MÀSTER EN COMPUTACIÓ

DEPARTAMENT D'INFORMÀTICA I MATEMÀTICA APLICADA

UNIVERSITAT DE GIRONA

The Viewpoint Channel: An Information-Theoretic Framework for Viewpoint Selection, Ambient Occlusion and Shape Similarity

Master Thesis

FRANCISCO GONZÁLEZ GARCÍA

Advisors: DR. MIQUEL FEIXAS
PROF. MATEU SBERT

Juliol 2007

The Viewpoint Channel: An Information-Theoretic Framework for Viewpoint Selection, Ambient Occlusion and Shape Similarity

Master Thesis, July 2007
Departament d'Informàtica i Matemàtica Aplicada
Grup de Gràfics de Girona
Universitat de Girona
17071 Girona
Spain
Tel. +34 - 972 41 84 17

Reproservice / Departament d'Informàtica i Matemàtica Aplicada
Girona, Spain 2007

The Viewpoint Channel: An Information-Theoretic Framework for Viewpoint Selection, Ambient Occlusion and Shape Similarity
Master Thesis in the Master in Computing
Francisco González García
Departament d'Informàtica i Matemàtica Aplicada
Grup de Gràfics de Girona
Universitat de Girona

Abstract

Viewpoint selection is an emerging area in computer graphics with applications in fields such as scene exploration, image-based modelling, and volume visualization. As an example, best view selection algorithms are used to obtain the minimum number of views (or images) to best understand or model an object or scene. In this master thesis, we present a new information-theoretic framework, created from the definition of an information channel between a set of viewpoints (input) and the set of polygons of an object (output), to deal with viewpoint selection, ambient occlusion and shape similarity. First, the mutual information of this channel is shown to be a powerful tool to deal with viewpoint selection, viewpoint similarity and scene exploration. Second, from the reversion of the channel, mesh information and mesh saliency are computed. Third, viewpoint mutual information is generalized in order to incorporate saliency as an importance factor to guide the viewpoint selection. Fourth, we use the mesh information as an ambient occlusion technique with applications to viewpoint enhancement and relighting for non-photorealistic rendering. Finally, a new view-based approach to compute shape similarity using a spherical registration process is presented. Although we use a sphere of viewpoints around an object, our framework is also valid for any set of viewpoints in a closed scene. A number of experiments demonstrate the robustness of the presented approach and the good behavior of the proposed measures.

Keywords: viewpoint selection, mesh saliency, information theory, ambient occlusion, shape similarity

Contents

Acknowledgements

Bueno, por fin hemos llegado a una de las primeras paradas que tiene la realización de un doctorado: mi master tesis!!! Me gustaría dar las gracias a tanta gente que seguro me olvidaré de la mayoría. Por favor pensad que mi memoria es a corto plazo y no mucho más duradera que la de dory :D

Perimero me gustaría agradecer a mi pareja toda la comprensión que ha tenido respecto a mi dedicación por mi trabajo. Sabes que me apasiona y que es importante para mi, y pese a que estas últimas semanas hemos pasado menos tiempo del que querríamos juntos, siempre ha salido de ti una palabra de ánimo, de apoyo, de cariño... gracies vida meva!!! T'estimo amb tot el meu cor i estic molt agraït de que deu ens hagi posat en el mateix camí!!! T'ESTIMO ANGELET!!!

Mi familia!!! Que haríamos sin la familia que te apoya siempre haga lo que hagas, elijas lo que elijas. Para mi mami que te quiero con locura, mi papi y mi hermana. Todos son importantísimos para mi!!! Se que sin vosotros jamás hubiera podido llegar al lugar y momento en el que me encuentro ahora. Gracias por vuestros sacrificios y por quererme!!! Papa, pienso en ti y se que me cuidas desde allí arriba!!! Jamás te olvides de nosotros!!!

Sílvix y dolors que no me olvido de vosotras!!! Para mi sois parte de mi familia y os quiero muchísimo. Sois muy especiales y doy gracias por teneros a mi lado!!! Ànims vacuneta que tindràs una super tesi!!!

A mis amigos y compis de tesi: Isma y Tere!!! Soys muy buenas personas y buenos amigos y os aprecio de corazón!!! Hemos compratido muchos momentos y siempre habéis estado ahí escuchando mis tonterias y agobios. Muchas gracias!!! Y a por el doctorado, que ya es nuestro!!! Bueno no ahora, pero si en cuatro años :P Eso si, como dice tere, al paso que vamos nos ponemos un chalet en la uni y a disfrutar!!! :D También a ti Verena, eres una gran persona y oajalá vuelvas por gerona una vez marches a alemania. Si no es así te hecharemos mucho de menos!!! Siempre tendrás amigos en Gerona!!! Te saco la lengua :D

Gus y familia mil millones de gracias por ser tan buenos conmigo. Gus eres una persona maravillosa, con un corazón que no te cabe en el pecho y creo que por eso la camisa te va un poco justita a veces :P Gracias por todos los momentos que me has dedicado, por el afecto que me tienes y por aconsejarme siempre!!! Ahora nos toca una aventura juntos!!! Qué ilusión me hace!!! Gracias de corazón!!!

Esta máster tesis supone para mi muchas cosas. Primero es el puente que permite cruzar el río para poder llegar al otro lado, que no es ni más ni menos que realizar el doctorado, y segundo y más importante, la oportunidad de haber podido trabajar con una grandísima persona como Miquel. Te estoy muy agradecido por todo el tiempo y esfuerzo que me has dedicado. He disfrutado mucho trabajando contigo estos dos últimos años, pero lo más importante es que me llevo el placer de haber podido conocer a la gran persona que eres. Muchas gracias por todo y espero que volvamos a colaborar juntos en un futuro no muy lejano!!!

A ti Mateu, gracias por la oportunidad que me diste de poder trabajar en gráficos dentro del grupo. Te estoy muy agradecido por ello!!!

También me gustaria no olvidarme de toda la buena gente que hay en el departamento y que hacen que me sienta feliz de estar trabajando en esta universidad: Sergi, Albert (alias maestro yoda), Carles, Francesc, Gonzalo, Marité, Narcís, Gonzalo, Jaume...a todos gracias!!! Especialmente, me gustaría agradecerte a ti Céline lo amable que has sido en todo momento conmigo desde tu llegada, por tu constante sonrisa y por la ayuda que me diste!!! You have an enormous heart and a nice soul!!!

Esta master tesis ha sido financiada por el TIN2004-07451-C03-01 del govierno español y el IST-2-004363 (*GameTools: Advanced Tools for Developing Highly Realistic Computer Games*) del sexto programa marco europeo.

1. Introduction

In this introductory chapter we reflect about the ideas behind viewpoint selection and we explain the motivations and objectives of our research work. Finally, we present an overview about the content of each chapter.

1.1. Viewpoint Selection

The basic question underlying the viewpoint selection study and application is "what is a 'good' scene viewpoint"? Obviously, this question does not have a unique answer. Depending on our objective, the best viewpoint can be, for instance, the most *representative* or the most *unstable* one, i.e., the one that maximally changes when it is moved within its close neighborhood [BS05]. Palmer et al. [PRC81] and Blanz et al. [BTB99] have presented different experiments demonstrating that observers prefer views (called *canonical views*) that avoid occlusions and that are off-axis (such as a three-quarter viewpoint), salient (the most significant characteristics of an object are visible), stable and with a large number of visible surfaces.

In computer graphics, several *viewpoint quality* measures have been applied in areas such as scene understanding [PB96, VFSH01, PPB*05], scene exploration [AVF04, SPT06], image-based modeling [VFSH03] and volume visualization [BS05, TFTN05, VFSG06, JS06]. In other areas, such as object recognition and mobile robotics, best view selection is also a fundamental task.

Many works have demonstrated that the recognition process is view-dependent [PRC81, BET95, TBZB97, BTB99]. In [TBZB97], the authors found that "visual recognition may be explained by a view-based theory in which viewpoint-specific representations encode both quantitative and qualitative features". In robotics, the simultaneous localization and mapping problem (SLAM) requires that the robot decides on its own the necessary motions to construct the most accurate map possible. In [GBL02] authors investigate map-building strategies for a mobile robot with imperfect control and sensing, proposing an algorithm to guide him through a series of 'good' positions, where 'good' refers to the expected amount and quality of the information that will be revealed at each new location.

1.2. Motivations and Objectives

Intuitively, we may state that a good view is the one which gives the highest amount of information of a model or scene. As an example, we can see in Figure 1.1 two different views of the same model. Though the image on the left gives us some clues in order to identify the object as a butterfly, the image on the right-hand side gives us a higher amount of information making the scene more understandable.

According to [Vaz03], to determine the goodness of a viewpoint we have to answer to basic questions:

1. **Which is the kind of information we want to measure?** In order to obtain a well-built measure of a quality of a view we need first to identify what is the information that has to be maximized. The *faces* that built up a 3D model or a scene are considered as the pieces of information we can deal with.

2. **How can we measure this information?** To address the problem of the measurement of the information we have borrowed some tools from *Information Theory*.

Extending the work initiated in [VFSH01, SPFG05], in this master thesis we present a new and robust framework to deal with viewpoint selection, ambient occlusion and shape similarity. Given a set of viewpoints surrounding the object, we define an *information channel* between the viewpoints and the polygons of the object. From this channel, the *viewpoint mutual information* is used to obtain the best views of an object, to calculate the stability of a viewpoint, and to guide the object exploration. Then, we reverse the channel and we compute both the information and the saliency associated with each polygon. From this polygonal saliency, we will show how to calculate how salient is a viewpoint and we incorporate it to viewpoint mutual information to drive the viewpoint selection. We also present a new information-theoretic ambient occlusion technique based on the information associated with each polygon. Finally, we introduce a new method for computing the shape dissimilarity between 3D polygonal models using the mutual information and saliency spheres as a shape signature of 3D objects.

1.3. Document Overview

The structure of this master thesis can be summarized as follows:

- **Chapter 1. Introduction.** We make a brief introduction and we give an overview of the master thesis.

(*a*) (*b*)

Figure 1.1.: Two different views from the same object. Left image (a) gives less information than the right (b) one.

- **Chapter 2. Background and Related Work.** We review some basic concepts of information theory, viewpoint quality measures and ambient occlusion, as well as some introduction about 3D shape retrieval.
- **Chapter 3. Viewpoint Channel.** We present the essence of the master thesis: the *Viewpoint Channel*. That is, an information channel between a set of viewpoints and the polygons of an object. From this channel, a new viewpoint quality measure based on mutual information is introduced. Finally, some mutual information properties are used to define viewpoint similarity and stability.
- **Chapter 4. Best View Selection and Object Exploration.** We present a new selection algorithm, also used for viewpoint clustering and object exploration.
- **Chapter 5. Mesh and Viewpoint Saliency.** We reverse the viewpoint channel and define the polygonal and viewpoint saliency. Our viewpoint quality measure is also extended by incorporating the saliency as an importance factor.
- **Chapter 6. Information-Theoretic Ambient Occlusion.** Using the reversed viewpoint channel, the information associated with each polygon is used as an occlusion measure and is applied as an obscurance or ambient occlusion map of the scene. Some interesting applications are also shown.
- **Chapter 7. View-based Shape Similarity.** We present a new spherical registration method to quantify the *shape similarity* between 3D models using mutual information and saliency spheres.
- **Chapter 8. Conclusions and Future Work.** We present our conclusions and possible future work.

2. Background and Related Work

In this chapter we review some basic concepts of information theory (see [CT91]), viewpoint quality, ambient occlusion and shape similarity.

2.1. Information-Theoretic Concepts

Let $\mathcal{X}$ be a finite set, let X be a random variable taking values x in $\mathcal{X}$ with distribution $p(x) = Pr[X = x]$. Likewise, let Y be a random variable taking values y in $\mathcal{Y}$. An information channel between two random variables (input X and output Y) is characterized by a *probability transition matrix* (composed of conditional probabilities) which determines the output distribution given the input.

The *Shannon entropy* $H(X)$ of a random variable X is defined by

$$H(X) = -\sum_{x \in \mathcal{X}} p(x) \log p(x). \tag{2.1}$$

It is also denoted by $H(p)$ and measures the average uncertainty of a random variable X. All logarithms are base 2 and entropy is expressed in bits. The convention that $0 \log 0 = 0$ is used. The *conditional entropy* is defined by

$$H(Y|X) = -\sum_{x \in \mathcal{X}} p(x) \sum_{y \in \mathcal{Y}} p(y|x) \log p(y|x), \tag{2.2}$$

where $p(y|x) = Pr[Y = y|X = x]$ is the conditional probability. The conditional entropy $H(Y|X)$ measures the average uncertainty associated with Y if we know the outcome of X. In general, $H(Y|X) \neq H(X|Y)$, and $H(X) \geq H(X|Y) \geq 0$.

The *mutual information* (MI) between X and Y is defined by

$$I(X,Y) = H(X) - H(X|Y) = \sum_{x \in \mathcal{X}} p(x) \sum_{y \in \mathcal{Y}} p(y|x) \log \frac{p(y|x)}{p(y)}. \tag{2.3}$$

It is a measure of the shared information between X and Y. It can be seen that $I(X,Y) = I(Y,X) \geq 0$. A fundamental property of MI is given by the *data processing inequality*

which can be expressed in the following way: if $X \rightarrow Y \rightarrow Z$ is a Markov chain, i.e., $p(x,y,z) = p(x)p(y|x)p(z|y)$, then

$$I(X,Y) \geq I(X,Z). \tag{2.4}$$

This result demonstrates that no processing of Y, deterministic or random, can increase the information that Y contains about X.

The *relative entropy* or *Kullback-Leibler distance* between two probability distributions $p = \{p(x)\}$ and $q = \{q(x)\}$ defined over $\mathcal{X}$ is given by

$$KL(p|q) = \sum_{x \in \mathcal{X}} p(x) \log \frac{p(x)}{q(x)}, \tag{2.5}$$

where, from continuity, we use the convention that $0 \log 0 = 0$, $p(x) \log \frac{p(x)}{0} = \infty$ if $p(x) > 0$, and $0 \log \frac{0}{0} = 0$. The relative entropy $KL(p|q)$ is a divergence measure between the *true* probability distribution p and the *target* probability distribution q. It can be proved that $KL(p|q) \geq 0$.

A convex function f on the interval $[a,b]$ fulfils the Jensen inequality: $\sum_{i=1}^{n} \lambda_i f(x_i) - f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \geq 0$, where $0 \leq \lambda \leq 1$, $\sum_{i=1}^{n} \lambda_i = 1$, and $x_i \in [a,b]$. For a concave function, the inequality is reversed. If f is substituted by the Shannon entropy, which is a concave function, we obtain the *Jensen-Shannon inequality* [BR82]:

$$JS(\pi_1, \pi_2, \ldots, \pi_N; p_1, p_2, \ldots, p_N) \equiv H\left(\sum_{i=1}^{N} \pi_i p_i\right) - \sum_{i=1}^{N} \pi_i H(p_i) \geq 0, \tag{2.6}$$

where $JS(\pi_1, \pi_2, \ldots, \pi_N; p_1, p_2, \ldots, p_N)$ is the *Jensen-Shannon divergence* of probability distributions $p_1, p_2, \ldots, p_N$ with prior probabilities or weights $\pi_1, \pi_2, \ldots, \pi_N$, fulfilling $\sum_{i=1}^{N} \pi_i = 1$. The JS-divergence measures how 'far' are the probabilities p_i from their likely joint source $\sum_{i=1}^{N} \pi_i p_i$ and equals zero if and only if all the p_i are equal. It is important to note that the JS-divergence is identical to $I(X,Y)$ when $\pi_i = p(x_i)$ and $p_i = p(Y|x_i)$ for each $x_i \in \mathcal{X}$, where $p(X) = \{p(x_i)\}$ is the input distribution, $p(Y|x_i) = \{p(y_1|x_i), p(y_2|x_i), \ldots, p(y_M|x_i)\}$, $N = |\mathcal{X}|$, and $M = |\mathcal{Y}|$ [BR82, ST00b].

2.2. Viewpoint Quality Measures

We review now some viewpoint quality measures for polygonal models. In [PB96], the quality of a viewpoint v of a scene is computed using the *heuristic measure* (HM) given by

$$C(v) = \frac{\sum_{i=1}^{n} \left\lceil \frac{P_i(v)}{P_i(v)+1} \right\rceil}{n} + \frac{\sum_{i=1}^{n} P_i(v)}{r} \ , \tag{2.7}$$

where $P_i(v)$ is the number of pixels corresponding to the polygon i in the image obtained from the viewpoint v, r is the total number of pixels of the image (resolution of the image), and n is the total number of polygons of the scene. In this formula, $\lceil x \rceil$ denotes the smallest integer, greater than or equal to x. The first term in (2.7) gives the fraction of visible surfaces with respect to the total number of surfaces, while the second term is the ratio between the projected area of the scene (or object) and the screen area (thus, its value is 1 for a closed scene).

From (2.1), the *viewpoint entropy* (VE) [VFSH01] has been defined from the relative area of the projected polygons over the sphere of directions centered at viewpoint v. Thus, the viewpoint entropy was defined by

$$H_v = -\sum_{i=0}^{N_f} \frac{a_i}{a_t} \log \frac{a_i}{a_t}, \tag{2.8}$$

where N_f is the number of polygons of the scene, a_i is the projected area of polygon i over the sphere, a_0 represents the projected area of background in open scenes, and $a_t = \sum_{i=0}^{N_f} a_i$ is the total area of the sphere. The maximum entropy is obtained when a certain viewpoint can see all the polygons with the same projected area. The best viewpoint is defined as the one that has maximum entropy. In molecular visualization, both maximum and minimum entropy views show relevant characteristics of a molecule [VFSL06].

From (2.5), a viewpoint quality measure, called *viewpoint Kullback-Leibler distance* (VKL) [SPFG05], has been defined by

$$KL_v = \sum_{i=1}^{N_f} \frac{a_i}{a_t} \log \frac{\frac{a_i}{a_t}}{\frac{A_i}{A_T}}, \tag{2.9}$$

where a_i is the projected area of polygon i, $a_t = \sum_{i=1}^{N_f} a_i$, A_i is the actual area of polygon i and $A_T = \sum_{i=1}^{N_f} A_i$ is the total area of the scene or object. The VKL measure is interpreted as the distance between the normalized distribution of projected areas and the 'ideal' projection, given by the normalized distribution of the actual areas. In this case, the background can not be taken into account. The minimum value 0 is obtained when the normalized distribution of projected areas is equal to the normalized distribution of actual areas. Thus, to select views of high quality means to minimize KL_v.

Apart from the previous references on viewpoint quality measures, [PPB*05] describe a number of different ways to measure the goodness of a view of an object. After analyzing different view descriptors, they conclude that no single descriptor does a perfect

job and possibly a combination of them would amplify the advantage that each one has. Given a sphere of viewpoints, [YSY*06] computes the similarity between each two disjoint views using Zernike moments analysis and obtain a similarity weighted spherical graph. A view is considered to be stable if all edges incident on its viewpoint in the spherical graph have high similarity weights. [AVF04] and [SPT06] present two different exploration algorithms guided by viewpoint entropy and the total curvature of a visible surface, respectively. In the volume rendering field, [BS05], [TFTN05] and [JS06] use an extended version of viewpoint entropy and [VFSG06] introduces the viewpoint mutual information. [CSCF07] uses viewpoint entropy as a perceptual measure for mesh simplification.

Based on the investigation on canonical views, [GRMS01] presents a new method for constructing images, where the viewpoint is chosen to be both off-axis and 'natural', and [LME06] obtains the viewing direction from the combination of factors such as saliency, occlusion, stability and familiarity. [LVJ05] has introduced the saliency as a measure for regional importance for graphics meshes and [KV06] presented a visual-saliency-based operator to enhance selected regions of a volume. [GCO06] introduced a method for partial matching of surfaces by using the abstraction of salient geometric features and a method to construct them.

