ADVANCED TOPICS IN SCIENCE AND TECHNOLOGY IN CHINA

Faxin Yu · Zheming Lu
Hao Luo · Pinghui Wang

Three-Dimensional Model Analysis and Processing

ZHEJIANG UNIVERSITY PRESS
浙江大学出版社

Springer

Faxin Yu
Zheming Lu
Hao Luo
Pinghui Wang

Three-Dimensional Model Analysis and Processing

With 134 figures

ZHEJIANG UNIVERSITY PRESS
浙江大学出版社

Springer

图书在版编目 (CIP) 数据

三维模型分析与处理=Three-Dimensional Model Analysis and Processing：英文 / 郁发新等著．—杭州：浙江大学出版社，2010.4
(中国科技进展丛书)
ISBN 978-7-308-07412-4

I. ①三… II. ①郁… III. ①三维—模型—计算机辅助设计—英文 IV. ①TP391.41

中国版本图书馆 CIP 数据核字(2010)第 034717 号

三维模型分析与处理

郁发新 陆哲明 罗 浩 王凭慧 著

责任编辑	伍秀芳
封面设计	俞亚彤
出版发行	浙江大学出版社 网址：http://www.zjupress.com Springer-Verlag GmbH 网址：http://www.springer.com
排　　版	杭州中大图文设计有限公司
印　　刷	杭州富春印务有限公司
开　　本	710mm×1000mm 1/16
印　　张	27.25
字　　数	785 千
版 印 次	2010 年 4 月第 1 版 2010 年 4 月第 1 次印刷
书　　号	ISBN 978-7-308-07412-4 (浙江大学出版社) ISBN 978-3-642-12650-5 (Springer-Verlag GmbH)
定　　价	176.00 元

Preface

With the increasing popularization of the Internet, together with the rapid development of 3D scanning technologies and modeling tools, 3D model databases have become more and more common in fields such as biology, chemistry, archaeology and geography. People can distribute their own 3D works over the Internet, search and download 3D model data, and also carry out electronic trade over the Internet. However, some serious issues are related to this as follows: (1) How to efficiently transmit and store huge 3D model data with limited bandwidth and storage capacity; (2) How to prevent 3D works from being pirated and tampered with; (3) How to search for the desired 3D models in huge multimedia databases. This book is devoted to partially solving the above issues.

Compression is useful because it helps reduce the consumption of expensive resources, such as hard disk space and transmission bandwidth. On the downside, compressed data must be decompressed to be used, and this extra processing may be detrimental to some applications. 3D polygonal mesh (with geometry, color, normal vector and texture coordinate information), as a common surface representation, is now heavily used in various multimedia applications such as computer games, animations and simulation applications. To maintain a convincing level of realism, many applications require highly detailed mesh models. However, such complex models demand broad network bandwidth and much storage capacity to transmit and store. To address these problems, 3D mesh compression is essential for reducing the size of 3D model representation.

Feature extraction is a special form of dimensionality reduction. When the input data to an algorithm is too large to be processed and is suspected to be notoriously redundant (much data, but not much information), the input data will be transformed into a reduced representation set of features (also named a feature vector). If the features extracted are carefully chosen, it is expected that the features set will extract the relevant information from the input data, in order to perform the desired task using this reduced representation instead of the full size input. Feature extraction is an essential step in content-based 3D model retrieval systems. In general, the shape of the 3D object is described by a feature vector that serves as a search key in the database. If an unsuitable feature extraction method has been used, the whole retrieval system will be unusable. We must realize that 3D objects can be saved in many representations, such as polyhedral meshes, volumetric data and parametric or implicit equations. The method of feature extraction should accept this fact and it should be independent of data representation. The method should also be invariant under transforms such as translation, rotation and scale of the 3D object. Perhaps this is the most important requirement, because the 3D objects are usually saved in various poses and on various scales. The 3D object can be obtained either from a 3D graphics program or from a 3D input device. The second way is more susceptible to some errors, therefore the feature extraction method should also be insensitive to noise. Perhaps the last requirement is that it has to be quick to compute and easy to index. The database may contain thousands of objects, so the agility of the system would also be one of the main requirements.

