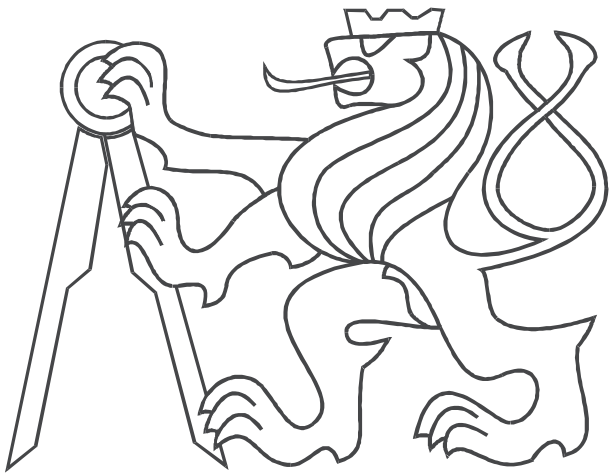
Czech Technical University in Prague



Doctoral thesis statement

Czech Technical University in Prague

Faculty of Electrical Engineering

Department of Telecommunication Engineering

Ing. Martin Nemčík

3D Model-Based Video Coding by Using Quadratic Triangular DMS-Surface

Ph.D. Programme: Electrical Engineering and Information Technology
Branch of study: Telecommunication Engineering

Doctoral thesis statement for obtaining the academic title of "Doctor", reviated to "Ph.D."

Prague, September 2011

The doctoral thesis was produced in full-time manner Ph.D. study at the Department of Telecommunication Engineering of the Faculty of Electrical Engineering of the CTU in Prague.

Candidate: Ing. Martin Nemčík
Department of Telecommunication Engineering
Technická 2
166 27 Prague 6 - Dejvice

Supervisor: Prof. Ing. Pavel Zahradník, CSc.
Department of Telecommunication Engineering
Technická 2
166 27 Prague 6 - Dejvice

Opponents: .....

.....

.....

The doctoral thesis statement was distributed on .....

The defence of the doctoral thesis will be held on ..... at ..... before the Board for the Defence of the Doctoral Thesis in the branch of study Telecommunication Engineering in the meeting room No. ..... of the Faculty of Electrical Engineering of the CTU in Prague.

Those interested may get acquainted with the doctoral thesis concerned at the Dean Office of the Faculty of Electrical Engineering of the CTU in Prague, at the Department for Science and Research, Technická 2, Prague 6, 166 27.

Chairman of the Board for the Defence of
the Doctoral Thesis in the branch of study
Telecommunication Engineering
Faculty of Electrical Engineering
of the CTU in Prague

1 Current situation of the studied problem

A better adaptation of multimedia systems to human visual perception must take into account all of these areas of research. The second most important task of multimedia systems is the effective transmission and storage of audiovisual information. Video compression enables more efficient use of transmission and storage resources. Even with constant advances in the storage and transmission capacity, compression is likely to be an essential component of multimedia services for many years to come. Signal compression can be described as removing signal redundancy and irrelevancy, because need only relevant and unique parts of the video signal are needed [1].

Model-based video coding

Almost all video compression algorithms operate by removing the redundancy in the temporal, spatial and/or frequency domains. Greater efficiency may be achieved by modelling the video's contents. That means the relevant objects in the video (i.e. the human head) are first identified and then analysed. General or specific models are used for this analysis. The most commonly used specific model is the 3D mesh model of the human head. The results of analysis using this model are parameters. These parameters are transmitted to the decoder and using the same model as the coder the output video is synthesized [1], [2].

Video telephony video scene

The typical video telephony scene [3] is a subset of real world videos with specific properties. This scene typically consists of a talking human head (bust) and background, which is not very significant. Utilization of this fact and empirical knowledge of human facial muscles allows for a significant reduction in the bit stream for transmission. Video of very good quality is possible even at 64 kbps [1].

MPEG-4 standard

The MPEG-4 standard was developed in response to the growing need for a coding method that can facilitate the access to visual objects in the natural and synthetic video and sound for various forms of wired or wireless communication [4]. ISO/IEC JTC1/SC29/ WG11 (MPEG) [5], [6] is currently working on this standardization including model-base video coding, compression of geometry, synchronization between Aural and Visual (A/V) objects, multiplexing of streamed A/V objects, and spatial-temporal integration of mixed media types.

