NORSIGD INFO

Nummer 1 2003

NORSK SAMARBEID INNEN GRAFISK DATABEHANDLING

ISSN 0803-8317

Aktivitetskalender

Hva skjer når og hvor?

Mai 2003

22–23 **IPT/EGVE 2003**, Joint 7th Immersive Projection Technology Workshop / 9th Eurographics Workshop on Virtual Environments, Zürich, Sveits. http://www.ipt-egve.ethz.ch/.

26–28 **VisSym 2003**, Symposium on Visualization, Grenoble, Frankrike. http://www.inrialpes.fr/VisSym03/.

Juni 2003

25–27 **EGRWS/EGSR 2003** – Eurographics Workshop/Symposium on Rendering 2003, Leuven, Belgia. http://www.egsr2003.org/.

23–25 **SGP 2003** – Symposium on Geometry Processing 2003, Aachen, Tyskland. http://www.geometryprocessing.org/.

Juli 2003

9–11 **CGI 2003** – Computer Graphics International 2003, Tokyo, Japan. http://www.img.cs.titech.ac.jp/~CGI2003/.

27–31 **SIGGRAPH 2003** – 30th International Conference on Computer Graphics and Interactive Techniques, San Diego, CA, USA. http://www.siggraph.org/.

September 2003

1–6 **EG 2003** – 24th annual conf. of the European Association for Computer Graphics (EUROGRAPHICS). Granada, Spania.. http://www.eg.org/eg2003/.

5–10 **GraphiCon 2003** – 13th Annual International Conference on Computer Graphics. Moskva, Russland. http://www.graphicon.ru/.

Februar 2004

28–29 **WCSG 2004**, 12th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision, Plzen, Tsjekkia. http://wscg.zcu.cz/wscg2004/wscg2004.htm.

Helwig's Conference Calender

Flere aktiviteter finner du på http://www.vrvis.at/ConfCal/.

Om forsiden
Bildet på forsiden viser en animert statue av fysikeren Max Planck. Metoden er forklart i bidraget på side 4 i denne utgaven.

Hilsen fra styret

Kjære medlemmer,

Datagrafikkseminaret på Fornebu i begynnelsen av mars 2003 var en stor suksess. Over femti deltagere fra forskjellige datagrafikkmiljøer i Norge hadde samlet seg for å fortelle hverandre hva som foregår innen vårt fagfelt. NORSIGD kommer til å fortsette med å støtte slike faglige utvekslinger av idéer. Det er planlagt å gjengi noen av temaene som ble presentert på seminaret i NORSIGD Info. Vi gir også et referat fra seminaret i denne utgaven.

Som faglige artikler presenterer vi en metode for animerte avatarer, som ble utviklet av en forskergruppe ved universitetet i Moskva. Metoden ble også presentert i en postersesjon ved Eurographics Konferansen i fjor.

I rekken av foredrag fra datagrafikkseminaret skisseres det hvordan billedbaserte datagrafikkmetoder kan spille sammen med digital TV.

Datagrafikk og bildebehandling har mye til felles. Vi har lånt en artikkel fra NOBIM om en applikasjon som ble presentert på forskningsdagene i fjor høst.

Hilsen,

Wolfgang Leister

NORSIGD Info
– medlemsblad for NORSIGD

Utgitt av:	NORSIGD
Ansvarlig:	Wolfgang Leister Norsk Regnesentral Postboks 114 Blindern 0314 OSLO
ISSN:	0803-8317
Utgivelser:	2003: 15/4 15/10
Annonsepriser:	Helside kr 5 000 Halvside kr 2 500
Layout:	Wolfgang Leister LaTeX2ε

Innhold

3D Talking Heads – Personalized and Animated

Denis Ivanov, Victor Lempitsky, and Yevgeniy Kuzmin
Dept. of Mathematics and Mechanics, Moscow State University[1]

In this paper we describe our experience in creating personalized realistically looking 3D models of a human head from digital images or video. The obtained models, being represented by textured polygonal meshes, can be efficiently animated and rendered utilizing modern GPU features. Our research is currently focused on making the complete process fully automatic, which will allow for use of the software by non-expert users.

Animated models of a human head are demanded in a large variety of modern applications, including among many others computer games, film production, and video conferencing. However, the problem of the effortless generation of a realistic looking, high quality model has been one of the most difficult in computer graphics, as no general, complete and efficient solution seems yet to be available.

On the one hand, being a solid object in a 3D world, the human head can be digitized using commercially available 3D scanning machinery based on laser range finders [14,16] or similar technologies. This approach allows the generation of a relatively accurate shape of the model and accompanying texture; however, the produced data are not directly suitable for animation, and proper adaptation usually requires a great deal of an effort.

On the other hand, using a priori knowledge about the underlying structure of a head enables one to obtain a better result. This idea is typically realized in the calibration strategies where an existing model of a generic head is adjusted so that it matches the available input data [7,10,13,17] or an existing set of models is considered as the formal basis in some vector space [2]. The resulting model is usually more suitable for animation, as it is based on a priori knowledge about the human head.

Head model calibration methods can be classified by the source data which they are based upon. As depth information is expensive to obtain, raster images are usually considered as the only input. Thus, there exist techniques of calibration from one image [2], a pair of orthogonal images [10,11], and a sequence of images or a video stream [4,6,17]. It is also important whether a special setup of the camera, lighting or viewing parameters is required, and having no special restrictions is obviously preferable.

The other property that characterizes calibration methods is to what extent the user participates in the process. Some methods are fully automatic [2,7], some require several points be selected on images [4,6], while others demand much additional effort from the user [11,12,17]. On average, manual input of some parameters allows a model of better quality to be produced, while automatic techniques are obviously more preferable from the customer's point of view.

The goal of our research was to develop a complete calibration pipeline that would allow for the adjustment of a polygonal model of a generic head based on a set of photographs. Textured polygonal models were taken into consideration because there exist several solutions for their animation [5,8,9] that take advantage of modern GPU features.