Content-based visual information retrieval (CBVIR) is the application of computer vision to the visual information retrieval problem, which solves the problem of searching for digital images/videos/3D models in large databases. "Content-based" means that the search will analyze the actual contents of the visual media. The term "content" in this context might refer to colors, shapes, textures, or any other information that can be derived from the visual media itself. Without the ability to examine visual media content, searches must rely on metadata such as captions and keywords, which may be laborious or expensive to produce. A common characteristic of all applications in multimedia databases (and in particular in 3D object databases) is that a query searches for similar objects instead of performing an exact search, as in traditional relational databases. Multimedia objects cannot be meaningfully queried in the classical sense (exact search), because the probability that two multimedia objects are identical is very low, unless they are digital copies from the same source. Instead, a query in a multimedia database system usually requests a number of objects most similar to a given query object or to a manually entered query specification. Therefore, one of the most important tasks in a multimedia retrieval system is to implement effective and efficient similarity search algorithms. Typically, the multimedia data are modeled as objects in a metric or vector space, where a distance function must be defined to compute the similarity between two objects. Thus, the similarity search problem is reduced to a search for close objects in the metric or vector space. The primary goal in a 3D similarity search is to design algorithms with the ability to effectively and efficiently execute similarity queries in 3D databases. Effectiveness is related to the ability to retrieve similar 3D objects while holding back non-similar ones, and efficiency is related to the cost of the search, measured e.g., in CPU or I/O time. But, first of all one should define how the similarity between 3D objects is computed.

Digital watermarking is a branch of data hiding (or information hiding). It is the process of embedding information into a digital signal. The signal may be audios, pictures, videos or 3D models. If the signal is copied, then the information is also carried in the copy. An important application of invisible watermarking is in copyright protection systems, which are intended to prevent or deter unauthorized copying of digital media. Another important application is to authenticate the content of multimedia works, where fragile watermarks are commonly used for tamper detection (integrity proof). Steganography is an application of digital watermarking, where two parties communicate a secret message embedded in the digital signal. Annotation of digital photographs with descriptive information is another application of invisible watermarking. While some file formats for digital media can contain additional information called metadata, digital watermarking is distinct in that the data is carried in the signal itself.

Reversible data hiding is a technique that enables images or 3D models to be authenticated and then restored to their original forms by removing the watermark and replacing the images or 3D data which had been overwritten. This would make the images or 3D models acceptable for legal purposes. Although reversible data hiding was first introduced for digital images, it has also wide application scenarios for hiding data in 3D models. For example, suppose there is a column on a 3D mechanical model obtained by CAD. The diameter of this column is changed with a given data hiding scheme. In some applications, it is not enough that the hidden content is accurately extracted, because the remaining watermarked model is still distorted. Even if the column diameter is increased or decreased by 1 mm, it may cause a severe effect for this mechanical model cannot be well assembled with other mechanical accessories. Therefore, it also has significance in the design of reversible data hiding methods for 3D models.

Based on the above background, this book is devoted to processing and analysis techniques for 3D models, i.e., compression techniques, feature extraction and retrieval techniques and watermarking techniques for 3D models. This book focuses on three main areas in 3D model processing and analysis, i.e., compression, content-based retrieval and data hiding, which are designed to reduce redundancy in 3D model representations, to extract the features from 3D models and retrieve similar models to the query model based on feature matching, to protect the copyright of 3D models and to authenticate the content of 3D models or hide information in 3D models. This book consists of six chapters. Chapter 1 introduces the background to three urgent issues confronting multimedia, i.e., storage and transmission, protection and authentication, and retrieval and recognition. Then the concepts, descriptions and research directions for the newly-developed digital media, 3D models, are presented. Based on three aspects of the technical requirements, the basic concepts and the commonly-used techniques for multimedia compression, multimedia watermarking, multimedia retrieval and multimedia perceptual hashing are then summarized. Chapter 2 introduces the background, basic concepts and algorithm classification of 3D mesh compression techniques. Then we discuss some typical methods used in connectivity compression and geometry compression for 3D meshes respectively. Chapter 3 focuses on the techniques of feature extraction from 3D models. First, the background, basic concepts and algorithm classification related to 3D model feature extraction are introduced. Then, typical 3D model feature extraction methods are classified into six categories and are, discussed in eight sections, respectively. Chapter 4 discusses the steps and techniques related to content-based 3D model retrieval systems. First, we introduce the background, performance evaluation criteria, the basic framework, challenges and several important issues related to content-based 3D model retrieval systems. Then we analyze and discuss several topics for content-based 3D model retrieval, including preprocessing, feature extraction, similarity matching and query interface. Chapter 5 starts with the description of general requirements for 3D watermarking, as well as the classification of 3D model watermarking algorithms. Then some typical spatial domain 3D mesh model watermarking schemes, typical transform-domain 3D mesh model watermarking schemes and watermarking algorithms for other types of 3D models are discussed respectively. Chapter 6 starts by introducing the background and performance evaluation metrics of 3D model reversible data hiding. Then some basic reversible data hiding schemes for digital images are briefly reviewed. Finally, three kinds of 3D model reversible data hiding techniques are extensively introduced, i.e., spatial domain based, compressed domain based and transform domain based methods.