DMS-surface

In computer graphics, a crucial aspect is to create the 3D scene using the smallest number of polygons, while maintaining a realistic scene. A small number of polygons is necessary because it allows for faster transmission of the given scene, more

precisely with low bit rate. However, the objects represented by a small number of polygons seem to be discontinuous (unrealistic). This disadvantage of discontinuity can be well-compensated for by the approximation of polygonal surfaces (smoothing) by employing various approximation methods. One of these methods is the DMS-spline function. This thesis describes this function and additional way of utilizing it in model-based video coding.

2 Aims of the doctoral thesis

1. Develop a new method of automatic generation of knotclouds and side control-points over arbitrary triangulation for DMS-spline surface approximation.

2. Apply computing of DMS-surface of human head model CANDIDE 3-1-6 and enable its effective animation.

3. Create an integrated program environment which allows for computation of the DMS surface over an arbitrary triangulation, an approximation of a surface area of 3D human head model CANDIDE 3-1-6, an interactive projection of a triangulation and an approximated surface, animation using animation units defined in codec MPEG-4 as well as an application of the texture of a human head on an approximated surface.

4. Define the possibilities of DMS approximation surface utilization in 3D model-based video coding and to point out potential fruitful and innovative directions of research and development in this area.

3 Working methods

In order to complete the given aims of the dissertation work I invented, after an in-depth theoretical research, a complex development environment in Borland© C++. The newly invented development environment (*projectHEAD*©) (Fig. 1) was set up to enable approximation of human head surface while using DMS-spline function and its animation. More precisely, the program can be used to compute placement of knot-clouds and side control points over the whole triangulation, approximated surface and its animation. The program allows to set up specific parameters and therefore to adjust placement of knotclouds and side control points which as a result influences the final surface. After computing of the final surface the result of this approximation is displayed in open graphics library OpenGL©. This new development environment

helped to verify the originally suggested algorithms for knotclouds and side control points placement.

While working on the solutions of the aims of the dissertation work, I invented and implemented in a program an environment for animation of approximated surface including implementation of real texture. Optimization of computation steps allows more effective computing of animation of approximated surface. The program also allows to manually create and save the final animation by using animation units AU or FAP defined in environment MPEG-4. In addition it can be used to open this animation from a file.
These particular steps present the whole synthesis part of model-oriented video coding and create an environment for its connection with the analysis part.

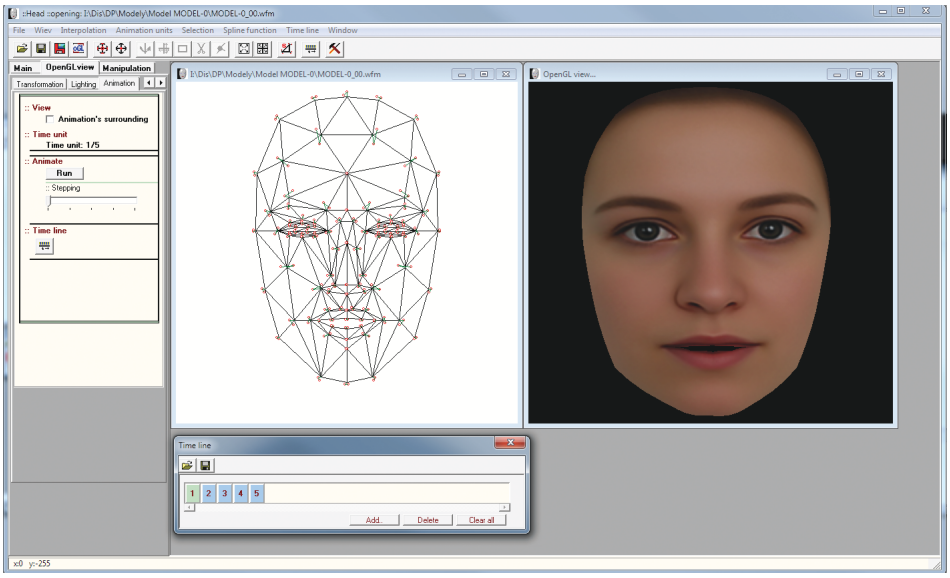


Figure 1: Program projectHEAD© screenshot.