We have started our research by considering two orthogonal views of front and profile as input data, which conceptually follows [10,11]. However, our analysis ended up with the conclusion that these two images with no other supporting data, such as, for example, images taken from other angles or a database of head models for reference, can not provide realistic model for high-resolution rendering. We will discuss this issue in more details in this paper.

Our further research focused on processing of several images, typically from 5 to 10, taken at different viewing angles by a still digital camera. Such approach allowed for reconstructing models of better quality in terms of their perceptual similarity compared to the real objects. However, we faced the necessity of selecting more facial features on the images, which might not be amenable to automatic selection.

In order to provide further improvement of resulting model quality, more suitable environment for facial feature detection and tracking (due to estimated temporal smoothness of head movements on video) and more user- friendly processing we are currently working on fully

[1] {denis,vitya,yevgeniy}@fit.com.ru

automatic creation of a complete personalized 3D model of a head from a video stream taken by consumer digital video camera.

Pipeline Overview

The head calibration process is organized as a pipeline comprising a sequence of separate stages, which are schematically shown in Figure 1. Each stage processes the data obtained from its predecessors and performs the corresponding operations. Below we describe the goals that are addressed by each stage.

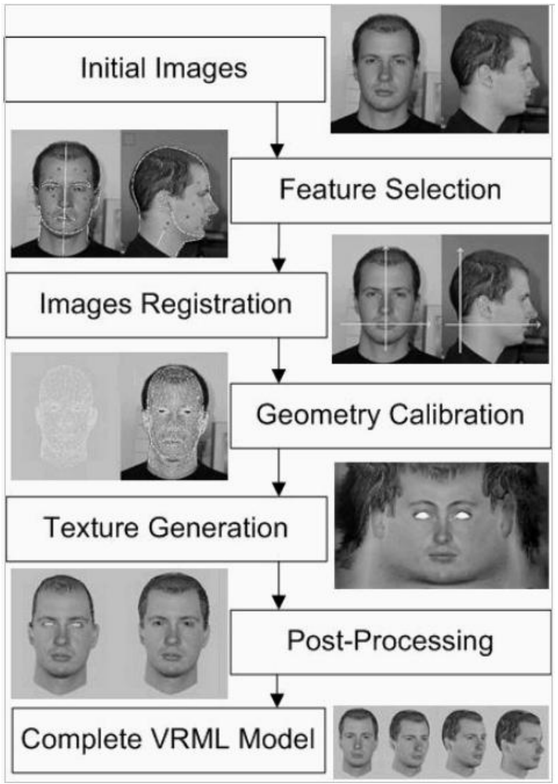


Figure 1. Overview of the head calibration pipeline.

Feature Selection. The first stage of the pipeline selects descriptive facial features on the given images or video frames. All features fall into several categories, which are (1) fixed feature points, e.g., nose tip or mouth corners; (2) individual feature points, e.g., moles or freckles; (3) feature contours, e.g., eye or mouth contour; and (4) silhouette lines, which are lines between face area and background. These elements are utilized by the succeeding stages of the pipeline for various purposes. Currently, we managed to select most of the features automatically; however, there are still some problems in this area since faces can vary significantly from one to the other.

Image Registration. The goal of this stage is to register available images within the coordinate system associated with the 3D head. In case of front and profile views we use orthogonal projection for camera model, which conceptually follows [10,11]. If more images are available, we estimate intrinsic (focal length) and extrinsic (position and orientation) parameters of perspective camera based on the point features extracted on the previous stage. For such estimation we use modification of the algorithm proposed in [17]. Obviously, perspective camera model allows achieving better overall result as it better represents the process that takes place during taking images; however, estimation of the required parameters implies solving a non-linear problem with no guarantee of finding exact solution. In practice, quite sufficient estimations are obtained if representative set of features is selected in images.

Geometry Calibration. In this stage, the shape of the generic model is adjusted so that it matches feature elements selected on the images. This operation can be implemented in several steps, the first of which adjusts global proportions of the head while the last is responsible for local fitting. We implemented and analyzed several solutions including Radial Basis Functions (RBF) [1,17], Dirichlet Free-Form Deformation (DFFD) [12], and point-based deformation [8]. However, our observation showed that most of these approaches are not applicable to models of high resolution as they have certain limitations (for example, RBF do not take surface shape into account as morphing of the whole 3D space actually takes place) which leads to noticeable artifacts. Thus, as an alternative we have implemented mesh relaxation scheme that deforms generic mesh so that it matches selected feature elements (including contours and silhouette lines) while preserving its authenticity (i.e. smoothness and consistency).

Texture Generation. This stage aims to generate an appropriate texture for the calibrated model. The strategy that we use for this stage comprises morphing of available images by inverse texture mapping and then merging areas obtained from the different views. This process can be conceptually viewed as composing final texture as a mosaic of fragments obtained from available bitmaps. Such fragments are merged using either weighted sum of the correspondent elements [4,17] or pyramidal decomposition based on Gaussian operator [3]. We have also proposed an original technique [15] that finds merging line between such fragments so that the resulting texture has maximum details possible. This approach allowed for constructing seamless texture of better quality as it preserved more details available on the original images.

Post-processing. Some applications may demand that special conditions be satisfied by the produced model. These conditions can be fulfilled in a post-processing stage. For example, eyes are required to be separate objects for proper animation; thus, corresponding geometry as well as texture should be generated, based on the initial images. Besides, hair style may vary significantly for different people, which requires representing hair as a separate object. We represent hair as a separate mesh which is deformed using silhouette lines selected on the given images. Such mesh is then seamlessly attached to the head model.

MPEG-4 Facial Feature Points

MPEG-4 standardizes a set of 84 feature points that characterize head shape (see [10] for an overview). These points are used for both personification (calibration) and animation processes. Some feature points are shown in Figure 2 as an example.