2.3. Obscurances and Ambient Occlusion

Let us take a look at the illumination of the objects in the real world. Imagine we are in an environment where the illumination is mostly diffuse and indirect, as if a clear white wall dominated the scene or we were in open air in a cloudy day. In these cases, the objects that are more hidden are seen darker, as the indirect light that comes from mostly everywhere is occluded by other objects. Modelling this effect, that we will call obscurances from now on, following [ZIK98, IKSZ03], is much more simple and much less costly than global illumination. In global illumination we need to simulate the interaction of light between all objects of our scene. The obscurances effect can be considered as a pure geometric property of every point in our scene: we just need to evaluate the hiddenness or occlusion of the point by considering the objects around it.

2.3.1. The Obscurances Illumination model

The Obscurances illumination model (see Figure 2.1) decouple direct and indirect illumination and were first introduced in the videogame context as a technique to allow fast editing of indirect illumination. The high quality of shadowing obtained made them later to be included in production replacing radiosity. In the obscurances model,

(*a*) (*b*) (*c*)

Figure 2.1.: Several models with obscurances: A car (a), a tank (b) and a cathedral (c) model. Thanks to Sergi Funtané and Nicolau Sunyer for the images.

obscurance W is given by

$$W(x) = \frac{1}{\pi} \int_{\Omega} \rho(d(x,\omega)) \cos\theta d\omega, \tag{2.10}$$

where ρ is a function of the distance $d(x,\omega)$ of the first intersection of a ray shot from x with direction ω, x is a surface point, θ is the angle between the normal vector at x and direction ω, and the integration is over the hemisphere oriented according to the surface normal.

The function ρ is a monotone increasing function of the distance d. It is defined for all positive values and results in 0 when d is 0. From 0 to a determined value $dmax$, the function increases from 0 to 1, and for values greater than $dmax$ the returned value is 1. This means that we only consider a limited environment around the point x and beyond it we will not take care of the occlusions. The shape of the function ρ is shown in Figure 2.2.

In this way, the integral function $W(x)$ captures the geometric properties of the environment of the point x. If we take a look at the extreme cases, an obscurance value of 1 means that the point is completely open (not occluded) and a value of 0 would mean that it is completely occluded. This would be a very strange case, there exist points with an obscurance value of 0, but as they are completely occluded, we can not see them. The $dmax$ (maximum distance) is the main parameter that controls the indirect illumination computed with obscurances. The user chooses the value of $dmax$ depending on the quantity of shadow she/he needs for the scene. It should be in concordance with the relative sizes of the objects with respect to the scene and with the size of the scene itself. It should be much larger, for example, if we compute a view of a stadium

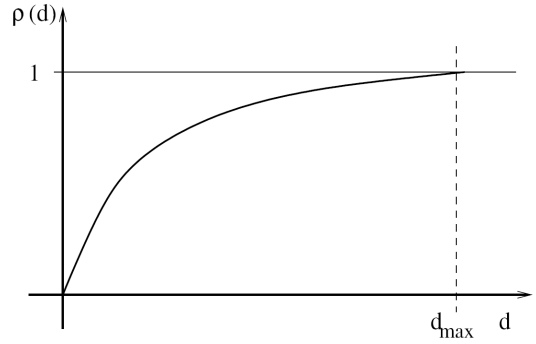


Figure 2.2.: Shape of funtion $\rho(d)$.

than if we compute a close-up of a foot and a ball.

The obscurances illumination model is thought to be used to simulate the indirect lighting caused by diffuse interreflections between objects in a fast and simple way. The direct lighting has to be computed apart and in an independent way. Fast, simple and known techniques to compute direct lighting are commonly used. We can take any of them and add the direct lighting results to the indirect lighting computation.

For more realistic results, the indirect lighting computed with obscurances has to correlate with physically realistic indirect lighting. In particular, an image of a scene computed with any global illumination technique, as path tracing, and an image of the same scene with the same camera and light sources computed with obscurances have to look similar, specially in average intensities. For this reason, the obscurance value of a point has to be used in the following way to obtain its indirect illumination:

$$I(x) = R(x) \times I_A \times W(x) \tag{2.11}$$

that is, the obscurance at point x is multiplied by the diffuse reflectivity ($R(x)$) at the point and by an average intensity value (I_A) of the whole scene.

The average ambient intensity (I_A) is computed assuming that light energy is distributed in a uniform way among the whole scene and illuminates all objects with the same intensity. In this way I_A is:

$$I_A = \frac{R_{avg}}{1 - R_{avg}} \times \frac{1}{A_{total}} \sum_{i=1}^{n} A_i \times E_i \tag{2.12}$$

where A_{total} and A_i are the total area and the area of each patch respectively, E_i is the emittance of the patch and R_{avg} is the average reflectivity of all patches weighted by the

area:

$$R_{avg} = \frac{1}{A_{total}} \sum_{i=1}^{n} A_i \times R_i \tag{2.13}$$

2.3.2. Ambient Occlusion

The obscurances, as described previously, have inspired a family of techniques that nowadays are implemented in most render software packages commonly used by the animation and videogame industry. It all started in a course about the widely known and used software package RenderMan at SIGGRAPH 2002, where two of the speakers talked about a technique that they called *ambient occlusion* and had been using each one in a different movie that same year. Hayden Landis [Lan02] from *Industrial Light & Magic* used the ambient occlusion for self-shadowing of objects to "attenuate properly the surfaces not fully exposed to the environment". The examples were the planes from the movie Pearl Harbor. They got the unoccluded lighting from a blurred environment map looked up in a direction that averaged the unoccluded zone of the surroundings of the hit point and called that direction bent normal.

In the same course and a few hours later, Bredow [Bre02] presented the use of the obscurance effect for the more hidden zones of the *Pishkin building*, a fictional skyscrapper in the movie Stuart Little 2, and computed it by putting a large area light source over the building and another one less intense under it, saving the results in textures to reuse it in different frames.

Ambient occlusion (see Figure 2.3)

$$A(x) = \frac{1}{\pi} \int_{\Omega} V(x, \omega) \cos\theta d\omega, \tag{2.14}$$

substitutes the ρ function in the obscurances (formula 2.10) by the visibility function $V(w, x)$ that has value zero when no geometry is visible in direction ω and one otherwise.

2.3.3. Obscurances vs. Ambient Occlusion

Ambient occlusion as presented in the three presentations at SIGGRAPH represents a simplification of the obscurances as understood in previous section in two ways. First, the distance of the ray intersection to the original point is not taken into account, and the function just results in 0 (intersects) or 1 (does not intersect). In a more conceptual way we could think of ambient occlusion as a simple percentage of openness of a point, and

(*a*) (*b*) (*c*)

Figure 2.3.: Several models with ambient occlusion: An indoor scene (a), a human model (b) and a molecule (c).

in obscurances the secondary interreflections are taken into account and some lighting intensity is added, that is, diffuse indirect lighting is actually taken into account (see Figure 2.4). Second, the ambient intensity and average reflectivity parameters are normally not used, and some ambient parameter is adjusted in an empirical way. It is easy to understand that physical accuracy is not really important in production rendering and even less in videogames, as we only need to find realistic and visually pleasant images. Only in other kind of applications as for example for architecture or interior design, accurate lighting is important. This is why ambient occlusion has become so popular in CGI and it is becoming widely used in videogames too.

Obscurances vs. ambient occlusion

- The use of a continuous function of the distance in obscurances to account for indirect lighting gives better results than ambient occlusion and allows the addition of color bleeding.
- The average ambient intensity and average reflectivity are used in obscurances to search for a more accurate physical lighting.
- Both work in open environments.
- Both illumination models are capable to simulate the ambient lighting distribution in an environment in a non-constant way and make possible to render realistic images.
- Ambient occlusion and obscurances techniques are faster than global illumination ones, but they not compute lighting in a physically based and accurated way.

(*a*) (*b*)

Figure 2.4.: A comparison between ambient occlusion (a) and obscurances with color bleeding (b). Thanks to Àlex Méndez for the images.

2.3.4. State of the art

Now we are going to introduce some of the most recent research done about obscurances and ambient occlusion. Méndez et al. [MSC03] introduced color bleeding, updated the obscurances dynamically in the presence of moving objects and dealt with the problem of important secondary reflectors. Later in [MSC*06] obscurances were computed in the GPU using the depth-peeling technique. Sattler et al. [SSZK04] compute the visibility from the vertices of the object to the vertices of an hemispherical mesh using the GPU. They also utilize the coherence in the visibility function to achieve interactive frame rates with deformable objects with illumination coming from point light sources at the vertices of the hemisphere. Bunnell [Bun05] approximates the mesh triangles of the scene using disks, and combines the occlusion of multiple disks heuristically. The visibility is approximated by an iterative algorithm. Kontkanen and Laine [KL05] precompute an ambient occlusion field around each rigid object. Recently, Kontkanen and Aila [KA06] apply ambient occlusion to animated characters by blending the textures obtained for the different positions of the character.

2.4. 3D Shape Retrieval

In this section we discuss several issues related to 3D Shape Retrieval.

2.4.1. 3D Shape Retrieval Framework

Recent developments in techniques for modelling, digitizing and visualizing 3D shapes has provoked an explosion in the number of available 3D models on the internet and in specific databases. This has led to the development of 3D shape retrieval systems (see

[JT04] for a survey) that, given a query object, retrieve similar 3D objects.

At conceptual level, a typical shape retrieval framework (see Figure 2.5) consists of a database with an index structure created offline and an online query engine. Each 3D model has to be identified with a shape descriptor, providing an overall description of its shape. The indexing data structure and the searching algorithm are used to search efficiently. The online query engine computes the query descriptor and the models similar to the query model are retrieved by matching descriptors to the query descriptor from the index structure of the database. The similarity between two descriptors is quantified by a dissimilarity measure.

According to [JT04] usually, 3D shape retrieval systems are usually evaluated with respect to several requirements of content based 3D retrieval, such as:

- **Shape representations requirements.** Most of the 3D models found on the World Wide Web are meshes defined in a file format supporting visual appearance. These models are represented by "polygon soups", consisting of unorganized sets of polygons.
- **Properties of dissimilarity measures.** In order to measure how similar two objects are, it is necessary to compute distances between pairs of descriptors using a dissimilarity measure.
- **Efficiency.** For large shape collections, it is inefficient to sequentially match all objects in the database with the query object. Because retrieval should be fast, efficient indexing search structures are needed to support efficient retrieval.
- **Discrimination abilities.** A shape descriptor should capture properties that discriminate objects well.
- **Ability to perform partial matching.** In contrast to global shape matching, partial matching finds a shape of which part is similar to a part of another shape.
- **Robustness.** It is often desirable that a shape descriptor is insensitive to noise and small extra features, and robust against arbitrary topological degeneracies. Also if the model is given in multiple levels of detail, these should not differ significantly from the original model.
- **Necesity of pose normalization.** It is common that models have arbitrary scale, orientation and position in the 3D space. Because not all dissimilarity measures are invariant under rotation and translation, it may be necessary to place the 3D models into a canonical coordinate system.

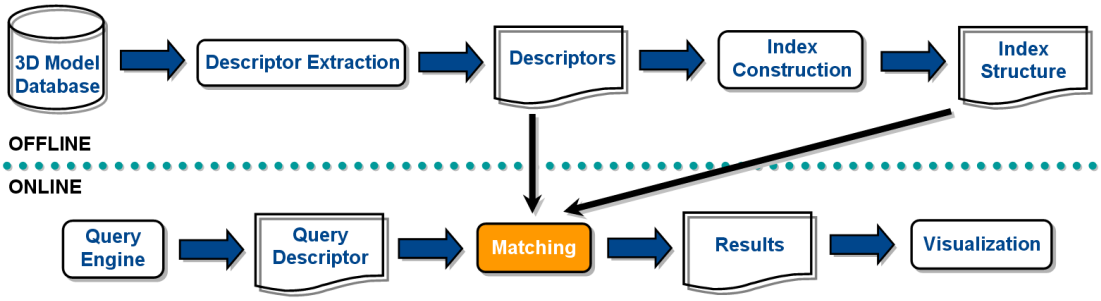


Figure 2.5.: Conceptual framework for shape retrieval.

2.4.2. Shape Matching Methods

Shape matching methods can be divided in three broad categories: *Feature-based methods*, *Graph based methods* and *View based methods*.

Feature based Methods

In the context of 3D shape matching, features denote geometric and topological properties of 3D shapes. So, 3D shapes can be discriminated by measuring and comparing their features. Feature based methods can be divided into four categories according to the type of shape features used: *Global Features*, *Global Feature distributions*, *Spatial Maps and Local Features*.

Feature-based methods from the first three categories represent features of a shape using a single descriptor consisting of a d-dimensional vector of values, where the dimension d is fixed for all shapes. The value of d can easily be a few hundred. The descriptor of a shape is a point in a high dimensional space, and two shapes are considered to be similar if they are close in this space. Retrieving the k best matches for a 3Dquery model is equivalent to solving the k nearest neighbor problem. Using the Euclidean distance, matching feature descriptors can be done efficiently in practice by searching in multiple 1D spaces to solve the approximate k nearest neighbor problem as shown by Indyk and Motwani [IM98]. In contrast with the feature based methods from the first three categories, local feature-based methods describe for a number of surface points the 3D shape around the point. For this purpose, for each surface point a descriptor is used instead of a single descriptor.

Graph based methods

In general, the feature-based methods take into account only the pure geometry of

the shape. In contrast, graph based methods attempt to extract a geometric meaning from a 3D shape using a graph showing how shape components are linked together. Graph based methods can be divided into three broad categories according to the type of graph used: *Model graphs*, *Reeb graphs*, and *Skeletons*. For an extensive discussion of Reeb graphs and skeletons we refer the reader to the paper of Biasotti et al. [SBP03]. Efficient computation of existing graph metrics for general graphs is not possible: computing the edit distance is *NP*-hard [KZS96] and computing the maximal common subgraph [GJ79] is even *NP*-complete. Polynomial solutions can be obtained for directed acyclic graphs such as shock graphs. Sebastian et al. [TBSK01] describe an approach to compute a pseudo-metric between shock graphs. It is obtained by exhaustively searching for the optimal deformation path between two 2D shapes, and using the cost of this path as a distance between two shapes. But the computation time of this method is too high for practical application, and it is not straightforwardly generalized to 3D.

View based methods

The main idea behind view-based similarity methods is that two 3D models are similar, if they look similar from all viewing angles. A natural application of this paradigm is the implementation of query interfaces based on defining a query by one or more sketches showing the query from different views.

In [LÖ0] authors apply view-based similarity to retrieve 3D models using a 2D query interface. In the preprocessing phase, for each 3D model a descriptor is obtained consisting of a number of binary images. In the query phase, a sketch or a 2D image is used as a query to retrieve a number of 3D models, whose images match the query. Also, Funkhouser et al. [FMK*02] apply view-based similarity to implement a 2D sketch query interface. In the preprocessing phase a descriptor of each 3D model is obtained by 13 thumbnail images of boundary contours of the 3D object as seen from 13 orthographic view directions. Then in the query phase the user defines a 3D shape query by drawing one or more sketches. 3D shape models are retrieved by comparing these sketches with the descriptors from the shapes in the database using image matching.

3. Viewpoint Channel

In this chapter, we introduce an information channel between a set of viewpoints and the set of polygons of an object to deal with viewpoint selection. Then we define the viewpoint mutual information to select the most representative views of an object and compare the behavior of this measure with the ones reviewed in Section 2.2. Finally, viewpoint stability is defined from the notion of dissimilarity between two viewpoints, which is given by the Jensen-Shannon divergence between their respective distributions.

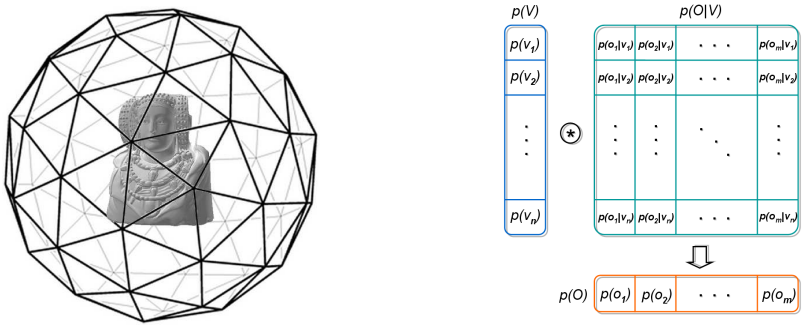
3.1. Viewpoint Mutual Information

Our viewpoint selection framework is constructed from an information channel $V \rightarrow O$ between the random variables V (input) and O (output), which represent, respectively, a set of viewpoints and the set of polygons of an object (see Figure 3.1(a)). This channel, which we call *viewpoint channel*, is defined by a conditional probability matrix obtained from the projected areas of polygons at each viewpoint. Viewpoints will be indexed by v and polygons by o. Throughout this document, the capital letters V and O as arguments of $p()$ will be used to denote probability distributions. For instance, while $p(v)$ will denote the probability of a single viewpoint v, $p(V)$ will represent the input distribution of the set of viewpoints.

The viewpoint channel can be interpreted as an *observation channel* where the conditional probabilities represent the probability of *seeing* a determined polygon from a given viewpoint (see Figure 3.1(b)). The three basic elements of this channel are:

- Conditional probability matrix $p(O|V)$, where each element $p(o|v) = \frac{a_o}{a_t}$ is defined by the normalized projected area of polygon o over the sphere of directions centered at viewpoint v. Conditional probabilities fulfil $\sum_{o \in \mathcal{O}} p(o|v) = 1$. In our approach, background is not taken into account but it could be considered as another polygon.

- Input distribution $p(V)$, which represents the probability of selecting a viewpoint. In our experiments, $p(V)$ will be obtained from the normalization of the projected area of the object at each viewpoint. This can be interpreted as the probability that



(*a*) Viewpoint sphere. (*b*) Probability distributions of channel $V \rightarrow O$.

Figure 3.1.: Viewpoint information channel.

a random ray originated at v hits (*sees*) the object. This assignation is consistent with the objective of selecting the viewpoints which *see* more projected area. Let us remember that this is a characteristic of a canonical view (see Chapter 1). The input distribution can also be interpreted as the *importance* assigned to each viewpoint v. For instance, the input distribution could also be defined by $p(v) = \frac{1}{N_v}$, where N_v is the number of viewpoints.

- Output distribution $p(O)$, defined by

$$p(o) = \sum_{v \in \mathcal{V}} p(v) p(o|v), \tag{3.1}$$

which represents the average projected area of polygon o, i.e., the probability of polygon o to be hit (*seen*) by a random ray cast from the viewpoint sphere.

From the previous definitions, the *conditional entropy* (2.2) is given by the average of all viewpoint entropies:

$$H(O|V) = -\sum_{v \in \mathcal{V}} p(v) \sum_{o \in \mathcal{O}} p(o|v) \log p(o|v) = \sum_{v \in \mathcal{V}} p(v) H(O|v), \tag{3.2}$$

where $H(O|v) = -\sum_{o \in \mathcal{O}} p(o|v) \log p(o|v)$ is the *viewpoint entropy* H_v (2.8) and measures the degree of uniformity of the projected area distribution at viewpoint v. Let us observe that H_v has been now rewritten in a different form. Both entropies $H(O|v)$ and $H(O|V)$ tend to infinity when polygons are infinitely refined. This makes these measures very sensitive to the discretisation of the object and in general not appropriate to evaluate the quality of a viewpoint.

We now devote our attention to the *mutual information* (2.3) between V and O, that expresses the degree of *dependence* or *correlation* between the set of viewpoints and the

object. From (2.3), mutual information is given by

$$I(V,O) = \sum_{v \in \mathcal{V}} p(v) \sum_{o \in \mathcal{O}} p(o|v) \log \frac{p(o|v)}{p(o)} = \sum_{v \in \mathcal{V}} p(v) I(v,O), \tag{3.3}$$

where we define

$$I(v,O) = \sum_{o \in \mathcal{O}} p(o|v) \log \frac{p(o|v)}{p(o)} \tag{3.4}$$

as the *viewpoint mutual information* (VMI), which gives us the degree of dependence between the viewpoint v and the set of polygons, and it is a measure of the *quality* of viewpoint v. Consequently, mutual information $I(V,O)$ can be interpreted as the average viewpoint quality. *Quality* is considered here equivalent to *representativeness*. It is also important to indicate that the level of resolution of the viewpoint sphere will determine the accuracy of the measures.