4 Result

Influence of Knotclouds Placement to Approximation Surface of Human Head

Knotclouds influence the approximated surface in particular triangulation triangles' supports intersection, where vertex having these knotclouds assigned is one of vertices. The knotclouds have the greatest influence in the area of the vertex to which they are assigned. After meeting conditions of knotclouds placement the approximated surface can be formed. Influence of knotclouds placement of approximated

surface is exemplified on displacement of knots t_{i1} and t_{i2} of vertex t_{i0} (Fig. 2) in the left eyebrow area of CANDIDE 3-1-6 model.

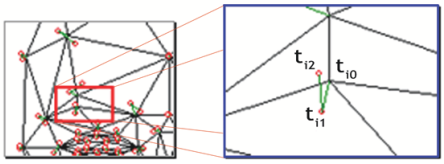


Figure 2: Triangulation with knotclouds placement. A: initial position of knots; B: Detail A - placement of knots t_{i1} and t_{i2} of vertex.

Figure (Fig. 3) shows knotclouds placement generated by the method which I propose in thesis.

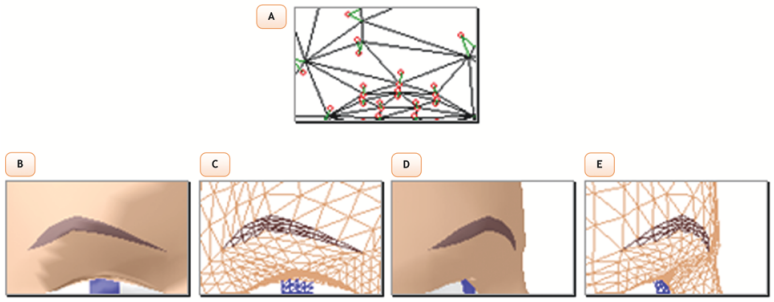


Figure 3: Initial position of knotclouds and approximated surface: A: left eyebrow area; B, C: approximated surface - front view filled (unfilled) triangulation after second interpolation and first DMS approximation with lighting and with averaging of normal vectors of particularly triangles. D, E: Approximation surface - left view filled (unfilled) triangulation after second interpolation and first DMS approximation with lighting and with averaging of normal vectors of particularly triangles.

Influence of side control point placement

Influence of side control point placement on the whole human head surface (CANDIDE 3-1-6) is demonstrated in figure (Fig. 4). Approximated surface with side control points generated by algorithm which I propose in thesis is gradually enhanced via displacement of side control point and subsequently via displacement of vertex control points.

Next (Fig. 5 is illustrated affect of vertices displacement by animation units *AUV 2 Lip stretcher (AU20)* to approximated surface:

In the figures (Fig. 5 - 3^{rd} column) animation effects of *AUV2 Lip stretcher* to the approximated surface is color-illustrated. The blue color means the least impact

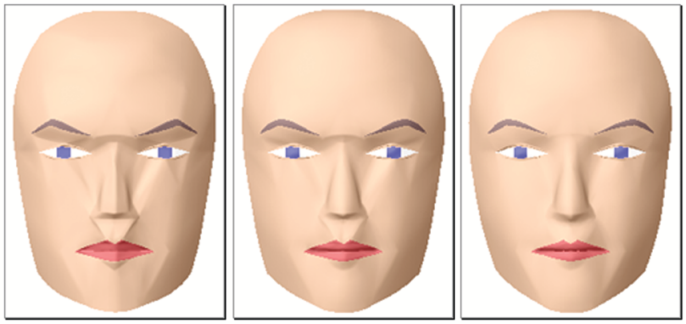


Figure 4: Influence of side control points placement on the whole head model. (CANDIDE 3-1-6 - front view)

and gradually over the green to red is in effect at the approximated surface is greater (cream color means that in given points the animation unit has not impact on the approximated surface). (Note: Difference is calculated as Euclidean distance for each image separately.)

Following figure (Fig. 6) shows process of sufrace approximation and texture mapping.