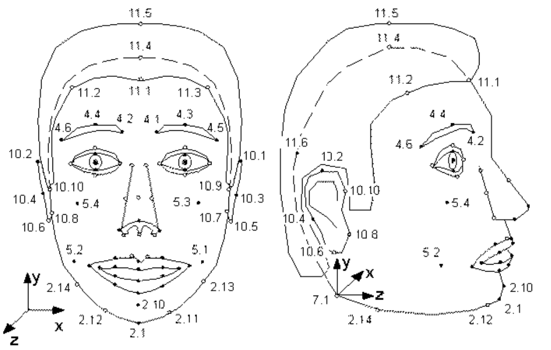


Figure 2. Example of MPEG-4 facial feature points.

In the beginning of our research we tried to use only these points, thus, following the standard. However, we soon realized that some of the points are not well-defined on images (for example, those belonging to the chicks) and, therefore, can not be reliably selected. In addition, some very useful points were missing (for example, bridge of the nose could not be shaped by available points).

The other problem that we observed during our further research while trying to match the geometry to the original images as precisely as possible is that in many cases points are useless compared to connected contour lines. Indeed, it has been already noted by some researchers [12] that feature points standardized by MPEG-4 do not provide enough information for a complete description of a head structure.

Finally, we had to improve the calibration pipeline so that it used well identifiable contour lines, such as, for example, outer contour of the lips, enhanced by a few feature points that were amenable to automatic labeling. Some of the feature contours that we use are shown in Figure 3.

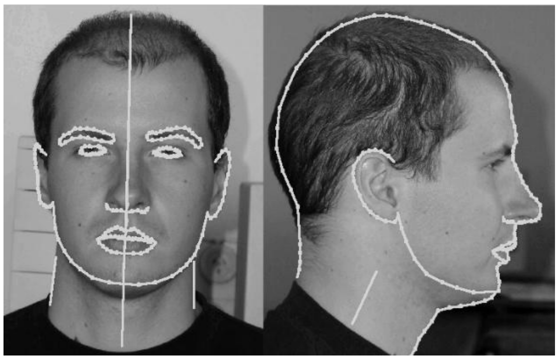


Figure 3. Example of facial feature contours.

More Images - Better Quality

The outcome of the first stage of our research was the calibration pipeline that takes 2 images (front and profile) as input data. Then it automatically generated the complete head model, including texture, from the feature points and contours as described above. Most of the features were selected manually at that time.

Analyzing models that we created we found out that the 2 proposed views, while being easy to acquire, still did not carry enough geometrical information with respect to the head shape. For example, the shape of the chicks could not be reliably estimated from available images. Therefore, we extended our calibration pipeline further by teaching it how to process images taken at arbitrary viewing angles. For such images we had to estimate camera parameters with respect to the head, which was done by considering correspondences of feature points for all views. Finally, we allowed silhouette lines be selected and improved calibration stage so that model geometry satisfied not only feature elements but also selected silhouette lines. Figure 4 demonstrates an example of selected silhouette lines on 45-degree images.

Figure 4. Example of selected silhouette lines.

Processing Video Stream

Processing of many images taken at arbitrary viewing angles required a large amount of feature elements and silhouette lines be selected. This process appeared to be much more complicated compared to selection of respective elements in front and profile views only. Thus, we decided to exploit spatial and temporal coherence of adjacent frames in a video stream, which was predicted to improve quality of selection process. Besides, considering video, we could use not only elements that can be selected for every head, but also individual points, such as moles or freckles tracking their flow on the video. By doing so the 3D position of these points could be estimated provided that camera parameters (position, orientation and focal length) were known. This procedure is conceptually similar to the technique described in [18]. An example of individual feature flow is shown in Figure 5.

Discussion and Future Work

In conclusion, starting from a task of creating personalized 3D head from two orthogonal views, i.e. front and profile, we have developed the calibration pipeline that creates ready for animation 3D model of a head from images taken at arbitrary views or a video stream.

Our first intention to use MPEG-4 facial feature points demonstrated its inability of defining personalized model of high quality at high resolution. Thus, we have enhanced the set of feature points, added feature contours and silhouette lines. We have also used feature flow tracking for estimating 3D positions of individual feature points.

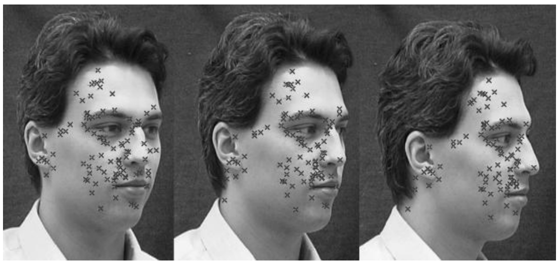


Figure 5. Example of individual feature flow.

Some examples of the calibrated heads are shown in Figure 6. It is worth noting here that the head shape is defined by selected points and contours only, which allowed us to build up a model of the well-known bust of Max Planck. In this case the required elements have been selected manually as selection algorithms greatly rely on texture properties. Going in this direction further, one of the possible extensions to the technique would be enabling 3D model construction for other subjects, for example, dogs or cats.

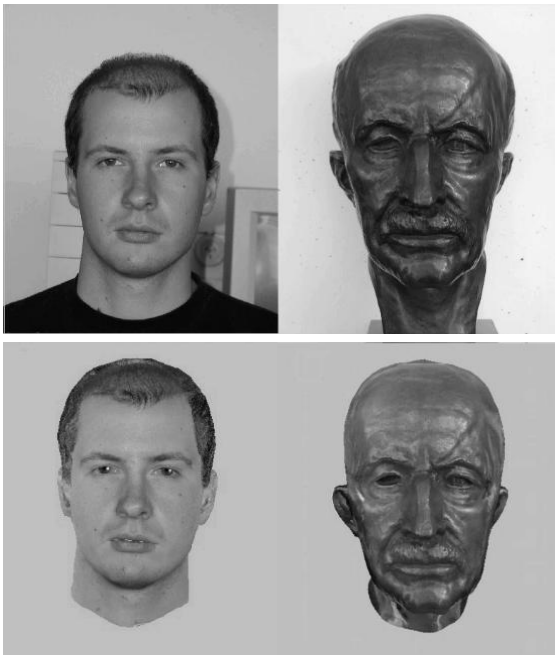


Figure 6. Examples of personalized 3D heads (bottom row).