In our framework, the best viewpoint is defined as the one that has *minimum* VMI. High values of the measure mean a high dependence between viewpoint v and the object, indicating a highly *coupled* view (for instance, between the viewpoint and a small number of polygons with low average visibility). On the other hand, the lowest values correspond to the most *representative* or *relevant* views, showing the maximum possible number of polygons in a balanced way.

3.2. Discussion

Note that $I(v,O) = KL(p(O|v)|p(O))$, where $p(O|v)$ is the conditional probability distribution between v and the object and $p(O)$ is the marginal probability distribution of O, which in our case corresponds to the distribution of the average of projected areas. It is worth observing that $p(O)$ plays the role of the *target* distribution in the KL distance and also the role of the *optimal* distribution since our objective is that $p(O|v)$ becomes similar to $p(O)$ to obtain the best views. On the other hand, this role agrees with intuition since $p(O)$ is the average visibility of polygon o over all viewpoints, i.e., the *mixed distribution* of all views, and we can think of $p(O)$ as representing, with a single distribution, the knowledge about the scene. Note the difference between VMI (3.4) and VKL (2.9), due to the fact that in the last case the distance is taken with respect to the actual areas.

In [VFSG06], it has been shown that the main advantage of VMI over VE is its robustness to deal with any type of discretisation or resolution of the volumetric dataset. The same advantage can be observed for polygonal data. Thus, while a highly refined mesh

will attract the attention of VE, VMI will be almost insensitive to changes in the mesh resolution. This behavior of both measures with respect to the discretization can be deduced from the mathematical analysis of VE and VMI. For instance, let us assume that a regular polygon o of the object is subdivided into two equal parts o_1 and o_2 such that $p(o_1|v) = p(o_2|v)$, $p(o_1) = p(o_2)$, $p(o|v) = p(o_1|v) + p(o_2|v)$ and $p(o) = p(o_1) + p(o_2)$. Assuming that only the term referred to polygon o can change in the formulas for VE (2.8) and VMI (3.4), we analyze their variation after the subdivision of o. The variation of VE is given by

$$\delta H(O|v) = -p(o_1|v)\log p(o_1|v) - p(o_2|v)\log p(o_2|v) - (-p(o|v)\log p(o|v)) = p(o|v).$$

Therefore, VE increases with a value $p(o|v)$ after the subdivision. On the other hand, the variation of VMI is given by

$$\delta I(v,O) = p(o_1|v)\log\frac{p(o_1|v)}{p(o_1)} + p(o_2|v)\log\frac{p(o_2|v)}{p(o_2)} - p(o|v)\log\frac{p(o|v)}{p(o)} = 0.$$

Thus, VMI remains invariant to the proposed subdivision. In general, if we compare both measures for different discretisations, mutual information will give similar results and VE will show an erratic behavior. Note that HM is also highly dependent on the discretisation, since the first term in (2.7) is given by the quotient between the number of visible polygons and the total number of polygons. The behavior of all these measures with respect to the discretisation will be experimentally shown in the next section.

3.3. Results

In this section, the behavior of VMI (3.4) is compared with the one of HM (2.7), VE (2.8), and VKL (2.9). To compute these viewpoint quality measures, we need a *preprocess step* to estimate the projected area of the visible polygons of the object at each viewpoint. Before projection, a different color is assigned to each polygon. The number of pixels with a given color divided by the total number of pixels projected by the object gives us the relative area of the polygon represented by this color (conditional probability $p(o|v)$). Although all these measures are sensitive to the size of the viewpoint sphere with respect to the object, in this master thesis we have not taken into account this parameter. For comparison purposes, all measures have been computed without taking into account the background.

In our experiments, all the objects are centered in a sphere of 642 viewpoints built from the recursive discretisation of an icosahedron and the camera is looking at the center of this sphere. The projection resolution adopted is 640x480 and increasing it would lead us to obtain a more accurate results with the drawback of a higher computational cost of the preprocess step. In Table 3.1 we show the number of polygons

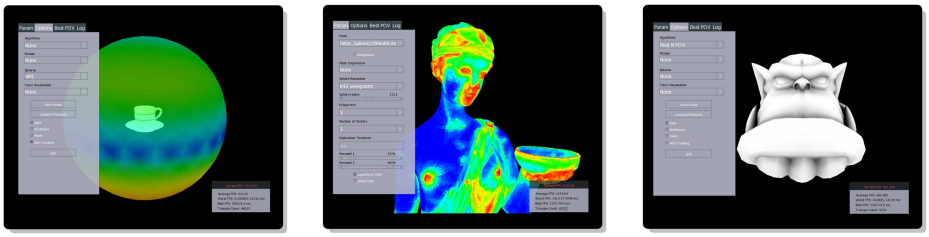


Figure 3.2.: The figure shows the interface of our viewpoint software.

	Cow	Coffee cup	Ship	Lady of Elche
Number of triangles	9593	43935	47365	51978
Computational cost	41 sec	81 sec	62 sec	80 sec

Table 3.1.: Number of triangles of the models used and computational cost of the preprocess step for each model.

of the models used in this section and the cost of the preprocess step, i.e., the cost of computing the probability distributions $p(V)$, $p(O|V)$ and $p(O)$. Even though a large number of viewpoints have been used, a high quality can be also achieved with much less viewpoints and the consequent reduction of timings. To show the behavior of the measures, the sphere of viewpoints is represented by a color map, where red and blue colors correspond respectively to the best and worst views. Note that a good viewpoint corresponds to a high value for both HM (2.7) and VE (2.8), and to a low value for both VKL (2.9) and VMI (3.4). Figure 3.2 shows the interface of our viewpoint software created using the 3D-rendering engine Ogre3D (`http://www.ogre3d.org`). Our tests were run on a Pentium IV 3GHz machine with 2 GB RAM and an Nvidia GeForce 8800 GTX with 768 MB.

(a) (b) (c)

Figure 3.3.: Cow, ship and lady of Elche wireframe models.

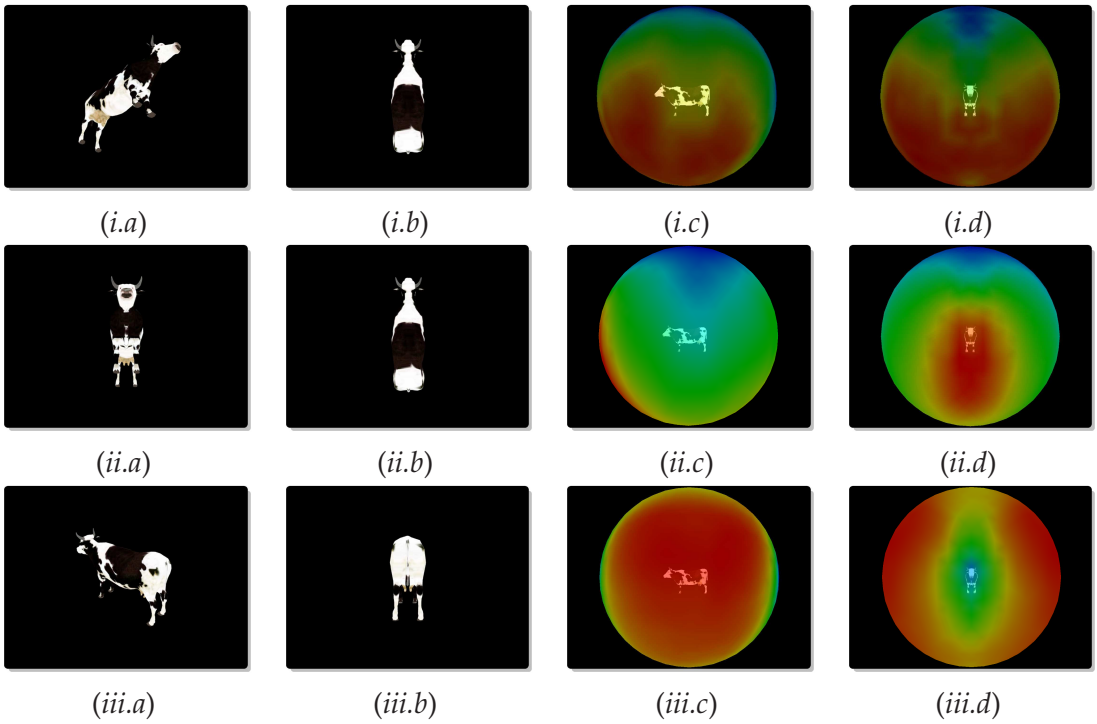


Figure 3.4.: (a) The most representative and (b) the most restricted views, and (c-d) the viewpoint spheres obtained respectively from the (i) HM, (ii) VE and (iii) VMI measures. Red colors on the sphere represent the highest quality views and blue colors represent the lowest quality views.

To evaluate the performance of the four viewpoint quality measures presented, five models have been used: a cow (Figure 3.3(a)), two coffee-cup-and-dish with two different discretisations of the dish (Figures 3.5(i.a) and 3.5(ii.a)), a ship (Figure 3.3(b)), and the lady of Elche (Figure 3.3(c)). Figure 3.4 has been organized as follows. Rows (i), (ii) and (iii) show, respectively, the behavior of HM, VE and VMI measures. Columns (a) and (b) show, respectively, the best and worst views, and columns (c) and (d) show two different projections of the viewpoint spheres. Figure 3.4 illustrates how VMI selects better views than both HM and VE. Observe how VE chooses to *see* the most highly discretised parts of the cow. The same occurs with HM, although this one also searches for a view with higher projected area. While the worst views for the HM and VE measures correspond to the ones that see the less discretised parts, in the VMI case a true restricted view is obtained.

Figure 3.5 shows the behavior of the HM, VE and VMI measures when the discretisation of the object varies outstandingly. Rows (i) and (ii) show the viewpoint spheres

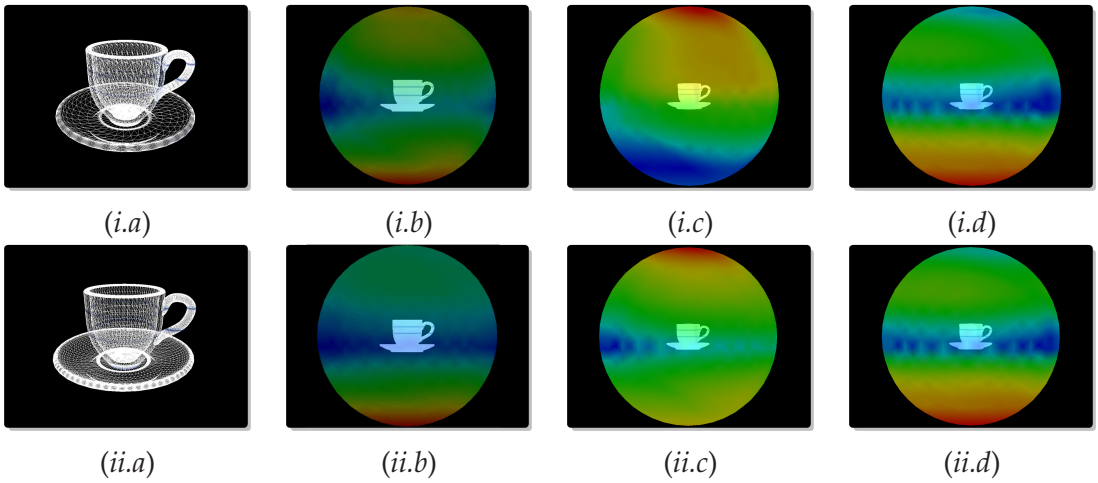


Figure 3.5.: Column (a) shows the models used to compute the viewpoint spheres. These are obtained respectively from (b) HM, (c) VE and (d) VMI measures.

computed respectively for the coffee-cup-and-dish model of Figure 3.5(i.a) and for the same model with a more refined dish (Figure 3.5(ii.a)). We can clearly observe how the spheres obtained from HM and VE change according to the discretisation variation, whereas VMI spheres are almost insensitive to this variation.

The different behavior between VKL (a-b) and VMI (c-d) is shown in Figure 3.6. Remember that the main difference between VMI and VKL is that while the former computes the distance between the projected areas of the polygons and the average area *seen* by the set of viewpoints, the later calculates the distance with respect to the *actual* areas of polygons. Due to this fact, the reliability of VKL is outstandingly affected by the existence of many non visible or poorly visible polygons, as in the case of the ship and lady of Elche models.

3.4. Viewpoint Similarity and Stability

As we have mentioned in Chapter 1, a basic property of a canonical view is its *stability* [BTB99]. That is, observers prefer a view which minimally changes when it is moved within its nearest neighborhood.

The use of Jensen-Shannon as a measure of *view similarity* has been previously proposed by [BS05] in the volume rendering field. In our approach, this measure appears naturally from the variation of the viewpoint quality (VMI).

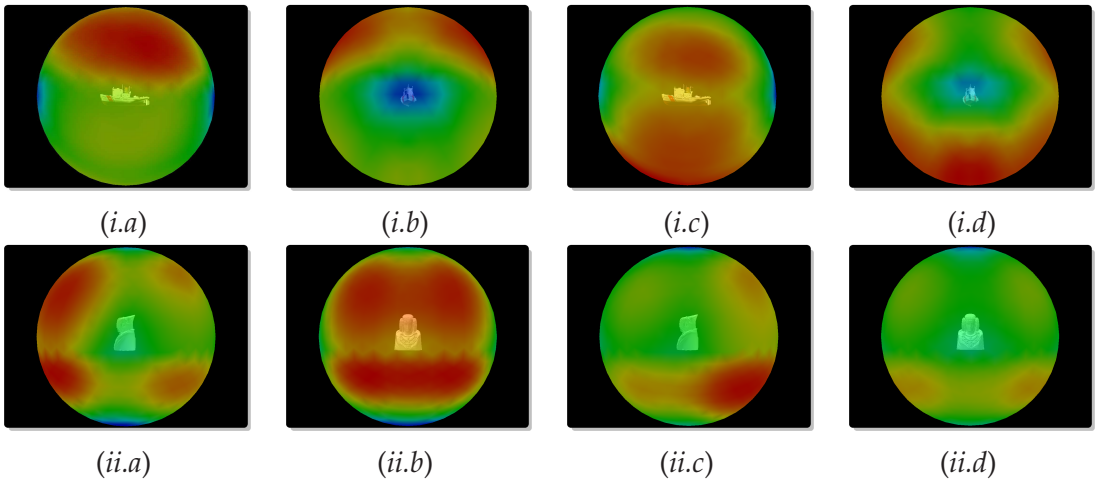


Figure 3.6.: Viewpoint spheres obtained respectively from the (a-b) VKL and (c-d) VMI measures.

If we apply the data processing inequality (2.4) to the channel $V \rightarrow O$, we find that any clustering over V or O, respectively denoted by $\widehat{V}$ and $\widehat{O}$, will reduce $I(V,O)$. Therefore, if neighbor viewpoints (or polygons) are clustered, then $I(\widehat{V},O) \leq I(V,O)$ (or $I(V,\widehat{O}) \leq I(V,O)$). The result of clustering (or merging) two viewpoints v_i and v_j is defined as a 'virtual' viewpoint $\widehat{v} \equiv v_i \oplus v_j$ such that

$$p(\widehat{v}) = p(v_i \oplus v_j) = p(v_i) + p(v_j) \tag{3.5}$$

and

$$p(o|\widehat{v}) = p(o|v_i \oplus v_j) = \frac{p(v_i)p(o|v_i) + p(v_j)p(o|v_j)}{p(\widehat{v})}. \tag{3.6}$$

The reduction of MI when two viewpoints v_i and v_j are merged is given by

$$\begin{aligned} \delta I(v_i,v_j) &= I(V,O) - I(\widehat{V},O) \\ &= \big(p(v_i)I(v_i,O) + p(v_j)I(v_j,O)\big) - p(\widehat{v})I(\widehat{v},O) \\ &= p(\widehat{v})\left(\frac{p(v_i)}{p(\widehat{v})}I(v_i,O) + \frac{p(v_j)}{p(\widehat{v})}I(v_j,O) - I(\widehat{v},O)\right) \\ &= p(\widehat{v})D(v_i,v_j), \end{aligned} \tag{3.7}$$

where we define

$$D(v_i,v_j) = \frac{p(v_i)}{p(\widehat{v})}I(v_i,O) + \frac{p(v_j)}{p(\widehat{v})}I(v_j,O) - I(\widehat{v},O) \tag{3.8}$$

as the *viewpoint dissimilarity* between v_i and v_j. That is, the loss of information when two viewpoints are merged is interpreted as the dissimilarity between them. Note that

the dissimilarity will be null when the two viewpoints capture the same distribution of projected areas: if $p(O|v_i) = p(O|v_j)$, then $\delta I_{v_i,v_j} = 0$.

It can be shown (see [ST00a]) that the viewpoint dissimilarity can also be written as

$$D(v_i, v_j) = JS\left(\frac{p(v_i)}{p(\widehat{v})}, \frac{p(v_j)}{p(\widehat{v})}; p(O|v_i), p(O|v_j)\right), \tag{3.9}$$

where the second term is the Jensen-Shannon divergence (2.6) between the distributions $p(O|v_i)$ and $p(O|v_j)$ captured by v_i and v_j with weights $\frac{p(v_i)}{p(\widehat{v})}$ and $\frac{p(v_j)}{p(\widehat{v})}$, respectively. If two views are very similar, i.e., the JS-divergence between them is small, the channel can be simplified by substituting these two viewpoints by their merging, without a significant loss of information.

Two interesting properties follow:

- It can be easily seen that the clustering of all viewpoints would give $\delta I = I(V, O)$ and, thus, $I(\widehat{V}, O) = 0$ (see Section 2.1).
- $H(O) = H(O|V) + I(V, O) = H(O|\widehat{V}) + I(\widehat{V}, O)$, where $H(O)$ is the entropy of $p(O)$. Note that if two viewpoints are clustered the decrease of $I(V, O)$ is equal to the increase of $H(O|V)$ since $H(O)$ remains constant (the discretisation of the object has not been changed).

View unstability was defined by [BS05] as the maximum change in view that occur when the camera position is small shifted within a neighborhood. Thus, a small change corresponds to a stable viewpoint and a large change to an unstable one. We now define the unstability of a viewpoint v as the average variation of dissimilarity between v and its neighbor viewpoints. That is, v_i is a stable viewpoint if $p(O|v_i)$ is close to the probability distributions $p(O|v_j)$ of its neighbors, where v_j stands for a neighbor of v_i. Thus, the *viewpoint unstability* of v_i is defined by

$$U(v_i) = \frac{1}{N_n} \sum_{j=1}^{N_n} D(v_i, v_j), \tag{3.10}$$

where v_j is a neighbor of v_i and N_n is the number of neighbors of v_i.

Figure 3.7 shows the behavior of the viewpoint stability measure for the coffee-cup-and-dish, cow and lady of Elche models. Observe how the results obtained agree with intuition.