This book embodies the following characteristics. Firstly, it has novelty. The content of this book covers the research hotspots and their recent progress in the field of 3D model processing and analysis. For example, in Chapter 6, reversible data hiding in 3D models is a very new research branch. Secondly it has completeness. Techniques for every research direction are comprehensively introduced. For example, in Chapter 3, feature extraction methods for 3D models are classified and introduced in detail. Thirdly it is theoretical. This book embodies many theories related to 3D models, such as topology, transform coding, data compression, multi-resolution analysis, neural networks, vector quantization, 3D modeling, statistics, machine learning, watermarking, data hiding, and so on. For example, in Chapter 2, several definitions related to 3D topology and geometry are introduced in detail in order to easily understand the content of later chapters. Fourthly it is practical. For each application, experimental results for typical methods are illustrated in detail. For example, in Chapter 6, three examples of typical reversible data hiding are illustrated with detailed steps and elaborate experiments.

In this book, Chapters 1, 4 and 5 were written by Prof. Zheming Lu, Chapters 2 and 3 were written by Prof. Faxin Yu, Chapter 6 was written by Dr. Hao Luo with the aid of student Hua Chen. The whole book was finalized by Prof. Faxin Yu. The research results of this book are based on the accumulated work of the authors over a long period of time. We would like to show our great appreciation for the assistance of other teachers and students in the Institute of Astronautics and Electronic Engineering of Zhejiang University. The work was partially supported by the National Natural Science Foundation of China, the foundation from the Ministry of Education in China for persons showing special ability in the new century, and the foundation from the Ministry of Education in China for the best national Ph.D dissertations. Due to our limited knowledge, it is inevitable that errors and defects will appear in this book and we invite our readers to comment.

The authors
Hangzhou, China
January, 2010

Contents

1

Introduction

The digitization of multimedia data, such as images, graphics, speech, text, audio, video and 3D models, has made the storage of multimedia more and more convenient, and has simultaneously improved the efficiency and accuracy of information representation. With the increasing popularization of the Internet, multimedia communication has reached an unprecedented level of depth and broadness, and multimedia distribution is becoming more and more manifold. People can distribute their own works over the Internet, search and download multimedia data, and also carry out electronic trade over the Internet. However, some serious issues accompany this as follows: (1) How can we efficiently transmit and store huge multimedia information with limited bandwidth and storage capacity? (2) How can we prevent multimedia works from being pirated and tampered with? (3) How can we search for the desired multimedia content in huge multimedia databases?

1.1 Background

We first introduce the background to three urgent issues for multimedia, i.e., (1) storage and transmission, (2) protection and authentication, (3) retrieval and recognition.

1.1.1 Technical Development Course of Multimedia

"Multimedia" [1] is a compound word composed of "multiple" and "media", which means "multiple media". Here, "media" is the plural form of the word "medium". In fact, the word "medium" has two kinds of meaning in the computer field: one stands for the entities for storing information, such as diskettes, CDs, magnetic tapes and semiconductor memorizers; the other stands for the carriers for transmitting information, such as digits, characters, audio clips, graphics and images. Here, the word "media" in multimedia technology means the latter. "Monomedia" is one (word) as opposed to "multimedia" and, literally, multimedia is composed of several "monomedia". People use various media during information communication, and multimedia is just the representation and transmission form for multiple information carriers. In other words, it is a technique to simultaneously acquire, process, edit, store and display more than two kinds of media, including text, audios, graphics, images, movies and videos, etc. In fact, it is the material development of computer and digital information processing technologies that enables people to process multimedia information and thus enables the realization of multimedia technology. Therefore, so-called "multimedia" stands no longer for multiple media themselves but for the whole series of techniques to deal with and apply them. In fact, "multimedia" has been viewed as a synonym of "multimedia technology". It is worth noting that multimedia technology nowadays is often associated with computer technology. The reason is that the computer's capability of digitization and interactive processing greatly promotes the development of multimedia technology. In general, people can view multimedia as the new technology or as product forming from the combination of advanced computer, video, audio and communication technologies.

The multimedia technique has been rapidly developed accompanied by the wide application of computer and network technologies, and computer network multimedia technology has become an area under rapid development and has gained research focus in the 21st century. As a rapidly developing all-round electronic information technology, multimedia technology has brought directional renovation to traditional computer systems and audio and video equipments, and will have a great effect on mass media. Since the mid to late 1980s, multimedia computer technology has become the focus of concern, and its definition is as follows: computers comprehensively process various kinds of multimedia information (text, graphics, images, audios and videos), which means various kinds of information is linked together to form a system with interactivity. Interactivity is one of the characteristics of multimedia computer technology, meaning the characteristic of interactive communication with users, which is the biggest difference from traditional media. Apart from providing users with solutions to problems on their own, such a change can help users learn and think with the aid of conversational communication and carry out systematical queries or statistical analysis in order to achieve the advancement of knowledge and the improvement of problem-solving ability. Multimedia computers will speed up the process of introducing computers to families and societies, and will bring a profound revolution to people's work, life and entertainment. Since the 1990s, the progress that the world has made towards an information society has been significantly expedited, in which the application of multimedia technology has been playing a vital role. Multimedia improves a human's information communication and shortens the communication path. The application of multimedia technology is a sign of the 1990s, and is a second revolution in the computer field.