Currently, all calibration procedures but feature elements selection are performed automatically. Thus, our present research activities are mainly focused on developing completely automatic head calibration environment that could be used by both professional and non-professional users.

Acknowledgements

This work was carried out by Computer Graphics Group at the Mathematics Department of MSU under a research agreement with Intel Nizhniy Novgorod Lab, which is a part of Intel Labs. We thank Valery Kuriakin (Intel, Corp.) for his constant interest in this work, and Tatiana Firsova, Elena Martinova, Victoria Zhislina and Konstantin Rodyushkin (Intel, Corp.) for their valuable comments.

References

1. L. Ambrosini, M. Costa, F. Lavagetto, and R. Pockaj. 3D Head Model Calibration Based on MPEG-4 Parameters. Proceedings of ISPACS'98, pp. 626-630, November 1998.
2. V. Blanz, and T. Vetter. A morphable model for the synthesis of 3D faces. Proceedings of SIGGRAPH'99, pp. 187-194, August 1999.
3. P. Burt, and E. Andelson. A Multiresolution Spline With Application to Image Mosaics. ACM Transactions on Graphics, 2(4), pp. 217-236, 1983.
4. M. Cohen, C. Jacobs, Z. Liu, and Z. Zhang. Rapid Modeling of Animated Faces From Video. Microsoft Technical Report MSR-TR-2000-11, 2000.
5. T. Firsova, V. Kuriakin, E. Martinova, O. Midlina, and V. Zhislina. MPEG-4 compliant 3D Face animation. Proceedings of Graphicon'2001, pp. 54-58, September 2001.
6. P. Fua. Regularized Bundle-Adjustment to Model Heads from Image Sequences without Calibration Data. International Journal of Computer Vision, 38(2), pp. 153-171, 2000.
7. T. Goto, S. Kshirsagar, and N. Magnenat-Thalmann. Automatic Face Cloning and Animation. IEEE Signal Processing Magazine, 18(3), pp. 17-25, May 2001.
8. S. Kshirsagar, S. Garchery, and N. Magnenat-Thalmann. Feature Point Based Mesh Deformation Applied to MPEG-4 Facial Animation. Proceedings of Deform'2000, pp. 23-34, November 2000.
9. V. Kuriakin et al. MPEG-4 Synthetic Video in real implementation. Proceedings of Graphicon'2001, pp. 203-207, September 2001.
10. F. Lavagetto, R. Pockaj, M. Costa. Smooth Surface Interpolation and Texture Adaptation for MPEG-4 Compliant Calibration of 3D Head Models. Image and Vision Computing Journal, 18(4), pp. 345-354, 2000.
11. W. Lee, and N. Magnenat-Thalmann. Fast Head Modeling for Animation. Image and Vision Computing Journal, 18(4), pp. 355-364, March 2000.
12. W. Lee, M. Escher, G. Sannier, and N. Magnenat-Thalmann. MPEG-4 Compatible Faces from Orthogonal Photos. Proceedings of CA'99, pp. 186-194, May 1999.
13. W. Lee, P. Kalra, and N. Magnenat-Thalmann. Model Based Face Reconstruction for Animation. Proceedings of MMM'97 (World Scientific Press), pp. 323-338, 1997.
14. Y. Lee, D. Terzopoulos, and K. Waters. Realistic Modeling for Facial Animation. Proceedings of SIGGRAPH'95, pp. 55-62, 1995.
15. V. Lempitsky, D. Ivanov and Ye. Kuzmin. Texturing Calibrated Head Model from Images. Short paper proceedings of EuroGraphics'2002, pp. 281-288, 2002.
16. S. Marschner, B. Guenter, and S. Raghupathy. Modeling and Rendering for Realistic Facial Animation. Proceedings of EGRW'2000, 2000.
17. F. Pighin, J. Hecker, D. Lischinski, R. Szeliski, D. Salesin. Synthesizing realistic facial expressions from photographs. Proceedings of SIGGRPAPH'98, pp. 75-84, 1998.
18. Z. Zhang et al. Robust and Rapid Generation of Animated Faces From Video Images: A Model-Based Modeling Approach. Technical Report MSR-TR-2001-101, Microsoft Research, 2001.

Virtually moving cameras for digital TV using IBR

Wolfgang Leister, Norsk Regnesentral
Henrik Olsen, Cosmo Scientific Consultant AS (csc.no)

We present an IBR method in order to see whether it is feasible to implement non-trivial graphical applications for use in a digital TV environment.

Television has been an information channel for the mass market for over fifty years. However, in recent years changes are emerging: (a) Analogue transmission technology will be phased out, while digital broadcast will take over. (b) Many viewers want to be more interactive, and new technology promises to allow that. For the broadcasters, the possible increase of bandwidth on the order of 6–8 times is most interesting.

The providers are less interested in introducing new possibilities. The reasons are high development costs for new products in a conservative market, where new technologies are only carefully introduced. Possible extensions to television are the subject of research, e.g., 3D-Television [1]. Today, proprietary systems and APIs within digital TV (e.g., MediaHighway, Beta) hinder interoperability, increase development costs for all platforms, and thus hinder a broader acceptance.