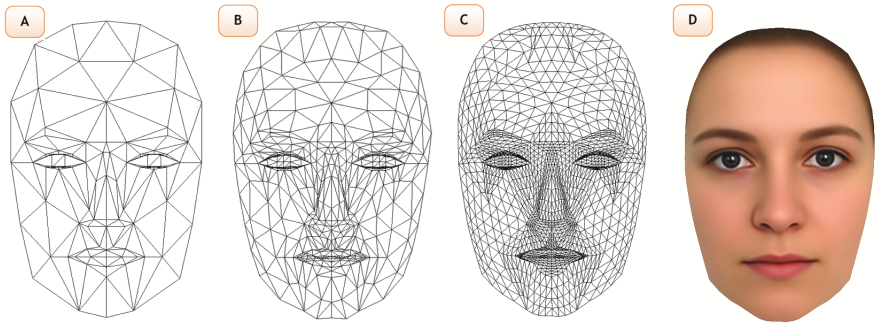


Figure 6: Process of approximation and texture mapping (front view). A: Basic triangulation; B: Control points mesh for quadratic DMS-surface; C: Approximated triangulation; D: Real human head texture mapped on a 3D model.

Animation of DMS-Surface

Animation of approximate a surface is based on basic triangulation vertex motion transfer. During this animation, control point network is modified (i.e. motion of vertex and side control points). The following figure (Fig. 7) show an example of animation of approximation human head model by animation units (AU) defined in MPEG-4 AVC environment.

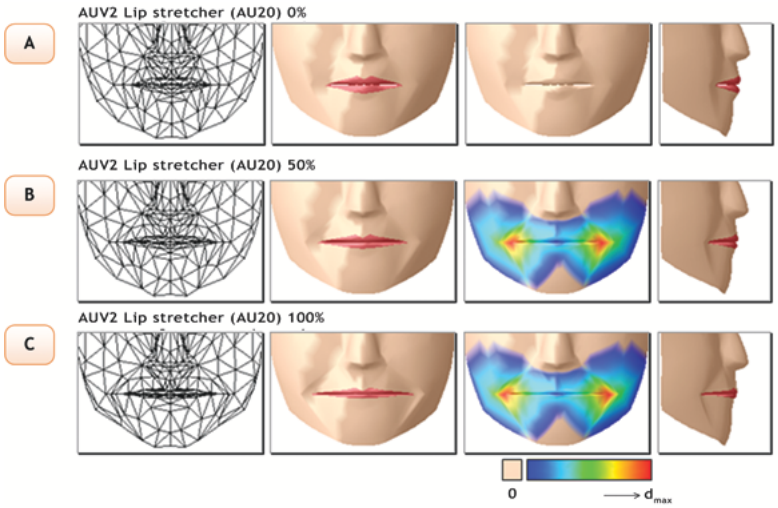


Figure 5: Influence of animation by *AUV 2 Lip stretcher (AU20)* to approximated surface. From left: Triangulation after approximation; Approximated surface with simple texture (front view); Range of influence (front view); Approximated surface with simple texture (left view).

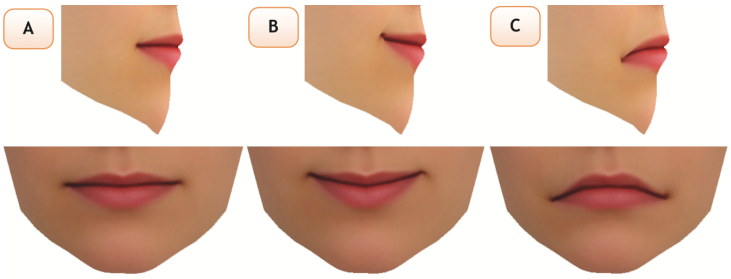


Figure 7: Example of animation of approximation human head model (CANDIDDE 3-1-6) with real texture. (Mouth detail)

Enhanced Model-based Video Coding

The model-based video coding extension, for of the public 3D head model database offers new possibilities. Figure (Fig. 8) illustrates interconnecting coder and decoder to a database. Although this extension of model-based video coding increases bit rate, it brings many useful multimedia features.

