The 17th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision

in co-operation with

EUROGRAPHICS

W S C G ' 2009

POSTER Proceedings

University of West Bohemia
Plzen
Czech Republic

February 2 - 5, 2009

Co-Chairs
Min Chen, Swansea University, United Kingdom
Vaclav Skala, University of West Bohemia, Czech Republic

Edited by
Min Chen
Vaclav Skala

WSCG'2009 POSTER Proceedings

Editor-in-Chief: Vaclav Skala
c/o University of West Bohemia, Univerzitni 8, Box 314
CZ 306 14 Plzen
Czech Republic
skala@kiv.zcu.cz

Managing Editor: Vaclav Skala

Author Service Department & Distribution:
Vaclav Skala - UNION Agency
Na Mazinách 9
CZ 322 00 Plzen
Czech Republic

Printed: University of West Bohemia, Plzeň, Czech Republic

Hardcopy: *ISBN 978-80-86943-95-4*

WSCG 2009

International Programme Committee

Adzhiev,V. (United Kingdom)
Baranoski,G. (Canada)
Bekaert,P. (Belgium)
Benes,B. (United States)
Bilbao,J. (Spain)
Biri,V. (France)
Bittner,J. (Czech Republic)
Bouatouch,K. (France)
Brodlie,K. (United Kingdom)
Buehler,K. (Austria)
Chen,M. (United Kingdom)
Chover,M. (Spain)
Coquillart,S. (France)
Crosnier,A. (France)
Csebfalvi,B. (Hungary)
Cunningham,S. (United States)
Daniel,M. (France)
de Geus,K. (Brazil)
Debelov,V. (Russia)
Feito,F. (Spain)
Ferguson,S. (United Kingdom)
Flaquer,J. (Spain)
Flusser,J. (Czech Republic)
Gallo,G. (Italy)
Groeller,E. (Austria)
Gudukbay,U. (Turkey)
Gutierrez,D. (Spain)
Havran,V. (Czech Republic)
Jansen,F. (Netherlands)
Kavan,L. (Ireland)
Klein,R. (Germany)
Klosowski,J. (United States)
Lee,T. (Taiwan)
Magnor,M. (Germany)
Mollá Vayá,R. (Spain)
Muller,H. (Germany)
Murtagh,F. (Ireland)
Myszkowski,K. (Germany)
Pasko,A. (United Kingdom)
Pedrini,H. (Brazil)
Peroche,B. (France)
Platis,N. (Greece)
Puppo,E. (Italy)
Purgathofer,W. (Austria)
Rojas-Sola,J. (Spain)
Rokita,P. (Poland)
Rosenhahn,B. (Germany)
Rossignac,J. (United States)
Rudomin,I. (Mexico)
Sakas,G. (Germany)
Schumann,H. (Germany)
Segura,R. (Spain)
Selim,B. (Turkey)
Skala,V. (Czech Republic)
Slavik,P. (Czech Republic)
Slusallek,P. (Germany)
Sramek,M. (Austria)
Stroud,I. (Switzerland)
Teschner,M. (Germany)
Theoharis,T. (Greece)
Triantafyllidis,G. (Greece)
Vergeest,J. (Netherlands)
Wu,E. (China)
Wuethrich,C. (Germany)
Zara,J. (Czech Republic)
Zemcik,P. (Czech Republic)

WSCG 2009
Board of Reviewers

Abas,M. (Malaysia)
Adzhiev,V. (United Kingdom)
Akleman,E. (United States)
Andreadis,I. (Greece)
Aspragathos,N. (Greece)
Aveneau,L. (France)
Baranoski,G. (Canada)
Battiato,S. (Italy)
Bekaert,P. (Belgium)
Bellon,O. (Brazil)
Benes,B. (United States)
Bilbao,J. (Spain)
Biri,V. (France)
Bittner,J. (Czech Republic)
Borchani,M. (France)
Bouatouch,K. (France)
Bouville,C. (France)
Brodlie,K. (United Kingdom)
Bruckner,S. (Austria)
Bruni,V. (Italy)
Brunnet,G. (Germany)
Buehler,K. (Austria)
CarmenJuan-Lizandra,M. (Spain)
Chaudhuri,D. (India)
Chen,M. (United Kingdom)
Chover,M. (Spain)
Chum,O. (Czech Republic)
Coquillart,S. (France)
Crosnier,A. (France)
Csebfalvi,B. (Hungary)
Cunningham,S. (United States)
Daniel,M. (France)
de Geus,K. (Brazil)
Debelov,V. (Russia)
Dingliana,J. (Ireland)
Drbohlav,O. (Czech Republic)
Duce,D. (United Kingdom)
Durikovic,R. (Slovakia)
Egges,A. (Netherlands)
Eisemann,M. (Germany)
Erbacher,R. (United States)
Feito,F. (Spain)
Ferguson,S. (United Kingdom)
Ferko,A. (Slovakia)
Fernandes,A. (Portugal)
Flaquer,J. (Spain)
Flusser,J. (Czech Republic)
Foufou,S. (France)
Gallo,G. (Italy)
Galo,M. (Brazil)
Ganovelli,F. (Italy)
Garcia-Alonso,A. (Spain)
Giannini,F. (Italy)
Gonzalez,P. (Spain)
Grammalidis,N. (Greece)
Groeller,E. (Austria)
Gudukbay,U. (Turkey)
Gumbau,J. (Spain)
Gupta,A. (United States)
Gutierrez,D. (Spain)
Hanak,I. (Czech Republic)
Haro,A. (United States)
Hasler,N. (Germany)
Havran,V. (Czech Republic)
Hernández,B. (Mexico)
Herout,A. (Czech Republic)
Horain,P. (France)
House,D. (United States)
Iwanowski,M. (Poland)
Janda,M. (Czech Republic)
Jansen,F. (Netherlands)
Joan-Arinyo,R. (Spain)
Jones,M. (United Kingdom)
Karabassi,E. (Greece)
Kavan,L. (Ireland)
Kimmel,B. (Canada)
Klein,R. (Germany)
Klosowski,J. (United States)

Knight,M. (United Kingdom)
Kohout,J. (Czech Republic)
Kolcun,A. (Czech Republic)
Krishnaswamy,A. (United States)
Lanquetin,S. (France)
Lee,T. (Taiwan)
Lewis,J. (New Zealand)
Lin,W. (Taiwan)
Liu,D. (Taiwan)
Maciel,A. (Brazil)
Magnor,M. (Germany)
Maierhofer,S. (Austria)
Mandl,T. (Germany)
Matey,L. (Spain)
Matkovic,K. (Austria)
Mattausch,O. (Austria)
Michoud,B. (France)
Mokhtari,M. (Canada)
Mollá Vayá,R. (Spain)
Montrucchio,B. (Italy)
Mudur,S. (Canada)
Muller,H. (Germany)
Murtagh,F. (Ireland)
Myszkowski,K. (Germany)
OliveiraJunior,P. (Brazil)
Papaioannou,G. (Greece)
Pasko,A. (United Kingdom)
Pasko,G. (Cyprus)
Paulin,M. (France)
Pedrini,H. (Brazil)
Peroche,B. (France)
Pettifer,S. (United Kingdom)
Platis,N. (Greece)
Přikryl,J. (Czech Republic)
Puig,A. (Spain)
Puppo,E. (Italy)
Purgathofer,W. (Austria)
Renaud,C. (France)
Ripolles,O. (Spain)
Ritschel,T. (Germany)
Rodeiro,J. (Spain)
Rojas-Sola,J. (Spain)
Rokita,P. (Poland)
Rosenhahn,B. (Germany)
Rossignac,J. (United States)
Rudomin,I. (Mexico)
Sakas,G. (Germany)
Sanna,A. (Italy)
Schumann,H. (Germany)
Segura,R. (Spain)
Selim,B. (Turkey)
Sellent,A. (Germany)
Sirakov,N. (United States)
Skala,V. (Czech Republic)
Slavik,P. (Czech Republic)
Slusallek,P. (Germany)
Solis,A. (Mexico)
Sousa,A. (Portugal)
Sramek,M. (Austria)
Stroud,I. (Switzerland)
SuarezRivero,J. (Spain)
Svoboda,T. (Czech Republic)
Teschner,M. (Germany)
Theoharis,T. (Greece)
Theußl,T. (Austria)
Torrens,F. (Spain)
Triantafyllidis,G. (Greece)
Tytkowski,K. (Poland)
Vanecek,P. (Czech Republic)
Vasa,L. (Czech Republic)
Veiga,L. (Portugal)
Vergeest,J. (Netherlands)
Viola,I. (Norway)
Vitulano,D. (Italy)
Wan,T. (United Kingdom)
Wu,E. (China)
Wuethrich,C. (Germany)
Yencharis,L. (United States)
You,S. (United States)
Zach,C. (United States)
Zachmann,G. (Germany)
Zalik,B. (Slovenia)
Zara,J. (Czech Republic)
Zemcik,P. (Czech Republic)
Zhu,Y. (United States)

Contents

POSTERS

Stoeva,M.T., Bozhikova,V.T.: Invariant to Transformations Image Retrieval from Image Databases using a Boundary Based Description of Object Shape 1

Kovacs,L.: Seamless NPR Video Painting 5

Kwon,J.H., Lee,B.G., Yoon,J., Lee,J.: Image Deformation using Radial Basis Function Interpolation 9

Sisojevs,A., Glazs,A.: Efficient approach to direct B-spline surface rendering by a ray tracing 13

Liu,Y., Laycock,S.D.: The Force-Map Haptic Rendering Algorithm for Drilling into Volume Data 17

Alizadehashrafi,B., Coors,V., Schulz,T.: POSTER: 3D Navigation Systems based on Synthetic Texturing 21

Prado,J. ,Lobo,J., Dias,J.: Robotic Visual and Inertial Gaze Control using Human Learning 27

Rachkovskaya,G., Kharabayev,Yu.: Geometric modeling and computer graphics of kinematic ruled surfaces on the base of complex moving one axoid along another (one-sheet hyperboloid of revolution as fixed and moving axoids) 31

Zinger,S., Ruijters,D., de With,P.H.N.: iGLANCE project: free-viewpoint 3D video 35

Choung,Y., Kang,K.-K., Kim, D.: Example Based Motion Generation with Efficient Control 39

Ji,S.-H., Park,S.-H., Ryu,D.-S., Cho,H.-G.: A Visualization and Management System for Chat Dialogues on 3D Virtual Avatars 43

Nishio,K., Takebayashi,Y., Teshima,Y., Kanaya,T., Kobori,K.: POSTER: Information Transformation for Point Cloud 47

Hwang,J., Lim,D., Paik,D.: A Composition-Based Image Retrieval Using Line Feature 53

POSTER: Invariant to Transformations Image Retrieval from Image Databases using a Boundary Based Description of Object Shape

Mariana Tsvetanova Stoeva
Technical University Varna
Studentska str. 1
Department of Computer Science and Technology, TU-Varna
Bulgaria (9010), Varna, Varna
mariana_stoeva@abv.bg

Violeta Todorova Bozhikova
Technical University Varna
Studentska str. 1
Department of Computer Science and Technology, TU-Varna
Bulgaria (9010), Varna, Varna
vbojikova2000@yahoo.com

ABSTRACT

The basic problem of the methods for shape description of image objects is their dependence on the graphical transformations. In this paper we give boundary based method for shape description based on a multi-step criterion. We present a new approach for similarity shape retrieval from image database that achieves invariance of results with respect to transformations. We define and use a criterion to achieve invariant image object representation with respect to rotation. The approach includes original description of image objects shape of the type of multi-metric attribute stored in image database and a similarity distance definition for image retrieval processing that is invariant to reflection too. Our approach achieves results that are invariant with respect to arbitrary compositions of graphical transformations.

Keywords

Object shape description, image transform, image database, retrieval by similarity, content-based retrieval.

1. INTRODUCTION

Content-Based Image Databases (CBIDB) have been one of the most vivid research areas in the field of computer vision over the last 10 years. The main features that are used to describe the content of the images in the databases are colour, texture and shape of image objects. The increasing sizes of image archives, the requirements of the new appearing appliances and the fact that a few content based systems process shape queries make the further investigations over the shape feature necessary. The requirements for invariance of objects shape description that is extracted from preliminary not normalized images, with respect to all transformations (translation, scaling, rotation and reflection) are basic in the contemporary investigations [Jou01a].

Different methods for objects shape description are used in the existing visual systems and one of the most well-known is the group of the boundary based methods [Con00a]. As a common defect they all have the disadvantage to depend on the graphical transformations [Con00b]. In this paper we present a new approach based on a Multi-Step Criterion (MSC) that we use to achieve invariant image representation with respect to rotation. This MSC determines synonymously a "Start" Point (SP) from the external contour of an image object. For the purpose of the criterion a Recursive Selective Function (RSF) is defined and it is repeatedly applied with different parameters. The presented approach includes definition of RSF, the built on its base MSC for synonymous determination of a SP from the external contour of an image object, original description of image objects shape, and a similarity distance definition for image retrieval processing. We present some of our experiments that confirm the invariance with respect to transformations, completeness, correctness, and sensitiveness of the results of the similarity retrieval by shape from approach application.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

2. AN APPROACH FOR SHAPE RETRIEVAL

The approach includes original description of image objects shape, based on multi-step criterion for determination of "start" for description of object shape point and a similarity distance definition that detects and finds out an image in the image database, that is equal in shape with the target image – query.

2.1 A Multi-Step Criterion for Determination of "Start" Point

Let for a given image object an ordered multitude of characteristic for its external contour points V({v_i}, i=1,…,n) is formed and let for these points a geometric measurement of feature p is made, that creates the ordered multitude P({p_i}, i=1,…,n), which is connected with V.

Let the function that obtains the multitude P containing the measurements of feature p from multitude V is indicated by $P=\varphi_p(V)$. And also let a selective function Max(V,P), that obtains the multitude $V_{max}\subseteq V$ by selecting those elements from the multitude V whose corresponding elements in P have value equal to the maximum one for the multitude P, is defined by (1).

$$V_{\max} = \mathrm{Max}(V,P) = \mathrm{Max}(V,\varphi_P(V)) = (\{v_i\}, p_i = p_{\max}) \qquad (1)$$

Definition 1: A Recursive Selective Function (RSF) $\rho(V,P)$ is defined over the multitudes V and P by dependence (2) and determines a sub-multitude $V_p\subseteq V$ for which its corresponding sub-multitude $P' \subseteq P$ contains only maximum values as evaluation of the geometric measurement P.

$$V_P = \rho(V,P) = \begin{cases} V, & ako \quad V = V_{\max} \\ \rho(V_{\max}, \varphi_P(V_{\max})), & ako \quad V \neq V_{\max}, \end{cases} \qquad (2)$$

$$where \quad V_{\max} = \mathrm{Max}(V,P) = (\{v_i\}, p_i = p_{\max})$$

The input of a recursion is imposed on the fact that if only one element from P is different from the others, the function Max(V,P) determines a sub-multitude with a smaller number of elements from that of the input multitude. For this selected sub-multitude the measurements of the feature P are already different as a rule. This imposes a pre-determination of the values of the geometric measurement for the points from the sub-multitude and applying of Max(V,P) again until it gives a result that is different from the input. This is repeated again and again until the function Max(V,P) stops to give a new result. In this way the whole potential of the function for a selection over a given measurement is utilized.

The recursive selective function is used for synonymous determination of the SP for description of image object's external contour. Such a description will be invariant to a possible rotation.

We suggest the choice of a start point from the multitude of points of the external contour to be made by introducing the MSC applied to the geometric features of the figure that will be applied consistently for three different measurements. If the criterion does not allow a synonymous choice after its third step, then it must be assumed that the figure is close to a regular polygon with number of top points equal to the number of points of the selected by the criterion sub-multitude. In this case as a "start" point is assumed the point from the selected sub-multitude that is closest to the axis $\overrightarrow{OX'}$. An example for the heaviest case for the criterion is the shape of the circle. Let $O(x_0,y_0)$ is the point - centroid of the points from the external contour ({C_{0i}},$1\le i\le k_0$). The number of the points is $k_0=\|C_0\|$.The applying of the criterion must determine one of them as a SP independently from the angle of arbitrary rotation around point O.

In order to be solved the task for synonymous determination of one SP (noted Vsp) from the external contour of the object, the MSC is suggested that consists of consistent applying to the multitude of points of the external contour the specially deduced for this purpose recursive selective function for three measurements P over the geometry of the external contour points. These measurements are: the radius-vectors (**r**) of the points from the contour; the angle (**Ω**) that makes each point's radius-vector with the radius-vector of the next point in anticlockwise direction; the area (**S**) of a sector of the figure determined by the point – centroid and the next point.