Digital TV services are meant to offer more than just viewing the broadcast content. Interactivity and applications (games, e-commerce, etc.) are an important issue for the future TV experience. With the venue of MHP (Media Home Platform) defined by DVB forum, and hardware that supports the new standard, some obstacles for the introduction of the new technology will disappear. In addition, the MHP standard has been defined for accessing additional content, navigating in the Electronic Programme Guide (EPG), accessing Internet content, and programming user interfaces. The MHP shall also be able to support more advanced applications like games.

In the following we present an application for digital TV that makes use of IBR methods in order to provide dynamic views from virtual moving cameras for preprocessed TV scenes. The viewer will be able to control the virtual camera using the remote control. This study is intended (1) as a feasibility study, to see whether hardware, software and technology within digital TV are ready for non-trivial applications; (2) to identify what kinds of applications are feasible; and (3) which potential extensions of APIs could be useful.

Digital TV

Digital TV in Europe is driven by DVB Forum (see `http://www.dvb.org`). The technology is mainly based on the definition of a set top box (STB), the MPEG standards (mostly MPEG-2 – ISO standard 13818), and the Media Home Platform (MHP) [2]. In MHP, standard applications are written as so-called "Xlets" in DVB-Java.

The STB with MHP implements a Java Virtual Machine (JVM). The Xlet is compiled to byte code, downloaded via the MPEG-TS, and installed by the class loader. The Xlet can access resources in the DVB stream.

An Xlet is like an applet, which is adjusted to the special needs for digital TV with regard to class loading from the DVB stream. The applications can access additional libraries for accessing DVB streams, providing user interfaces, and other functionalities defined by HAVi, DAVIC, and DVB. More on digital TV can be found in a separate report [3].

Rendering

Here, we suggest an application where the viewer can interactively walk around in the scene, which she can control via a remote control. The application is implemented using a view interpolation method as described by Chen and Williams [4]; see also [5].[2]

In our view interpolation method we use depth maps in order to create coloured reliefs in space, which can be viewed from other positions. We have also experimented with disparity maps. These depth maps could be provided

[2]The purpose of this work is not to present or implement a new IBR method. The IBR method used is well-known and too simple for a real-world application. Therefore, the resulting images are of lower quality and contain artifacts, as shown in [5]. The intention of this work is to see whether the use of an IBR method is feasible in a digital TV context.

by e.g., a camera with special depth sensors; such products are available, see e.g. 3DV (`http://www.3dsystems.com`). For our experiments we use the output of a raytracer.

After the content has arrived on the STB, the image for another viewpoint is rendered from one or several sources and presented to the viewer. Note that the rendering process has to be done for each frame, thus the rendering algorithm limits the achievable frame rate.

Transmission

The input of the view interpolation method is a raster image, a depth image, and information about the camera parameters; this inormation together with the images must be transmitted in the DVB stream. Data are encoded into the MPEG-TS, which can be done in several ways. For images the MPEG standard offers compression using the DCT method; see [6].

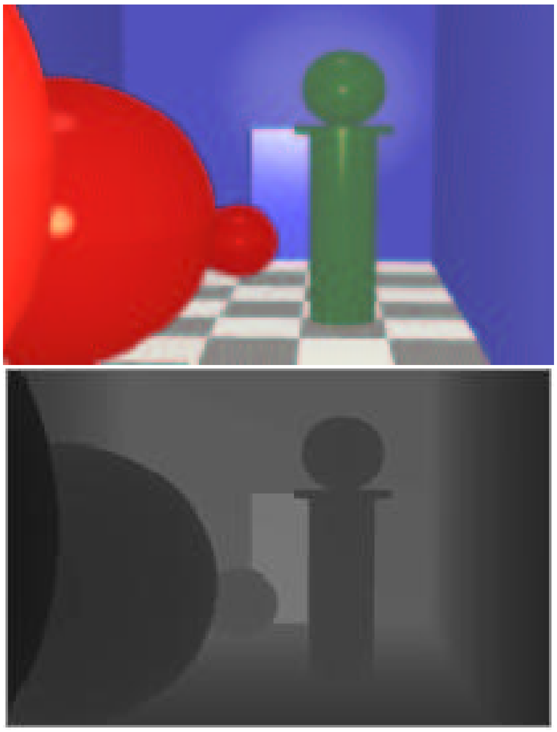


Figure 1: The original colour image and depth map

Tests

Image size on TV is normally 780×576 pixels for the PAL standard. Our test images were of the size 200 × 130 in order to get a suitable frame rate on the PC for our study. In our experiments, we could achieve a frame rate of about 12 frames per second on a PC (ca. 800 MHz), while the Nokia Mediaterminal only reached 1 frame per second. This rate is far too low. However, the development of faster processors makes us believe that we could reach the necessary frame rate in the near future.

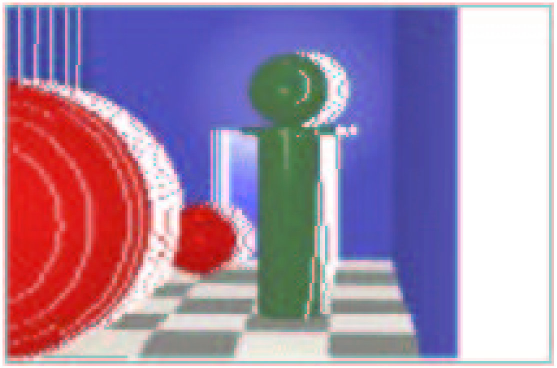


Figure 2: The reconstructed image using a simple IBR method

We are aware that the choice of the programming language Java might not be optimal with respect to processing time. However, the digital TV platform makes it necessary to use DVB Java, and one intention of this report is to make conclusions about whether it is possible to use DVB Java for non-trivial functionality (i.e. more than handling menus on the screen).