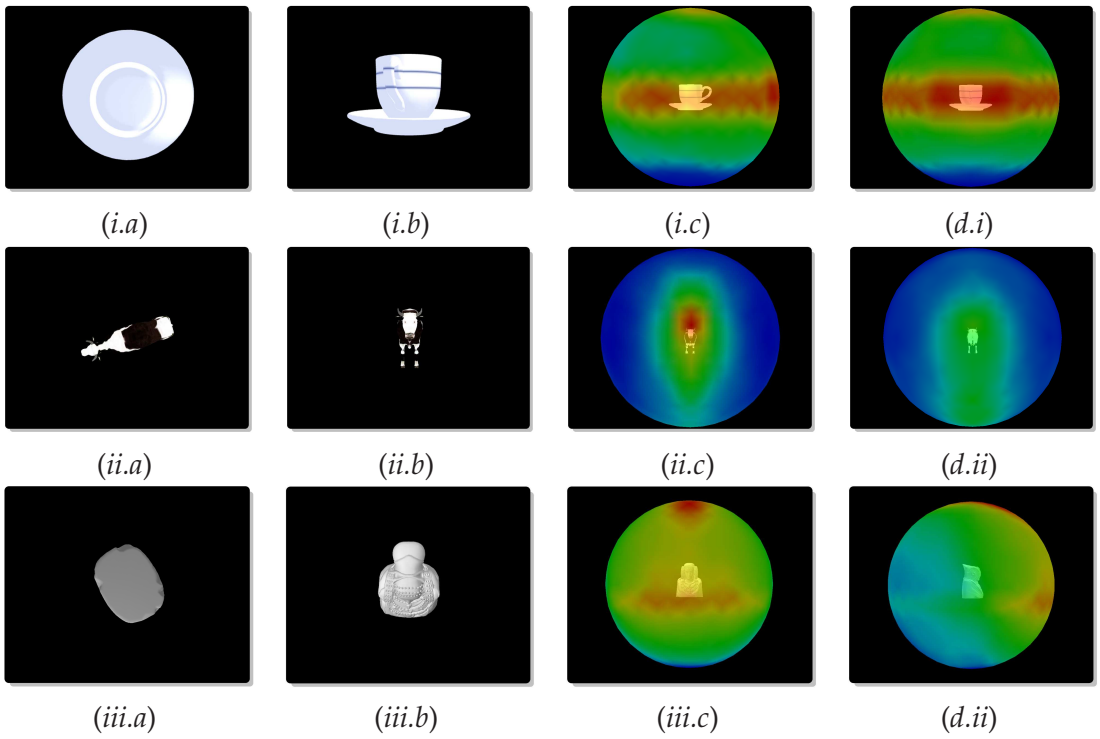


Figure 3.7.: The (a) most stable and (b) most unstable viewpoints, and (c-d) unstability spheres obtained for the (i) coffee-cup-and-dish, (ii) cow and (iii) lady of Elche models. Red colors on the sphere represent high unstability values, blue colors represent low unstability values.

4. Best View Selection and Object Exploration

In order to understand or model an object, we are interested in selecting a set of representative views. This set has to provide a complete representation of the object, i.e., a simplified representation of the information provided by all viewpoints. In this chapter, new algorithms based on the concepts introduced in previous chapters (see Chapter 3) are applied to both the selection of the N best representative views and object exploration.

4.1. Selection of N Best Views

With the goal of obtaining the best representation of the object using the minimum number of views, a new viewpoint selection algorithm based on VMI is presented. If we look for a good set of views within the set of viewpoints, we will obtain the most representative set by selecting the views such that their mixing (merging) minimizes the distance to the target distribution $p(O)$. We consider that this mixing provide us with a *balanced* view of the object.

Thus, our selection algorithm will select the N viewpoints so that their merging $\widehat{v}$ minimizes the viewpoint mutual information $I(\widehat{v}, O)$. Due to the fact that this optimization algorithm is NP-complete, we adopt a greedy strategy by selecting successive viewpoints that minimize $I(\widehat{v}, O)$. That is, at each merging step we aim to maximize the JS-divergence between the set of previously merged viewpoints and the new viewpoint to be selected. This algorithm permits us to find in an automated and efficient way a minimal set of views which represent the object or scene.

The algorithm proceeds as follows. First, we select the best viewpoint v_1 with distribution $p(O|v_1)$ corresponding to the minimum $I(v, O)$. Next, we select v_2 such that the mixed distribution $p(v_1)p(O|v_1) + p(v_2)p(O|v_2)$ will minimize $I(\widehat{v}, O)$, where $\widehat{v}$ represents the clustering of v_1 and v_2. At each step, a new mixed distribution $p(v_1)p(O|v_1) + p(v_2)p(O|v_2) + \ldots + p(v_n)p(O|v_n)$ is produced until the VMI-ratio given by $\frac{I(\widehat{v},O)}{I(V,O)}$ is lower than a given threshold or a fixed number of views is achieved. This ratio is the percentage of information about the object that we still lack to capture, and therefore

	Coffe Cup		Dragon		Lady of Elche		Armadillo	
Best view	**VMI**	**Ratio**	**VMI**	**Ratio**	**VMI**	**Ratio**	**VMI**	**Ratio**
1	1.471	0.730	2.124	0.903	1.355	0.703	1.791	0.850
2	0.692	0.343	1.134	0.482	0.644	0.334	0.837	0.397
3	0.346	0.172	0.725	0.308	0.458	0.237	0.616	0.292
4	0.262	0.130	0.550	0.234	0.275	0.143	0.416	0.197
5	0.207	0.130	0.479	0.203	0.219	0.113	0.310	0.147
6	0.190	0.095	0.378	0.160	0.153	0.079	0.238	0.113
Cost	36 sec		61 sec		38 sec		77 sec	

Table 4.1.: For each model, we show the VMI values of the merging of all selected viewpoints and the corresponding VMI-ratio.

can be interpreted as a measure of the goodness or representativeness of the selected viewpoints.

Figure 4.1 show the six best views obtained with our selection algorithm for four different models. In Table 4.1, for each new viewpoint selected we show the VMI of the clustering of selected viewpoints ($I(\widehat{v}, O)$) and the corresponding VMI-ratio. For instance, to achieve a degree of representativeness given by a VMI-ratio lower than 0.1, six views are needed for the coffee-cup-and-dish and lady of Elche models but more for the dragon and armadillo models. In addition, Table 4.1 shows the computation cost of selecting the six best views. The behavior of our algorithm is also shown in Figure 4.2, where we observe how the VMI values obtained from the successive mixed distributions (corresponding to the views of Figures 4.1(i), 4.1(ii), 4.1(iii), and 4.1(iv)) converge asymptotically to zero. It is important to note that the best views for the coffee-cup-and-dish, lady of Elche and armadillo models, shown respectively in Figures 4.1(i.a), 4.1(iii.a) and 4.1(iv.a), are not perceptually pleasant. This is due to the fact that, from a purely geometric approach, the best views of Figure 4.1 correspond to the viewpoints that their projected area distribution is more *similar* (in the KL sense) to the average projected area distribution (target distribution). This problem will be tackled in the next chapters, introducing perceptual criteria to select the best views.

From the N best representative views, a simple greedy clustering algorithm which partitions the sphere of viewpoints assigning each viewpoint to the 'nearest' best viewpoint is proposed. This assignation is determined by the minimum JS-divergence between the viewpoint to be clustered and the best views. Using this method, the *centroids* of the respective clusters are given by the most representative viewpoints. In Figure 4.3, we show the behavior of this clustering algorithm for the (i) coffee-cup, (ii) dragon, (iii)

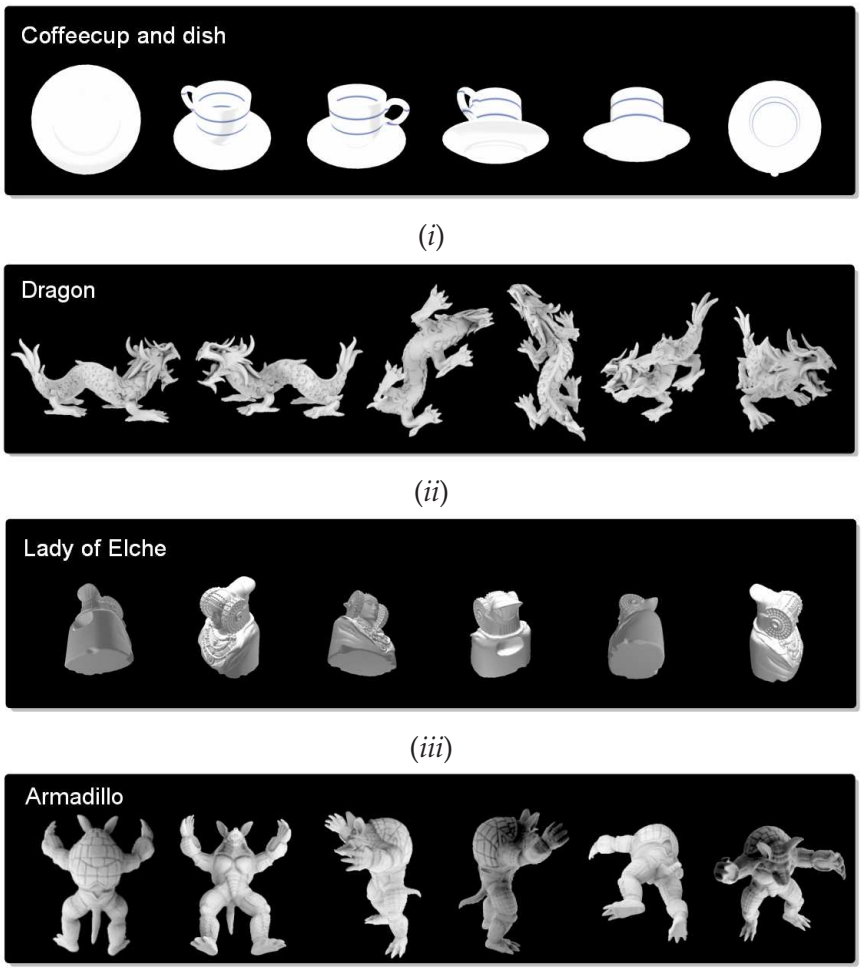


Figure 4.1.: The six most representative views selected by the VMI algorithm for the (i) coffeecup-and-dish, (ii) dragon, (iii) lady of Elche (iv) and armadillo models.

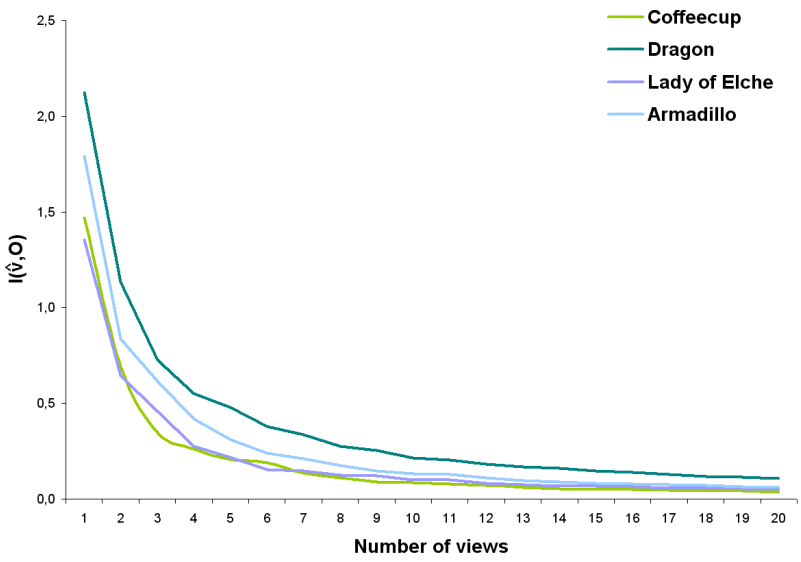
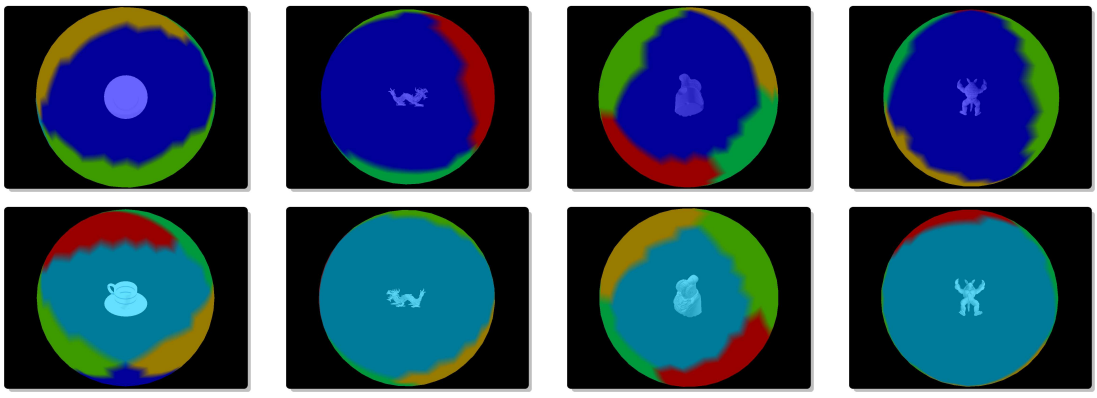


Figure 4.2.: VMI values obtained from the successive mixed distributions corresponding to the views of Figures 4.1(i), 4.1(ii), 4.1(iii), and 4.1(iv).



(*a*) (*b*) (*c*) (*d*)

Figure 4.3.: Viewpoint clustering spheres with six clusters for the (a) coffee-cup-and-dish, (b) dragon, (c) lady of Elche and (d) Armadillo models.

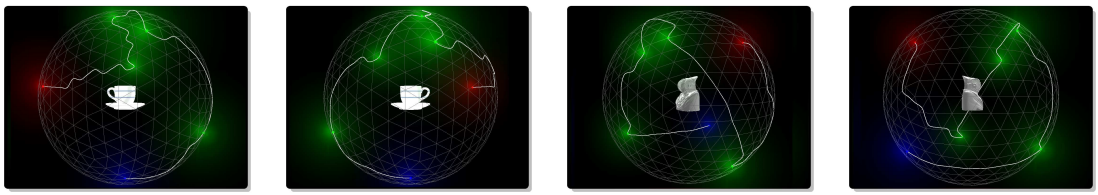
lady of Elche and (iv) armadillo models.

4.2. Object Exploration

In this section, two greedy algorithms are presented to explore the object. In both cases, the best viewpoint (minimum VMI) is the starting point of the object exploration. In the first algorithm (*guided tour*), the path visits a set of N preselected best views which ensure a good exploration of the object. In the second algorithm (*exploratory tour*), the successive viewpoints are selected using the maximum *novelty* criterion with respect to the parts seen of the object.

4.2.1. Guided tour

First, we obtain the list of the N best viewpoints. Then, the algorithm starts at the best viewpoint and visits all the other best viewpoints following the *minimum path*. This is obtained as follows. From the best viewpoint, we find the *nearest* (with minimum JS-divergence) best viewpoint in the list. This is now the *target* viewpoint. Thus, from the best viewpoint, successive neighbor viewpoints will be selected so that, without any viewpoint repetition, their *distance* to the target viewpoint is minimum. The distance between two viewpoints is always calculated from the JS-divergence. When the first target viewpoint is achieved, we select a new target one among the rest of best viewpoints in the list. Then we proceed in the same way until the last best view is reached or the cycle is completed arriving at the initial best viewpoint. Figure 4.4(i) shows the exploration of the coffee-cup-and-dish and the lady of Elche models from the six best views obtained in each case (the blue, yellow and red light points correspond to the starting, intermediate and ending viewpoints, respectively). Two different projections of the sphere are shown to see the path better.

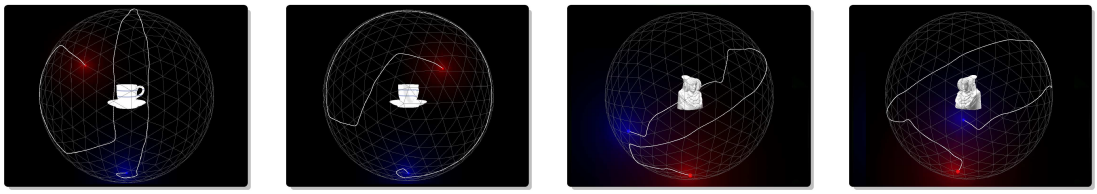


(a) (b) (c) (d)

Figure 4.4.: Guided tour around the coffee-cup-and-dish (a-b) and ship (c-d) models, respectively.

4.2.2. Exploratory tour

From [IB05], we know that maximum novelty or surprise attracts the attention of an observer. Following this principle, the exploratory tour algorithm selects the best viewpoint and then successively visits the (non-visited) neighbor viewpoints that minimize the $I(\widehat{v}, O)$ of all visited viewpoints. This means that at each step we select the viewpoint that maximizes its JS-divergence with respect to all visited viewpoints and, consequently, the most dissimilar (*surprising*) viewpoint is selected. This procedure stops whether the VMI-ratio is lower than a given threshold. Figure 4.5(ii) shows the result of the exploration of the coffee-cup-and-dish and the lady of Elche models.



(*a*) (*b*) (*c*) (*d*)

Figure 4.5.: Exploratory tour around the coffee-cup-and-dish (a-b) and ship (c-d) models, respectively.

5. Mesh and Viewpoint Saliency

In this chapter we reverse the viewpoint channel presented previously (see Chapter 3) to define the information associated to each polygon of the model. Then, this information is used to compute the saliency of both, the polygons and the viewpoints. Finally, we generalize the VMI to introduce the saliency of the triangles as an importance factor to drive the viewpoint selection (see Chapter 4).

5.1. Polygonal Mutual Information

As we have seen in Chapter 3, the information associated with each viewpoint has been obtained from the definition of the channel between the sphere of viewpoints and the polygons of the object. Now, the *information associated with a polygon* will be defined as the contribution of this polygon to the MI of this channel. To illustrate this new approach, the reversed channel $O \rightarrow V$ is considered, so that O is the input and V the output.

From the Bayes theorem $p(v,o) = p(v)p(o|v) = p(o)p(v|o)$, the mutual information (3.3) can be rewritten as

$$\begin{aligned} I(O,V) &= \sum_{o\in\mathcal{O}} p(o) \sum_{v\in\mathcal{V}} p(v|o) \log \frac{p(v|o)}{p(v)} \\ &= \sum_{o\in\mathcal{O}} p(o) I(o,V), \end{aligned} \tag{5.1}$$

where we define

$$I(o,V) = \sum_{v\in\mathcal{V}} p(v|o) \log \frac{p(v|o)}{p(v)} \tag{5.2}$$

as the *polygonal mutual information* (PMI), which represents the degree of correlation between the polygon o and the set of viewpoints, and can be interpreted as the *information* associated with polygon o. Analogous to VMI, low values of PMI correspond to polygons that 'see' the maximum number of viewpoints in a balanced way, i.e., $p(V|o)$ is close to $p(V)$. The opposite happens for high values. Let us remind that MI is invariant to the reversion of the channel since $I(V,O) = I(O,V)$. The application of PMI will be shown in Chapter 6.

5.2. Mesh Saliency

Itti et al. [IKN98] maintained that visual attention is saliency-dependent and use a saliency map to represent the conspicuity or saliency at every location in the visual field by a scalar quantity and to guide the selection of attended locations. In [LVJ05], mesh saliency is captured from surface curvatures and is considered as a perception-inspired measure of *regional importance* and has been used in graphics applications such as mesh simplification and viewpoint selection. We now propose a new definition of mesh saliency based on PMI.

Analogous to the view unstability (Section 3.4), defined from the dissimilarity between two views, we now define the view-based mesh saliency from the dissimilarity between two polygons, which is given by the variation of polygonal mutual information when two polygons are clustered. In this approach, mesh saliency is formulated in terms of how the polygons 'see' the set of viewpoints. Thus, the saliency of a polygon is defined as the average dissimilarity between this polygon and its neighbors.