Definition 2: The multi-step criterion for synonymous determination of one SP from the external contour of the object is given by the general record (3) of the consistent calls of the selective function.

$$\begin{aligned} V_S = \rho\{[&\rho(\rho(C_0,\varphi_r(C_0)), \\ &\varphi_\Omega(\rho(C_0,\varphi_r(C_0))))], \\ [&\varphi_S(\rho(\rho(C_0,\varphi_r(C_0)), \\ &\varphi_\Omega(\rho(C_0,\varphi_r(C_0)))))]\} \end{aligned} \qquad (3)$$

If the MSC gives as a result a multitude of only one element (k_s=1), then it determines the Vsp=Vs for the external contour. In the other cases when k_s>1 it is assumed that the figure is a regular k_s-gon and as a SP for the contour the point that is closest to the axis $\overrightarrow{OX}$ is chosen. Some more interesting examples of shapes which have numerous points that meet all three conditions for maximum of the measurements in the criterion are illustrated in Figure 1.

The aim of determining these parameters in dependence of object shape by different

 ISBN 978-80-86943-95-4

measurements over the contours describing it, is the preparation of a concrete composition of transformations of the input data.

Figure 1. Examples for SP determination: a/Irregular shape, k_r=2, Vsp=V_{r2}; b/Irregular shape, k_r=3, Vsp=V_{r2}; c/For shape – rectangle, Vsp=V_{r1}; d/For shape – regular octagon Vsp=V_{r1};

2.2 Description of Objects Shape

From the external contour coordinates of an object F from a black-and-white image we determine the composition parameters: the external contour points centroid – point $O(x_0,y_0)$, the maximum Euclid distance from the centroid to the external contour points $r_{0max.}$, and a SP for the external contour. Invariance of figure description with respect to the angle of arbitrary rotation is achieved by synonymous determination of a SP by the MSC. The input data information prepares a normalized representation of the object by dependences (4):

$$x'_{ji} = \frac{1}{r_{0\max}}(x_{ji} - x_0)\cos\alpha + \frac{1}{r_{0\max}}(y_{ji} - y_0)\sin\alpha \qquad (4)$$
$$y'_{ji} = -\frac{1}{r_{0\max}}(x_{ji} - x_0)\sin\alpha + \frac{1}{r_{0\max}}(y_{ji} - y_0)\cos\alpha$$

where: (x_{ji}, y_{ji}) – Decart coordinates of the i-point in the j-contour, (x_0,y_0) – coordinates of the external contour centroid, r_{0max} – the maximum Euclid distance from the centroid to the external contour, α - the angle equal to the one between the radius-vector of the contour "start" pixel and the positive primitive axis X.

Image	Contours	Histogram description

Table 1. Illustration of obtaining an object shape description with L=96:

The histogram description of object shape that forms the multi-size index F= $((F_{\theta i}), 1 \le i \le l)$ is obtained from the contour coordinates transformed this way. Let a straight line start at the center of the coordinate system and let it have θ_i angle with the positive direction of X axis. This line crosses the contours C_j of the object in points with polar coordinates $F \cap \theta_i = ((r_{j1i}, r_{j2i}), 0 \le j < n)$, where : i is the serial number of the cross-points of the crossing axis ($1 \le i \le l$), l is the general number of the crossing axes ($L=3l$), r_{j1i}, r_{j2i} are the radius-vectors modules of the cross-points, $\theta_i = \theta_{i-1} + \pi/l$ and r_{0max} – the maximum Euclid distance from the centroid to the external contour.

Definition 3: The shape description $F(f_1,f_2,\ldots,f_{3l})$ is given by the dependence (5) and forms L- size index for shape that is stored in IDB.

$$F_{\theta i}(f_i, f_{i+1}, f_{i+2l}) = \begin{cases} f_i = \dfrac{1}{r_{0max}}\max(r_{01i}) \\ f_{i+l} = \dfrac{1}{r_{0max}}\min(r_{02i}) \\ f_{i+2l} = \dfrac{1}{r_{0max}}\sum_{j=0}^{n-1}(-1)^j(r_{j1i} - r_{j2i}) \end{cases} \qquad (5)$$

Table 1 presents the states of obtaining the shape description in form of multi-size attribute of medical object $F(f_1,f_2,\ldots,f_{L=96})$.

2.3 Image Retrieval from CBIDB

The technique for image retrieval from CBIDB as an answer of query by image example of object shape is consistent with the typical methodology for similarity query processing and includes the definition of a quadratic similarity histogram distance. The similarity distance presents the range of similarity between the normalized histogram descriptions of the query image and the next object image stored in CBIDB, evaluates the $D(Q,F) \in [0;1)$, and accounts value 0 for objects equal in shape. A similarity shape order may be obtained on the base of the similarity distance values of every image to the others.

$$D(Q,F) = \frac{1}{C^2 L}\min\begin{cases} \sum_{i=1}^{L}(q_i - f_i)^2 \\ (q_1 - f_1)^2 + (q_{L/3+1} - f_{L/3+1})^2 + \\ (q_{2L/3+1} - f_{2L/3+1})^2 + \\ \sum_{i=2}^{L/3}\left[(q_i - f_{2L/3+2-i})^2 + (q_{L/3+i} - f_{L/3+2-i})^2 + (q_{2L/3+i} - f_{3L/3+2-i})^2\right] \end{cases} \qquad (6)$$

Definition 4: Let the shape query is transformed into an image histogram description $Q(q_1,q_2,\ldots,q_n)$, and the image in the database has histogram description $F(f_1,f_2,\ldots,f_n)$, where q_i, f_i are histograms with size L and C is a presentation constant. The similarity distance between Q and F for the examined retrieval model is determined by Equation (6).

3. EXPERIMENTAL RESULTS

Our algorithms are implemented in MatlabR12 and C+ + and are evaluated on test database of 2000 binary images from an image collection. The experimental collection includes also some modified by us images of the same object, transformed by composition of translations and intentionally added changes in the objects contours. Their corresponding attributes, stored in IDB that contain description of object shape, have dimension L=96 and L=12 for l = 32 and l=4. Experiments had been implemented that investigated the invariance, completeness,

correctness, and sensitiveness of the results from the similarity retrieval by shape. Some representative examples are illustrated below.

Query and **Result:**

Figure 2. Similarity ordering of retrieved images

Figure 2 illustrates a test demonstration of numerous experiments whose purpose is the evaluation of the approach stability with respect to transformations both in extraction of image shape description and image retrieval from CBIDB. A set of objects are used that are obtained from transformation variants of the original. An expert evaluation should account as equal all images in the set. The original image with target value of the parameter L=12 is used as a query. It accounts the similarity distance between the query and each one image from the set and an ordering of the variants by similarity follows. The results – the values of the similarity distance $d \in [0, 10^{-7}]$ prove the approach stability with respect to arbitrary combined compositions of transformations.

Query and **Result:**

Result					
Similarity Distance D×10^{-4}	0	0	42	45	45
Result					
Similarity distance D×10^{-4}	145	175	388	434	

Table 2. Similarity ordering for L=96.

The detected minimal differences in the similarity distance values in the range of 10^{-8} present the "strength" of the similarity distance, i.e. it is a low bounding sensitive object similarity distance.

The next experiment presented with Table. 2 is focused on the similarity strength of the similarity distance. By this experiment, we aim to ascertain the similarity distance behavior when smaller or bigger errors in the shape of the internal and external contours of the transformed images are introduced. We use some modified images of the same object transformed by translation and rotation and/or with changed internal and external contours. The replies of the k-query for k = 9 are obtained for value of the parameter L = 12. The experiments show good sensibility of the similarity distance to recognize big as well as small differences of objects shapes.

The correctness of the results was investigated in comparison also with the results of other no invariant methods, such as the one of The Axially-parallel Rectangles of [Con00b]. Copies of the images published in [Con00b] are used for the experiment.

Query **and Results:**

a/

b/

Figure 3. Similarity ordering according [Con00b] and our approach.

Fig.3.a/ presents the results - answer of the query from [Con00b]. Fig3.b/ presents the results of our approach for L=96. $R_{norm}(sys)$ [Jou01b] measure for quantity evaluation of the similarity ordering results is used to evaluate the result of this experiment. In this case for the method a/ $R_{norm}(sys)_б = 0.7031$, and for our approach $R_{norm}(sys)_б = 0.7812$, that means that our results are closer to the user's ordering.

4. CONCLUSIONS

The presented approach is investigated in details by use by use of various evaluations of the results of the implemented experiments. Our conclusions are:

- The developed approach is stable with respect to arbitrary compositions of transformations;
- The approach application is efficient as it achieves completeness, correctness, and sensitiveness of the results of the similarity retrieval by shape;
- The approach achieves a very good effectiveness of information storage in IDB and a good effectiveness of query processing to IDB.

5. REFERENCES

[Con00a] Alferez, Y., Wang F. Image indexing and retrieval using image-derived, geometrically and illumination invariant features. *IEEE Int. Conf. Multimedia Comp. Systems*, 1, 177–182, 1999

[Con00b] Jagadish, H.V.A retrieval technique for similar shapes. In ACM Proc. Int. Conf. Manag. Data (SIGMOD), pp. 208-217, February 1997

[Jou01a] Muller H. Michoux N., Bandom D. Geissbuhler A. A review of CBIR systems in medical applications – clinical benefits and future directions. *International Journal of Medical Informatics*, 73.1, pp.1-23, 2004

[Jou01b].V. N. Gudivada and V. V. Raghavan, Design and evaluation of algorithms for image retrieval by spatial similarity. ACM Trans. on Information Systems, vol. 13 no. 2, pp.115-141, April 1995

Seamless NPR Video Painting

Levente Kovács

Distributed Events Analysis Research Group
Computer and Automation Research Institute, Hungarian Academy of Sciences
Kende u. 13-17,
H-1111, Budapest, Hungary

levente.kovacs@sztaki.hu - http://web.eee.sztaki.hu

ABSTRACT

The paper presents a method for automatic transformation of ordinary videos into non-photorealistic (NPR) graphical representations, for achieving painterly effects. The primary goal is the non-interactive transformation of videos, with the intended use as a service in community video databases (as an extension to content based retrieval). There are no restrictions on the types of videos. The main steps involve scene separation, color segmentation, foreground and focus area extraction, and vector graphical transformation of selected frames.

Keywords

Non-photorealistic rendering, image/video painting, painterly rendering, visualization.

1. INTRODUCTION

Methods for creating painterly, cartoonish effects on ordinary videos are already present in the NPR literature, although their number is fairly limited. Most of these approaches are manual or at most semi-automatic in nature. A number of them is based on stroke simulation, with different optimization techniques to reduce the computational time. Also, the movie and game development industries know and use cartoon renderings, which can produce good results, although they need manual interaction.

Our goal is to create an automatic NPR video rendering method which is light in computational complexity, yet provides nice results. Since we are also working on creating content based video retrieval applications, we intend to provide additional services beside retrieval, one of which is automatic painterly effect rendering on stored videos.

Thus one of the primary requirements is that we create an automatic video renderer. Another requirement is to create the painterly effects on videos by such feature descriptors that are also used when processing images and videos for database retrieval, which are already available in the database, and have a light computational complexity.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

One of the first video NPR methods was that of [Lit97a], where strokes with a given center, length, radius and orientation are generated, adding random perturbations and variation; [Hay04a] is a painting method based on generating meshes of brush strokes with adapting spatio-temporal properties. Like in [Kov02a] optic flow data is used to paint frames, in this case by propagating the stroke objects along motion trajectories. [Col05a] is also a semi-automatic NPR animation method based on stroke simulation and propagation, creating various styles, and being highly versatile. [Kim05a] presents a semi-automatic video transformation method which produces cartoonish motion cues on the images, based on segmentation and tracking. [Wan04a] provides a method based on manual object selection and temporal Mean Shift tracking.

The main steps of the approach presented here are the following. We detect the shot changes in the video; then, the renderer creates a two-layer transformation of the shot frames, a background layer which has lower complexity, and a foreground layer, with higher level of detail. The foreground is extracted by a Stauffer-Grimson based foreground detector combined with an automatic focus area extracting method. The color segmentation necessary for the two layers is obtained by using a Mean Shift color segmenting step. Between the frames, only those areas will be updated where the focus and

foreground change detector signals a change in content.

Our approach differs from other NPR video creating methods in that it is not based on stroke simulation, is automatic, does not use computationally highly intensive optimizations, generates the least possible layers, and only uses features which are already available in the video database. Thus we will be able to provide an automatic effect generator service as an extension to already available content based retrieval modules in a video retrieval application.

2. NON-PHOTOREALISTIC EFFECTS ON VIDEOS

As we mentioned above, the goal is to build a non-photorealistic video renderer which uses features that are generally (and specifically, in our case) available in a video database. Such features include shot/scene, color, texture, edge, invariant feature, motion descriptors. These features are not randomly selected, but chosen based on how their combination works better for retrieving specific classes of videos [Szl08a]. We tried to reduce the amount of necessary steps and used features to a minimum, to be able to create an efficient video NPR method. The basic elements one needs for a method not based on stroke simulation is to extract moving areas (foreground), which in itself will help in identifying the background layer. Then, we need to apply color reduction which produces a painterly effect. Since among the image and video features we almost always have some sort of color segmentation, it seems natural to use the already available results in producing the NPR effects. In our case we chose a Mean Shift [Com02a] based segmenter to generate lower and higher detail layers.

In the following subsections we will present the three main steps of the method.

Scene changes

For detecting scene changes, we use a motion based approach, similar to what is used in a variety of video encoders. We obtain local motion estimation on frames, and calculate the error of the estimations. We build a statistics of these errors, and when they cross a threshold, we signal a shot change. First we calculate a block matching based motion estimation on small frame blocks. For each block b_l (8x8 pixel size) from a frame (t) we search around its location r_l (with a step size of 4 pixels, in the maximum range of 4 locations in each direction) on the following frame ($t+1$) to find the same size block which matches best, by calculating a local error minima:

$$D = \min\left\| b_l^t(r_l), b_i^{t+1}(r_i) \right\| \qquad i = \overline{1,n}$$

Thus we obtain a motion direction vector for each frame block b_l, i.e. (v_x,v_y). Then, we calculate the estimation error on the frame blocks, based on the motion vector estimates above:

$$E_l = \left\| b_l^t(x_l, y_l), b^{t+1}(x_l + v_x, y_l + v_y) \right\|$$

and sum them up for the whole frame $E^t = \sum_{l=1}^{m} E_l$,

where m is the total number of blocks. If the total error crosses a threshold then we signal a shot change position. Currently we have a database with over 6000 correctly segmented scenes and shots.

Foreground

For extracting the foreground, i.e. the changing areas between consecutive frames, we use a combination of a Stauffer-Grimson [Sta00a] foreground extractor and a focus area extractor.

We consider each pixel s as a separate process, which generates an observed pixel value sequence over time (t is the time index): $\left\{x^{[1]}(s), x^{[2]}(s), \ldots x^{[t]}(s)\right\}$. To model the recent history of the pixels, [Sta00a] suggested a mixture of K Gaussians distributions:

$$P\left(x^{[t]}(s)\right) = \sum_{k=1}^{K} w_k^{[t]}(s) \cdot \eta\left(x^{[t]}(s), \mu_k^{[t]}(s), \sigma_k^{[t]}(s)\right)$$

where $k = 1, \ldots, K$ are unique and in time static ID-s of the mixture components, while $\eta(.)$ is a Gaussian density function, with given μ mean and σ deviation. The idea is to model the image/frame pixels as of being a mixture of the Gaussians, which are then evaluated over time to determine which of them is more likely the background and which is the foreground model. Figure 1 (top right image) shows an example.