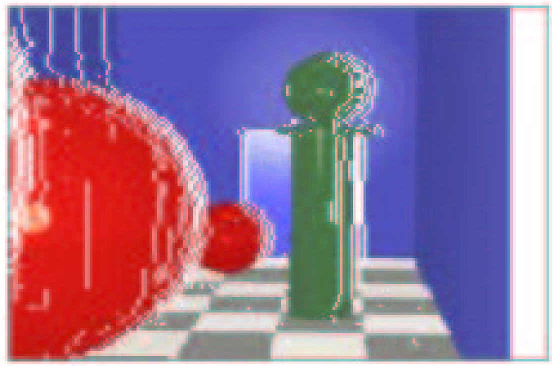


Figure 3: The reconstructed image using compressed depth map.

In our example, the image in Figure 2 is reconstructed from the image and depth map in Figure 1. The white parts in the reconstructed image are areas where no information about depth or colour is available, as they are occluded by some objects. Some dots and lines in the reconstructed image are caused by the algorithm, when the original view is "stretched". These problems can be eliminated as outlined by Chen and Williams [5]. However, this is beyond the scope of this presentation.

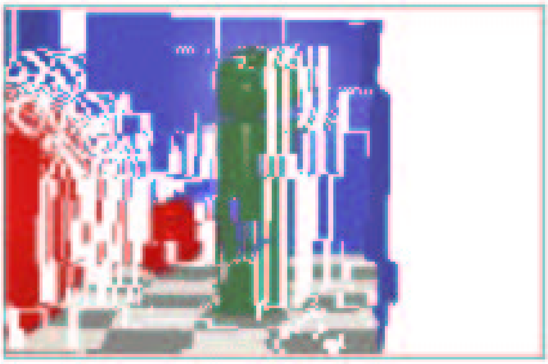


Figure 4: The reconstructed image using compressed disparity map (right)

For transmission the MPEG standard defines the DCT method as a compression method, in order to increase bandwidth. The DCT method is lossy, and therefore has a visible impact on the image quality. The compression of the raster image reduces colour quality in the same manner as MPEG or JPEG does.

Due to the compression of the depth map, the geometry of the relief gets distorted; therefore the reconstructed image is visibly reduced in quality. However the problems are moderate, as shown in Figure 3. The impact of compression methods in a disparity map is much larger, which means that the disparity data are much more sensitive to information loss (see Figure 4). Therefore, our conclusion is that the use of depth maps is preferable over disparity maps, since quality and execution time are improved.

More suitable library functions of the MHP libraries could accelerate the rendering, but MHP is not optimised for 3D graphics at all.

References

[1] C. Fehn, P. Kauff, M. Op de Beeck, F. Ernst, W. IJsselsteijn, M. Pollefeys, L. Van Gool, E. Ofek, and I. Sexton. An Evolutionary and Optimised Approach on 3D-TV. In *Proc. IBC 2002, Amsterdam.* available on `http://bs.hhi.de/~fehn/publications/IBC02.pdf`, 2002.

[2] DVB. Multimedia Home Platform 1.0.1. DVB Blue Book A057 Rev.1, 2001.

[3] W. Leister, S. Boudko, O. Aamot, and P. Holmes. Digital TV — an overview. report 988, Norsk Regnesentral, Oslo, 2002.

[4] Shenchang Eric Chen and Lance Williams. View interpolation for image synthesis. *Computer Graphics*, 27(Annual Conference Series):279–288, 1993.

[5] A. Watt and F. Policarpo. *The Computer Image.* Addison-Wesley, 1998. chapter 21.

[6] J. Mitchell, W. Pennebaker, C. Fogg, and D. LeGall. *MPEG Video Compression Standard.* Chapman Hall, New York, 1997.

I kø for billedanalyse

Line Eikvil, Norsk Regnesentral

Det er ikke ofte vi forskere i bildeanalyse får oppleve at almuen står i kø for våre tjenester, men slik var det på Forskningstorget under fjorårets Forskningsdager. Da var det lang kø for å prøve Norsk Regnesentrals demo basert på bildeanalyse som skulle kunne fortelle folk hvilken kjendis de lignet på.[3]

Forskningsdagene er en landsomfattende festival der folk flest får møte forskere og forskning på en ny måte. Ett av de største arrangementene under fjorårets festival var Forskningstorget på Universitetsplassen i Oslo 20. og 21. september. Her var det mange forskjellige forskningsinstitusjoner som var på plass, og publikum kunne oppleve forskning i litt mer popularisert innpakning enn vanlig og for eksempel få innsikt i hvordan det er å bo i en glassboble, få vite hvorfor dinosauren plutselig døde ut, trille egne piller eller ta en titt på stjernehimmelen.

Blant mange andre forskningsinstitusjoner var Norsk Regnesentral også på plass, og i vår

[3]Artikkelen er tatt fra PIKSEL'N, årgang 20, nr. 1, mars 2003 og gjengitt med tillatelse fra NOBIM (Norsk forening for bildebehandling og mønstergjenkjenning).

bod var temaet "Hvilken kjendis ligner du på?". Her tok vi bilder av dem som kom innom med et kamera koblet til en PC. Ved hjelp av teknikker som vanligvis brukes til ansiktsgjenkjenning, ble bildene sammenlignet med bilder av kjendiser fra en database. Dermed kunne man få svar på hvilken av disse kjendisene man lignet mest. Det var ingen tvil om at demoen slo godt an, spesielt blant de unge. I løpet av de to dagene arrangementet varte, hadde vi mellom 600 og 700 personer innom for å finne ut hvem de lignet på. Mange fikk seg en overraskelse da de fikk vite at de lignet kjendiser som Mette-Marit, Tommy Lee Jones og Kurt Russell. Statsråd Kristin Clemet var også innom, men hvem hun lignet på, vil vi ikke røpe her.