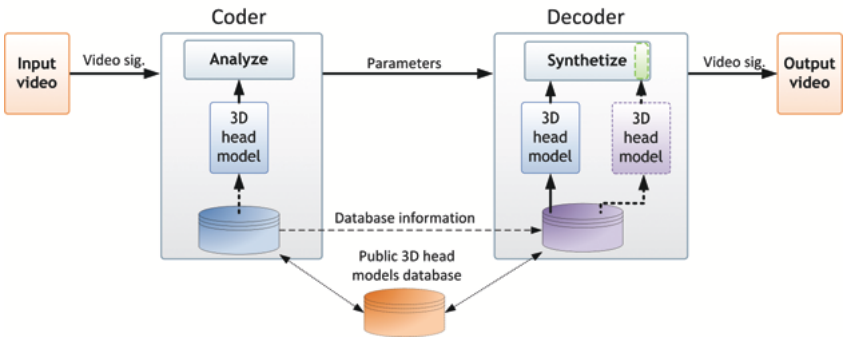


Figure 8: The model-based video coding with the customers' database.

5 Main Results

Thesis present usage of quadratic triangular DMS-spline Surface by enhancing 3D model-based video coding algorithm.

5.1 Accomplishment of Objectives of Dissertation Thesis

This section describes the ways in which the objectives of dissertation thesis were accomplished.

1. **Goal**
   To develop a new method of automatic generation of knotclouds and side control points over arbitrary triangulation for DMS-spline surface approximation.

   Achievement
   Automatic generation of knotclouds and side control points over arbitrary triangulation is a necessary condition of an autonomic system which can approximate model surface based on input animation parameters. I developed an algorithm which complies with this method and which I described in thesis. I used this algorithm for animation of CANDIDE 3-1-6 and/or M-4A© human head models.

2. **Goal**
   To apply computing of DMS-surface of CANDIDE 3-1-6 human head model and invent its effective animation.
   Achievement
   Choice of an effective approach to computing of DMS-surface for animation plays an important role given the nowadays efficiency of videophone terminals. I simplified the complexity of computing by:

   - using a symmetric model,
   - selecting knotclouds and side control points which are necessary to be recomputed for each specific animation part.

   The above-mentioned approaches are described in thesis and used in the *projectHEAD*© application.

3. **Goal**
   To create an integrated program environment which would allow computing of

the DMS surface over an arbitrary triangulation, an approximation of a surface of CANDIDE 3-1-6 3D human head model, an interactive projection of a triangulation and an approximated surface, an animation by using animation units defined in MPEG-4 codec as well as an application of a texture of a human head on an approximated surface.

Achievement

Together with the development of the algorithm for the automatic knotclouds and side control points placement, I implemented an algorithm for computing a DMS-surface in the Borland C++ integrated development environment. The *projectHEAD*© application which I developed, is described in appendix B and allows import of a triangulation (including CANDIDE 3-1-6 and M-4A© model), generation of knotcloud and side control points and is able to compute DMS-surface over this triangulation. This application allows animation of a human head model based on a FAPs sets defined in MPEG-4 environment. The application uses OpenGL© library [7], [8] for 3D rendering of a final scene. The *projectHEAD*© program presents a real life verification - of the theoretical assumptions properly described in this thesis. More precisely, these theoretical assumptions include especially and most importantly methods of an automatic generation of knotclouds and side control points placement.

4. **Goal**

   To define possibilities of DMS-spline approximation surface utilization in 3D model-based video coding and to point out potential promising and innovative directions of research and development in this area.

   Achievement

   Combination of global smoothness of linear splines and possibilities of local control of DMS-spline function's surface, seems to have good application in 3D model-based video coding. The possibility of local animation of human head surface model (CANDIDE 3-1-6 and/or M-4A©) using FAP units (defined in MPEG-4) allows to apply DMS-spline approximation in model-based video coding. Furthermore, using DMS-spline functions improves the quality (more precisely approximation of real human head) and after enhancing local animation brings in addition flexibility and output video scene's increased computing speed.

As follows from the above, all tasks assigned to this dissertation have been met.

6 List of literature used in the thesis statement

[1] Mihalík J. *Kódovanie obrazu vo videokomunikáciach.* ISBN: 80–89061–47–8. Mercury – Smekl, Košice, 1st edition, 2001. In Slovak.