The focus area extractor builds upon a deconvolution-based region classification approach [Kov07a]. Basically, it is a method for automatically classifying image areas relative to each other based on the local blur/focus on the image, without a priori knowledge. The estimation of local blur and reconstructed region is an iterative approach based on [Ric72a], for estimating the local blur and the original unblurred region:

$$\begin{cases} f_{k+1}(r) = f_k(r)\left[h_k(r) * \dfrac{g}{g_k}(r)\right] \\ h_{k+1}(r) = \dfrac{h_k(r)}{\gamma}\left[f_k(r) * \dfrac{g}{g_k}(r)\right] \end{cases}$$

where k is the current iteration, γ is a weighting factor, f is the unknown original undistorted image (or region), g is the observed image region, h is the

unknown point spread function which caused the distortion, thus $g = h * f$, meaning g is the convolution of some original signal with a distortion function. The classification is based on a local reconstruction error measure:

$$E_r(g, g_k) = \left| arcsin \frac{< g - g_k, g >}{|g - g_k| \cdot |g|} \right| \cdot \frac{C_r(g_r)}{\max_r \{C_r(g_r)\}}$$

where r is the location vector in the image, g is the observation, index k denotes the actual reconstruction iteration, and the second part of the function stands for the local contrast. The values obtained are fed into a classifier. A sample result is in Figure 1 (bottom left) and Figure 2.

The final foreground is a logical or combination of the two sub-steps above, and samples are presented in Figure 1 (bottom right) and transformed foregrounds are shown in Figure 4.

The colorization is produced by reduced resolution color segmentation, for which we use Mean Shift [Com02a] segmentation. Larger homogenous areas can be produced quickly, giving the sensation of a low-detail blurred background. This provides an easy way for specifying the level of detail of the background and the foreground. A resulting rendered frame is shown in Figures 4, 5.

Figure 1. Input frame (top left), foreground mask (top right). Focus region mask (bottom left), final combined mask (bottom right).

Figure 2. Focus area extraction sample.

Figure 3. Input frames (left column) and backgrounds (right column).

Background

The background is generated as a low detail version of the current frame. For this purpose we use a Mean Shift color segmentation of the frame, with a high area setting and a low color count. The idea is to create a coarse representation of the frames, with larger consolidated areas. Figure 3 shows some examples. These generated segmentations will serve as the background of each frame.

Results and Application

As we do not use tracking, the question of inter-frame coherence arises. Since we transform every frame of the videos (or single frames, as requested) separately, we need to take care that color vibration does not cause disturbing artifacts. As a prevention step, we provide the possibility of adding a temporal blur to the resulting frame sequences, with a varying radius setting. Thus eventual color vibrations become unnoticeable and the computation complexity does not increase noticeably.

Using the above presented feature extractions, the final videos will be the result of foreground extraction and color segmentation steps. The specific steps have been chosen so that they should be already available features in a content based video retrieval application, mostly for speed and reaction time considerations.

As an additional possibility, we provide a step to create scalable vector graphical (SVG) versions of the rendered frames (one frame at a time). To do this, first, we extract the perimeter points of the background areas with a color area extracting algorithm (based on contour tracking with color similarity checking) and describe them as *path* SVG elements. Thus each region becomes a closed path, with specified stroke and fill color settings.

3. FUTURE WORK

We intend to include the described method into a web service which combines content based video retrieval with NPR video creation. Videos or segments/shots and frames will be available for transformation. The service will provide visually interesting versions of the uploaded video content, and ways to create content for enthusiasts and videographers. We intend to provide a style transfer method, i.e. transport a style of a rendering onto another input material.

4. CONCLUSIONS

We presented a method for automatic video NPR, which uses fairly simple features already available in most video databases intended for content based retrieval. The computational complexity of the approach is fairly low, yet the produced effect is visually appealing. Reproduction of the method is also quite easy, yet it provides a plethora of new possibilities for video database applications besides "simple" content based retrieval features.

5. ACKNOWLEDGMENTS

This work is partially supported by the JUMAS EU FP7 project.

6. REFERENCES

[Sta00a] Stauffer, C., Grimson, E. Learning patterns of activity using real-time tracking. IEEE Tr. on Pattern Analysis and Machine Intelligence, 22(8), pp. 747-757, 2000.

[Kov07a] Kovács, L., Szirányi, T. Focus area extraction by blind deconvolution for defining regions of interest. IEEE Tr. on Pattern Analysis and Machine Intelligence, 29(6), pp. 1080-1085, 2007.

[Lit97a] Litwinowicz, P. Processing Images and Video for An Impressionist Effect. Proc. of ACM SIGGRAPH 97, pp. 407-414, 1997.

[Kov02a] Kovács, L., Szirányi, T. Creating Animations Combining Stochastic Paintbrush Transformation and Motion Detection, In Proc. of 16th ICPR, pp. 1090-1093, 2002.

[Com02a] Comaniciu, D., Meer, P. Mean shift: A robust approach toward feature space analysis. IEEE Tr. on Pattern Analysis and Machine Intelligence 24(5), pp. 603–619, 2002.

[Hay04a] Hays, J., Essa, I. Image and video based painterly animation. In Proc. of NPAR, pp. 113–120, 2004.

[Col05a] Collomosse, J.P., Hall, P.M. Video Paintbox: The fine art of video painting, Computers & Graphics vol. 29, pp .862–870, 2005.

[Kim05a] Kim, B., Essa, I. Video-based nonphotorealistic and expressive illustration of motion. Computer Graphics International, pp. 32-35, 2005.

[Wan04a] Wang, J., Xu, Y.Q.., Shum, H.Y., Cohen, M.F. Video tooning. In Proc. of SIGGRAPH 2004, pp. 574-683, 2004.

[Ric72a] Richardson, W. Bayesian-based iterative method of image restoration. Journal of the Opt. Soc. of America vol. 62, pp. 55–59, 1972.

[Szl08a] Szlávik, Z., Kovács, L., Havasi, L., Benedek, Cs., Petrás, I., Utasi, Á., Licsár, A., Czúni, L., Szirányi, T. Behavior and event detection for annotation and surveillance. In Proc. of CBMI 2008, pp. 117-124, 2008.

Figure 4. Rendered foreground layers (top). Transformed frames (bottom).

Figure 5. Transformed frames.

Image Deformation using Radial Basis Function Interpolation

Jung Hye Kwon, Byung Gook Lee
Dept. of Visual Contents, Dongseo University.
San 69-1 Churye-2Dong,Sasang-Gu, 617-716, Busan, Korea.
kjh@dit.dongseo.ac.kr
lbg@dongseo.ac.kr

Jungho Yoon
Dept. of Math.
Ewha University.
11-1 Daehyun-Dong Seodaemun-Gu ,120-750, Seoul, Korea.
yoon@ewha.ac.kr

JoonJae Lee
Dept. of Game Mobile Contents.
Keimyung University.
2800 Dalgubeoldaero, Dalseo-Gu, 704-701, Daegu, Korea.
joonlee@kmu.ac.kr

ABSTRACT

Image deformation technique is widely used in the field of computer animation, image editing, medical imaging, and other applications in 2D and 3D computer graphics. All the algorithms aim to provide simple user interface, most of which need the user to drag the control points, lines or polygon. The deformation process and the final position of the controlling points should be smooth and precise respectively, and it should also run in real-time.

This paper provides a simple image deformation method using the radial basis function interpolation in approximation theory. Radial basis function is a very popular and convenient tool for data representation problems. The proposed method using radial basis function is fast and easy to use more than previous deformation methods. Experiments indicate that the algorithm is stable and well performed.

Keywords

Image deformation, Radial basis functions, Interpolation /approximation, Image editing, Computer animation

1. INTRODUCTION

Image deformation has a number of uses from computer animation to morphing. It gives the user the ability to manipulate the object as if the user was handling it in real life. The image is given a set of handles in which the user can intuitively control its position and orientation.

Although there are many ways to deform image to make the desired shape, to perform these deformations the user selects a set of handles to control the deformation. These handles may take the form of points, lines, or even polygon grids. As the user modifies the position and orientation of these handles, the image should deform in an intuitive fashion.

The deformation techniques divide into two parts. One is to deform the shape using mesh based method, the other is to deform the shape using approximation method.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
Copyright UNION Agency – Science Press, Plzen, Czech Republic.

Igarashi et al. [Iga05a] and Weng et.al. [Wen06a] belongs to the first parts. Igarshi et al. presented an interactive system that allows the user to deform a 2D shape by manipulating a few points. The shape is represented by a triangle mesh and the user moves several vertices of the mesh as constrained handles. The system then computes the positions of the remaining free vertices by minimizing the distortion of each triangle. To make the problem linear, they present a two-step closed-form algorithm. The first step is to compute the rotation and the second step to compute the scale. This divides the problem into two least-squares minimization problems that can be solved sequentially. Weng et.al. presents a novel 2D shape deformation algorithm based on nonlinear least squares optimization. The algorithm have to preserve two local shape properties: Laplacian coordinates of the boundary curve and local area of the shape interior, which are together represented in a non-quadratic energy function. The result is an interactive shape deformation system that can achieve physically plausible results.

In approximation method, Arad et al. [Ara94a] has introduced radial basis function for image warping and Schaefer et al. [Sch06a] provided an image deformation based on moving least squares using various classes of linear functions including affine, similarity and rigid transformations. These

deformations are realistic and give the user the impression of manipulation real-world objects.

Our paper builds on a recent paper by Arad et al. Therefore our paper provides a simple image deformation technology using the radial basis function interpolation in approximation theory. The radial basis function method is easier use, quicker to calculate, than the previous deformation method. Experiments indicate that the algorithm is stable and well performed.

The paper is organized as follows: chapter 2 is devoted to introduce the general theories on the radial basis function interpolation. In chapter 3, we explain image deformation using RBF. In chapter 4, we illustrate some resulting image deformation experiments, and conclude.

2. Radial Basis Function

Radial basis function interpolation [Lee06a, For99a, Toi08a] is a very popular and convenient tool for scattered data approximation problems. This chapter introduces the general setting and basic theory of radial basis function interpolation [Lee06a]. Suppose that a continuous function $f:\mathbb{R}^{d_1}\to\mathbb{R}^{d_2}$ is known only at a set of discrete points $U:=\{u_1,u_2,\cdots,u_n\}$ and desired function values $V:=\{v_1,v_2,\cdots,v_n\}$. The radial basis function interpolation to f on U starts with choosing a basis function ϕ, and then it defines an interpolant by

$$S_{f,U}(u)=\sum_{i=1}^{n}\alpha_i\phi(\|u-u_i\|)+\sum_{j=1}^{m}\beta_j p_j(u), \quad (1)$$

where $p_1,p_2,\cdots,p_m$ is a basis for $\Pi_r^{d_1}$ and the coefficients α_i are chosen so that

$$\sum_{i=1}^{n}\alpha_i p_j(u_i)=0, \qquad j=1,2,\cdots,m. \quad (2)$$

Here, Π_r^d is the space of polynomial of total degree r in d spatial dimensions, $m=\dim(\Pi_r^d)=\dfrac{(d+r)!}{d!r!}$.

For a wide choice of function ϕ and polynomials in Π_r^d, the coefficients of $S_{f,U}(u)$ that are required to satisfy the $(n+m)\times(n+m)$ system of linear equations, which can be written in a matrix form as

$$\begin{pmatrix} A & P \\ P^T & 0 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix}=\begin{pmatrix} \mathrm{v} \\ 0 \end{pmatrix}, \quad (3)$$

where A and P are the $n\times n$ and $n\times m$ matrices that have the elements $A_{ij}=\phi(\|u_i-u_j\|)$ and $P_{ij}=p_j(u_i)$ respectively.

This can be written as

$$A=\begin{pmatrix} \phi(\|u_1-u_1\|) & \phi(\|u_1-u_2\|) & \dots & \phi(\|u_1-u_n\|) \\ \phi(\|u_2-u_1\|) & \phi(\|u_2-u_2\|) & \dots & \phi(\|u_2-u_n\|) \\ \vdots & \vdots & \ddots & \vdots \\ \phi(\|u_n-u_1\|) & \phi(\|u_n-u_2\|) & \dots & \phi(\|u_n-u_n\|) \end{pmatrix}$$

and $P=\begin{pmatrix} p_1(u_1) & p_2(u_1) & \dots & p_m(u_1) \\ p_1(u_2) & p_2(u_2) & \dots & p_m(u_2) \\ \vdots & \vdots & \ddots & \vdots \\ p_m(u_n) & p_m(u_n) & \dots & p_m(u_n) \end{pmatrix}$.

Further $\alpha\in\mathbb{R}^{n\times d_1}$ and $\beta\in\mathbb{R}^{m\times d_2}$ are the vectors of coefficients of $S_{f,U}(u)$ and components of v are the data v_j with $j=1,2,\cdots,n$.

In a fundamental paper by Micchelli [Mic86a], the existence and uniqueness of the solution of the linear system is ensured when the basis function ϕ is a conditionally positive definite function.

3. Image Deformation using RBF

We construct a deformation function $S_{f,U}(u)$ relates the u_i points in the set of control points to their v_i-u_i to the counter parts in the deformed position where v_i-u_i is difference vector (Figure 1).

If we have given two sets data $U:=\{u_1,u_2,\cdots,u_n\}$ and $V:=\{v_1-u_1,v_2-u_2,\cdots,v_n-u_n\}$.

Given a point u in the image, we solve for the RBF interpolation $S_{f,U}(u)$, satisfying

$$S_{f,U}(u_i)=v_i-u_i, i=1,2,\cdots n. \quad (4)$$

Finally, we obtain a deformed position v of u

$$v=u+S_{f,U}(u), \quad (5)$$

Figure 1. An Image Deformation Data Set.

 ISBN 978-80-86943-95-4

Figure 2. shows the deformation result according to polynomial degree. (a) shows a change by using RBF without polynomial terms when one control-points were moved. In case of constant term($m = 1$), if we click on the shape to place one control-point and then drag the control-point to its new position, only translation occurs from old position to new position as shown in (b). In case of a linear term($m = 3$), deformation occurs when over 3 points are selected. If they are 3, the result with both rotation and shearing is obtained in (c) .In case of a quadratic terms($m = 6$), the use of over 6 points makes deformation, resulting in curved surface as shown in (d) because it has the characteristics of quadratic function.

4. Experimental Result

We have tested some of the images using Gaussian and Wendland's [Wen95a] functions as basis functions, respectively, to evaluate the proposed algorithm. Both characteristic feature of the function monotonically decreases with distance from the centre.

Wendland's functions $\phi_{3,1}(r) = (1-\frac{r}{c})_+^4 + (4\frac{r}{c}+1)$ are known as basis functions with compact-support basis function. To help illustrate the behavior of a number of the basis functions consider the cat image shown in Figure 3. The same locality parameter was used for each of the localized basis functions. Figure 3. (b) and (c) are images using Gaussian function and Wendland's basis function, respectively. To expend only the cat's ears, that is, to deform image locally, the Gaussian function have to use 5 control-points while Wendland's one does only 2-control points. That's due to compactly supported property.

Figure 4 shows the comparison results of our algorithm (b) and Schaefer et al of algorithm (c). (c) have to set many control-points on the image in order to avoid image translation. As a result, deformation brings the lack of smoothness while our algorithm can get smoother result than Figure 4. (c) only with fewer control points.

The deformation of fish image is shown in Figure 5 and its continuous frames made by moving a tail can take effect as if the fish swim.

5. Conclusion

In this paper, we implemented the deformation based on RBF. As you saw in experimental results, the various results of deformation were obtained by a term of controlling polynomial of degree. The proposed method is faster by simple calculation and its result is better than the previous methods. The term of controlling polynomial degree has many possibilities for various application fields. Especially, in field of animation, we can select it according to global or partial changes. In this paper, we used two basis functions. Further research will be extended to uses of other basis functions and line or curve segments instead of points as control attribute.

6. Reference

[Ara94a] N. Arad, N. Dyn, D.Reisfeld, Y. Yeshurun,"Image warping by radial basis functions :Application to facial expressions", Computer Vision Graphics and Image Processing , p.p 161-172 (1994).

[For99a] M. Fornefett, K. Rohr, H. Siegfried Stiehl, "Elastic registration of medical images using radial basis functions with compact support", CVPR, p.p402-407 (1999).

[Iga05a] T. Igarashi, T. Moscovich, and J. F. Hughes, "As-rigid-as-possible shape manipulation.", ACM Trans. Graph 2005, 24, 3, pp 1134-1141 (2005).

[Lee06a] Byung-Gook Lee, Yeon Ju Lee, and Jungho Yoon, "Stationary binary subdivision schemes using radial basis function interpolation.", Advances in Computational Mathematics, pp.57-72 (2006).

[Mic86a] C. A. Micchelli, " Interpolation of scattered data : distance matrices and conditionally positive definite functions", Constr. Approx., pp. 11-22 (1986).

[Sch06a] S. Schaefer, T. McPhail, J. Warren, "Image deformation using moving least squares.", Proceedings of ACM SIGGRAPH , pp. 533-540 (2006).

