing, Visualization and Transmission Symposium (3DPVT 2004)*, Thessalonika, Greece, September 2004.

[AS02] Michael Ashikhmin and Peter Shirley. Steerable illumination textures. *ACM Trans. Graph.*, 21(1):1–19, 2002.

[BBA+08] Joao Barbosa, Paulo Bernardes, Matheus Almeida, Pedro Gomes, Ricardo Goncalves, and Alberto Proenca. A technology cocktail for a 3d photo-realistic model of a i century roman fountain: Range scanning, rti and physically based rendering. In *The 9th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST (2008)*, 2008.

[BFMS05] R.S. Berns, F.S. Frey, M.R.Rosen, and E.P.M. Smoyer. Direct digital capture of cultural heritage benchmarking american museum practices and defining future needs - project report. Technical report, MCSL Technical Report, April 2005.

[Bli77] James F. Blinn. Models of light reflection for computer synthesized pictures. *SIGGRAPH Comput. Graph.*, 11(2):192–198, 1977.

[BP09a] Joao Barbosa and Alberto Poenca. Rtibuilder specification. Technical report, Universidade do Minho, April 2009.

[BP09b] Joao Barbosa and Alberto Poenca. Xml carrier specification. Technical report, Universidade do Minho, April 2009.

[BSP07] Joao Barbosa, Joao Luis Sobral, and Alberto Poenca. Imaging techniques to simplify the ptm generation of a bas-relief. In *The 8th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST (2007)*, November 2007.

[CDMR04] O. G. Cula, K. J. Dana, F. P. Murphy, and B. K. Rao. Bidirectional imaging and modeling of skin texture. *Biomedical Engineering, IEEE Transactions on*, 51(12), Dec 2004.

[Cov00] Robin Cover. The XML Cover Pages. WWW page, 2000. `http://www.oasis-open.org/cover/xml.html`.

[CT82] R. L. Cook and K. E. Torrance. A reflectance model for computer graphics. *ACM Trans. Graph.*, 1(1):7–24, 1982.

[CU] COLUMBIA-UTRECH. Columbia-utrecht reflectance and texture database:.

[DCCS06] M. Dellepiane, M. Corsini, M. Callieri, and R. Scopigno. High quality ptm acquisition: Reflection transformation imaging for large objects. In *VAST'06: Proc.7th Int. Symp. Virtual Reality, Archaeology and Cultural Heritage*, Cyprus, 2006.

[DHT+00] Paul Debevec, Tim Hawkins, Chris Tchou, Haarm-Pieter Duiker, Westley Sarokin, and Mark Sagar. Acquiring the reflectance field of a human face. In *SIGGRAPH '00: Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 145–156, New York, NY, USA, 2000. ACM Press/Addison-Wesley Publishing Co.

[DLHS01] Katja Daubert, Hendrik P. A. Lensch, Wolfgang Heidrich, and Hans-Peter Seidel. Efficient cloth modeling and rendering. In *Proceedings of the 12th Eurographics Workshop on Rendering Techniques*, pages 63–70, London, UK, 2001. Springer-Verlag.

[Doe05] Martin Doerr. The cidoc crm, an ontological approach to schema heterogeneity. In Y. Kalfoglou, M. Schorlemmer, A. Sheth, S. Staab, and M. Uschold, editors, *Semantic Interoperability and Integration*, number 04391 in Dagstuhl Seminar Proceedings, Dagstuhl, Germany, 2005. Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany.

[DvGNK99] Kristin J. Dana, Bram van Ginneken, Shree K. Nayar, and Jan J. Koenderink. Reflectance and texture of real-world surfaces. *ACM Transactions on Graphics*, 18(1):1–34, 1999.

[DWT+02] Paul Debevec, Andreas Wenger, Chris Tchou, Andrew Gardner, Jamie Waese, and Tim Hawkins. A lighting reproduction approach to live-action compositing. In *SIGGRAPH '02: Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, pages 547–556, New York, NY, USA, 2002. ACM.

[FBM+06a] T. Freeth, Y. Bitsakis, X. Moussas, J. H. Seiradakis, A. Tselikas, H. Mangou, M. Zafeiropoulou, R. Hadland, D. Bate, A. Ramsey, M. Allen, A. Crawley, P. Hockley, T. Malzbender, D. Gelb, W. Ambrisco, and M. G. Edmunds. Decoding the ancient greek astronomical calculator known as the antikythera mechanism. *Nature*, 444:587–591, November 2006.