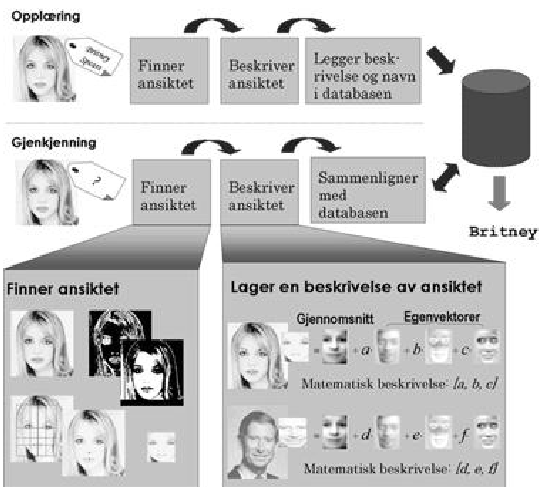
Bilde 1: Det ble etterhvert en lang kø av folk som ønsket å finne ut hvilken kjendis de lignet på. Foto: Norsk Regnesentral.

Metodikken bak det hele baserer seg altså på teknikker som benyttes til ansiktsgjenkjenning. Her fins det flere ulike metoder; både globale som bruker ansiktet som en helhet og lokale som bruker noen regioner som øyne, nese og munn. Metoden som ble brukt i demoen, er en global metode som er basert på såkalte eigenfaces.

Fra et portrettbilde skiller programmet automatisk ut ansiktet fra bakgrunnen og lokaliserer øyne, ansikt og munn. Deretter gjøres det en normalisering med hensyn på størrelse og lys, slik at en sitter igjen med et normalisert bildeutsnitt av bare ansiktet. Utsnittet av ansiktene finnes for alle som skal ligge i databasen. Deretter lages det en matematisk beskrivelse av hvert ansikt som beregnes ved at det først gjøres en prinsipal komponentanalyse på hele settet av ansikter i databasen for å finne egenvektorene og egenverdiene. De M egenvektorene med høyest egenverdi plukkes ut som eigenfaces. Disse beskriver hovedkomponenten for ansiktene, og alle ansiktene i databasen vil kunne tilnærmes ved en lineær kombinasjon av gjennomsnittsansiktet og disse M eigenfaces. Hvert ansikt kan dermed representeres ved et sett av vekter som sier hva en må multiplisere hvert eigenface med for å få ansiktet.

Bilde 2: Kristin Clemet var også innom for å finne ut hvem hun lignet på. Resultatet er en hemmelighet, men hun ser i hvert fall fornøyd ut. Foto: Jon Solberg, Forskningsdagene.

Når et nytt ansikt skal sammenlignes med dem som ligger i databasen, beregnes disse vektene for det nye ansiktet. Vektene finner man ved å projisere bildet ned på settet av eigenfaces. Deretter finner systemet den som ligner mest ved å finne det ansiktet hvis vekter ligger nærmest det som er beregnet for det nye bildet.



Bilde 3: Metoden bak ansiktsgjenkjenning forklart grafisk.

I databasen ligger det bilder av ca. 100 kjendiser som man blir sammenlignet med. Men det er jo slett ikke sikkert at man egentlig ligner på noen av dem. Algoritmen vil likevel plukke ut den man ligner mest på, selv om denne likheten er svært liten. Derfor kan det nok hende at enkelte ikke syns de lignet så mye. Når likheten er så liten, vil også lysforholdene kunne ha stor innvirkning på resultatet. Derfor kan det også hende at et litt annet bilde av samme person, kan gi et annet resultat. Men for mange var likheten absolutt tilstede, selv om de kanskje ikke lignet den de helst ville.

For dem som har lyst til å prøve seg, fins det en nettversjon av programmet: `ttp://www.nr.no/documents/samba/research_areas/BAMG/Demos/NorMatch.html`.

Lykke til, hvis du tør!

Inntrykk fra seminar i Grafisk Databehandling på Fornebu

Wolfgang Leister, Norsk Regnesentral

Institutt for informatikk, Universitetet i Oslo, Simula Research Laboratory AS i samarbeid med Norsigd inviterte til seminar i grafisk databehandling 3.–4. mars 2003 hos Simula Research Laboratory AS på Fornebu.

Over femti deltagere hadde funnet veien til Fornebu, der seminaret ble avholdt i den tidligere flyplassbygningen som nå huser Simula Research Laboratory. 16 bidrag ble presentert på halvannen dager. Det var mange forskjellige temaer, applikasjonsområder og forskningsmiljøer representert. Dessuten fikk de fremmøtte oppleve et invitert foredrag av Richard Riesenfeld ved Universitetet i Utah om utfordringer innen datagrafikk.

Temaene var veldig mangfoldige, både praktiske og teoretiske emner var med. Vi har planer om å gjengi flere av presentasjonene og temaene på seminaret i kommende utgaver av NORSIGD Info. Under seminaret ble det bare delt ut en kort oversikt over hvert foredrag, så en langversjon i NORSIGD Info kan være et supplement til seminaret.

Organisatorene med Morten Dæhlen og Tom Lyche i spissen gjorde en formidabel jobb med å organisere seminaret. Seminaret ble organisert i samarbeid med NORSIGD, og derfor får vår formann Ketil Aamnes æren med å presentere NORSIGDs arbeid under åpningen.

På et seminar som dette er middagen et viktig element, der man fikk anledning i å prate med hverandre utenom det annonserte programmet, og knytte kontakter. Middagen ble avholdt i Radisson SAS Park Hotel på Fornebu, og den inkluderte stand-up komikk ved Morten Dæhlen ;-)

I løpet av seminaret fikk vi en fin oversikt over hva som foregår innen datagrafikk i Norge. NORSIGD takker organisatorene for to begivenhetsrike dager, og vi håper at vi alle kan bygge videre på seminarets suksess.

Bildene på denne siden gir noen visuelle inntrykk fra seminaret.

Øverst til venstre ser vi Richard Riesenfeld under foredraget sitt. De påfølgende bildene viser seminardeltagerne under middagen. Nederst til høyre ser vi organisasjonskommitéen med Morten Dæhlen og Tom Lyche (som har tatt de fleste av bildene).