[Toi08a] Wilna du Toit, "Radial Basis Function Interpolation .", the degree of Master of Science at the University of Stellenbosch (2008).

[Wen95a] H. Wendland, "Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree.", Advances in Computational Mathematics, pp.389-396 (1995).

[Wen06a] Y. Weng, W. Xu, Y. Wu, K. Zhou, B. Guo, "2D shape deformation using nonlinear least squares Optimization .", The visual computer , pp.653-660 (2006).

(a) without polynomial (b) with constant (c) with linear (d) with quadratic

Figure 2. Comparison result between different polynomial degree about moving the control-points.

(a) Original image (b) Gaussian basis (c) Wendland's basis

Figure 3. The effects of different radial basis functions.

(a) Original image (b) Our algorithm (c) Schaefer et al [Sch08a].

Figure 4. Comparison between our algorithm and [Schaefer et al].

Figure 5. Deformation of a fish image. Left : original image ; Middle, Right : deformation results

POSTER: Efficient approach to direct B–spline surface rendering by a ray tracing

Aleksandrs Sisojevs
Riga Technical University
Meza Str. 1/3 – 304
LV-1048, Riga, Latvia
alexiv@inbox.lv

Aleksandrs Glazs
Riga Technical University
Meza Str. 1/3 – 304
LV-1048, Riga, Latvia
glaz@egle.cs.rtu.lv

ABSTRACT
There are different approaches to B-spline surface visualization by ray tracing methods. But these methods of computer graphics do not solve this task at an expected level. Therefore in this paper the authors propose to use a new approach for B-spline surface visualization by ray tracing method to render objects with high accuracy and quality.

Keywords
B-spline surface, free-form surface, gradient methods, optimization, ray tracing, rendering.

1. INTRODUCTION
The ray tracing method is a popular technique for rendering high quality images. This paper presents a new approach for high-quality visualization of non-rational B-spline.

There are many approaches to solving this problem.

Martin et al. [Mar00a] describes a framework for ray-tracing of trimmed NURBS using Newton's method. A nonlinear equation system is used for finding the intersection point. The Newton's method is used for solving this equation system.

Nishita et al. [Nis90a] describes an iterative algorithm called Bézier Clipping, used to compute intersections between a ray and a Bézier patch by identifying and cutting away regions of the patch that are known not to intersect with the ray. There are several disadvantages in this approach, like unstable work for some surfaces and problems in solving the case of multiple equivalent intersection points.

Efremov et al. [Efr05a] proposed some modification to Nishita's algorithm for Bézier and NURBS surfaces. But in this case, the number of patches and calculations increases.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

In this work we propose a new approach for finding the values of unknown parameters, which are necessary for visualization of a B - spline surface. The method is based on transforming the problem of finding the roots of equation to the task of optimization. In this case the optimization task is divided into two parts: preliminary search and optimization.

2. PROPOSED APPROACH
The mathematical task of finding an intersection point between the ray and a parametric surface can be described as a nonlinear equations system [Mar00a, Sis08a]:

$$\begin{cases} Q_X(u,v) = Q_X(t) \\ Q_Y(u,v) = Q_Y(t) \\ Q_Z(u,v) = Q_Z(t) \end{cases}, \quad (1)$$

where:

$Q_X(u,v)$, $Q_Y(u,v)$, $Q_Z(u,v)$, – surface equations,

$Q_X(t)$, $Q_Y(t)$, $Q_Z(t)$, – ray equations.

A non-rational B-spline surface can be formulated as [Rog90b, Hea04b]:

$$Q(u,v) = \sum_{i=0}^{n}\sum_{j=0}^{m} P_{i,j} \cdot N_{i,p}(u) \cdot N_{j,q}(v), \quad (2)$$

where:

$P_{i,j}$ – control points,

$N_{i,p}(u)$, $N_{j,q}(v)$ – the B-spline polynomials;

u, v – parameters.

Analogically, a ray can be represented as:

$$Q(t) = \sum_{i=0}^{1} P_i \cdot N_{i,1}(t)\,, \qquad (3)$$

where:

$P_{i,}$ – control points,

$N_{i,1}(t)$ – the B-spline polynomials of the 1st power;

t – parameter.

For calculation of unknown parameters u, v and t, the above equation system should be considered as a case when the ray coincides with one of coordinate axes (Oz). In this case the system (1) is defined as follows [Sis08a]:

$$\begin{cases} Q_X(u,v) = 0 \\ Q_Y(u,v) = 0 \\ Q_Z(u,v) = Q_Z(t) \end{cases}, \qquad (4)$$

There are many approaches for solving the equation system. In our work we transform the task of equation system solving to the optimization task. In order to find the parameters u and v only the first 2 equations should be considered, therefore the function of minimization can be described as follows:

$$w = \left(Q_X(u,v)\right)^2 + \left(Q_Y(u,v)\right)^2 \to \min_{u,v}\,, \qquad (5)$$

We divided the task of function w optimization into two parts: preliminary search and the optimization of the parameters (roots finding).

2.1 Preliminary search

In order to obtain the preliminary values of parameters u and v in every pixel, we propose to create a map of preliminary values. The map consists of the patch data and is coded in RGB channels. The Red channel contains the number of the patch (r=*number*+1). The Green and Blue channels contain the uniform gradient texture fill of the patches. Meaning that in the point $P_{0,0}$ the color value is ($r;0;0$), $P_{n,0}$ – ($r;255;0$), $P_{0,m}$ – ($r;0;255$) , $P_{n,m}$ – ($r;255;255$) and the rest of the gradient texture color values are evenly distributed on the patch surface.

In this case the preliminary values of parameters can be described as is follows:

$$\begin{cases} u_0 = (u_{\max} - u_{\min}) \cdot \dfrac{g}{255} + u_{\min} \\ v_0 = (v_{\max} - v_{\min}) \cdot \dfrac{b}{255} + v_{\min} \end{cases}, \qquad (6)$$

The map is coded using OpenGL graphics library.

The example of preliminary values map is shown in Figure 1.

Figure 1. The example of preliminary values map.

2.2 Root finding

For optimization we chose a gradient method [Him72b]. The values for these parameters are defined at each search step as follows [Sis08a]:

$$\begin{cases} u^{new} = u^{old} - h\dfrac{\partial w}{\partial u} \\ v^{new} = v^{old} - h\dfrac{\partial w}{\partial v} \end{cases}, \qquad (7)$$

where: h – weight working step.

In matrix form this equation system (7) can be formulated as is follows:

$$\begin{bmatrix} u^{new} \\ v^{new} \end{bmatrix} = \begin{bmatrix} u^{old} \\ v^{old} \end{bmatrix} - 2 \cdot h \cdot \begin{bmatrix} \dfrac{\partial Q_X(u,v)}{\partial u} & \dfrac{\partial Q_Y(u,v)}{\partial u} \\ \dfrac{\partial Q_X(u,v)}{\partial v} & \dfrac{\partial Q_Y(u,v)}{\partial v} \end{bmatrix} \cdot \begin{bmatrix} Q_X(u,v) \\ Q_Y(u,v) \end{bmatrix}, \qquad (8)$$

The optimization is carried out until the following condition is met:

$$w \le \varepsilon\,, \qquad (9)$$

where ε – is a small number (in this paper we assume ε<<0,0001).

 ISBN 978-80-86943-95-4

2.3 Fast root finding (modification)

The approach described in section 2.2 is effective for the task of optimization. However, such approach gives an overflow of calculations. This is connected to the necessity of calculating a gradient during each iteration step. At this time it is possible to use the previous direction of a gradient (in case of the function *w* value decrease). In a pseudo-code it can be described as follows:

$w_{old}=w(u_{old}, v_{old})$

$u_{new}=u_{old}-u_{grad}$

$v_{new}=v_{old}-v_{grad}$

$w_{new}=w(u_{new}, v_{new})$

if $w_{new}>w_{old}$ then

(

$u_{grad}=grad_u(u_{old}, v_{old})$

$v_{grad}=grad_v(u_{old}, v_{old})$

$u_{new}=u_{old}-u_{grad}$

$v_{new}=v_{old}-v_{grad}$

)

In order to find an intersection point of a surface with any ray it is sufficient to transform such a ray and the control points of a surface with the help of elementary geometrical transformations in such a way that the ray would coincide with a necessary coordinate axis (in our example - axis Z).

3. EXPERIMENTAL RESULTS

In this work the proposed method, as well as the methods suggested by Martin et al were implemented.

In order to visualize a scene the 1 ray/pixel approach was used. The size of the images in experiment is 512*512 pixels. The experiments were carried out on the computer with CPU Intel Xeon 3,2 GHz, RAM 2 GB. The received images are shown in a Figure 2-7.

As it is possible to see from these figures the proposed methods gives an advantage on quality of the images (there's no distortion on the borders of patches). Image rendering time is shown in Table 1.

Method	"Column"	"Water-lily"
Proposed	74,906	180,234
Proposed (modif.)	*31,156*	*98,313*
Martin et al.	60,750	205,600

Table 1. Image rendering time in seconds

As seen from the table the unmodified proposed method gives unstable results on rendering time (compared to method Martin et all.). While, the modified proposed method results in the fastest rendering time in our experiments.

For comparison we shall consider the time of visualization in percentage, by taking earlier known method (Martin et al) for 100%. The results are shown in Table 2. As seen from the table, the modified proposed method works 2 times faster than the other methods.

Method	"Column"	"Water-lily"
Proposed	123,3	87,7
Proposed (modif.)	*51,3*	*47,8*
Martin et al.	100	100

Table 2. Images rendering time in percentage (Martin et al. – 100%)

4. CONCLUSION

The results (in Figures 2 – 7) seen, that:

- As seen from comparison of Figure 2 (or 3) with Figure 4, and Figure 5 (or 6) with Figure 7, the use of the modified proposed method results in better quality of the image than Martin et al. method. The proposed method has no distortions on the image, while method Martin et al has some faults (like distortion on border of patches).
- The modified proposed method results in faster image rendering time. This method works approximately 2 times faster than other methods.

5. REFERENCES

[Nis90a] Nishita. T., Sederberg. T.W., and Kakimoto M. Ray tracing trimmed rational surface patches, Computer Graphics, vol.24, no. 4, pp. 337–345, 1990.

[Efr05a] Efremov, A., Havran, V. & Seidel, H.-P. 'Robust and numerically stable Bézier clipping method for ray tracing NURBS surfaces.' In Proceedings of the 21st Spring Conference on Computer Graphics SCCG'05:127 - 135. 2005

[Mar00a] Martin, W., Cohen, E., Fish, R., & Shirley, P. (2000). 'Practical Ray Tracing of Trimmed NURBS Surfaces.' JGT 5(1): 27-52. 2000.

[Rog90b] Rogers, D.F. and Adams, J.A. Mathematical Elements for Computer Graphic, 2nd Ed., McGraw-Hill, Boston, MA, 1990..

[Hea04b] Hearn, D. and Baker, M.P. Computer Graphics with OpenGL, 3rd Ed., Prentice Hall, 2004.

[Him72b] Himmelblau, D. Applied Nonlinear Programming, – McGraw-Hill Book Company, 1972.

[Sis08a] Sisojevs A, Glazs A An new approach of visualization of free-form surfaces by a ray tracing method, IEEE MELECON Proceedings, 2008, pp 872-875

Figure 2. The image of an object obtained using the proposed method.

Figure 3. The image of an object obtained using the proposed method (modification).

Figure 4. The image of an object obtained using the method Martin et al.

Figure 5. The image of an object obtained using the proposed method.

Figure 6. The image of an object obtained using the proposed method (modification).

Figure 7. The image of an object obtained using the method Martin et al.

 ISBN 978-80-86943-95-4

The Force-Map Haptic Rendering Algorithm for Drilling into Volume

Y.Liu
School of Computing Sciences,
University of East Anglia,
NR4 7TJ, Norwich, UK
Yu.liu@uea.ac.uk

S.D.Laycock
School of Computing Sciences,
University of East Anglia,
NR4 7TJ, Norwich, UK
S.Laycock@uea.ac.uk

ABSTRACT

With the developments of volume visualization technology for complex data sets comes new challenges in terms of user interaction and information extraction. Volume haptics has proven itself to be an effective way of extracting valuable information by providing an extra sense from which to perceive three dimensional data. The work presented in this paper introduces a Force-Map method that combines the benefits of the indirect and the direct haptic rendering approaches. Users can select from a variety of virtual tools to gain continuous and smooth force feedback during the drilling of volumetric data which increases the applicability of the approach.

Keywords

Volume haptics, Direct volume haptics, Marching cubes, Force-Map haptic rendering.

1. INTRODUCTION

Volume visualization has become widely utilized for many applications. The ability to visualize volume data directly is particularly important for medical applications where a correct anatomical view of the patient can prove vital for surgical planning. Recent developments in graphics accelerator cards have enabled systems to render large and complex volumetric data sets in a variety of different rendering styles aiding the observer's perception of the data. As early as 1993, these visualizations were linked with haptic feedback devices to enable the user to touch the volumetric data. Iwata and Noma used their approach for the haptic interaction of data produced in Computational Fluid Dynamics. In this case a force could be mapped to the velocity and torque mapped to the vorticity [Iwa93]. Virtual Sculpting systems linked to haptic feedback devices have been available for many years; however, these often do not ensure the modified data remains faithful to the characteristics of the original volumetric data. In this paper, a novel approach to drill into surfaces based on the volumetric data is presented.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

One of the major considerations of any application incorporating haptic feedback is the rate at which the calculations must be performed. Based on the results of analyzing human factors, an update rate of 1KHz is required in order for a user to perceive stable and smooth haptic feedback from the system. This is in contrast to the visualization which must update at approximately 30Hz. If the haptic update frequency is lower than 1KHz, an obvious vibration can be felt from the haptic device. One objective of this work is to create a system which can accurately render volume data at sufficient rates for both the visualization and the haptics. To achieve this goal, there are a number of key points that must be addressed. Firstly, the algorithm used to update the volume data must be fast, especially considering the fact that the surface representation of the volume data may be constructed from millions of triangles.

Secondly, the haptic feedback should be rendered such that when the probe point is moving across the voxel boundaries a continuous force is returned to the user. Lastly, since the haptics and visualization calculations will be performed in separate threads, mechanisms are required to ensure that each thread can be updated in a safe manner.

2. PREVIOUS WORK

Volume haptic rendering techniques can be categorized mainly into two classes. The first is indirect haptic rendering, which typically extracts an intermediate geometric representation or iso-surface from the volumetric dataset. The second class is direct volume haptic rendering, which computes the force feedback according to the information stored in the voxels. In this work, the Marching Cubes algorithm [Lor87] is used for extracting a surface from volume data. Once a surface pattern has been identified, for each voxel, a complete triangulation of the volume data will be obtained. There have been several attempts to utilize a localized Marching Cubes algorithm for haptic applications [Kob99]. However, when used in conjunction with haptics force discontinuities often result.

The first category of haptic rendering approaches utilizes a surface-based haptic rendering technique for the interaction with the volume data. In this case an intermediate geometric representation of the volumetric datasets, such as an iso-surface, is constructed. The geometry of the surface can then be easily haptically rendered by utilizing a standard constraint-based method [Zil95]. However, this suffers from stability problems which occur when the surface is updated. The direct volume haptic rendering approach is capable of providing a way to generate force feedback directly from the volume data without extracting an intermediate representation. Even though it is able to represent the force at any position in the volume data, the haptic feedback generated by this method suffers from force instabilities since it is difficult to properly decide the rendering parameters in the force function. This is especially the case when the function is changing during the process, such as when drilling or milling, in real applications. Moreover, forces may vary significantly in strength and direction which sometimes can not be represented by a simple mapping method.

Many previous efforts have been done on these issues which limit the usability of direct volume haptics. The Voxmap Point-Shell algorithm was introduced by McNeely et al. [Mcn99] in which the three dimensional volume data is represented using a Voxmap. The virtual three-dimensional object is encoded into the Voxmap using voxels that identify which locations in space are occupied. The haptic device is modelled by a number of points that sit around the surface. A force response is then calculated for the haptic device based on the voxel-point intersections. However, the force calculation method suffers force discontinuities at voxel boundaries. Pflesser and Petersik et al. [Pet02] also proposed a haptic system for virtual temporal bone surgery which uses a modified version of the Voxmap-Pointshell algorithm [Mcn99]. Their approaches sample the surface of the drilling instrument and then generate appropriate forces at each sampled point. A number of samples are distributed around the drill and a ray-tracing approach is then employed to calculate the force vectors towards the tool centre, which can subsequently be combined to generate the overall force returned to the haptic feedback device. The ray-tracing algorithm has the potential to miss voxel data located between two rays due to an insufficient sampling.