Similarly to (3.7), the reduction of MI when two polygons o_i and o_j are clustered is given by

$$\begin{aligned}
\delta I(o_i, o_j) &= I(V, O) - I(V, \widehat{O}) \\
&= (p(o_i)I(o_i, V) + p(o_j)I(o_j, V)) - p(\widehat{o})I(\widehat{o}, V) \\
&= p(\widehat{o}) \left(\frac{p(o_i)}{p(\widehat{o})} I(o_i, V) + \frac{p(o_j)}{p(\widehat{o})} I(o_j, V) - I(\widehat{o}, V) \right) \\
&= p(\widehat{o}) D(o_i, o_j), \qquad (5.3)
\end{aligned}$$

where $\widehat{o} = o_i \oplus o_j$ is the result of clustering o_i and o_j and the *polygonal dissimilarity* between o_i and o_j is defined by

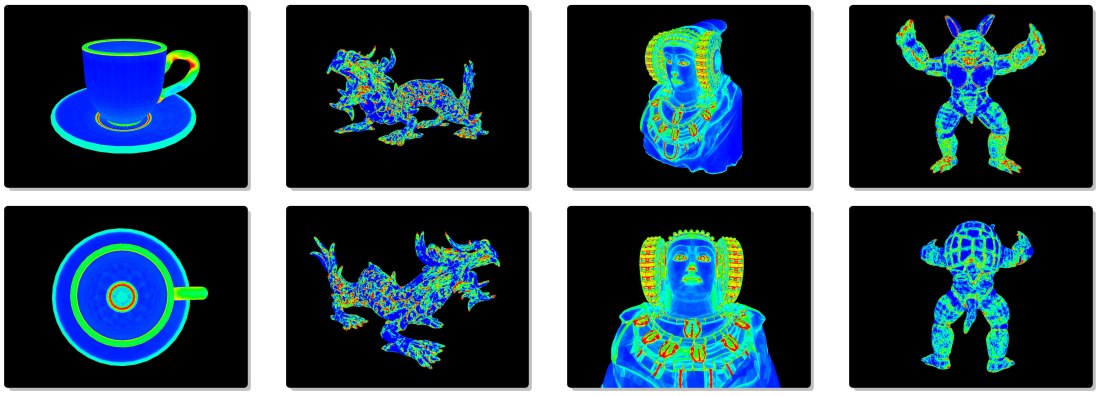
$$D(o_i, o_j) = \frac{p(o_i)}{p(\widehat{o})} I(o_i, V) + \frac{p(o_j)}{p(\widehat{o})} I(o_j, V) - I(\widehat{o}, V). \qquad (5.4)$$

This dissimilarity can also be written as

$$D(o_i, o_j) = JS \left(\frac{p(o_i)}{p(\widehat{o})}, \frac{p(o_j)}{p(\widehat{o})}; p(V|o_i), p(V|o_j) \right), \qquad (5.5)$$

where the second term is the Jensen-Shannon divergence (2.6) between $p(V|o_i)$ and $p(V|o_j)$ with weights $\frac{p(o_i)}{p(\widehat{o})}$ and $\frac{p(o_j)}{p(\widehat{o})}$, respectively. Hence, two polygons are 'similar' when the JS-divergence between them is small.

Some interesting properties follow:



(*a*) (*b*) (*c*) (*d*)

Figure 5.1.: Mesh saliency for the (a) coffee-cup-and-dish, (b) dragon, (c) lady of Elche, and (d) armadillo models.

- If two polygons are very 'similar', their clustering involves a small loss of mutual information. If $p(V|o_i) = p(V|o_j)$, then $\delta I(o_i, o_j) = 0$.
- It can be easily seen that the clustering of all polygons would give $\delta I = I(V, O)$ and, thus, $I(\widehat{O}, V) = 0$.
- $H(V) = H(V|O) + I(O, V) = H(V|\widehat{O}) + I(\widehat{O}, V)$, where $H(V)$ is the entropy of $p(V)$, $H(V|O) = -\sum_{o\in\mathcal{O}} p(o) \sum_{v\in\mathcal{V}} p(v|o) \log p(v|o) = \sum_{o\in\mathcal{O}} p(o) H(V|o)$ and $H(V|o)$ can be interpreted as the uncertainty that polygon o "sees" the set of viewpoints. The reduction of $I(O, V)$ is equal to the increase of $H(V|O)$ since $H(V)$ remains constant (the input distribution of V is not changed).

Similarly to the unstability of a viewpoint (3.10), the *polygonal saliency* of o_i is defined by

$$S(o_i, V) = \frac{1}{N_o} \sum_{j=1}^{N_o} D(o_i, o_j) \geq 0, \tag{5.6}$$

where o_j is a neighbor polygon of o_i and N_o is the number of neighbor polygons of o_i. Thus, a polygon o will be salient if the average of JS-divergences between o and its neighbors is high. On the other hand, a polygon at the center of a smooth region will have probably low saliency since the polygons of this region will present small visibility differences with respect to the set of viewpoints.

Figure 5.1 shows the behavior of our saliency measure. The most salient parts are represented in red colors and the least salient ones in blue. For instance, the handle of the coffee-cup or the nose, mouth and eyes of the models are the most salient surfaces.

5.3. Viewpoint Saliency

Similarly to [LVJ05], where mesh saliency was used to select the best views, we propose a method to calculate the saliency of a viewpoint. Up to now we have calculated the saliency of a polygon, however we can convey this information to the sphere of viewpoints, using the conditional probabilities of the reversed channel. Hence, the *viewpoint saliency* is defined by

$$S(v, O) = \sum_{o \in \mathcal{O}} S(o, V) p(v|o). \tag{5.7}$$

Figure 5.2 show the viewpoint saliency for the coffee-cup-and-dish, Hebe and lady of Elche models. Columns (a) and (b) illustrate the most salient view and the least one, respectively. Columns (c) and (d) show two different projections of the corresponding saliency spheres. Observe how the most salient views show us the most salient parts of each object.

5.4. Importance-driven Viewpoint Selection

As we have mentioned in Chapter 1, it is desirable that a canonical view of an object shows its most salient parts and also the largest number of visible surfaces [PRC81, BTB99]. However, the measures introduced up to now in this master thesis only consider the geometric relation between the object and the set of viewpoints. Therefore, we can not expect that in general the best views fulfill the desired properties for a canonical view. This fact motivates us to investigate how perceptual criteria such as saliency can be introduced into our viewpoint selection framework in order to improve the automatic selection of good views.

In the previous section we have presented a method to compute how salient is a viewpoint, but we aim now to incorporate the polygonal saliency to the viewpoint mutual information in order to take into account different factors concerning, respectively, the amount of projected area, the geometric representativeness and the saliency of a polygon.

First, we demonstrate how the importance can be introduced into the object space by modifying directly the target distribution $p(O)$. Second, we show the results obtained by the use of the polygonal saliency as an importance factor in the viewpoint mutual information measure.

Due to the fact that VMI represents the distance between the projected visibility distribution $p(O|v)$ from viewpoint v and the target distribution $p(O)$, VMI can be extended by weighting the target distribution with an importance factor. Thus, adding

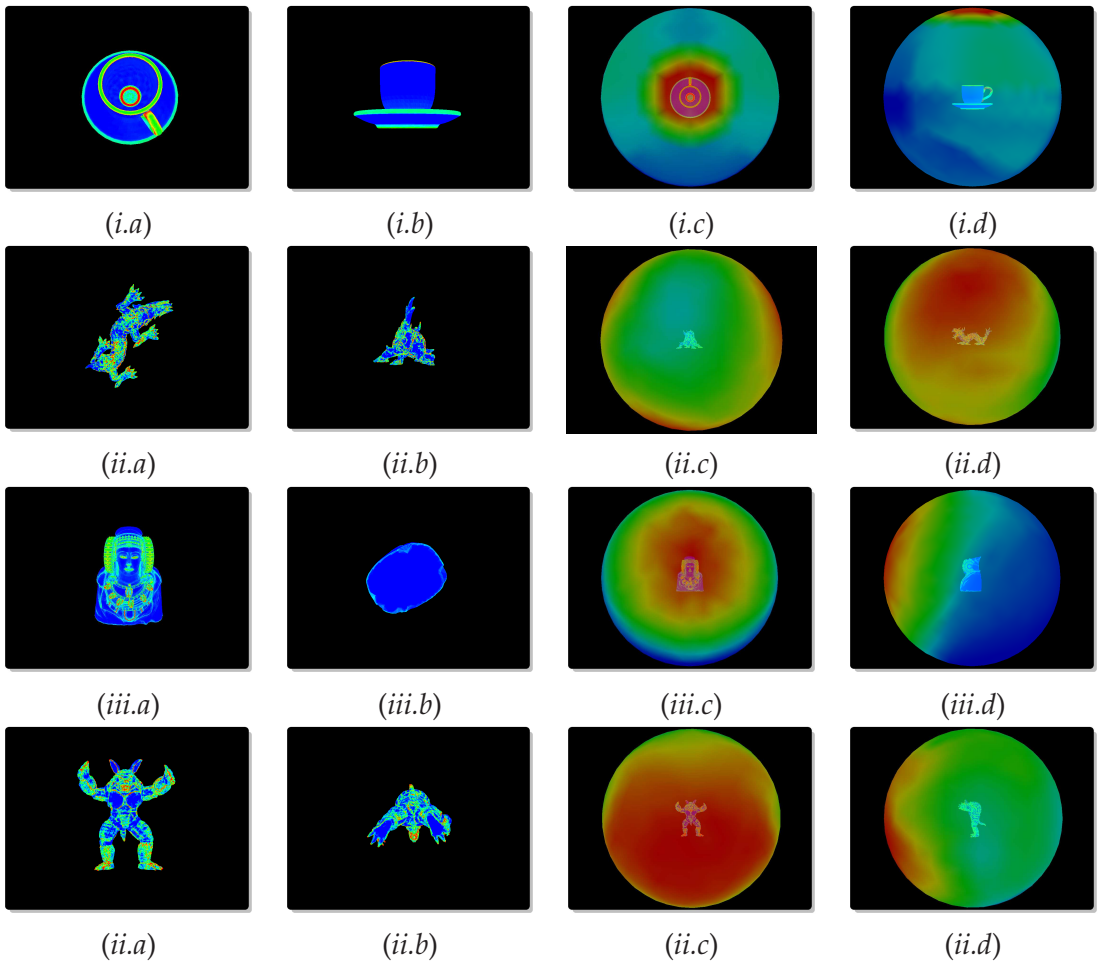
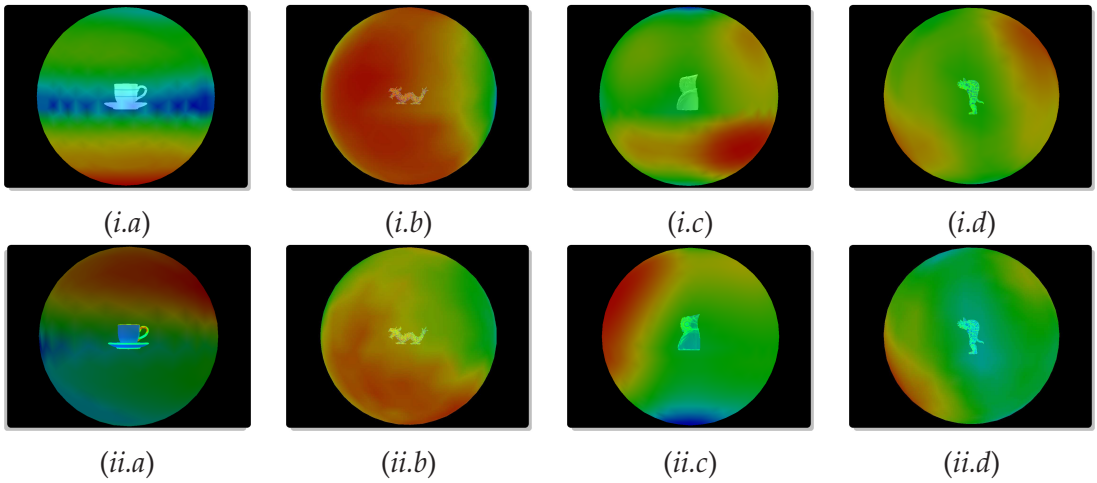


Figure 5.2.: The (a) most salient and (b) least salient views, and (c-d) saliency spheres obtained for the (i) coffee-cup-and-dish, (ii) dragon, (iii) lady of Elche and (iv) armadillo models. Red colors on the sphere represent high saliency values, blue colors represent low saliency values.



(*i.a*) (*i.b*) (*i.c*) (*i.d*)

(*ii.a*) (*ii.b*) (*ii.c*) (*ii.d*)

Figure 5.3.: VMI (i) and saliency-based VMI (ii) spheres for the coffee-cup-and-dish (a), dragon (b), lady of Elche (c) and armadillo (d) models.

importance to our scheme means simply weighting the original target distribution by an importance factor in order to obtain the new target distribution. The optimal viewpoint would be the one viewing every polygon proportional to its average projected area multiplied by its importance. Hence, the *extended viewpoint mutual information* (EVMI) is given by

$$I'(v,O) = \sum_{o \in \mathcal{O}} p(o|v) \log \frac{p(o|v)}{p'(o)}, \tag{5.8}$$

where

$$p'(o) = \frac{p(o)i(o)}{\sum_{o \in \mathcal{O}} p(o)i(o)} \tag{5.9}$$

and $i(o)$ is the importance of polygon o. In the experiments of this section, $i(o)$ has been substituted by the polygonal saliency $S(o)$. We follow the convention that if $S(o) = 0$ then polygon o is not taken into account. Other features, such as illumination, could be introduced as importance factors in the EVMI. In [VFSG06], the object importance has been used to calculate the best views for a volumetric dataset.

The effects of incorporating saliency in our viewpoint selection framework are depicted in Figures 5.3 and 5.4, which show for the coffee-cup-and-dish, dragon, lady of Elche and armadillo models the saliency-based VMI spheres and the six most representative views, obtained with the best view selection algorithm (Chapter 4) using EVMI. The saliency-based VMI spheres of Figure 5.3 show the perceptual improvement obtained with respect to the VMI spheres shown in Figure 3.5(i.d) and Figures 3.6(ii.c-d), respectively. For instance, whereas the VMI-based best view of the coffee-cup-and-dish

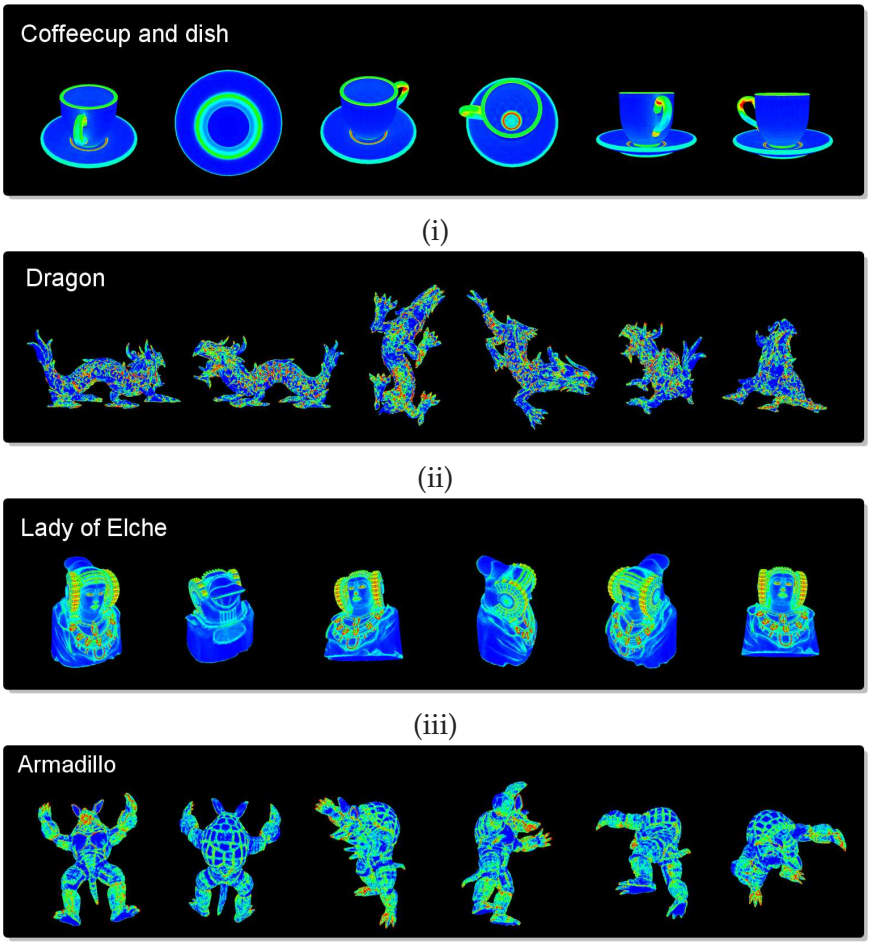


Figure 5.4.: The six most representative saliency-based views for the coffeecup-and-dish (i), dragon (ii), lady of Elche (iii) and armadillo (iv) models.

shows the bottom of the dish (Figure 4.1(i)), the best view based on EVMI with saliency shows a lateral view of the coffee cup (Figure 5.4(i)) which is perceptually much better than the one of Figure 4.1(i). Similarly, the same conclusion can be obtained for the lady of Elche and armadillo models (see the respective best views shown in Figure 4.1 and Figure 5.4). Note that, the dragon model has a good selection of views taking into account a purely geometric approach (see Chapter 4) but also including saliency as an importance factor.

6. Information-Theoretic Ambient Occlusion

Ambient occlusion is a powerful technique that mimics indirect global illumination at a fraction of the cost. In this chapter, we introduce a new ambient occlusion technique based on the mutual information (see Section 5.1) of a polygon with respect to all viewpoints, which is dependent on the importance assigned to each viewpoint and helps to enhance features such as the most salient parts. Further, the assignation of color to each viewpoint combined with the polygonal information produces a nice visualization of the object. Examples are given with coloroid palettes and non-photorealistic rendering.

6.1. Information-Theoretic Ambient Occlusion

As we have seen in Section 5.1, the *information associated with a polygon* (PMI) appears from the reversed channel $O \rightarrow V$, so that O is the input and V the output (see Figure 6.1), and is defined as

$$I(o, V) = \sum_{v \in \mathcal{V}} p(v|o) \log \frac{p(v|o)}{p(v)} \tag{6.1}$$

The PMI represents the degree of correlation between a certain polygon and the set of viewpoints. Low values of PMI correspond to polygons that "see" the maximum number of viewpoints in a balanced way and the opposite happens for high values. Now, we can interpret PMI as a measure of the degree of occlusion of each triangle of the model. That is, the more occluded a polygon, the less number of viewpoints it will "see".

To obtain the ambient occlusion of a model, the PMI of all polygons has to be normalized between 0 and 1 and subtracted from 1, because low values of PMI, represented in the grey-map by values near 1, correspond to non-occluded or visible (from many viewpoints) polygons, while high values of PMI, represented in the grey-map by values near 0, correspond to occluded polygons.

In Figures 6.2(ii-iii) we show the information maps corresponding to the models shown in Figure 6.2(i). In Figure 6.2(ii) the polygonal information values are computed from the center of each polygon, while in Figure 6.2(iii) these values have been linearly

interpolated at the vertexes of the triangles. From now on, all images presented are obtained from the interpolated values at the vertexes. Also, in Figure 6.2(iv) we show the results of applying classic ambient occlusion. We want to stand out that, the projection resolution used in the preprocess step (see Chapter 3) to obtain PMI and the texture resolution used in the computation of ambient occlusion models are 5120x3840 and 4096x4096, respectively.

In [SSZK04], a similar approach was used to compute ambient occlusion. A matrix M_{ij} is computed as $n_i l_j$, where n_i is the normal to the vertex of the object and l_j is the direction of a virtual light source placed at a bounding sphere. A number of virtual light sources is used to approximate ambient lighting. The final contribution to the vertex i is given by the sum for all visible light sources of $M_{ij} I_j$, where I_j is the intensity of the source.

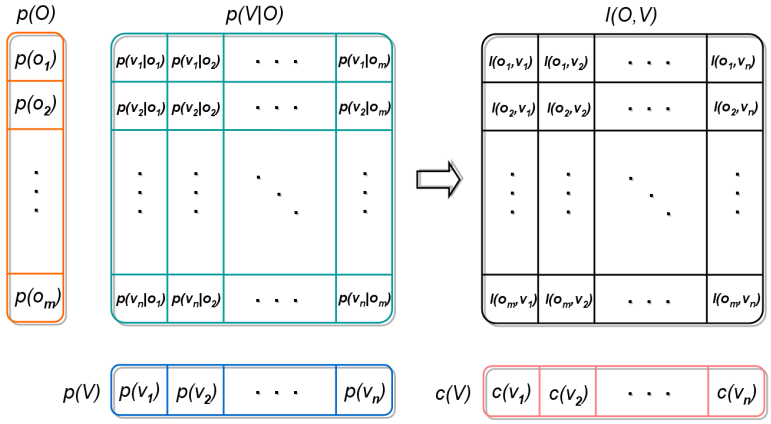


Figure 6.1.: Probability distributions of channel $O \rightarrow V$, elements of matrix $I(O,V)$ and color distribution $c(V)$ assigned to the viewpoint sphere.