Hva er NORSIGD?

NORSIGD – Norsk samarbeid innen grafisk databehandling – ble stiftet 10. januar 1974. NORSIGD er en ikke-kommersiell forening med formål *å fremme bruken av, øke interessen for, og øke kunnskapen om grafisk databehandling i Norge.*

Foreningen er åpen for alle enkeltpersoner, bedrifter og institusjoner som har interesse for grafisk datbehandling. NORSIGD har per januar 2003 27 institusjons-, 37 personlige og 5 EG-medlemmer. Medlemskontingenten er 1.000 kr per år for institusjoner. Institusjonsmedlemmene er stemmeberettiget på foreningens årsmøte, og kan derigjennom påvirke bruken av foreningens midler.

Personlig medlemskap koster 250 kr per år. Personlige medlemmer får tilsendt medlemsbladet *NORSIGD Info*. Kontingenten er redusert til 150 kr ved samtidig medlemskap i vår europeiske samarbeidsorganisasjon *Eurographics*.

Alle medlemmer får tilsendt medlemsbladet *NORSIGD Info* 2–4 ganger per år. NORSIGD har tilrettelagt informasjon om foreningen på Internett på adressen `http://www.norsigd.no`. Der finnes det også informasjon om GPGS, samt tidligere utgaver av *NORSIGD Info*.

Interesseområder

NORSIGD er et forum for alle som er opptatt av grafiske brukergrensesnitt og grafisk presentasjon, uavhengig av om basisen er *The X window System*, *Microsoft Windows* eller andre systemer. NORSIGD arrangerer møter og seminarer, formidler informasjon fra internasjonale fora og distribuerer fritt tilgjengelig programvare. I tillegg formidles kontakt mellom brukere og kommersielle programvareleverandører.

NORSIGD har lang tradisjon for å støtte opp om bruk av datagrafikk. Foreningen bidrar til spredning av informasjon ved å arrangere møter, seminarer og kurs for brukere og systemutviklere.

GPGS

GPGS er en 2D- og 3D grafisk subrutinepakke. GPGS er maskin- og utstyrsuavhengig. Det vil si at et program utviklet for et operativsystem med f.eks. bruk av plotter, kan flyttes til en annen maskin hvor plotteren er erstattet av en grafisk skjerm uten endringer i de grafiske rutinekallene. Det er definert grensesnitt for bruk av GPGS fra FORTRAN og C.

Det finnes versjoner av GPGS for en rekke forskjellige maskinplattformer, fra stormaskiner til Unix arbeidsstasjoner og PC. GPGS har drivere for over femti forskjellige typer utsyr (plottere, skjermer o.l.). G-PGS støtter mange grafikkstandarder slik som Postscript, HPGL/2 og CGM. GPGS er fortsatt under utvikling og støtter stadig nye standarder.

GPGS eies av NORSIGD, og leies ut til foreningens medlemmer.

Eurographics

NORSIGD samarbeider med Eurographics. Personlige medlemmer i NORSIGD får 20 SFr rabatt på medlemskap i Eurographics, og vi formidler informasjon om aktuelle aktiviteter og arrangementer som avholdes i Eurographics-regi. Tilsvarende får Eurographics medlemmer kr 100 i rabatt på medlemskap i NORSIGD.

Eurographics ble grunnlagt i 1981 og har medlemmer over hele verden. Organisasjonen utgir et av verdens fremste fagtidsskrifter innen grafisk databehandling, *Computer Graphics Forum*. *Forum* sendes medlemmene annen hver måned. Eurographics konferansen arrangeres årlig med seminarer, utstilling, kurs og arbeidgrupper.

NORSIGD
v/ Reidar Rekdal
Postboks 290
1301 Sandvika

Returadresse:
NORSIGD v/ Reidar Rekdal
Postboks 290
1301 Sandvika

Styret i NORSIGD 2002

Funksjon	Adresse	Telefon	email
Leder	Ketil Aamnes Telenor Business Solutions PB 16517 Etterstad 0606 OSLO	22 78 05 56 (direkte) 23 16 64 70 (fax)	Ketil.Aamnes @telenor.com
Fagansvarlig	Wolfgang Leister Norsk Regnesentral Postboks 114 Blindern 0314 OSLO	22 85 25 78 (direkte) 22 85 25 00 (sentralbord) 22 69 76 60 (fax)	leister@online.no
Sekretær	Reidar Rekdal Norsigd Postboks 290 1301 Sandvika	67 57 73 18 (direkte) 67 57 72 50 (sentralbord) 67 57 72 72 (fax)	reidar.rekdal @dnv.com
Styremedlem	Gisle Fiksdal LODIC AS Pirsenteret 7462 TRONDHEIM	73 83 17 34 (direkte) 73 83 17 35 (fax)	Gisle.Fiksdal @lodic.no
Varamedlem	Svein Taksdal Norges Vassdrags- og Energiselskap Hydrologisk Avdeling, Seksjon data Postboks 5091, Majorstua 0301 OSLO	22 95 92 86 (direkte) 22 95 92 01 (fax)	svein.taksdal @nve.no
Varamedlem	Magnar Granhaug ProxyCom AS Klæbuvn. 194 7037 Trondheim	73 95 25 00 97 72 76 98 (mobil) 73 95 25 09 (fax)	Magnar.Granhaug @proxycom.no

Svarkupong

- ◯ Innmelding – institusjonsmedlem (Kr 1000)
- ◯ Innmelding – personlig medlem (Kr 250)
- ◯ Innmelding – Eurographics medlem (Kr 150)
- ◯ Ny kontaktperson
- ◯ Adresseforandring

Navn: ....
Firma: ....
Gateadresse: ....
....
Postadresse: ....
....
Postnummer/sted: ....
....
Telefon: ....
Telefaks: ....
email: ....