[FBM+06b] T. Freeth, Y. Bitsakis, X. Moussas, J. H. Seiradakis, A. Tselikas, H. Mangou, M. Zafeiropoulou, R. Hadland, D. Bate, A. Ramsey, M. Allen, A. Crawley, P. Hockley, T. Malzbender, D. Gelb, W. Ambrisco, and M. G. Edmunds. Decoding the ancient greek astronomical calculator known as the antikythera mechanism. *Nature*, 444(7119):587–591, November 2006.

[Foo97] Sing-Choong Foo. A gonioreflectometer for measuring the bidirectional reflectance of material for use in illumination compu-

tation. Master's thesis, Cornell University, Ithaca, NY, August 1997.

[GBK99] A. Georghiades, P. Belhumeur, and D. Kriegman. Illumination-based image synthesis: Creating novel images of human faces under differing pose and lighting, 1999.

[GGSC96] Steven J. Gortler, Radek Grzeszczuk, Richard Szeliski, and Michael F. Cohen. The lumigraph. In *SIGGRAPH '96: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 43–54, New York, NY, USA, 1996. ACM.

[HBMG02] Oyvind Hammer, Stefan Bengtson, Tom Malzbender, and Dan Gelb. Imaging fossils using reflectance transformation and interactive manipulation of virtual light sources. In *Manipulation of Virtual Light Sources, Palaeontologia Electronica*, page 2002, 2002.

[HEE+02] M. Hauth, O. Etzmuss, B. Eberhardt, Reinhard Klein, Ralf Sarlette, Mirko Sattler, K. Daubert, and J. Kautz. Cloth animation and rendering. In *Proceedings of Eurographics 2002 Tutorials*. The Eurographics Association, September 2002. Tutorial.

[HNP06] Sorin Hermon, Joanna Nikodem, and Cinzia Perlingieri. Deconstructing the vr - data transparency, quantified uncertainty and reliability of 3d models. In *VAST 2006: 7th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage*, pages 123–130, October 2006.

[HP03] Jefferson Y. Han and Ken Perlin. Measuring bidirectional texture reflectance with a kaleidoscope. *ACM Trans. Graph.*, 22(3):741–748, 2003.

[LFTG97] Eric P. F. Lafortune, Sing-Choong Foo, Kenneth E. Torrance, and Donald P. Greenberg. Non-linear approximation of reflectance functions. In *SIGGRAPH '97: Proceedings of the*

24th annual conference on Computer graphics and interactive techniques, pages 117–126, New York, NY, USA, 1997. ACM Press/Addison-Wesley Publishing Co.

[LH96] Marc Levoy and Pat Hanrahan. Light field rendering. In *SIGGRAPH '96: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 31–42, New York, NY, USA, 1996. ACM.

[LHZ+04] Xinguo Liu, Yaohua Hu, Jingdan Zhang, Xin Tong, Baining Guo, and Heung-Yeung Shum. Synthesis and rendering of bidirectional texture functions on arbitrary surfaces. *IEEE Transactions on Visualization and Computer Graphics*, 10(3):278–289, 2004.

[MAA01] Michael D. McCool, Jason Ang, and Anis Ahmad. Homomorphic factorization of brdfs for high-performance rendering. In *SIGGRAPH '01: Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 171–178, New York, NY, USA, 2001. ACM.

[MCS90] J.F. Murray-Colemann and A.M. Smith. The automated measurement of brdfs and their application to luminaire modeling, 1990.

[MGW01] Tom Malzbender, Dan Gelb, and Hans Wolters. Polynomial texture maps. In *SIGGRAPH'01: Proc. 28th Conf. Computer Graphics and Interactive Techniques*, pages 519–528, USA, 2001.

[MLH02] David K. McAllister, Anselmo Lastra, and Wolfgang Heidrich. Efficient rendering of spatial bi-directional reflectance distribution functions. In *HWWS '02: Proceedings of the ACM SIGGRAPH/EUROGRAPHICS conference on Graphics hardware*, pages 79–88, Aire-la-Ville, Switzerland, Switzerland, 2002. Eurographics Association.

[MMC+08] Mark Mudge, Tom Malzbender, Alan Chalmers, Roberto Scopigno, James Davis, Oliver Wang, Prabath Gunawardane,

Michael Ashley, Martin Doerr, Alberto Proenca, and João Barbosa. Image-based empirical information acquisitionimage-based empirical information acquisition, scientific. reliability, and long-term digital preservation for the natural sciences and cultural heritage. In *Eurographics 2008 Tutorials*. Eurographics Association, 2008.