[2] Al mualla M. E., Canagajarah C. N., and Bull D. R. *Video Coding for Mobile Communications.* ISBN: 0-12-053079-1. Academic press, China, Beijing, 2002.

[3] Kauff P., Schreer O., and Sikora T. *3D Videocommunication: Algorithms, concepts and real-time systems in human centred communication.* ISBN:-13 978-0-470-02271-9. John Wiley & Sons, Ltd, 1st edition, 2005.

[4] Won-Sook L., Escher M., and Sannier G. *MPEG-4 compatible faces from orthogonal photos.* ISBN: 0-7695-0167-2. MIRA Lab., Geneva Univ., Geneva , Switzerland, 1999.

[5] Koenen R. (Editor). *MPEG-4 System Goup, WG11(MPEG), Coding of moving picture and audio.* ISO/IEC JTC1/SC29/WG11 N3536, China, Beijing, 2000.

[6] Pereira F. *MPEG-4 System Group, MPEG-4 Requirements, version 14 (Beijing version), Coding of moving picture and audio.* ISO/IEC JTC1/SC29/WG11 N3534, China, Beijing, 2000.

[7] OpenGL tutorials, 2011.
`http://nehe.gamedev.net/`, 12. 01. 2011.

[8] Shreiner D., Neider J. Woo M., and Davis T. *OpenGl - Průvodce programátora.* ISBN: 80-251-1275-6. Computer Press, Praha, CZ, 2006. In Czech.

7 List of Author's Publications (Related to the Dissertation)

Impact Journal

Reviewed Journal

[1] Nemčík, M.: *Model-based Video Compression in the Videophone Calls and Benefit of 3D Bust Models Database of Videoconference Customers.* In *Digital Technologies 2006 - 3rd International Workshop* [CD-ROM]. Žilina: University of Žilina, Faculty of electrical engineering, 2006, vol. 1, ISBN 80-8070-637-9.

[2] Nemčík, M.: *Creation of 3D Head Videoconferencing Customers Database and its Benefits in Model-Based Video Coding.* In *RTT 2007* [CD-ROM]. Žilina: University of Žilina, 2007, p. 277-280. ISBN 978-80-8070-735-4.

[3] Nemčík, M.: *Model-Based Video Compression.* In *Proceedings CD-ROM of Digital Technologies 2007* [CD-ROM]. Žilina: Slovenská elektrotechnická spoločnosť, 2007, ISBN 978-80-8070-786-6.

[4] Nemčík, M. - Svítek, J.: *Lab Station for Video Signal Processing.* In *RTT 2007* [CD-ROM]. Žilina: University of Žilina, Faculty of electrical engineering, ISBN 978-80-8070-735-4, 2007, [CD-ROM]. (Participation 50%)

[5] Nemčík, M. - Fišerová, V.: *Spatial Laser Scanning "By Hand".* Computer Design. 2007, č. 2/2007, s. 72-73. ISSN 1212-4389. (in Czech). (Participation 50%)

[6] Nemčík, M.: *Automatic Knotclouds Placement Algorithm for Quadratic DMS-spline Function. Access server* [online]. 2011, roč. 9, č. 2011090001, Internet: http://access.feld.cvut.cz/view.php?cisloclanku=2011090001. ISSN 1214-9675.

[7] Nemčík, M.: *Automatic Generation of Side Control Points over Arbitrary Triangulation for Quadratic DMS-spline Function. Access server* [online]. 2011, roč. 9, č. 2011090002, Internet: http://access.feld.cvut.cz/view.php?cisloclanku=2011090002. ISSN 1214-9675.

WOS

Patent

Registered software

[8] Nemčík, M.: *projectHEAD* . [Authorized software] Czech Technical University in Prague. FEE. 2011.

Grants

[9] Nemčík, M.: *Development of a System for Creating 3D Polygonal Human Head Model of a Videotelephony Customer by a System of Cameras 07–07.* 2007. CTU CTU0705513.

[10] Nemčík, M. - Svítek, J.: *Preparation of New Laboratory Exercises Dealing with Video Signal Processing 07–07.* 2007, RV 1786G1. (Participation 50%)

No response and reviews.