On the whole, multimedia technology is nowadays developing in the following two directions.

One is networking, which means that, combined with wide-band network communication technology, multimedia technology enters areas such as scientific research, designing, enterprise management, office automation, remote education, telemedicine, retrieval, entertainment and automatic testing. In some recent films, we can often see a very personalized computer that can talk with humans and provide any information they want to know. It can play any music they want to listen to. If there is any accident anywhere in the world, it can report to them in time. It can monitor the status of all the apparatus at home, and can help to receive phone calls and remind humans what to do, and even transmit messages to their friends living far away. Today, because of the development of multimedia, all of the above dreams will come true.

The other direction is componentization together with intelligentization and embeddability of the multimedia terminal, which means improving the multimedia performance of computer systems to develop intelligent household appliances. The current household television system cannot be called a multimedia system, because although existing televisions also provide "sound, graphics, text" information, people can do nothing but select different channels, and people cannot interfere or change them but passively receive the programs from TV stations. This process is not two-way but one-way. However, we can forecast that, in the near future, the household television system will definitely be a multimedia system, which will combine many functions, such as entertainment, education, communication and consultation, all in one.

In summary, the birth of multimedia technology will definitely bring a revolution to the computer field once more. It indicates computers will not only be used in offices and laboratories but also be used in the household, in commerce, for travel, amusement, education and art, etc., i.e., in nearly all areas of daily life. At the same time, it means computers can be developed in the most ideal way for humans, i.e., with the integration of seeing and hearing, which completely plays down the human-computer interface.

1.1.2 Information Explosion

Real human civilization starts from the Internet. In fact, we are living with all kinds of networks, such as electrical networks, telephone networks, broadcast/television networks, commercial networks and traffic networks. However, all these networks are very different from the Internet, which has affected so many governments, enterprises and individuals in such a short time. Nowadays, the network has become a substitutable noun for the Internet. In the past few years, with the rapid development of computer and network techniques, the scale of the Internet has been suddenly expanded. The Internet technique breaks the traditional borderline, which makes the world smaller and smaller, while making the market larger and larger. The wide world is like a global village, where the global economy and information networking promote and depend on each other. The Internet makes the speed and scale of information acquisition and transmission reach an unprecedented level. In the era of information networking, the Internet should be considered for any product or technique. Network information systems are playing more and more important roles in politics, military affairs, finance, commerce, transportation, telecommunication, culture and education. Modern communication and transmission techniques have greatly improved the speed and extent of information transmission. The technical means include broadcasts, television, satellite communication and computer communication using microwave and optical fiber communication networks, which overcome traditional obstacles in space and time and further unite the whole world. However, the accompanying issues and side effects are as follows: A surge of information overwhelms people, and it is very hard to retrieve accurately and rapidly the information most needed from the tremendous amount of information. This phenomenon is called the information explosion [2], also called "information overload" or "knowledge bombing".

The information explosion describes the rapid development in the amount of information or human knowledge in recent years, whose speed is like a bomb engulfing all the world. With regard to the phrase "information explosion", it can date back to the 1980s. At that time, besides broadcasting, television, telephone, newspapers and various publications, new means of communication, i.e., computers and communication satellites emerged, making the amount of information increase suddenly like an explosion. Statistics show that over the past decade the amount of information all over the world doubled every 20 months. During the 1990s, the amount of information continued to increase dramatically. At the end of the 1990s, due to the emergence of the Internet, information distribution and transmission got out of control, and a great deal of false or useless information was generated, resulting in the pollution of information environments and the birth of "waste messages". Because everyone can freely air his opinion over the Internet, and the distribution cost can be ignored, in a sense everyone can become an information manufacturer on the global level, and thus information really starts to explode. As times go by, the information explosion manifests itself mainly in five aspects: (1) the rapid increase in the amount of news; (2) the dramatic increase in the amount of amusement information; (3) a barrage of advertisements; (4) the rapid increase in scientific and technical information; (5) the overloading of our personal receptiveness. However, faced with the inflated amount of information and the enormous pressure of "chaotic information space" and "information surplus", people out of the blue become hesitant in their urgent pursuit and expectation of information. Even if we take 24 hours every day to read information, we cannot take it all in, and besides, there is a great deal of useless or false information. Useful information can increase economic benefits and promote the development of human society, but if the information increases in a disorderly fashion and even runs out of control, it will bring about various social problems such as information crime and information pollution. People on the one hand are enjoying the convenience brought about by abundant information over the Internet; on the other hand they are suffering from annoyance due to the "information