Morris et al. [Mor04] simplifies the computations for drilling through the use of another point-shell method to compute haptic interactions and bone erosion for spherical drill bits. In contrast to the work of Pflesser et al. [Pet02], Morris et al. use the data within the spherical tool to perform bone removal as opposed to sampling points on the tool's surface. Both of these approaches limit the user to drilling with a spherical drill. Eriksson et al. [Eri05] proposed a haptic milling surgery simulator using a localized Marching Cubes algorithm for the visualization. To improve the stability they employed a direct haptic rendering method with mechanisms to remove fall-through issues. The data inside the virtual drill is set to a vector pointing to the centre of the voxel. The output force is the sum of all those vectors. This approach works well when the drilling tool moves in a small area, but a "kicking" would result when the haptic test points move across the cubes' boundaries.

To overcome the stability issues discussed in the previous papers, the work presented in this paper utilizes a fast local Marching Cubes algorithm and a Force-Map haptic rendering procedure to gain smooth and stable force feedback during drilling with more varied drilling tools.

3. SURFACE EXTRACTION AND MODIFICATION

The volume-based representation is a natural choice for rendering a collection of digital images produced by medical scanning technologies such as Magnetic Resonance Imaging (MRI) or Computed

Tomography (CT). There are a variety of graphical rendering techniques for visualizing the three dimensional data, often with options to display the material properties such as density and viscosity within the voxels. This has the potential to greatly enhance a user's performance in medical and scientific three dimensional data exploration.

When using the Marching Cubes algorithm [Lor87], a volume can be interpreted by generating polygons representing the surface, typically constrained to a specified value of the data. To handle large data sets, an Octree-based structure is employed which enables the data to be changed dynamically in an efficient manner.

The efficiency of the volume data modification step is dependent on the volume of the tool. As the user manipulates the haptic device the tool moves accordingly through the three dimensional data. A bounding box around the tool is constructed, which is aligned to the coordinate system of the volume data. The data points within the bounding box are tested to determine if they are inside the tool's volume. If the points are interior to the volume then their density value will be reduced based on the distance to the centre axis of the cylinder or a point in the case of the sphere. This gradual data value adjustment permits the smooth removal of volume data. Once adjusted, the Bounding box volume will be updated using the local Marching Cubes algorithm to generate a new surface.

4. HAPTIC RENDERING

The approach described by Eriksson et al. [Eri05] suffers force discontinuities when the tool moves between the encoded cubes. Sample points in this work are tested for contact with the volume data.
Given a sample point position, a vector calculated from the occupancy force-map can be output. By using this method, the force feedback is stable and smooth even though it has a similar force cube encoding system. The force vectors stored in the data are calculated based on the local surface, which also benefits from the advantages of the surface-based haptic rendering approach. The synchronisation of updating the graphic and haptic loops enhances the fidelity of the virtual visual-haptic system when applied to real applications. The three steps below outline the Force-Map haptic rendering method adopted for a surface representation of dynamically changing voxel data.

Initialization and construction

Initially all the normals of the triangles contained in each octree leaf node (voxel) are averaged to result in a single force vector representing the data in the voxel. The larger the voxel is, the more volume data points lie within it. Additionally, only the data inside the voxel is assigned to a force vector while others are set to none. After this initialisation step, all the data near to the surface is set to a force vector which approximately equals the closest surface normal.

Update and reconstruction

When the surface is updated in the haptics thread the data points that are found to lie inside the new voxel are set to a force value based on the triangle's face normal. If there is more than one triangle in the voxel, the averaged face normal will be used. Some force values in the old surface might also need to be updated since the triangles forming the surface in the voxel may have changed.

Calculating the force feedback

The force vectors stored in the data must be combined appropriately before being returned to the haptic device. When the virtual drilling tool moves into the volume data, a haptic test point checks the surrounding eight data values in the three dimensional space. These eight data values are referred to as the force cube in this work. The corners of the force cube contain the force vectors stored in the data. Tri-linear interpolation is employed here to enable an interpolated force vector to be calculated for any position inside the force cube. Another advantage of using the tri-linear interpolation method is that the haptic test point can be smoothly moved from one force cube to another without any force discontinuities occurring between them.

(a) (b)

Figure 1. Force-Map haptic rendering, the red arrows represent the force vector. (a) The yellow square indicates one force cube represented in two dimensions. (b) The same single force cube in three dimensions.

For any real application, drilling with a single point does not lead to a realistic result. An approach involving multiple test points approximating the drilling tool is usually preferred. In this work, a number of hapic points are distributed approximately

 ISBN 978-80-86943-95-4

around the surface of the drilling tool. At each time step, each haptic point is tested in the constructed Force-Map to calculate the contribution to the overall haptic force. In many applications, the properties of the simulated materials differ depending on the location being drilled. This is particularly the case in medical and dental applications where the material properties of each voxel must be considered. For example drilling through soft tissue should be very different to drilling through rigid bone. The Force-Map method can easily incorporate this issue by simply setting a scaled force vector where the scaling factor is related to the neighbouring voxel data.

5. RESULTS

Figure 2 illustrates a procedurally generated sphere along side a surface representation of a human pelvis. The surface was extracted from 87 CT slices obtained at the Norfolk and Norwich University Hospital, UK.

Figure 2. The left sphere-like object is created procedurally whilst the right hand image was extracted from 87 CT image slices. Each slice contains 256 x 256 pixels.

The volume on the right of Figure 2 has been calculated from a data set comprising of 256 x 256 x 87 data points. For most drill tools tested, the haptic rendering loop is updated at a rate which exceeds 1000Hz. Figure 3 shows the time required to perform the surface modification and Force-Map updates during rendering. Table 1 illustrates the average update times during drilling for the selection of tools. The sphere drilling tool is the most efficient since its volume occupies on average the fewest voxels.

Figure 3. A graph presenting the time taken to update the surface during drilling with a sphere

Sphere	Cylinder	Combination	Dental Drill
1550Hz	877Hz	770Hz	1378Hz

Table 1. A table showing the average haptic rates whilst drilling with a selection of tools.

The work has been tested on a Quad Core 2.4GHz processor PC with a NVIDIA Geforce 8800GTX graphics card. To provide haptic feedback a PHANToM Omni device has been employed.

6. CONCLUSIONS

In this paper a Force-Map haptic rendering algorithm is proposed to achieve real-time drilling of volumetric objects. In order to gain more realistic force feedback for drilling applications, arbitrary tool model selection has been implemented in this work, for tools based on implicit equations. The future work will implement the tangential force which is an important property of the drilling application by using the Force-Map haptic rendering algorithm.

7. REFERENCES

[Eri05] Eriksson, M., Flemmer, H., and Wikander, J. A Haptic and Virtual Reality Skull Bone Surgery Simulator. Proc. WorldHaptics, Italy, 2005.

[Iwa93] Iwata H., Noma H. Volume haptization. In Proc. of IEEE Symp. On Research Frontiers in Virtual Reality, pp. 16–23, 1993.

[Kob99] Kobbelt L, Botsch M, Schwanecke U, Seidel P. Haptic volume rendering with an intermediate local representation. pp. 79–84.

[Lor87] Lorenson W.E., Cline H.E. Marching cubes:A high resolution 3d surface construction algorithm. SIGGRAPH. 21, pp. 163–169, 1987.

[Mor04] Morris D., Sewell C., Blevins N., Barbagli F., and Salisbury K. A collaborative virtual environment for the simulation of temporal bone surgery. In MICCAI, 2004.

[Mcn99] McNeely WA., Puterbaugh KD., and Troy JJ. Six degree-of-freedom haptic rendering using voxel sampling. Proceedings of ACM SIGGRAPH, pp. 401-408, 1999.

[Pet02] Petersik A., Pflesser B., Tiede U., Hohne K. H., and Leuwer R. Haptic volume interaction with anatomic models at sub-voxel resolution. In IEEE Symp. Haptic Interfaces for Virtual Environment and Teleoperator Systems, pp. 66–72, 2002.

[Zil95] Zilles C., Salisbury J. A constraint based god-object method for haptic display. In Proc. of the IEEE Conference on Intelligent Robots and Systems, pp. 146–151, 1995.

3D Navigation Systems based on Synthetic Texturing

[1]Behnam Alizadehashrafi
[2]Alias Abdul Rahman
Department of Geoinformatics
Faculty of Geoinformation Science and Engineering University Teknologi Malaysia 81300
[1]behnam5us@gmail.com
[2]alias@utm.my

Volker Coors
Stuttgart University of Applied Science
Schellingstr. 24
70023 Stuttgart, Germany
volker.coors@hft-stuttgart.de

Thorsten Schulz
Stuttgart University of Applied Science
Schellingstr. 24
70023 Stuttgart, Germany
thorsten.schulz@hft-stuttgart.de

ABSTRACT
Navigation system is the order of the day to pilot people to their destination. This paper focuses on the uses of mobile devices such as PDA's and smart phones without additional hardware to direct pedestrians based on synthetic texturing along with 3D modeling. Each facade is reconstructed by arraying small sized textures in respect to their geometries in different layers. In the process of texture generation cropping, rectifying, removing disturbing objects and exposure setting should be done in advance. Unlimited number of layers with different priorities and their horizontal and vertical pulse functions and texture files can be utilized for creating a simple square looking facade. Each 3D model is created by mapping the synthetic textures on the 3D geometries of each building's 3D model. The processes to create the synthetic textures as well as their usage in mobiles' context are described in detail in this paper.

Keywords: Synthetic texturing, pulse function, 3D urban modeling, navigation system, lightweight geometry

Introduction
Different kinds of road presentations in navigation system devices are in use. They are based on 2D map, text, voice and pictures. People like to identify their world in the navigation system display to compare with their position and environment. MoNa3D[1] is carrying out a project by use of synthetic texturing and pulse functions for 3D navigation purpose. The two main goals of the project are providing a cognitive semantic route description by using landmarks in 3D, which allow context dependent personalized navigation support and developing an approach to create suitable non-photorealistic building textures using image processing methods and synthetic textures along with a corresponding compression system for an efficient storage and transfer of 3D building models (Coors and Zipf, 2007). In particular, the generation and usage of synthetic textures are addressed in this paper.

The Motivation
The idea of having 3D navigation software on smart phones and PDA's along with limited resources on these devices, leads us to generate a small program with less overhead and high performance by means of combining photo realistic with pure synthetic approaches. For this issue an efficient rule of thumb is repeating a small and high quality texture in arbitrary vertical and horizontal direction. The texture image file even can be a pixel which is actually a color map. The impression of having unlimited number of layer with different priorities,

[1] MObile NAvigation 3Dimentional

gives us the opportunity to generate any kind of complex and historical facades. The aim can be achieved from the behavior of raster based applications.

Related works

Using landmarks in mobile navigation systems could improve way finding approach. In this field enriching way finding instructions by means of local landmarks is a reliable method. Simple instructions are defined like geometric data for street network and shape, color, visibility for a determinate facade as visual attraction parameters. Semantic attraction parameters of outstanding landmarks are defined as cultural and historical aspects, Level of importance and explicit marks e.g. street signs. By using these definitions and information, landmarks are extracted from dataset and provide Point of Interest (POI). In fact the POI is a hard coded and predefined data which is geo coded in spatial datasets see Raubal and Winter (2002).

The decision for implementing 3D content into navigation system rose at the end of the nineties after a huge progress in hardware designs. First of all, generating 3D model of the urban area is necessary which can be achieved by creating wire frames out of filtered point clouds. Nowadays this procedure can be done semi-automatically out of laser data or stereo image pairs and photogrammetric methods (e.g. Epipolar match points and so on). High quality and rectified photographs from geometries are necessary for mapping on the wire frames. For this aim, disturbing objects must be removed from shadow less photographs. Finally mapping process can be done see Schulze and Horsel.

Figure 1: Mapping the images after rectification, obstacles removal and radiometric adjustment on the light weight geometry of Building 2" of Stuttgart University of Applied Science in VRML 2.0

The problems of conventional photogrammetry methods are disturbing objects, leaning geometries, shadows, and reflection, time consuming, expensive and "*heavy"* models, which make it unfit for the intention of navigation system devices. Another project for 3D navigation system purpose on smart phones, is M-LOMA[1] Nurminen, A and Tuominen, J (2008). This application which is programmed in C++, can be installed on smart phones and PDA's for running VRML file. The visible part of the model can be rendered on the screen. Moreover lightweight geometries are used for the 3D modeling. In this issue VRML parsing is the first step in implementation process for only visible area on the screen. Texture processing for different LOD's comes afterwards which are created and stored separately. In the next step the visibility calculations based on PVS[2] algorithm is done. Then the visibility list encoding can be followed by geometry files packaging and compressing them in binary format.

Our aim is to generate a noble solution e.g. by enabling the creation of high quality and easy to describe facades for each object in respect to low memory and storage requirements.

Concept of synthetic textures and pulse functions Parish & Müller (2002)

Each building consists of polygons and for each polygon there are some pulse functions in x and y axes which controls the process of texture generation. Pulse function can be easily defined with the logical values. In intersection of X and Y axes, TRUE means insert and FALSE means not to insert the texture (Coors and Zipf, 2007). In order to increase the flexibility, the number of layers of textures and pulse functions are unlimited and each of them has its own priority in respect to others. In the process of creating database for synthetic texturing method, we have to define lots of fields like number of facades which are generating an identical building, texture file name and path or directory; pulse functions and their related parameters and so on. In MoNa3D, for this issue we have defined an XML schema and we applied the same schema in this work Bauer et al (2008). Furthermore, we have created a user friend interface by means of JavaScript and ActiveX objects of Microsoft windows which can receive the parameters and generates XML file fitting to this schema. By using this JavaScript program (see Figure5(c), the distances between similar textures in horizontal and vertical direction are computed semi automatically in respect to the real geometry for pulse functions. The number of the layers for each facade can be defied by operators in respect to the

[2] Mobile LOcation Aware Messaging Application

[2] potentially visible set

requested quality or number of different and sparse textures. Figure 2 shows the generation of facade (in five layers).

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

Figure 2: (a) illustrates 4 different layers of windows along with their related pulse functions and (b) is the real resized and rectified image of the facade used for measuring process and (c) is the final model without leaning problem and (d) is the rectified form of real image of facade.

Synthetic texturing

Identifying and describing the texture

We can extract texture heuristically according to the form of the facade. Normally it is possible to find a section on every building facade which is repeating exactly with the same size, form and shape. This part should repeat in horizontal or vertical direction with the same distance and often equal size. It depends on the taste of the operator and there is no strict rule and automatic method for selecting textures. It can be a vertical slice which is repeating horizontally or horizontal slice which is repeating vertically with a distance equal to zero or any calculated value in respect to the generated XML schema. Texture of the wall can be stretched in vertical direction or horizontal direction or can be a color map or a texture which is repeating in both directions [Bau06]. The texture of the window has also the same behaviors except stretching in both directions(see Figure3 (b)).

(a) (b) (c)

Figure 3: (a) is the rectified and resized facade (c) is the texture which is extracted from upper part of (a) and repeated in the upper part of (b) in vertical and horizontal direction to generate the model.

Cropping and rectifying

After deciding which part can be extracted as our texture we have to follow cropping and rectifying procedure. For instance Figure3 (c) is the window texture which is extracted from upper part of Figure 3 (a) and repeated in both directions in order to generate the model (see Figure3 (b)). In many cases we need rectangle texture and the single point perspective effect of the camera can be removed by means of cropping along with rectification of the rectangle texture shape. In order to avoid leaning problems, it is also possible to crop and rectify the facade of the building and then crop the selected texture which is as perpendicular as possible to the exposure angle. For instance the window texture in Figure 3(c) was cropped from the rectified and cropped facade which is as perpendicular as possible to the exposure angel (see Figure 3(a)).

Removing the disturbing objects

To have a high quality texture we need to remove disturbing objects and undesired parts of the texture. For example the flying birds, moving cars, flags, signs and moving pedestrians should be omitted from the selected texture.

(a) (b)

Figure4: (a) is a vertical slice texture which is cropped and rectified from the perpendicular picture of the facade and (b) is the same texture after removing the disturbing objects (e.g. flying birds, moving cars and moving pedestrians and so on). To reduce the complexity, the inside geometries were neglected as well.