8 List of Author's Publications (Others)

Impact Journal

Reviewed Journal

[1] Nemčík, M.: *An Objective Method for End-to-End Speech Quality Assessment. POSTER 2006* [CD-ROM]. Prague: CTU, 2006.

[2] Nemčík, M. - Levák, M.: *An Objective Method for End-to-End Speech Quality Assessment Using PESQ Algorithm.* In *Digital Technologies 2006 - 3rd International Workshop* [CD-ROM]. Žilina: University of Žilina, Faculty of electrical engineering, 2006, vol. 1, ISBN 80-8070-637-9.

[3] Nemčík, M. - Pravda, M.: *Voice Transmission Quality Tester.* In *Applied Electronic 2010.* Pilsen: University of West Bohemia, 2010, p. 277-280. ISBN 978-80-7043-865-7. (Participation 50%)

[4] Nemčík, M. - Pravda, M. - Zavrtálek, J.: *Voice Transmission Quality Testing and the Design of Measuring Unit.* In *TSP 2010 - 33rd International Conference on Telecommunications and Signal Processing* [CD-ROM]. Budapest: Asszisztencia Szervezo Kft., 2010, p. 397-401. ISBN 978-963-88981-0-4. (Participation 33%)

[5] Nemčík, M.: *Evaluation of Voice Transmission Quality. Access server* [online]. 2006, roč. 4, č. 2006050402, Internet: http://access.feld.cvut.cz/view.php?cisloclanku=2006050402. ISSN 1214-9675. (in Slovak).

[6] Nemčík, M.: *Speech Quality Assessment. Akustick listy.* 2006, roč. 12, č. 1, s. 7-13. ISSN 1212-4702. (in Slovak).

[7] Nemčík, M. - Pravda, M. - Vodrážka, J.: *Voice Transmission Quality Tester. Access server* [online]. 2010, roč. 8, č. 1, s. 1-4. Internet: http://access.feld.cvut.cz/view.php?cisloclanku=2010010001. ISSN 1214-9675. (in Czech). (Participation 33%)

WOS

Patent

Unpublished Lectures

[8] Nemčík, M.: *Voice Quality Assessment - End-to-End VQA for Telecommunication Networks*. [1 day training]. Prague. Education Support Centre - Trainingpoint. 2008.

First Functional Prototype

[9] Nemčík, M. - Pravda, M.: *Voice Transmission Quality Tester*. Czech Technical University in Prague. FEE. 2009.

No response and reviews.

No response and reviews.

9 Summary and Conclusions

This thesis focuses on possibilities of quadratic DMS-spline function's utilization in model-based video coding. In this thesis I propose a solution for the problem of automatic knotclouds placement with regard to fulfilment of condition of normality over the whole triangulation. Furthemore, I propose a solution for additional determination of side control points of triangulation. This solution is especially needed when quadratic form of a spline function is used. It uses neighbour triangles to compute position of the given side control point with regard to realistic appearance of a model. DMS-surface combines global smoothness of linear splines with possibility of local control. Automatic side control point placement and knotclouds placement allows to use a quadratic form of DMS-surfaces while animating triangulation of a human head model. For these features and the possibility of effective animation of approximation surface this system presents very convenient tool for a model-oriented video compression. Extraction of definition and animation parameters from input frame represents nowadays the main focus of many research teams. Based on extraction of definition and animation parameters of an input frame a human head model is shaped and animated. It is the same model which is then approximated. I used CANDIDE 3-1-6 model which in addition to basic triangulation, contains also definition and animation vector. I also suggest enhancement of this model (triangulation), more precisely I add the back part of human head which I call M-4A©. Definition and animation parameters are unified in the MPEG-4 AVC videocodec which ensures general utilization of this system. As a part of this thesis I developed a special *projectHEAD*© development framework. It allowed me experimental verification of suggested algorithms and simulation of decoding part of model-based video coding, including real texture mapping. Mathematical complexity of quadratic DMS-surface computing and absence of algorithm of automatic knotclouds and side control points placement have been limiting factor of very convenient function for video coding utilization.Nowadays the increased computing power of equipments with videophone features (SmartPhone, PDA, PC) significantly supports the progress of usage of more complex model and compression algorithm.It is the utilization of quadratic DMS-spline function in a model-based videocompression which seems to be the way which would make the current conventional videophone much more effective. In addition, it would also enrich it with new multimedia features.