In Figures 6.2 and 6.3 we can observe our technique compared with classic ambient occlusion. In the latter case, there is only a discrete set of possible values, since it is computed as a proportion of hits. On the other hand, the quality of our ambient occlusion measure depends on the resolution used to project the polygons of the scene (see Chapter 3). If the resolution is not enough, the small triangles won't be projected correctly. We can observe this effect in Figure 6.6 and Figure 6.7, where we show a four by four grid of the ambient occlusion computed for the Lady of Elche model. The rows represents the resolution of the projection (640x480, 1280x960, 2560x1920 and 5120x3840) while the columns represents the number of viewpoints (12, 42, 162, 642) used. See how the artifacts (the darker small spherical bumps of the Lady of Elche garments) tend to dissapear as the resultion and/or the number of viewpoints are increased. Of course, the higher the resolution of projection, the more time-consuming the preprocess

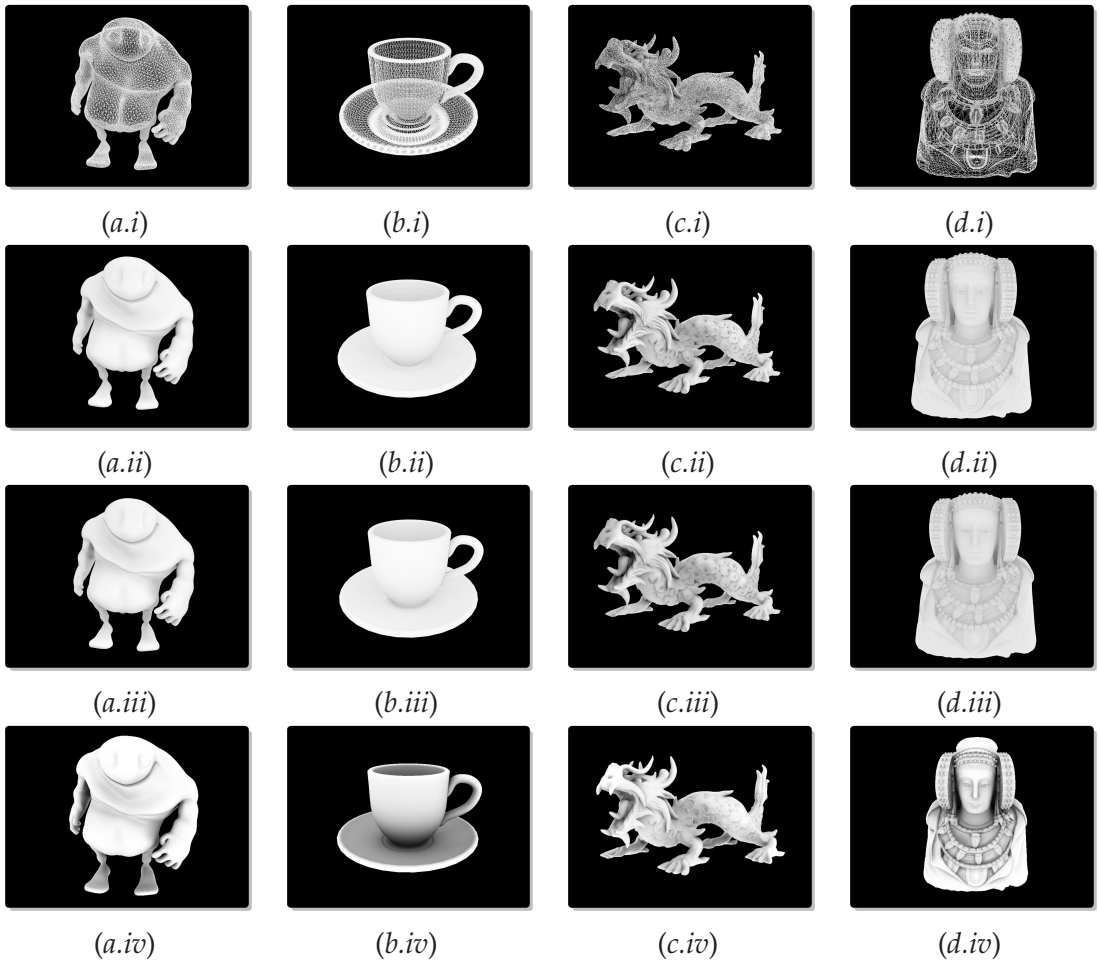
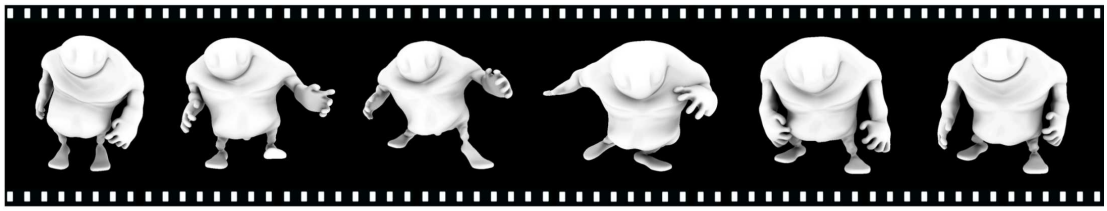
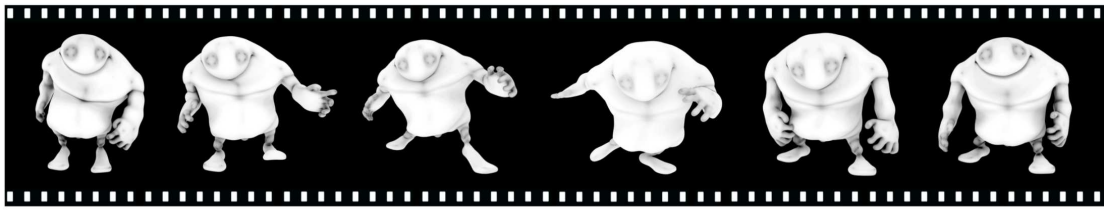


Figure 6.2.: (i) Wireframe models. (ii) Polygonal mutual information maps. (iii) Values from (ii) interpolated at the vertices. (iv) Ambient occlusion maps.

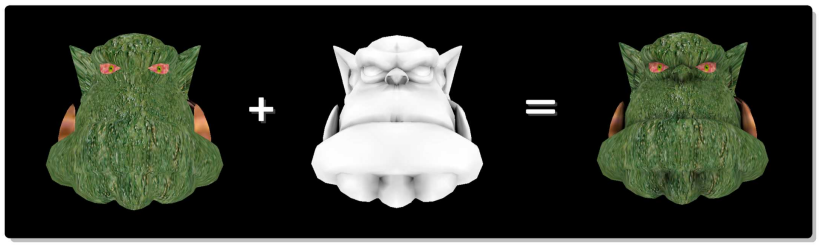


(*i*)

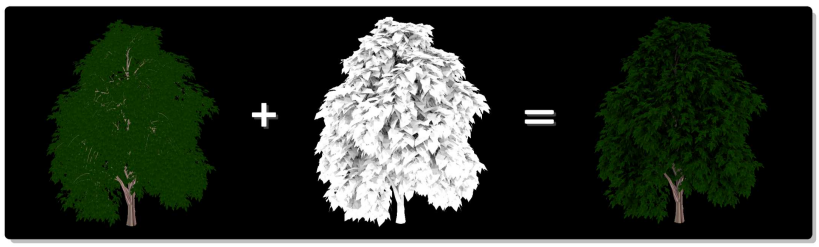


(*ii*)

Figure 6.3.: Different frames of an animation using ambient occlusion (top row) and our method (bottom row). The projection resolution used by our method is 640x480. The classic ambient occlusion have been computed using a texture of 1042x1024. See how our ambient occlusion model is good enough, in spite of the low resolution used.



(*i*)



(*ii*)

Figure 6.4.: Composition of our information-theoretic ambient occlusion and the textures of the Ogre (i) and tree models (ii).

Figure 6.5.: Snapshots of an animation showing the use of our ambient occlusion technique on trees.

needed to compute PMI. We also want to note that we have experimentally verified that the increase in the number of viewpoints above 642 viewpoints does not provoke visual difference in the obtained results.

The models used in our examples come from Nvidia dynamic ambient occlusion demo (Figure 6.2(a)), Xfrog public plants (Figure 6.5), De Espona 3D encyclopedia (Figure 6.2(b), Figure 6.2(d) and Figure 6.11), and The Stanford 3D Scanning Repository (Figure 6.2(c) and Figure 6.12).

In Figure 6.12 we show a snapshot from a scene with our ambient occlusion technique applied to the lucy model. Also, in Figures 6.4 and 6.5 we show several examples of the use of polygonal information as ambient occlusion, where this is added to a textured model.

6.2. Applications

As we have shown above, our polygonal mutual information can be used as an ambient occlusion technique. In this section, PMI is also extended to enhance the most important viewpoints or to reflect the color of the environment, a sort of color bleeding. Both extensions are explained below.

6.2.1. Viewpoint Importance

From (6.1), importance can be introduced into the viewpoint space by modifying the input distribution $p(V)$ according to the importance we want to assign to each viewpoint. The polygonal information will be modified accordingly. The effect can be observed in Figure 6.8. For the three models shown, the range of images go from assigning almost all importance to the best viewpoint in the first image, to assign equal importance to

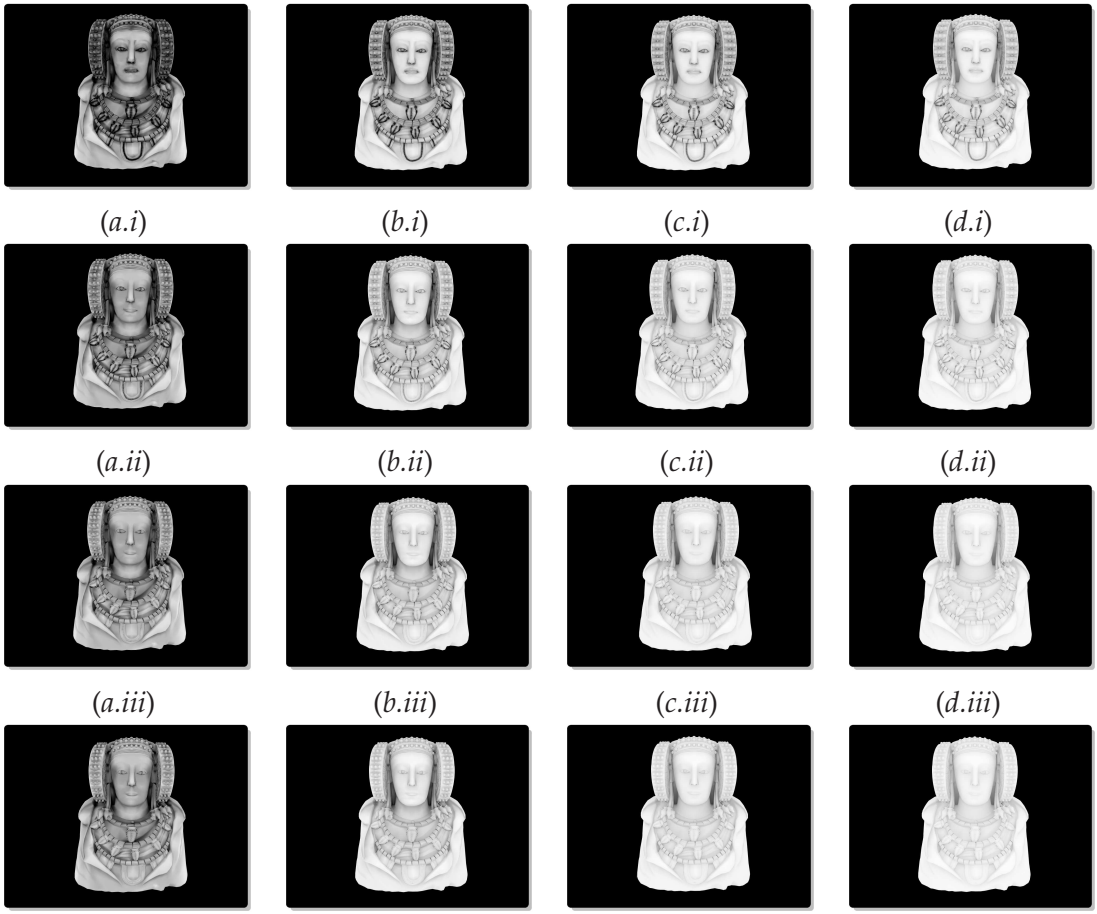


Figure 6.6.: Several ambient occlusion models for the Lady of Elche. For each row we increase the projection resolution twice the row before, that is: 640x480 (i), 1280x960 (ii), 2560x1920 (iii) and 5120x3840 (iv). On the other hand, each column increase the number of viewpoints in the following way: 12 (a), 42 (b), 162 (c) and 642 (d).

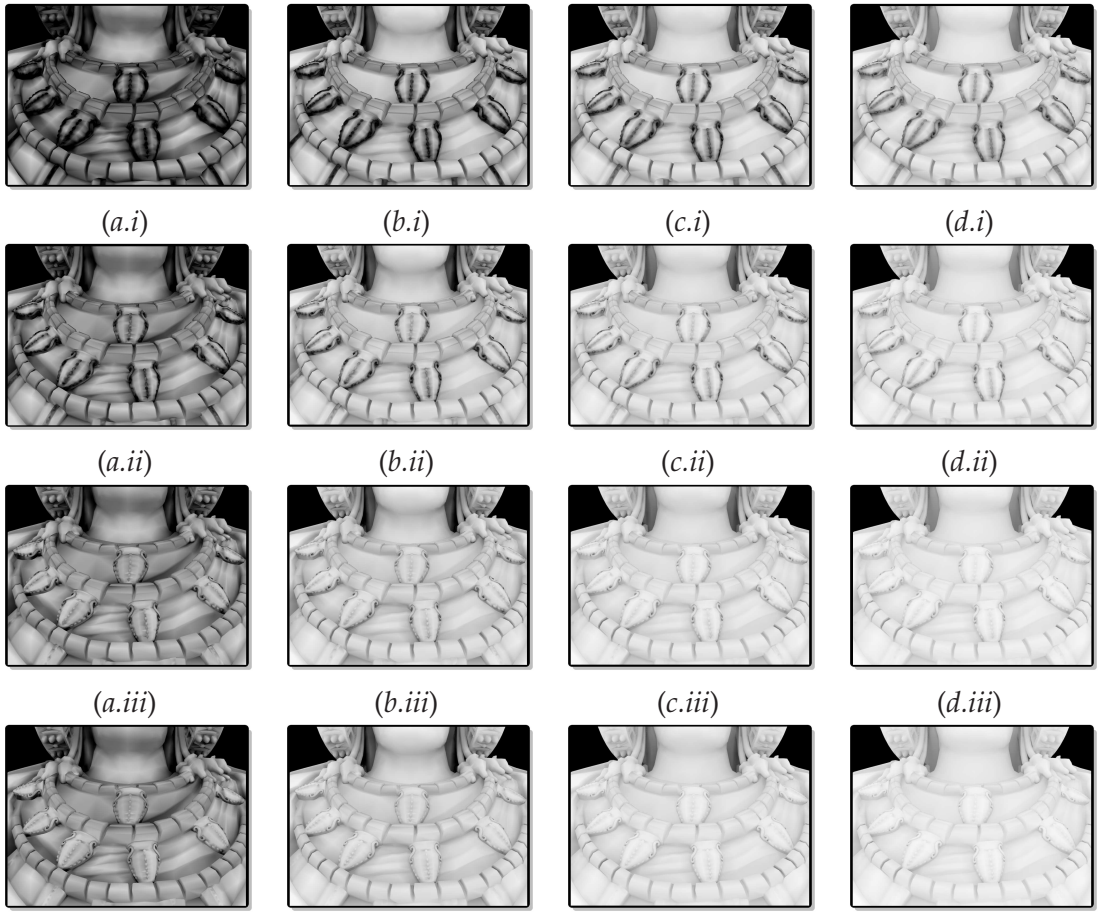
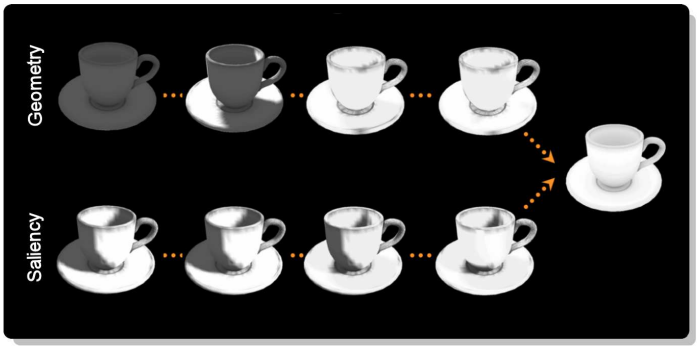
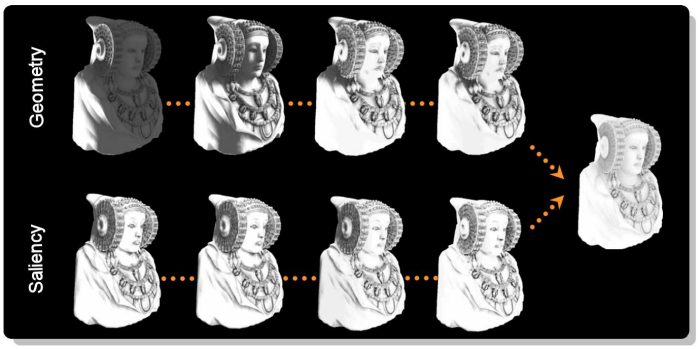


Figure 6.7.: A close-up from several ambient occlusion models for the Lady of Elche. For each row we increase the projection resolution twice the row before, that is: 640x480 (i), 1280x960 (ii), 2560x1920 (iii) and 5120x3840 (iv). On the other hand, each column increase the number of viewpoints in the following way: 12 (a), 42 (b), 162 (c) and 642 (d).



(*i*)



(*ii*)

Figure 6.8.: Effect of assigning importance to the best viewpoint (first left image), plus second best (second left image), plus third best (third left image), plus fourth best (fourth best image) for the coffe cup (i) and Lady of Elche (ii) models. Upper row viewpoints are selected according to geometry, in lower row according to saliency. Last image (both upper and lower row) corresponds to equal importance for all the viewpoints of the sphere.

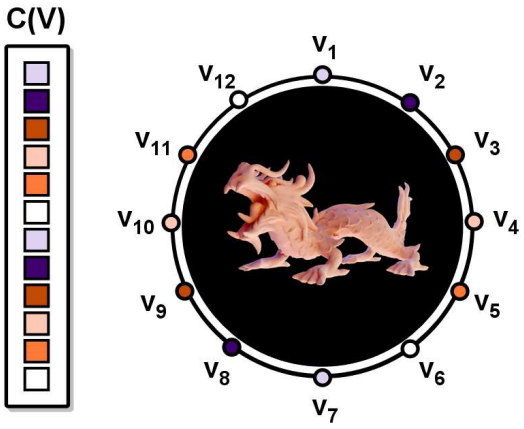


Figure 6.9.: The result of warping a color texture to the viewpoint sphere for computing the color ambient occlusion model.

the two best viewpoints in the second image, till assigning equal importance to the best 4 points in the fourth image. Last image is obtained assigning equal importance to all viewpoints in the sphere. For each model, in the upper row we have considered the viewpoints obtained from the best view selection algorithm presented in [FSG07], while in the lower row the best viewpoints have been selected using the same algorithm driven by the saliency of polygons (see also [FSG07]). Observe the improvement of the images obtained when the most important viewpoints are the most salient ones.

6.2.2. Relighting for Non-Photorealistic rendering

Color ambient occlusion is obtained from the scalar product of a matrix row of $I(O,V)$ and the complementary of a color vector $c(V)$:

$$I_\alpha(o,V) = \sum_{v \in \mathcal{V}} I(o,v)(1 - c_\alpha(v)), \tag{6.2}$$

where α stands for each color channel, $c_\alpha(v)$ is the normalized vector for channel α and $I(o,v)$ is a matrix element of $I(O,V)$ (see Figure 6.1). After computing the polygonal mutual information for each channel, the final color ambient occlusion is given by the combination of the channels. We can get a color vector by warping a color texture to the sphere of viewpoints. In this way, a color is assigned to each viewpoint (see Figure 6.9). In Figure 6.10 and 6.11 we show the combination of this kind of relighting technique with an NPR technique [LMHB00, Lak], where the several color palettes used are Coloroid ones [Nem80]. Observe the nice effects obtained by this combination of techniques.