Radiometric adjustment and texture size settings

 ISBN 978-80-86943-95-4

In order to increase the quality of the model some additional exposure settings, radiometric adjustments, color settings, contrast settings and also texture size settings, were carried out. The texture size is an important issue for generating output image file for different applications. Our Java program generates a square image file according to our Pulse Function and it is possible to change the size of output image file in Java program for each façade even to any shape. In order to have a simple and user-friendly Java program, the size of output image file is defined as a fixed size (256*256 for popular mobile navigation systems and 512*512 for the test on computer screen and VRML.2 environment) in the respective utilized hardware and software environment.

Generating the facade semi automatically

Among the issues in generating facade semi automatically such as the geometries of the selected textures, two different issues can be preceded. Both issues depend on the form of textures and taste of the operator for choosing the texture. In the following examples we will illustrated these aspects in detailed.

(a) (b) (c)

Figure 5: (a) is bird's view option of Microsoft Virtual Earth from main station in Stuttgart. (b) is the generated model of front part without knowing the Geometry of the texture in respect to the façade and (c) is the JavaScript user friendly interface for calculating the parameters and generation XML file which can be run in internet explorer

In fact, the generated texture in Figure 4 is repeated 25 times in horizontal direction. The distance between them are equal to zero. The out put image file is created without having knowledge of any parameters about the selected texture in respect to the facade.

In the next example in order to measure the parameters of texture in respect to the facade, we need to rectify and resize the original picture of the facade taken from any angle (see Figure 6 (b)) without having quality and completeness as shown in Figure 6 (a). Lets assume that window is our texture, now we can use different applications like Photoshop to measure the parameters like width and height of the window, starting point from left side, from upside, vertical and horizontal distance between them, starting point of the door from upside, from left side and so on. Figure 7 shows the comparison of the generated model of Figure 5 (see Figure 7(b)) with bird's view option of Microsoft Virtual Earth (Figure 7 (a)).

(a) (b)

Figure 6: Right image (b) is a picture of facade from any angel and the left image (a) is rectified and resized form of (b) for measuring purpose.

(a) (b)

Figure 7: The comparison between the models of the buildings of University of Stuttgart (b) with bird's view option of Microsoft Virtual Earth (a).

Generating lightweight geometry

For heavy geometry on each surface we need to map a texture. To decrease system overhead for mobile navigation system, generating lightweight geometry is necessary. The geometries as the roofs and floors are less important and in many cases we do not need to map any texture on them except some special cases. For instance The *Mercedes isometric Star* of the main station which is rotating could also be helpful to identify the building (see Figure 9(b)). Normally, a small texture can be repeated or stretched or even one pixel color map would be sufficient to generate the output image file to map on the light weighted geometry of roof. Removing unnecessary geometries is based on the taste of the operator which might be done heuristically. This is the main reason that we couldn't define any

semiautomatic method for this issue and this step has been done manually. For instance for the front part of the main station (in our existing VRML file) there were 20 points for the walls and we reduced them to four points in outside which are enough for the mapping purpose (all the models are for outdoor 3D navigation systems and indoor geometries should be omitted). In addition, after removing the small geometries we can represent them with different radiometric adjustment and exposure settings (in this case the user can "feel" them as a small geometry (see Figure 8)).

Figure 8: The representation of small geometries with different exposure settings and radiometric adjustments.

Mapping the output image on the geometry

In the process of mapping output image on the geometry, we have defined the output image file of Pulse Function as a square shape. We noticed that if the height of building is less than half of the width of the building or width of the building is less than half of the height of the building, we will face with deformation and lack of quality in the process of mapping square output image on rectangle geometry. In order to deal with this problem we generated the output image for the geometry in a part of output image file and then used that section for the mapping and stretching on the geometry (Figure 9). In addition to above mentioned method we can also use different resolutions and shapes or transparent in our Java program to deal with the problem of deformation or lack of quality and strict square shapes.

Figure 9: The mapping and stretching a part of output image for the geometry.

Test environment

25 buildings were modeled from the Stuttgart University of Applied Science to the main train station in Stuttgart and represented in VRML2 as are illustrated in Figure 10 (buildings around white line as a shortest path). The walls of some of models are generated with just one pixel or one color map which is repeated in vertical and horizontal directions. Many of other historical buildings like "Building 1" of the Stuttgart University of Applied Science was generated with a vertical slice of texture (which was repeated in horizontal directions (see Figure 11). In some of buildings' models the pictures of the facades are resized and placed or mapped on the related geometry with respect to the drawing rules and level of details.

Figure 10: The white line represents the path of the 25 buildings, modeled from main station to the Stuttgart University of Applied Science.

Conclusions

Synthetic texturing is one of the outstanding methods for the aim of 3D modeling of urban area which has been proven in this procedure. The pulse functions are very useful in texturing the facades in respect to their own priority in the representation for 3D mobile navigation system devices with high quality and small size for the 3D models. The concept of having unlimited number of layers with their own pulse functions and priorities, provides high flexibility and easy to generate any kind of desired façade for each geometry in the 3D model. Furthermore, this method is not time consuming compared to conventional photogrammetric methods.

Figure 11: The model for "Building 1" of Stuttgart University of Applied Science via VRML2 and used textures.

ACKNOWLEDGMENTS

We appreciate the Stadtmessungsamt Stuttgart for providing the data and Federal Ministry of Education and Research for enabling the project. We also appreciate the staff of Stuttgart University of Applied Science that they supported us during this study.

REFERENCES

Bauer, M. & V,Coors. & T.Schulz &A.Zipf (2008) Zur Nutzung von 3D-Stadtmodellen für mobile Navigationssysteme

Coors, V. & A.Zipf (2007): MONA 3D -- Mobile Navigation Using 3D City Models

Nurminen, A. & J.Tuominen (2007): Mobile LOcation-aware Messaging Application

Helsinki University of Technology Industrial Information Technology Laboratory-P.O. Box 9204, 02015 HUT, Finland-MSc

Parish, Y.I.H. & P. Müller (2002) "Procedural Modeling of Cities," *Proc. Siggraph,* ACM Press, 2001, pp. 301–308.

Raubal, Martin &S. Winter (2002): Enriching Wayfinding Instructions with Local Landmarks - Institute for Geoinformatics, University of Münster - Vienna University of Technology, Austria

Schulze, M. –Horsel (2007): 3D Landmarks– Generation, Characteristics and Applications. ISPRS Conference – 3D Architecture 2007.

POSTER: Robotic Visual and Inertial Gaze Control using Human Learning

José Prado
University of Coimbra, Portugal
jaugusto@isr.uc.pt

Jorge Lobo
University of Coimbra, Portugal
jlobo@isr.uc.pt

Jorge Dias
University of Coimbra, Portugal
jorge@isr.uc.pt

Figure 1: Robotic Head

ABSTRACT

Humans make use of inertial and vision cues to determine ego-motion. Bayesian models can be used to represent the human behaviour to be used in a robot. An environment may be composed by an infinite number of variables and humans deal with some of them each time a motor decision needs to be taken.

1 INTRODUCTION

One of our main challenges was finding variables that model the environment in way similar to humans. Of course it was needed to start from accepting some assumptions. In this work we are only considering visuo-inertial information. The visual information that we use is the mean of sparse optical flow based on feature tracking. We show that the mean flow is strongly related with the robotic ego-motion but alone it cannot deal with ambiguous situations that arise. Further we concluded that the fusion between inertial and visual information solves ambiguity problems.

Ambiguity is another problem that happens even on humans, if we look only to a black wall and someone shake our head, we could not know that the head was moving only by using the vision, but we are sure our head is being shake because of our vestibular system. In our robotic platform, the same ambiguity happens, when turning off the lights of the room the robot will act only with the inertial information.

Inertial sensors explore intrinsic properties of body motion. From the principle of generalised relativity of Einstein we known that only the specific force on one point and the angular instantaneous velocity.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

No other quantity concerning motion and orientation with respect to the rest of the universe, can be measured from physical experiments inside an isolated closed system. Therefore from inertial measurements one can only determine an estimate for linear acceleration and angular velocity. Linear velocity and position, and angular position, can be obtained by integration.

2 RELATED WORK

Fusion of inertial and visual information was done in [1] and good results were obtained. Unit sphere projection camera model was used, providing a simple model for inertial data integration. Using the vertical reference provided by the inertial sensors, the image horizon line could be determined. Using just one vanishing point, it is possible to recover the camera's focal distance. In a typical indoor corridor scene, the vanishing point can also provide an external bearing for the system's navigation frame of reference.

The integration of inertial sensors can reduce ambiguities and improve robustness of structure from motion methods. The dual problem of motion estimation from observed structure has long been pursued. It is also said in [2] that these sensors have useful complementarities, each able to cover the limitations and deficiencies of the other. [3]

3 PROPOSED APPROACH

Our robotic implementation of gaze control basically perform visuo-inertial servoing. But the sensory inputs are being processed using Bayesian inference using the current instantaneous data and a previous probabilistic learned table.

The image optical flow provides motion data from the visual sensor. We analysed and implemented a variation of Lucas Kanade sparse optical flow building upon the ideas in [4]. Thus, we also decided to rely in the following assumption: the information from the image motion which is strongly related with the human decision when reacting to some stimulus is the mean flow.

It is intuitively preferable to have a real time application, since the idea was to show the flow arrows in the screen and also the inertial data on the fly to influence the human during the learning phase. Thus, performance challenge was instantaneously added to our project idea together with the purpose of using a robust

 ISBN 978-80-86943-95-4

flow algorithm when dealing to changes of lighting, size or image motion.

3.1 Sparse optical flow

We used an algorithm based on a pyramidal implementation of the classical Lucas-Kanade algorithm. An iterative implementation of the Lucas-Kanade optical flow computation provides sufficient local tracking accuracy. RANSAC was implemented to remove the outlier vectors and then the meanflow was calculated with the remaining flow vectors (fig. 2).

3.2 Bayesian Model of Visuo-vestibular-Based Gaze Control

The contribution in the mean flow calculation allows a better characterisation of the human decision.

The purpose is about reaching to a feasible Bayesian model for a robotic gaze control following the ideas presented on[5] and [6]. From the camera image we extract Fd^t and Fa^t which are the direction and the amplitude of the mean flow. From the IMU(Inertial Measurement Unit) we get the angles (Roll, Pitch and Yaw) that are further shown in this paper as R^t (for Roll), P^t (for Pitch), and Y^t (for Yaw).The actuator control acts based on current angular position and on instantaneous flow information of the system. The pan-tilt unit are controlled with combined commands for target position and velocity. The motors move to the desired target position with the selected velocity and stop. The motor model takes this into account by having the current motor command depending on the current state and also on the probabilistic table filled out with the human reaction information.

The following variables are used:

- S_t: is a tuple with the following four variables transformed in possible motor reactions
  - R^t: (roll) angle of the human-reaction for a given state
  - Y^t: (yaw) angle of the human-reaction for a given state
  - Fd^t: direction of the vector of the mean flow (comes from vectored product between u and v) (Radians)
  - Fa^t: amplitude of the vector of the mean flow
- M_t: is a movement variable with the following scope (UP, DOWN, LEFT, RIGHT, STOP)
  - The five states of M_t are concluded by doing atomisation of the raw values in the following variables
    - pan motor velocity: $\mathscr{P}_\omega$ — pan motor target position: $\mathscr{P}_\theta$
    - tilt motor velocity: $\mathscr{T}_\omega$— tilt motor target position: $\mathscr{T}_\theta$
- H_t: is the human reaction to be learned (UP, DOWN, LEFT, RIGHT, STOP)

To simplify notation, state variables are grouped in a vector $S = (R^{0:t}, P^{0:t}, Y^{0:t}, Fd^{0:t}, Fa^{0:t})$ and motor variables are considering to be in the range U,D,L,R,S after atomisation from $M = (\mathscr{P}_\omega, \mathscr{P}_\theta, \mathscr{T}_\omega, \mathscr{T}_\theta)$. The Bayesian program that show the relation between these variables is shown in (fig. 3).

4 EXPERIMENTAL PARADIGM AND PROTOCOLS

4.1 Apparatus and Stimuli

We want to use a HMD (Head Mounted Device) to pass the visual stimulus from the robot to the human subject. However currently we used the input from the keyboard reflexes when the subject looks to the screen. Our robotic head (fig.1) is a common platform consisted by many sensors, but basically those that we are using are a stereo camera, a pan-tilt unit and a inertial sensor. Those sensors are attached in the same structure. Thus, every motor command sent to the pan-tilt unit will reflect on IMU (Inertial Measurement Unit) and also in the camera images. Consequently the sent motor commands will also have a direct influence on the calculated optical flow and inertial data.

.

Attitude Heading Reference System (AHRS) are 3-axis sensors that provide heading, attitude and yaw information for aircraft. AHRS are designed to replace traditional mechanical gyroscopic flight instruments and provide superior reliability and accuracy. We used a low grade AHRS that is a Xsens MTi , which uses a combination of 3-axes accelerometers, gyroscopes and magnetic sensors to output estimates of its own orientation in geo-referenced coordinates. The AHRS orientation can be given in the form of a rotation matrix ${}^{\{\mathscr{W}\}}\mathbf{R}_{AHRS}|_i$ which register the AHRS sensor axes with the north-east-up axes. The camera is a videre STH-MDCS3-9cm stereo camera. Videre Design makes stereo imaging hardware for real-time 3D imaging using the triangulation principle. Stereo images are transferred to a PC using the IEEE 1394 (FireWire) bus.

4.2 Subjects

Five human subjects with normal working visual and vestibular systems. In those subjects with visual distortion this should be compensated by using glasses or lens, thus the distortion perform no impact to the experiment. We will mix male and female according to availability of the persons. The subjects should be at least three naive and two authors.

4.3 Protocol

By using the HMD to give the robot images to the human eyes the visual connection becomes direct between the robot and the human. We can not inject artificial inertial sensor data into human brain. Thus, what is possible to be done is a indirect correlation where during the tests, the human will use it's own vestibular system while the robot will use the artificial inertial system. Gray scale images are the input, with several visual detectable features on the environment. Visual Features are necessary by the human brain to have notion of motion. If a human is moving sideways in front of a perfectly white wall, once the acceleration stabilises subject will have no sensation of motion. However if this same wall is full of visual detectable features, human will naturally detect the motion only by the visual influence. We have the same response on artificial optical flow algorithms, and that's why we are interested on considering the optical mean flow as a artificial visual ego-motion notion measurement variable.

5 RESULTS AND CONCLUSION

A first version of human-learning was implemented, using keyboard to control the robot while monitored by a human (human in the loop way of learning like in [7]). We still want to improve our way of learning by using a helmet equipped with camera and IMU and then detecting real human neck movements.

Consider that we numbered our random variables as follows:

1. is $Roll^t$, sub-variable of Imu^t variable
2. is $Pitch^t$, sub-variable of Imu^t variable
3. is Yaw^t, sub-variable of Imu^t variable
4. is Fd^t, the Flow Direction
5. is Fa^t, the Flow Amplitude

The learned table (fig. 4) is a 4D probability table with dimensions [36x5x10x5] in our test, we plotted this in five 3D graphics in order to be possible to visualise them.

It is possible to observe that for the UP, DOWN, LEFT and RIGHT movements, the main categorising random variable is Fd^t, in the other hand for the STOP movement Fd^t is very confusing, thus Imu^t will be much more useful categorising this decision.

Testing the reaction of the system

Fake stimulus were injected into the system to measure if the robot's reaction would be like expected for that stimulus. As human trained the system, we know (approximately) which stimulus to create and which reaction to expect. For example if we train a walking robot

Figure 2: Sparse Flow Experimental Test

not to fall from a step, we can put a step in front of it and our expectation will be that the robot do not fall. In our case we trained the head to be centred and then we give stimulus simulating that the head would moving to one or another side "forever" during each test. We also gave stimulus for the system to believe the head was flying up like a rocket and also falling down in free fall. It was performed 100 trials with different stimulus for each expected reaction.

Program
- Description
  - Specification
    - Variables:
    - $S^{0:t} = (R^{0:t}, Y^{0:t}, Fd^{0:t}, Fa^{0:t})$: state variable
    - $M^{0:t}$: motor movement variables, scope {U,D,L,R,S}
    - $H^{0:t}$: human reaction variables, scope {U,D,L,R,S}
    - $Fd^{0:t}$: flow direction (Rad)
    - $Fa^{0:t}$: flow magnitude (pixels)
    - Decomposition:
    - $P(S^t \wedge M^t \wedge H^t \wedge Fd^t \wedge Fa^t =$
    - $P(R^t.P(y)^t)$
    - $.P(H|R \wedge y)$
    - $.P(Fd|H \wedge R \wedge y)$
    - $.P(Fa|Fd \wedge H \wedge R \wedge y)$
    - $.P(M|Fa \wedge Fd \wedge H \wedge R \wedge y)$
    - Parametric forms:
    - $P(M^t) = P(H^{0:t}|S^t)$
  - Identification:
  - Bayesian human-learned gaze control.
- Question:
- $P(M^t|S^t \wedge H^{0:t})$

Figure 3: BayesianProgram

One stimulus for each reaction would be the trivial case to categorise, but for this preliminary results we had approximately 98% of correct decisions in 500 different stimulus to be categorised in 5 movements.