I introduced an enhancement of the model-based video coding method by utilizing the database of the 3D model of the human head. This enhancement brings new useful features e.g. real video scene coding, speech-driven and text-driven facial animation,

3D video visualization, etc.

The result of the thesis can be used in shifting from the classical 2D video communication to more realistic 3D video communication (Fig. 9). Such a type of video communication uses 3D mesh model, which is able to simulate the natural human mime. Transmitting only facial expression parameters drastically reduces the bandwidth requirements to a few kbps. In the same way, face animations or new human-computer interfaces can be realized with low demands on storage capacities. On the high quality end, film productions may get new impacts for animation, realistic facial expression, and motion capture without the use of numerous sensors that interfere with the actor. Last, but not least, information about motion and symmetry of facial features can be used in medical diagnosis and therapy. All these applications have in common the requirement of accurate information about 3D motion deformation and facial expressions in common.

Figure 9: 3D Videocommunication.

Concerning the financial arrangements, this work was supported by the following grants: *Development of a System for Creating 3D Polygonal Human Head Model of a Videotelephony Customer by a System of Cameras 07–07, CTU CTU0705513* and *Preparation of New Laboratory Exercises Dealing with Video Signal Processing 07–07, FRV 1786G1.*

Résumé

Efektívnejší spôsob kódovania video telefónneho hovoru poskytuje algoritmus modelovo-orientovanej videokompresie. Tento algoritmus využíva pre dosiahnutie velmi nízkej prenosovej rýchlosti informácie o tvare a texúre účastníka všeobecný 3D model ľudskej hlavy. Na zjednodušenie extrakcie príznakových bodov a prenosu sa pre modelo-orienotovanú videokompresiu využíva model s relatívne málo vrcholmi a trojuholníkmi. Na prijímacej strane sa tento jednoduchý model aproximuje. Autor v prací navrhuje efektívny spôsob, akým je možné aproximovat povrch pomocou DMS kvadratických splajnových funkcii. Rieši problematiku automatického umiestnovania prídavných a stranových kontrolných bodov pre dosiahnutie maximálnej hladkosti povrchu a pri zachovaní normality nad celou trianguláciou. Ďalej navrhuje spôsob animácie aproximovaného povrchu ľudskej hlavy použitím kvadratického DMS-povrchu pomocou animačných jednotiek definovaných v prostredí videokodeku MPEG-4 AVC. V práci je využitý trojuholníkový mriežkový model CANDIDE 3-1-6, ktorý autor rozširuje o zadnú cast s názvom M-4A© (*Model for Animation in MPEG-4 environment*).

Autor ďalej vytvoril komplexné vývojové prostredie pre aproximáciu povrchu ľudskej hlavy pomocou DMS-splajnových funkcií a jeho animáciu. Program je určený na výpočet pozície prídavných a stranových kontrolných bodov nad celou trianguláciou, výpočet aproximovaného povrchu, a jeho animáciu. Umožňuje nastavovať parametre, ktoré ovplyvňujú polohu prídavných a stranových kontrolných bodov a tak ovplyvňova výsledný povrch. Po prevedení výpočtov aproximovaného povrchu je tento povrch následne v prostredí openGL zobrazený. Vytvorené vývojové prostredie pomohlo potvrdiť navrhovane postupy pre umiestňovanie prídavných a stranových kontrolných bodov. V rámci riešenia zadaných cieľov dizertačnej práce autor navrhol a vytvoril priestor pre animáciu aproximovaného povrchu ako aj, implementácie reálnej textúry. Optimalizácia výpočtových krokov umožnila zefektívnenie výpočtov animácie aproximovaného povrchu. Program umožňuje manuálne vytvoriť a uložiť animáciu pomocou animačných jednotiek AU resp. FAP definovaných v prostredí MPEG-4. A ďalej načítať tuto animáciu zo súboru. Tieto jednotlivé kroky predstavujú kompletnú syntetizačnu časť modelovo-orientovanej videkompresie čím a vytvárajú priestor pre prepojenie s analyzanou časťou.