Figure 6.10.: Combination of information-theoretic ambient occlusion with a non-photorealistic technique using Coloroid color palettes (left).

Figure 6.11.: Snapshots of an animation using our information-theoretic ambient occlusion and the first row Coloroid palette shown in Figure 6.10, with a non-photorealistic technique on a boat model.

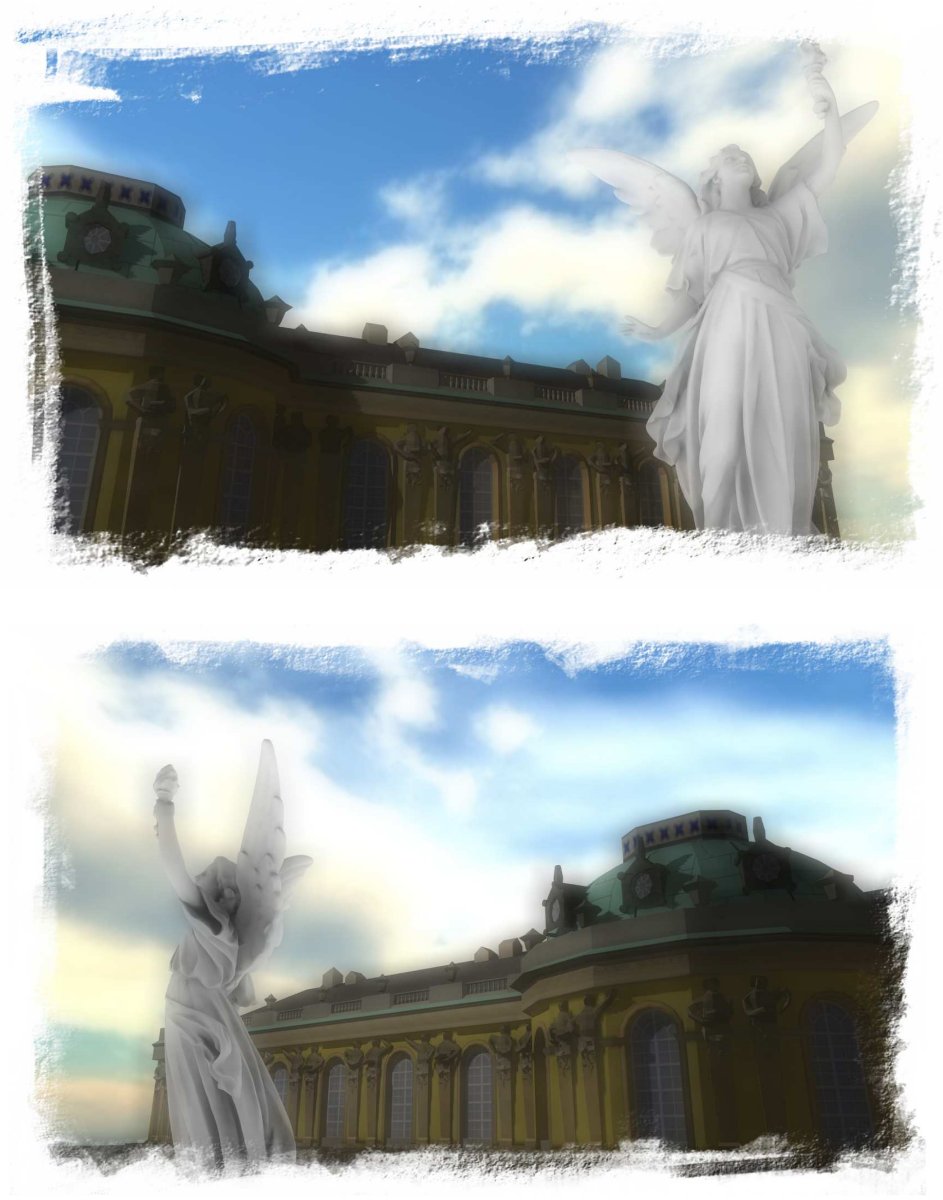


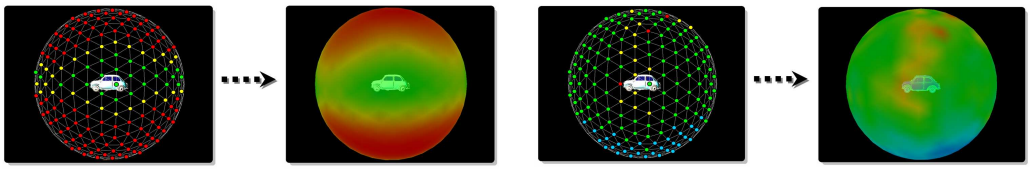
Figure 6.12.: Snapshots of a video demo using our information-theoretic ambient occlusion on the lucy model.

7. View-based Shape Similarity

In this chapter we present a new method for computing the shape similarity between 3D polygonal models using the information-theoretic viewpoint selection framework presented previously. Given a 3D model, a sphere of viewpoints surrounding this model is used to obtain its shape signature from the mutual information and saliency of each viewpoint. These signatures represent the essence of the shape from a view-based approach. Then, in order to quantify the dissimilarity between two models, their mutual information and saliency spheres are registered, respectively, by minimizing the L2 distance between them. Several experiments show the discrimination capabilities of our approach and its potential suitability for object recognition.

7.1. Introduction

As we have seen in the previous chapters (see Chapter 3 and Chapter 5), while VMI measures the degree of correlation between a viewpoint and the model, the viewpoint saliency determines the amount of mesh saliency "seen" by each viewpoint. The VMI and saliency spheres (Figure 7.1) are obtained from the mutual information and saliency of each viewpoint and are now interpreted as *shape signatures* that capture the essence of the shape from a view-based approach. So, the VMI and saliency spheres can be considered as n-dimensional shape descriptors, that is, a shape descriptor composed by n-values, being that values the VMI or saliency values corresponding to each viewpoint of the VMI and saliency spheres.



(a) (b)

Figure 7.1.: The VMI (a) and saliency (b) sphere of the first car model shown in Figure 7.3. The VMI and saliency spheres shown on the right-hand side of (a) and (b), have been obtained by linear interpolation of the VMI and saliency values at viewpoint positions (left image of (a) and (b)).

Our shape descriptors, VMI and saliency spheres, fulfill the following properties:

1. **Discriminative power.** A good shape descriptor should capture properties that discriminate objects well. In the case of the VMI sphere, we capture the dependence between the viewpoints and the polygons, while the saliency sphere measures the amount of saliency information coming or perceived from the model to each viewpoint.

2. **Insensitive to noise.** A shape descriptor is desirable to be insensitive to noise and small extra features and robust against arbitrary topological degeneracies, e.g. if it is obtained by laser scanning. Also, different resolutions of the same model should not modify the description given by the same shape descriptor. Because of the values given by the VMI and saliency spheres are computed using a view-based approach only the most important details from the objects are captured. Also, in chapter 3, we have demonstrated the insensitiveness of VMI to the discretisation variation of the models. On the contrary, the saliency sphere is sensitive to the resolution of the 3D objects.

3. **Rotation, translation and scale invariant.** It is common that 3D models have arbitrary orientation, position and scale. Due to this fact, it may be necessary to place the 3D models into a canonical coordinate system and then compute the shape descriptor. As we have shown in previous chapters (see chapter 3 and chapter 5) the VMI and saliency spheres have been computed placing the object into a canonical coordinate system. That is, we place the model into the origin of coordinates, then we scale it to fit into a unitary bounding sphere and compute VMI and saliency for each viewpoint placed over a bounding sphere three times the unitary one. The orientation of the model does not affect the our similarity method because as we will explain later, we perform a spherical registration which explores all the sphere positions.

Taking into account these properties and how VMI and saliency spheres verify them, we can conclude that both spheres are suitable to be considered as good shape descriptors.

7.2. View-based Shape Similarity

The goal of the registration between two VMI or two saliency spheres, respectively, is to find the transformation that brings one sphere (floating) into the best possible spatial correspondence with the other one (fixed) by minimizing a dissimilarity metric. To achieve that objective we propose a registration method where the main components and their interconnections are shown in Figure 7.2. The basic input data to the

registration process are two VMI or two saliency spheres. The transform component represents the spatial mapping of points from the fixed sphere space to points in the floating sphere space. The interpolator component is used to evaluate floating sphere values at non-viewpoint positions and, finally, the metric module provides a measure of how well the fixed sphere is matched by the transformed floating sphere.

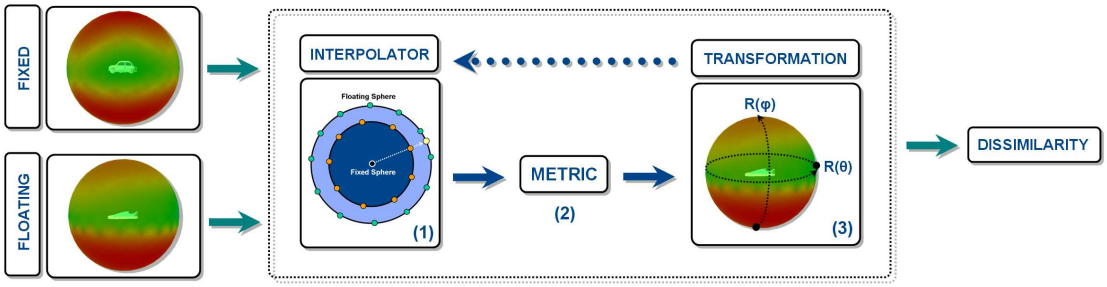


Figure 7.2.: Registration of VMI or saliency spheres and its main components.

The steps followed by our method to achieve the best matching between the fixed and the floating sphere are:

1. **Interpolation**. The discrete nature of our VMI and saliency spheres implies the need of having an interpolator component. The nearest neighbor interpolator has been used. This means that when we need to evaluate values at non-viewpoint positions on the floating sphere we will use the VMI or saliency value of the closest viewpoint.

2. **Comparison**. To quantify the quality of the alignment between the fixed and the floating sphere we need a dissimilarity metric. In our method we have adopted the L2 distance between the VMI

$$D_{VMI}(S_1, S_2) = \sqrt{\sum_{v \in \mathcal{V}} (I(v, O_1) - I(v, O_2))^2}. \tag{7.1}$$

   or saliency

$$D_{saliency}(S_1, S_2) = \sqrt{\sum_{v \in \mathcal{V}} (S(v, O_1) - S(v, O_2))^2}. \tag{7.2}$$

   values of the spheres S_1 and S_2 corresponding to models O_1 and O_2, respectively.

3. **Transformation**. We need two transformation parameters (degrees of freedom): R(θ) and R(φ), defined respectively as the rotation around Z and Y axis. These two

parameters take values in the range [0°,360°] for the first parameter and [0°,180°] for the second.

When all the possible registration positions (dependent on the values of the transformation parameters) have been analyzed, the correct matching is given by the minimum dissimilarity. That is, the position that minimize the L2 distance between the fixed and the floating sphere values.

In our current implementation, running on a Pentium IV 3GHz machine with 2GB RAM and an NVidia GeForce 8800 GTX, a single registration takes approximately two minutes when the transformation parameters are increased in steps of five degrees. Obviously, at smaller steps more time-consuming the method will perform. However, the cost of this registration process could be considerably improved by using numerical optimizers. The memory space consumption required by our method can be considered negligible: we need a list of n float numbers for each of both spheres involved in the registration, being n the number of viewpoint of the spheres. In our experiments shown in next section we have used $n = 642$, but with less viewpoints we could obtain also good results.

7.3. Results

The view-based shape matching described in the previous section has been incorporated into our viewpoint software using the Ogre3D rendering engine (`http://www.ogre3d.org`). In our experiments, the viewpoint sphere is built from the smallest bounding sphere of the model. The radius of the viewpoint sphere is three times the radius of the bounding one.

In order to demonstrate the performance of our approach we have used four families of models (fishes, chairs, cars and shoes) where each one is composed by four different samples (Figure 7.3). Our registration method has been applied to all pairs of models obtaining the dissimilarity between the VMI and the saliency spheres. The transformation parameters R(θ) and R(φ) take values in intervals of five degrees.

From the dissimilarity values obtained with the spherical registration of both, the VMI and the saliency spheres, we have built the dissimilarity map shown in Figure 7.4 for the VMI and in Figure 7.6 for the saliency sphere. Each row and column of the maps represents an object and the color given to the intersection between them is the resulting dissimilarity (not intersected regions between models have been linearly interpolated). Red and blue colors represent dissimilar and similar objects, respectively. Note the blue regions along the diagonal as well as the predominance of warm colors

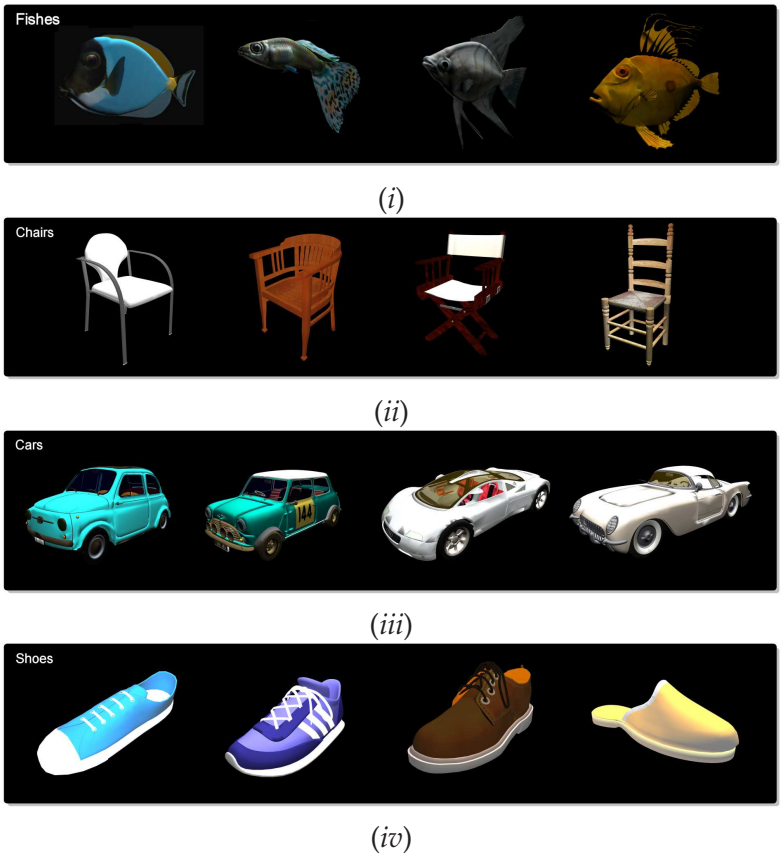


Figure 7.3.: Models used in our experiments. Each row shows a family of models: fishes (i), chairs (ii), cars (iii) and shoes (iv).

while moving away. Let us note the large blue and green area surrounding the car family region and the dissimilarity of fishes with respect to the rest of the models in both spherical registrations.

In Figure 7.5 and 7.7 we show the shape similarity between the first model of each family and the rest of models. The list of models has been ordered according to the dissimilarity obtained with the VMI or saliency spherical registration. Observe the perfect matching in the fish and car families and the good behavior of the chairs and shoes for both shape descriptors.

As can be seen, the registration of VMI or saliency spheres give us good results but not the perfect ones (not all the family models are between the first four in the similarity lists shown). This is due to the fact that we use only one shape descriptor for each model (the VMI or the Saliency sphere) and then, we compute the dissimilarity. Now,

the idea is to combine the goodness given by the results of both spherical registrations to improve the similarity given. So, we obtain a new dissimilarity value, combining the dissimilarity value given by the VMI and the saliency spherical registration:

$$D(S_1, S_2) = \alpha_1 D_{VMI}(S_1, S_2) + (1 - \alpha_1) D_{saliency}(S_1, S_2). \tag{7.3}$$

We have used value $\alpha_1 = \frac{1}{2}$, so we are giving the same importance to both shape dissimilarities provided by both shape descriptors, the VMI and saliency spheres.

In Figure 7.8 we show the dissimilarity map created using the combination of VMI and saliency dissimilarities between models. In Figure 7.9 we can observe the improvement of the results compared to the VMI or saliency ones. See how the first four models obtained for each family are models belonging to the family of the target one.

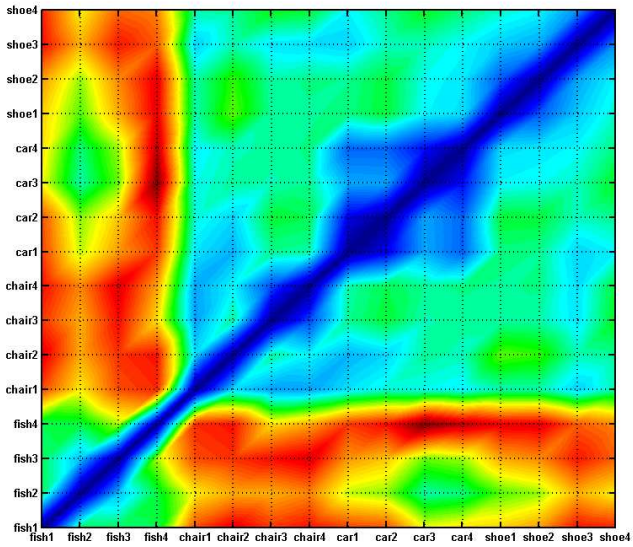


Figure 7.4.: Dissimilarity map using VMI spheres. Blue and red values correspond to the most similar and dissimilar models respectively.

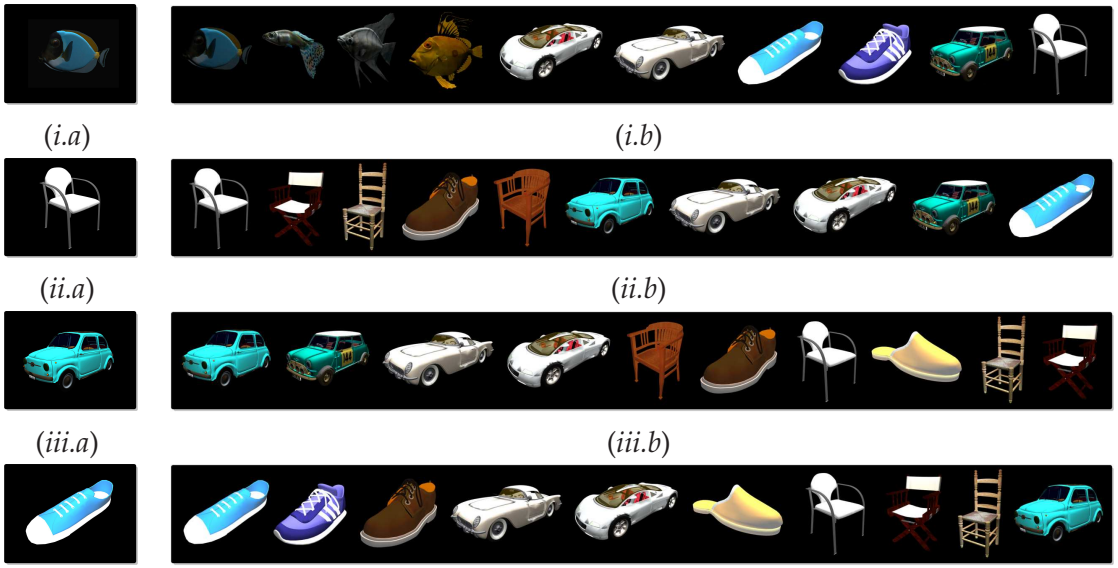


Figure 7.5.: Spherical registration using VMI spheres. In column (a) we show the first model of each family and in column (b) the list of the first ten objects sorted by its similarity with respect to the target model (a).