6 FUTURE WORK

In this work we performed a learning based on human reactions, and the Bayesian program we proposed efficiently models the ego-motion with visuo-inertial information. Adding non-representative variables to the Bayesian model may not only waste performance, rather it may also cause confusion to the algorithm when trying to categorise the correct reaction to take.

The system is running real time at 3fps, we did not use a real time operational system, so this frequency

(a) UP (b) DOWN

(c) LEFT (d) RIGTH

(e) STOP

Figure 4: Probability Table - Learned data -

(a) Falling down in free fall (simulation)

(b) Launched up as a "forever" rocket to the sky "forever" (simulation)

(c) Translating to left side "forever" (simulation)

(d) Translating to right side "forever" (simulation)

(e) Stopped

Figure 5: System reaction (Imu^t vary even stopped because of Mag^t mainly)

may vary a little according to our non real time UNIX system. Learning of motor behaviours was done by our Bayesian program and it allowed fusion between inertial and visual input.

For future work, we intend to perform more experiments with different environments. Restrict the environment, for example a white point in a black background would be a way to push the human to pay attention in something near to the flow, rather then in other things that we cannot predict. Another desired set up is adding more moving objects into the scene it will be interesting to verify what happens if we try to exceed the gross errors that our RANSAC implementation is able to filter. Also we want to improve the speed of the system motor reaction, a new hardware platform is already being constructed. This new platform will have a binocular active vision system witch will allow stereo vergence with an adjustable base line, with common head tilt and neck pan, mimicking the human degrees of freedom.

ACKNOWLEDGMENTS: The authors gratefully acknowledge support from EC-contract number BACS FP6-IST-027140, the contribution of the Institute of Systems and Robotics at Coimbra University and reviewers' comments

REFERENCES

[1] Jorge Lobo and Jorge Dias. Vision and inertial sensor cooperation using gravity as a vertical reference. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(12):1597–1608, December 2003.

[2] Peter Corke, Jorge Lobo, and Jorge Dias. An introduction to inertial and visual sensing. *The International Journal of Robotics Research (IJRR) Special Issue from the 2nd Workshop on Integration of Vision and Inertial Sensors.*, 26(6):519–535, June 2007.

[3] Jorge Lobo and Jorge Dias. Relative pose calibration between visual and inertial sensors. *The International Journal of Robotics Research (IJRR) Special Issue from the 2nd Workshop on Integration of Vision and Inertial Sensors.*, pages 561–575, 2007.

[4] Andrew Lookingbill, David Lieb, David Stavens, and Sebastian Thrun. Learning activity-based ground models from a moving helicopter platform. *In Proceedings of the International Conference on Robotics and Automation*, 5:3948–3953, 2005.

[5] Pierre Bessiere, Juan-Manuel Ahuactzin, Kamel Mekhnacha, and Emmanuel Mazer. *Bayesian Programming Book*. Springer, 2006.

[6] Pierre Bessiere, Christian Laugier, and Roland Siegwart. *Probabilistic Reasoning and Decision Making in Sensory-Motor Systems*. Springer, 2008.

[7] Xavier Perrin, Ricardo Chavarriaga, Roland Siegwart, and Jose del R. Millan. Bayesian controller for a novel semi-autonomous navigation concept. *ECMR*, 2007.

Geometric modeling and computer graphics of kinematic ruled surfaces on the base of *complex moving* one axoid along another

(one-sheet hyperboloid of revolution as fixed and moving axoids)

Galina S. Rachkovskaya
107 Pushkinskaya Str. # 25
344007, Rostov-on-Don, Russia
galina.rachkovskaya@verizon.net

Yuriy N. Kharabayev
1Б Budennovskiy Ave. # 46
344008, Rostov-on-Don, Russia
kharabayev@aaanet.ru

ABSTRACT

In concordance with logical structure "Geometrical Model of Interaction of two contacted surfaces during the movement of one ruled surface along another – Corresponding Mathematical Transformations of Surfaces – New Kinematic Ruled Surfaces" [1-3] a new geometrical model of *complex moving* one axoid along another for the case of one-sheet hyperboloid of revolution as fixed and moving axoids has been proposed. The main condition of constructing kinematic ruled surfaces is that moving axoid contact with fixed axoid along one their common generating line in each of their positions during *complex moving* one axoid along another. A case when the axes of fixed and moving axoids are crossed (Fig. 1,2,3), has been considered in this research. Analytical development and computer graphics of the new kinematic surfaces are realized for three types of *complex moving*. (1)The outside surface of the fixed axoid is revolved *slipping-free* by the outside surface of the corresponding moving axoid (Fig. 1). (2)The interior surface of the fixed axoid is revolved *slipping-free* by the outside surface of the corresponding moving axoid (Fig. 2). (3)The outside surface of the fixed axoid is revolved *slipping-free* by the interior surface of the corresponding moving axoid (Fig. 3). Computer graphics of the constructed surfaces (Fig. 1*a*,2*a*,3*a*) have been performed by the previously developed software application [4].

Figure 1. Figure 1*a*.

Figure 2. Figure 2*a*.

Figure 3. Figure 3*a*.

Keywords
Geometric Modeling, Computer Graphics, Kinematic Surfaces.

1. INTRODUCTION

Kinematic ruled surfaces are constructed by the movement of a generating line of one (moving) axoid during its moving along another (fixed) axoid [1]. The main condition of constructing kinematic ruled surfaces is that moving axoid contact with fixed axoid along one their common generating line in each of their positions during movement of one axoid along another. In some cases such moving as rolling one axoid along another is sufficient to meet this main condition. The examples of these cases are well-known kinematic ruled surfaces constructed on the base of rolling one cylinder along another cylinder or one cone along another cone [1]. At the same time such moving as rolling one axoid along another in the case of one-sheet hyperboloid of revolution as fixed and moving axoids is insufficient to meet the main condition of constructing kinematic ruled surfaces. However, as follow from this research

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

in this case task of constructing kinematic ruled surfaces move to feasible solution on the base of *complex moving* one axoid along another. *Complex moving* can be represented as a combination of concerted movement such as rolling one axoid along another and translational movement of the moving axoid along the common generating line of both axoids The similar case of geometrical modeling of *complex moving* one axoid along another on the example of moving a cone along a torse was described earlier [2, 3]. In that case *complex moving* was represented as a superposition of three interrelated elementary movements: rotational movement of the cone around its axis; turn of the cone axis; translational movement of the cone vertex along the torse edge-of-regression. Similar task of constructing geometrical model of *complex moving* one-sheet hyperboloid of revolution along another one as the base for generating new kinematic ruled surfaces was solved in this research.

2. PRINCIPAL GEOMETRICAL MODEL OF *COMPLEX MOVING*

One-sheet hyperboloid surface of revolution is a ruled surface, which can be constructed by the straight line rotation about the axis, if the straight line is crossed with the axis of rotation. A case of two interacting axoids contacted along mutual ruling when the axoids' axis are crossed (Fig. 1, 3, 5), has been considered in this research. In this case, *complex moving* one axoid along another can be represented as a superposition of three interrelated elementary movements: rotational movement of the moving axoid around its axis *OZ* in the moving coordinate system *OXYZ* connected with the moving axoid, rotational movement of the moving axoid axis *OZ* around the fixed axoid axis *oz* in the fixed coordinate system *oxyz* connected with the fixed axoid, and translational movement of the moving axoid along the common generating line of both axoids. The origin of the fixed coordinate system *oxyz* is located in the center of the waist circle of the fixed axoid. The origin of the moving coordinate system *OXYZ* is located in the center of the waist circle of the moving axoid.

3. GEOMETRICAL MODEL №1 (ONE AXOID LOCATE ON THE OUTSIDE OF ANOTHER AXOID)

Geometrical model №1 is related to the case if the outside surface of the fixed axoid is revolved *slipping-free* by the outside surface of the corresponding moving axoid (Fig. 1). Model №1A is related to the case if moving axoid is the same as fixed axoid. Model №1B is related to the case of two contacted different axoids.

3.1 Geometrical Model №1A

Geometrical model №1A is related to the case if moving axoid (2) is the same as fixed axoid (1), i.e. $a_1 = a_2 = a$, $c_1 = c_2 = c$, where

a_1, c_1 – parameters of the fixed axoid,

a_2, c_2 – parameters of the moving axoid.

Parameters a, c – parameters of the canonical equation of one-sheet hyperboloid surface of revolution in the coordinate system *oxyz*:

$\frac{x^2}{a^2} + \frac{y^2}{a^2} - \frac{z^2}{c^2} = 1$, where a – radius of waist circle.

Parametric equations of surface generated by one of the ruling of the moving axoid in the system *OXYZ*:

$X = R(v\cos(\varphi+\alpha) - (v-1)\cos(\varphi-\alpha))$;

$Y = R(v\sin(\varphi+\alpha) - (v-1)\sin(\varphi-\alpha))$;

$Z = -h(2v-1)$,

where R – radius of circular section of one-sheet hyperboloid of revolution at h distance from the waist section: $R = a\sqrt{1+(h/c)^2}$; $\alpha = arctg(h/c)$;

φ – current value of the angle of rotation of the moving axoid around its axis, v – input parameter.

The equations of transition from the coordinate system *OXYZ* to the coordinate system *oxyz*:

$x = X\cos\varphi - (Y\cos\theta - Z\sin\theta)\sin\varphi + 2a\cos\varphi$

$y = X\sin\varphi + (Y\cos\theta - Z\sin\theta)\cos\varphi + 2a\sin\varphi$

$z = Y\sin\theta + Z\cos\theta$,

where $\theta = 2arctg(a/c)$.

The resulting parametric equations of the kinematic ruled surface in the fixed coordinate system *oxyz*:

$x = (A+2a)\cos\varphi - (B\cos\theta + C\sin\theta)\sin\varphi$;

$y = (A+2a)\sin\varphi + (B\cos\theta + C\sin\theta)\cos\varphi$;

$z = B\sin\theta - C\cos\theta$,

where

$A = R(v\cos(\varphi+\alpha) - (v-1)\cos(\varphi-\alpha))$;

$B = R(v\sin(\varphi+\alpha) - (v-1)\sin(\varphi-\alpha))$;

$C = h(2v-1)$.

Pair of contacted axoids (Fig. 4)) and corresponding kinematic ruled surface (Fig. 4*a*) are shown below:

Figure 4.

Figure 4*a*.

Figures of pairs of contacted axoids and kinematic ruled surfaces have been constructed with the help of the previously developed AMG ("ArtMathGraph") software application [3, 4].

 ISBN 978-80-86943-95-4

3.2 Geometrical Model №1B

Geometrical model №1B is related to the case of two different axoids ($a_1 \neq a_2$, $c_1 \neq c_2$). As it will be shown below parametric condition for two contacted different axoids is $a_1^2 + c_1^2 = a_2^2 + c_2^2$.

This formula follows from model №1B (Fig. 5).

Figure 5.

In the developed geometrical model one axoid is revolved *slipping-free* by the another axoid. In this case any circle (waist circle and all others) of the fixed axoid (1) is revolved *slipping-free* by the corresponding circle of the moving axoid (2). Hence it follows that a relation length of each pairs of contacted circles of contacted axoids must be equal. The formula of dependence $R(l)$ was obtained on the base of model №1B (Fig. 5), where l – length of the segment of the common generating line of both axoids from the waist circle to the circle of radius R:

$$R = a\sqrt{1 + \frac{l^2}{a^2 + c^2}} \ .$$

As stated above the relation

$$\frac{R_2(l)}{R_1(l)} = \frac{a_2\sqrt{1 + \frac{l^2}{a_2^2 + c_2^2}}}{a_1\sqrt{1 + \frac{l^2}{a_1^2 + c_1^2}}}$$ must be l -independent.

It is possible if $a_1^2 + c_1^2 = a_2^2 + c_2^2$.

The result parametric equations of the kinematic ruled surface have been derived in the system oxyz:

$$x = (A + a_1 + a_2)\cos(\varphi / n) -$$
$$-(B\cos\theta + C\sin\theta)\sin(\varphi / n);$$
$$y = (A + a_1 + a_2)\sin(\varphi / n) +$$
$$+(B\cos\theta + C\sin\theta)\cos(\varphi / n);$$
$$z = B\sin\theta - C\cos\theta,$$ where

$$A = R(v\cos(\varphi + \alpha) - (v-1)\cos(\varphi - \alpha));$$
$$B = R(v\sin(\varphi + \alpha) - (v-1)\sin(\varphi - \alpha));$$
$$C = h(2v - 1);$$
$$n = a_2 / a_1; \ \theta = arctg(a_1 / c_1) + arctg(a_2 / c_2);$$
$$R = a_2\sqrt{1 + (h / c_2)^2}; \ \alpha = arctg(h / c_2).$$

Some matched pairs of contacted different axoids (Fig. 1, 6, 7) and corresponding generated kinematic ruled surfaces (Fig. 1*a*, 6*a*, 7*a*) are shown. (Figure 6: $(a_1 / a_2) = 2/1$; Figure 7: $(a_1 / a_2) = 1/2$).

Figure 6.

Figure 6*a*.

Figure 7.

Figure 7*a*.

4. GEOMETRICAL MODEL №2 (ONE AXOID LOCATE IN THE INTERIOR OF ANOTHER AXOID)

Geometrical model №2 of *complex moving* one axoid along another is related to the two cases:

- the interior surface of the fixed axoid is revolved *slipping-free* by the outside surface of the corresponding moving axoid (model №2A),

- the outside surface of the fixed axoid is revolved *slipping-free* by the interior surface of the corresponding moving axoid (model №2B).

Using the geometric model №2 the result parametric equations of the kinematic ruled surface have been derived in the coordinate system *oxyz*:

$$x = (A + |a_1 - a_2|)\cos(\varphi / n) +$$
$$+(B\cos\theta + C\sin\theta)\sin(\varphi / n);$$
$$y = -(A + |a_1 - a_2|)\sin(\varphi / n) +$$
$$+(B\cos\theta + C\sin\theta)\cos(\varphi / n);$$
$$z = B\sin\theta - C\cos\theta,$$ where

$$A = R(v\cos(\varphi + \alpha) - (v-1)\cos(\varphi - \alpha));$$
$$B = R(v\sin(\varphi + \alpha) - (v-1)\sin(\varphi - \alpha));$$
$$C = h(2v - 1);$$
$$n = a_2 / a_1; \ \theta = |arctg(a_1 / c_1) - arctg(a_2 / c_2)|;$$
$$R = a_2\sqrt{1 + (h / c_2)^2}; \ \alpha = arctg(h / c_2).$$

4.1 Geometrical Model №2A

For example, the kinematic ruled surfaces (Fig. 2*a*, 8*a*) for matched pairs of contacted axoids ($(a_1 / a_2) = 6/1$) are constructed. The axis of the fixed axoid (1) is located vertically on the Fig. 2, 8.

Figure 8.

Figure 8*a*.

4.2 Geometrical Model №2B

The types of kinematic ruled surfaces constructed on the base of the model №2A or model №2B are different (Fig. 8*a* or 9*a*). In this case the type of surfaces is defined by the proportion (a_1 / a_2). The surfaces shown in Fig. 2*a*, 8*a* are corresponded to the proportion $(a_1 / a_2) = 6/1$, and the surfaces shown in Fig. 3*a*, 9*a* are corresponded to the proportion $(a_1 / a_2) = 1/6$. The axis of the fixed axoid (1) is located vertically in Fig. 3, 9.

Figure 9.

Figure 9*a*.

It is necessary to note that all described above matched pairs of contacted axoids fulfill geometric condition of overlap-free *complex moving* one axoid along another. However, fulfillment of this condition is optional for constructing kinematic ruled surfaces.

5. MODEL OF *COMPLEX MOVING* WITH OVERLAPPING TWO CONTACTED AXOIDS

The example of matched pair of axoids (Fig. 10) and corresponding kinematic ruled surface (Fig. 10*a*) as the case of *complex moving* one axoid along another with overlapping two contacted axoids is shown.