[MMK03] Jan Meseth, Gero Müller, and Reinhard Klein. Preserving realism in real-time rendering of bidirectional texture functions. In *OpenSG Symposium 2003*, pages 89–96. Eurographics Association, Switzerland, April 2003.

[MMK04] Jan Meseth, Gero Müller, and Reinhard Klein. Reflectance field based real-time, high-quality rendering of bidirectional texture functions. *Computers and Graphics*, 28(1):103–112, February 2004.

[MMS+04] Gero Müller, Jan Meseth, Mirko Sattler, Ralf Sarlette, and Reinhard Klein. Acquisition, synthesis and rendering of bidirectional texture functions. In Christophe Schlick and Werner Purgathofer, editors, *Eurographics 2004, State of the Art Reports*, pages 69–94. INRIA and Eurographics Association, September 2004.

[MMSL06] Mark Mudge, Tom Malzbender, Carla Schroer, and Marlin Lum. New reflection transformation imaging methods for rock art and multiple-viewpoint display. In *VAST 2006: 7th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage*, pages 195–202, October 2006.

[MPBM03] Wojciech Matusik, Hanspeter Pfister, Matthew Brand, and Leonard McMillan. Efficient isotropic brdf measurement. In *EGRW '03: Proceedings of the 14th Eurographics workshop on Rendering*, pages 241–247, Aire-la-Ville, Switzerland, Switzerland, 2003. Eurographics Association.

[MVSL05] Mark Mudge, Jean-Pierre Voutaz, Carla Schroer, and Marlin Lum. Reflection transformation imaging and virtual representa-

tions of coins from the hospice of the grand st. bernard. In *VAST 2005: 6th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage*, pages 29–40, November 2005.

[MWGA06] Tom Malzbender, Bennett Wilburn, Dan Gelb, and Bill Ambrisco. Surface enhancement using real-time photometric stereo and reflectance transformation. In *Eurographics Symposium on Rendering 2006*, Nicosia, Cyprus, 2006. Eurographics Association.

[MWL+99] Stephen R. Marschner, Stephen H. Westin, Eric P. Lafortune, Kenneth E. Torrance, and Donald P. Greenberg. Image-based brdf measurement including human skin. In Dani Lischinski and Gregory Ward Larson, editors, *Rendering Techniques*, pages 131–144. Springer, 1999.

[Nic76] F.E. Nicodemus. Self-study manual on optical radiation measurements: Part i—concepts, chapters 1 to 3, 1976.

[NRH+77] F.E. Nicodemus, J.C. Richmond, J.J. Hsia, W.I. Ginsberg, and T. Limperis. Geometrical considerations and nomenclature for reflectance. *Applied Optics*, 9:1474–1475, 1977.

[oB] University of Bonn. Btf database bonn.

[Pho73] Bui Tuong Phong. *Illumination for computer-generated images.* PhD thesis, The University of Utah, 1973.

[RM] S. RUSINKIEWICZ and S. MARSCHNER. Measurement i - brdfs. Script of Course CS448C: Topics in Computer Graphics, held at Stanford University.

[Ros94] RossMcCluney. Introduction to radiometry and photometry, 1994.

[SP09] Adam Spring and Caradoc Peters. The current climate for data capture, processing and archiving. *Geoinformatics*, 12:53–56, 2009.

[SSK03] Mirko Sattler, Ralf Sarlette, and Reinhard Klein. Efficient and realistic visualization of cloth. In *Eurographics Symposium on Rendering 2003*, June 2003.

[TSE+04] Chris Tchou, Jessi Stumpfel, Per Einarsson, Marcos Fajardo, and Paul Debevec. Unlighting the parthenon. In *SIGGRAPH '04: ACM SIGGRAPH 2004 Sketches*, page 80, USA, 2004. ACM Press.

[WAA+00] Daniel N. Wood, Daniel I. Azuma, Ken Aldinger, Brian Curless, Tom Duchamp, David H. Salesin, and Werner Stuetzle. Surface light fields for 3d photography. In ., pages 287–296, 2000.

[War92] Gregory J. Ward. Measuring and modeling anisotropic reflection. *SIGGRAPH Comput. Graph.*, 26(2):265–272, 1992.