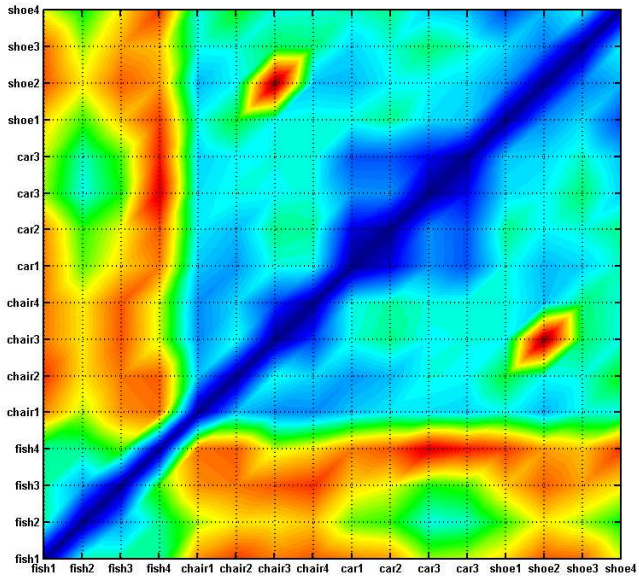


Figure 7.6.: Dissimilarity map using saliency spheres. Blue and red values correspond to the most similar and dissimilar models respectively.

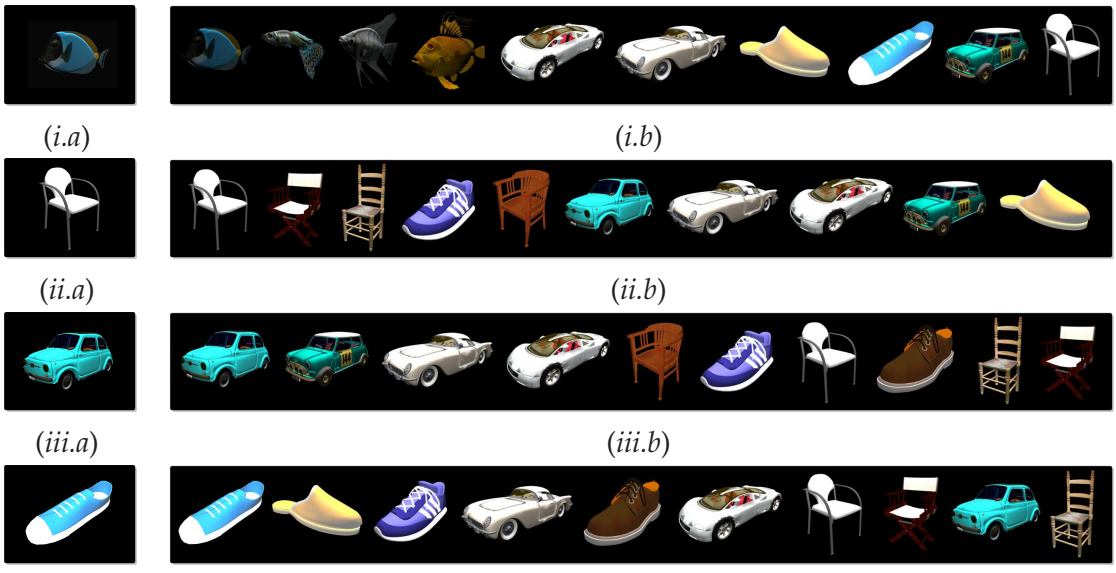


Figure 7.7.: Spherical registration using saliency spheres. In column (a) we show the first model of each family and in column (b) the list of the first ten objects sorted by its similarity with respect to the target model (a).

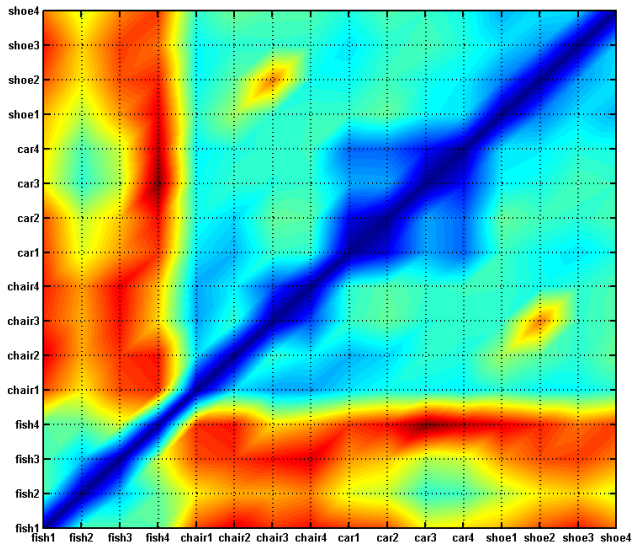


Figure 7.8.: Dissimilarity map combining VMI and saliency spheres registration. Blue and red values correspond to the most similar and dissimilar models respectively.

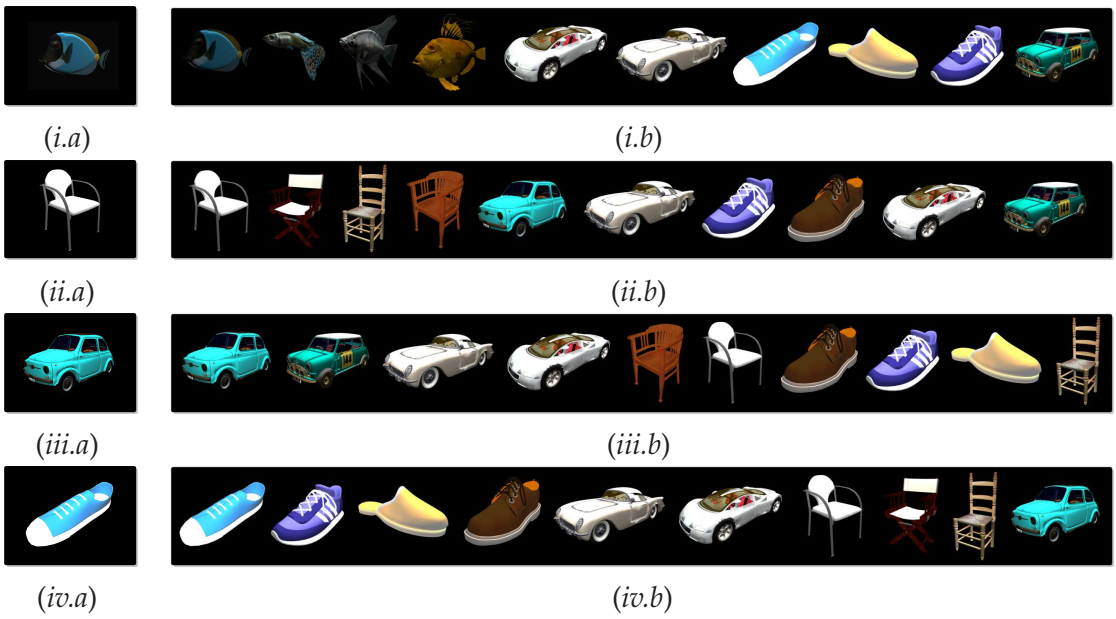


Figure 7.9.: Spherical registration using VMI and saliency spheres. In column (a) we show the first model of each family and in column (b) the list of the first ten objects sorted by its similarity with respect to the target model (a).

8. Conclusions and Future Work

We have defined a unified framework for viewpoint selection and mesh saliency based on an information channel between a set of viewpoints and the set of polygons of an object. A new viewpoint quality measure, the viewpoint mutual information, has been introduced to quantify the representativeness of a view and has been used to compute viewpoint stability, to select the N best views and to explore the object. From the reversion of the information channel, we have defined both the information and the saliency associated with each polygon, and we have also calculated the saliency of a viewpoint. Then, the viewpoint mutual information has been extended by incorporating the saliency as an importance factor. We have also presented a new information-theoretic approach to ambient occlusion. Our technique is based on the reversed viewpoint channel using the mutual information associated with each polygon with applications such as model enhancement (The important viewpoints can modulate the obtained ambient occlusion values), and as a relighting technique in combination with an NPR technique. Finally, we have made a first step in exploring the possibilities of our information-theoretic viewpoint selection framework to quantify the shape dissimilarity between 3D polygonal models. The presented approach is based on a spherical registration process considerating the VMI and saliency spheres as shape descriptors.

There are many issues of future research worth exploring. First, we plan to extend our viewpoint framework to indoor scenes. Second, we wish to analyze how our saliency approach can operate at multiple scales and can guide the object exploration. Third, we will investigate the integration of importance (for example, obtained from saliency or lighting) to the input distribution. Fourth, we will explore how the best viewpoint selection can be fine-tuned with other perceptual characteristics such as stability. Fifth, research will be addressed to investigate the quality of the ambient occlusion obtained with generalized Tsallis-Havrda-Charvat mutual information and to obtain a GPU implementation of our technique. Sixth, we will study how to apply our framework to obtain a *Polygonal Mutual Information-based* refinement criteria of models to represent, accurately, the occlusion details while avoiding unnecessary refinements that would increase the computational cost. Seventh, we will explore the use of stability spheres (see [FSG07]) as shape signatures in combination with VMI and saliency spheres, other metrics for spherical registration and dissimilarity quantification, numerical optimizers to speed up the registration process, several interpolators

and different resolutions of both the viewpoint sphere and the model mesh. Finally, we will study the creation of a 3D shape retrieval system based on the presented shape descriptors, the VMI and saliency spheres.

Bibliography

[AVF04] ANDÚJAR C., VÁZQUEZ P. P., FAIRÉN M.: Way-finder: guided tours through complex walthrough models. *Computer Graphics Forum (Eurographics 2004)* (2004).

[BET95] BÜLTHOFF H., EDELMAN S., TARR M.: How are three-dimensional objects represented in the brain? *Cerebral Cortex 5* (1995), 247–260.

[BR82] BURBEA J., RAO C. R.: On the convexity of some divergence measures based on entropy functions. *IEEE Transactions on Information Theory 28*, 3 (May 1982), 489–495.

[Bre02] BREDOW R.: Renderman in production, 2002.

[BS05] BORDOLOI U. D., SHEN H.-W.: Viewpoint evaluation for volume rendering. In *Visualization, IEEE 2005* (May 2005), pp. 62–62.

[BTB99] BLANZ V., TARR M., BÜLTHOFF H.: What object attributes determine canonical views? *Perception 28* (1999), 575–599.

[Bun05] BUNNELL M.: Dynamic ambient occlusion and indirect lighting. *GPU Gems 2* (2005), 223–233.

[CSCF07] CASTELLÓ P., SBERT M., CHOVER M., FEIXAS M.: Viewpoint entropy-driven simplification. In *Proceedings of WSCG 2007* (2007).

[CT91] COVER T. M., THOMAS J. A.: *Elements of Information Theory*. Wiley Series in Telecommunications, 1991.

[FMK*02] FUNKHOUSER T., MIN P., KAZHDAN M., CHEN J., HALDERMAN A., DOBKIN D., JACOBS D.: A search engine for 3d models. *ACM Transacions on Graphics* (2002).

[FSG07] FEIXAS M., SBERT M., GONZÁLEZ F.: *A Unified Information-Theoretic Framework for Viewpoint Selection and Mesh Saliency.* Research Report IIiA-07-03-RR, IIiA - Institut d'Informàtica i Aplicacions, Universitat de Girona (Girona, Spain), 2007.

[GBL02] GONZÁLEZ-BAÑOS H. H., LATOMBE J.-C.: Navigation strategies for exploring indoor environments. *I. J. Robotic Res. 21*, 10-11 (2002), 829–848.

[GCO06] GAL R., COHEN-OR D.: Salient geometric features for partial shape matching and similarity. *ACM Trans. Graph. 25*, 1 (2006), 130–150.

[GJ79] GAREY M. R., JOHNSON D. S.: Computers and intractability: A guide to the theory of np-completeness. *W.H. Freeman and Company* (1979).

[GRMS01] GOOCH B., REINHARD E., MOULDING C., SHIRLEY P.: Artistic composition for image creation. In *Rendering Techniques* (2001), pp. 83–88.

[IB05] ITTI L., BALDI P.: Bayesian surprise attracts human attention, Sep 2005.

[IKN98] ITTI L., KOCH C., NIEBUR E.: A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence 20*, 11 (Nov 1998), 1254–1259.

[IKSZ03] IONES A., KRUPKIN A., SBERT M., ZHUKOV S.: Fast, realistic lighting for video games. *IEEE Computer Graphics and Applications 23*, 3 (2003), 54–64.

[IM98] INDYK P., MOTWANI R.: Approximate nearest neighbors: Towards removing the curse of dimensionality. *In 30th Symposium on Theory of Computing* (1998), 604–613.

[JS06] JI G., SHEN H.-W.: Dynamic view selection for time-varying volumes. *IEEE Trans. Vis. Comput. Graph. 12*, 5 (2006), 1109–1116.

[JT04] J. TANGELDER R. V.: A survey of content based 3d shape retrieval methods. *Proceedings of Shape Modeling International* (2004), 145–156.

[KA06] KONTKANEN J., AILA T.: Ambient occlusion for animated characters. In *Eurographics Symposium on Rendering* (2006), Akenine-Möller T., Heidrich W., (Eds.).

[KL05] KONTKANEN J., LAINE S.: Ambient occlusion fields. In *Proceedings of I3D* (2005).

[KV06] KIM Y., VARSHNEY A.: Saliency-guided enhancement for volume visualization. *IEEE Trans. Vis. Comput. Graph. 12*, 5 (2006), 925–932.

[KZS96] K. ZHANG J. T. L. W., SHASHA D.: On the editing distance between undericted acyclic graphs. *Int'l J. Foundations of Computer Science* (1996), 43–57.

[LÖ0] LÖFFLER J.: Content-based retrieval of 3d models in distributed web databases by visual shape information. *In IV2000* (2000).

[Lak] LAKE A.: Cartoon rendering using texture mapping and programable vertex shaders. *Game Programming Gems 2*.

[Lan02] LANDIS H.: Renderman in production, 2002.

[LME06] LU A., MACIEJEWSKI R., EBERT D. S.: Volume composition using eye tracking data. In *Eurographics/ IEEE-VGTC Symposium on Visualization* (2006), Ertl T., Joy K., Santos B., (Eds.).

[LMHB00] LAKE A., MARSHALL C., HARRIS M., BLACKSTEIN M.: Stylized rendering techniques for scalable real-time animation. In *Proceedings of the 1st International Symposium on Non-photorealistic Animation and Rendering (NPAR)* (2000).

[LVJ05] LEE C. H., VARSHNEY A., JACOBS D. W.: Mesh saliency. *Computer Graphics (Proceedings of SIGGRAPH'05)* (July - August 2005). Held in Los Angeles, USA.

[MSC03] MÉNDEZ A., SBERT M., CATÀ J.: Real-time obscurances with color bleeding. In *Proceedings of the 19th spring conference on Computer graphics* (2003).

[MSC*06] MÉNDEZ A., SBERT M., CATÀ J., SUNYER N., FUNTANÉ S.: Real-time obscurances with color bleeding (gpu obscurances with depth peeling). *ShaderX 4* (2006), 1196–1201.

[Nem80] NEMCSICS A.: The role of aesthetical uniformicity in the color harmony (in hungarian). In *In Proceeding of the XII. Color Symposium* (1980), pp. 24–29.

[PB96] PLEMENOS D., BENEYADA M.: Intelligent display techniques in scene modelling. new techniques to automatically compute good views. In *International Conference GraphiCon'96* (July 1996). Held in St. Petersburg, Russia.

[PPB*05] POLONSKY O., PATANÈ G., BIASOTTI S., GOTSMAN C., SPAGNUOLO M.: What's in an image? *The Visual Computer 21*, 8-10 (2005), 840–847.

[PRC81] PALMER S., ROSCH E., CHASE P.: Canonical perspective and the perception of objects. *Attention and Performance IX* (1981), 135–151.

[SBP03] S. BIASOTTI S. MARINI M. M., PATANÉ G.: An overview of properties and efficacy of topological skeletons in shape modelling. *In SMI* (2003), 245–254.

[SPFG05] SBERT M., PLEMENOS D., FEIXAS M., GONZÁLEZ F.: Viewpoint quality: Measures and applications. In *Proceedings of 1st Computational Aesthetics in Graphics, Visualization and Imaging* (May 2005). Held in Girona, Spain.

[SPT06] SOKOLOV D., PLEMENOS D., TAMINE K.: Methods and data structures for virtual world exploration. *The Visual Computer 22*, 7 (2006), 506–516.

[SSZK04] SATTLER M., SARLETTE R., ZACHMANN G., KLEIN R.: Hardware-accelerated ambient occlusion computation. In *Proceedings of International Fall Workshop on Vision, Modeling, and Visualization* (2004).

[ST00a] SLONIM N., TISHBY N.: Agglomerative information bottleneck. In *Proceedings of NIPS-12 (Neural Information Processing Systems)* (2000), MIT Press, pp. 617–623.

[ST00b] SLONIM N., TISHBY N.: Document clustering using word clusters via the information bottleneck method. In *Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (2000), ACM Press, pp. 208–215. Held in Athens, Greece.

[TBSK01] T. B. SEBASTIAN P. N. K., KIMIA B. B.: Recognition of shapes by editing shock graphs. *In ICCV* (2001), 755–762.

[TBZB97] TARR M., BÜLTHOFF H., ZABINSKI M., BLANZ V.: To what extent do unique parts influence recognition across changes in viewpoint? *Psychological Science 8*, 4 (1997), 282–289.

[TFTN05] TAKAHASHI S., FUJISHIRO I., TAKESHIMA Y., NISHITA T.: Locating optimal viewpoints for volume visualization. In *Visualization, IEEE 2005* (May 2005).

[Vaz03] VAZQUEZ P. P.: *On the Selection of Good Views and its Application to Computer Graphics*. PhD thesis, Technical University of Catalonia, Barcelona, Spain, 2003.

[VFSG06] VIOLA I., FEIXAS M., SBERT M., GRÖLLER M. E.: Importance-driven focus of attention. *IEEE Trans. Vis. Comput. Graph. 12*, 5 (2006), 933–940.

[VFSH01] VÁZQUEZ P. P., FEIXAS M., SBERT M., HEIDRICH W.: Viewpoint selection using viewpoint entropy. In *Proceedings of Vision, Modeling, and Visualization 2001* (Stuttgart, Germany, November 2001), Ertl T., Girod B., Greiner G., Niemann H., Seidel H.-P., (Eds.), pp. 273–280. Held in Stuttgart, Germany.

[VFSH03] VÁZQUEZ P. P., FEIXAS M., SBERT M., HEIDRICH W.: Automatic view selection using viewpoint entropy and its application to image-based modeling. *Computer Graphics Forum* (Desember 2003).

[VFSL06] VÁZQUEZ P. P., FEIXAS M., SBERT M., LLOBET A.: Realtime automatic selection of good molecular views. *Computers & Graphics 30*, 1 (February 2006), 98–110.

[YSY*06] YAMAUCHI H., SALEEM W., YOSHIZAWA S., KARNI Z., BELYAEV A. G., SEIDEL H.-P.: Towards stable and salient multi-view representation of 3d shapes. In *SMI* (2006), p. 40.

[ZIK98] ZHUKOV S., IONES A., KRONIN G.: An ambient light illumination model. In *Rendering Techniques* (1998), pp. 45–56.

A. Publications

Publications that support the contents of this master thesis:

- M.Sbert, D.Plemenos, M.Feixas and F.González. Viewpoint quality: measures and applications. *Eurographics Workshop on Computational Aesthetics in Graphics, Visualization and Imaging*, Girona (Spain), 2005.
- Miquel Feixas, Mateu Sbert and Francisco González. A Unified Information-Theoretic Framework for Viewpoint Selection and Mesh Saliency. IIiA-07-3-RR, Institut d'Informàtica i Aplicacions, Universitat de Girona (Girona, Spain). Extended version of IIiA-06-06-RR. Submitted to *ACM Transactions on Applied Perception.*
- Francisco González, Mateu Sbert and Miquel Feixas. An Information-Theoretic Ambient Occlusion. *International Symposium on Computational Aesthetics in Graphics, Visualization, and Imaging*, Banff (Canada), 2007. Eurographics.
- Francisco González, Miquel Feixas and Mateu Sbert. View-based Shape Similarity using Mutual Information Spheres. *Eurographics'07 Short Paper*, Prague, Czech Republic, September 3-7, 2007.