Figure 10.

Figure 10*a*.

At the first sight the shown above pair of axoids (Fig. 10) looks like an "inappropriate pair" of contacted axoids. Nevertheless, this axoids pair was selected in correspondence with condition $a_1^2 + c_1^2 = a_2^2 + c_2^2$ as well. The axial angle of the contacted axoids is θ:

$$\theta = arctg(a_1 / c_1) + arctg(a_2 / c_2) = 102^\circ .$$

In correspondence with value $\theta = 102^\circ$ fixed and moving axoids are located in Fig. 10.

It should be emphasized, that the mathematical transformations as well as the resulting parametric equations of the kinematic ruled surfaces are independent of the type of *complex moving* - with or without overlapping two contacted axoids.

6. CONCLUSIONS

Thus, the proposed principal geometrical model of *complex moving* one axoid along another for the case of one-sheet hyperboloid of revolution as fixed and moving axoids has been assumed as a basis for analytical development of the new logical mean for constructing kinematic ruled surfaces. The main result of this analytical treatment in the combination with graphic ability of the previously developed software application gives a good opportunity for computer search of desirable kinematic surfaces.

REFERENCES

[Kri06a] Krivoshapko, S.N., Ivanov, V.N., Khalabi, S.M. Analytical Surfaces. M, Nauka, 2006, 544 p.

[Con04a] Rachkovskaya, G.S., Kharabayev, Yu.N., and Rachkovskaya, N.S. The computer modelling of kinematic linear surfaces (based on the *complex moving* a cone along a torse). Proceedings of the International Conference on Computing, Communications and Control Technologies (CCCT), Austin (Texas), USA. Vol.1. P. 107-111, 2004.

[Con04a] Rachkovskaya, G.S., Kharabayev, Yu.N., and Rachkovskaya, N.S. Computer graphics of kinematic surfaces. Proceedings of the 12-th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision 2004, Plzen, Czech Republic, p.p.141-144.

[Con07a] Rachkovskaya, G.S., Kharabayev, Yu.N., and Rachkovskaya, N.S. Computer composition of the transformed classical surfaces as the ways and means of the construction of visual models of realistic objects (The new software application "ArtMathGraph") Proceedings of the 15-th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision 2007, Plzen, Czech Republic, p.p. 29-32.

POSTER: iGLANCE project: free-viewpoint 3D video

Svitlana Zinger
Eindhoven University of Technology
P.O. Box 513
5600MB Eindhoven, Netherlands
s.zinger@tue.nl

Daniel Ruijters
Philips Healthcare
Veenpluis 6
5680DA Best, Netherlands
danny.ruijters@philips.com

Peter H. N. de With
Eindhoven University of Technology
P.O. Box 513
5600MB Eindhoven, Netherlands
P.H.N.de.With@tue.nl

ABSTRACT

The iGLANCE project aims at researching the methods of receiving and rendering of free-view 3D-TV sequences. We describe the ongoing iGLANCE project and introduce the iGLANCE decoding and rendering system. It consists of free viewpoint video interpolation and coding. One of the key components of this system is view interpolation and its related challenges. There are two main application fields within iGLANCE: consumer and medical systems. In this article we discuss an example use case for a medical application that illustrates the need for the research and development in the iGLANCE framework.

Keywords

3D video, view interpolation, autostereoscopic display, free-viewpoint.

1. INTRODUCTION

Multi-view autostereoscopic displays allow depth perception of displayed scenes without the aid of additional goggles. The stereoscopic effect is obtained by presenting a slightly different image of a 3D scene to the left and the right eye of a viewer (see Figure 1). Since a multi-view display emits more than two views, a viewer can be positioned anywhere within the stereoscopic range, and multiple viewers can perceive depth at the same display.

The combination of autostereoscopic displays with free-viewpoint video creates opportunities with respect to interaction between the user and the broadcasted content. Furthermore, the view interpolation techniques allow reducing the amount of information that has to be coded and transmitted. To achieve these goals, new algorithms for image and video processing are required.

The iGLANCE project has started in October 2008 and aims at investigating both the algorithmic solutions as well as the hardware design for the objectives described above. The project challenge originates from the use of multiple views of the same scene (multiple cameras) and the generation of arbitrary in-between views. This involves advanced coding, compression and rendering of 3D video in HD resolution.

Figure 1. The multi-view autostereoscopic display.

The generating process of free-viewpoint 3D video, which is a research topic in computer vision, enables the user to view a scene from a freely chosen position. For example, a possible use case of iGLANCE is watching a concert or football game from the preferred viewpoint. The possibility to interactively choose a viewpoint in-between the broadcasted 3D video views distinguishes this projects from other research projects [Red02, Onu06] and present commercial 3D-TV solutions.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

3D Video processing is emerging for both medical purposes and consumer electronics. In the medical domain, 3D data is generated by several imaging modalities, and after processing, visualization of this data on an autostereoscopic screen with free-viewpoint option will help to optimally navigate instruments during interventions. This type of data display is especially useful in minimally invasive interventional procedures.

2. MULTI-VIEW AUTOSTEREO-SCOPIC DISPLAY

A multi-view autostereoscopic lenticular display consists of a cover sheet of cylindrical lenses (lenticulars) placed on top of an LCD, in such a way that the LCD image plane is located at the focal plane of the lenses [Ber99, Rui06]. The effect of this arrangement is that LCD pixels located at different positions underneath the lenticulars fill the lenses when viewed from different directions. Provided that these pixels are loaded with suitable stereo information, a 3D stereo effect is obtained, in which the left and right eye see different, but corresponding, information. A commercially available display offers nine distinct views, but our technology will be applicable to any number of views.

The fact that the different LCD pixels are assigned to different views (spatial multiplex), leads to a lower effective resolution per view than the intrinsic resolution of the LCD grid [Dod97]. In order to distribute this reduction of resolution over the horizontal and vertical axis, the lenticular cylindrical lenses are not placed vertically and parallel to the LCD column, but slanted at a small angle.

The resulting assignment of a set of LCD pixels is illustrated in Figure 2. Note that the red, green and blue color channels of a single pixel are depicted in different views, and that 9 different views are covered within a column of 3 pixels.

Figure 2. The cylindrical lenses depict every sub-pixel in a different view. The numbers in the sub-pixels indicate in which view they are visible.

The frustums that result from the different focal spot positions, are illustrated in Figure 3. The viewing directions of the frustums are not parallel to the normal of the screen, except for the central one. Therefore the corresponding frustums are asymmetric [Mau05]. The viewpoint coordinates of such parallel axis, asymmetric frustum perspective projections are determined by

$$\left(\frac{(x - n \cdot d) \cdot f}{f - z} + n \cdot d \;,\; \frac{y \cdot f}{f - z} \right), \qquad (1)$$

where f denotes the focal distance, n the view number and d the distance between the view cameras.

Figure 4 illustrates the images that result from rendering the scene from focal spot positions with an offset to the center of the screen.

Figure 3. The frustums resulting from three different viewpoints.

Figure 4. The same scene rendered from the most left and most right viewpoint.

3. FREE VIEWPOINT RENDERING

A key algorithmic development within iGLANCE concentrates on the creation of intermediate 3D video views, given several predetermined views of the same 3D scene. These algorithms include multi-view image interpolation, view warping and depth estimation [Mor08]. The input views are supposed to encode the conventional image (texture) and a depth or disparity map. Obtaining such depth images for real-world video sequences [Red02] is not an explicit objective of the iGLANCE project; the availability of such information is assumed as a precondition and can be obtained with off-line computations. For full

3D data, such as in virtual reality, computer games, or medical volumetric data, the generation of the depth information is more trivial.

The fundamental challenge that remains is displaying occluded parts of the scene [Vaz08]: in interpolated views, a part of the scene may become visible that has been hidden in any of the transmitted views, see Figure 5.

Figure 5. Illustration of occlusion: the area O of object B can only be seen from focus 2, thus is occluded from focus 1.

4. HARDWARE DESIGN

The hardware that is being developed for the iGLANCE project focuses on the reception and decoding of the multi-view video streams and performing the free-viewpoint interpolation.

Each video stream frame contains 2D texture and its respective depth map. Figure 6 defines the iGLANCE transmission chain, showing its main components.

Our research on view interpolation aims at finding the optimal methods for generating new views. The questions we are addressing include search for an optimal rendering method that produces good results in reasonable time, choice of the existing views (in time and space) for producing a new view, dealing with occlusions.

The outcome of this research will be beneficial for both consumer and medical applications. For the consumer applications our results will provide basis for free-viewpoint 3D-TV: the user will choose the viewpoint, and the iGLANCE system will make sure that only the necessary for view interpolation texture and depth are transmitted. When they are received and decoded, the iGLANCE system will ensure view interpolation in order to generate a view from the viewpoint chosen by the user. The new view will contain both texture and depth so that the autostereoscopic screen can generate the other textures necessary for creating a stereoscopic effect.

Figure 6. Multi-view transmission chain with iGLANCE decoding and processing.

5. MEDICAL USE CASE

In the case of medical applications, image artifacts due to occlusions may lead to misdiagnosis. Therefore, artifacts should be avoided as much as possible. In order to achieve this, all the multi-view textures are explicitly generated and fed to the stereoscopic display, rather than relying on a single combination of a texture and depth image. Medical 3D imaging will profit from the iGLANCE system by saving bandwidth and time: fewer views will be transmitted compared to the number of views needed for the autostereoscopic display. For each transmitted 3D view (combination of texture and depth), textures are generated only for the missing (multiple) views required by the stereoscopic display. This leads to a lower artifact level, since only a few views are interpolated, so that the occluded areas are inherently much smaller.

Given a number of n views that are being transmitted, and a number of m views that are required by the stereoscopic display, there is multitude of configurations possible, as is illustrated by Figure 7. Since the virtual camera setup can be controlled by the system, the asymmetric right configuration is preferred over the symmetric left one in this example, since it contains fewer interpolated views and will therefore lead to an overall lower artifact level.

Figure 7. Two configurations for 4 transmitted views, and 9 displayed views. Solid black: transmitted views that can be mapped directly on an output view. Dashed: transmitted views that cannot be mapped on an output view. Light blue: interpolated view.

Lossy compression is an additional source for image artifacts. Both lossy compression and view interpolation allow to reduce the required bandwidth, and the overall system therefore has to be optimized with respect to performance to maintain a low artifact level, combined with moderate bandwidth usage.

Another challenge concerns the latency, which has to be sufficiently low to enable interactive usage in the operating room. At this moment, we define 250 ms latency as the absolute maximum allowed, but the preferred latency is 100 ms.

Contrary to the consumer electronics scenario, the requirements regarding the frame rate for the medical use case are relatively modest; the target frame rate amounts 24 frames per second.

6. CONCLUSIONS

This paper has discussed the use of multiple views of the same scene and the generation of arbitrary in-between views, involving advanced 3D coding and rendering.

The iGLANCE project concentrates on the decoding and rendering of multi-view video with a free-viewpoint and considers also the hardware design of such a receiving platform. Analysis of 3D rendering with multi-view interpolation in a medical use case has shown that it potentially offers lower artifacts and bandwidth reduction.

Working in close collaboration with internationally renowned industrial partners and research institutions, while developing software and hardware simultaneously, the iGLANCE project intends to achieve not only scientific but also relevant results for the industry.

7. REFERENCES

[Ber99] van Berkel, C. Image Preparation for 3D-LCD. Proc. SPIE, Stereoscopic Displays and Virtual Reality Systems VI, Vol. 3639, pp.84-91, 1999.

[Dod97] Dodgson, N.A. Autostereo displays: 3D without glasses. EID: Electronic Information Displays, 1997.

[Mau05] Maupu, D., van Horn, M.H., Weeks, S., Bullit, E. 3D Stereo Interactive Medical Visualization. IEEE Computer Graphics and Applications, Vol. 25, No. 5, pp.67-71, 2005.

[Mor08] Morvan, Y., Farin, D., de With, P.H.N. System architecture for free-viewpoint video and 3D-TV. IEEE Transactions on Consumer Electronics, Vol. 54, pp.925-932, 2008.

[Onu06] Onural L., Ozaktas H.M., Stoykova E., Gotchev A., Watson J. An Overview of the Holographic Display Related Tasks within the European 3DTV Project, Proc. SPIE, Vol 6187, Photon Management II, Strasbourg, France, pp.61870T1-61870T10, 2006.

[Red02] Redert, A., op de Beeck, M., Fehn, C., IJsselsteijn, W., Pollefeys, M., van Gool, L., Ofek, E., Sexton, I., Surman, P. Advanced three-dimensional television system technologies. Proceedings of First International Symposium on 3D Data Processing Visualization and Transmission, pp.313-319, 2002.

[Rui06] Ruijters, D. GPU-Accelerated Volume Rendering to 3D Lenticular Displays. IEEE Workshop on Content Generation and Coding for 3D-Television, Eindhoven, the Netherlands, Vol. 1, 2006.

[Vaz08] Vázquez, C., Tam, W.J. 3D-TV: Coding of Disocclusions for 2D+Depth Representation of MultiView Images. In Proceedings of Tenth IASTED International Conference on Computer Graphics and Imaging (CGIM), Innsbruck, Austria, pp.26-32, 2008.

POSTER: Example Based Motion Generation with Efficient Control

Yujean Choung
Soongsil University
MAGIC Lab., IT building 602
Sangdo-dong, Dongjak-gu
156-743, Seoul, Korea
yjchoung@magiclab.kr

Kyung-Kyu Kang
Soongsil University
MAGIC Lab., IT building 602
Sangdo-dong, Dongjak-gu
156-743, Seoul, Korea
rcrookie@magiclab.kr

Dongho Kim
Soongsil University
MAGIC Lab., IT building 602
Sangdo-dong, Dongjak-gu
156-743, Seoul, Korea
prof@magiclab.kr

ABSTRACT

This paper presents a motion generation technique for arbitrary morphologies with the user defined correspondences between joints. Users can define the controlling part in the source character and the part to be controlled in the target character in our system. To remove the restriction in the morphology of the target character, we use the pair of example posture sets. In our system, in order to provide the correspondence regardless of the number of joints, the deformed part in the target character is simplified into the direction vector. The final postures are then generated with the weighted sum of the examples. Our experimental results demonstrate that our approach can generate motions for various target characters and can control the user defined joints. Exploiting this correspondence between joints, our system can reflect the intention of animators efficiently and reduce their manual efforts.

Keywords

Character animation, Motion generation, Motion blending, Motion retargeting, Motion synthesis

1. INTRODUCTION

Motion capture technology has been used in movies and animations for long time. Many motion retargeting algorithms have been developed to apply the motions to 3D virtual characters. Most of the retargeting researches can generate target characters which have the same structure as the source character but different only in the segment lengths [Gle98, SLSG01]. Recently, there has been growing interest in the motion generation research which supports different skeletal structures between the source and target characters. Kulpa et al. [KMA05] proposed a method that can apply motions without limitation in the number of bones by defining a new human skeletal structure using the links between the joints of the body. To remove the restriction in the morphology, there have been two kinds of approaches. One is an illustrative way that creates the motion data expressed in novel contexts [HRE+08], and the other way is the example based approach [PS04].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

The example based methods usually generate the final postures through the weighted sum of the given set of examples [PS04, LLC+07]. Park and Shin [PS04] presented Motion cloning method. It uses the pair of examples where source and target characters are related one by one. They compute the weights for the source character examples which have the identical structure to the input character. Then they generate the motion of target character by applying weights for the corresponding pairs. So this method does not have any limitation in morphologies of the targets. But they need an additional pair of examples for the whole body even if users want to create a posture for only some part of the character. This may result in the motions which the users did not expect.

This paper uses the pair of examples because our system is based on Motion cloning technique to set the target character's morphologies without limitation in the number of bones. But our system enables the users to control parts of the character freely with existing examples. We aimed to reflect the intention of the animator without additional manual efforts.

Figure 1. Overview of the motion generation system

2. Overview

Motion generation system

Figure 1 illustrates our system. Our motion generation system is similar to the motion synthesis process in the motion cloning method. But we developed the process for the user defined correspondence between joints. The system generates the posture of the target character corresponding to the input posture per frame in real-time. It is generated by the weighted sum of example postures.

The example postures are prepared in the preprocessing stage. Let P_i^s and P_i^t $(1 \le i \le n_{example})$ denote the set of source example postures and that of their corresponding target example postures, where i is the index of the examples and $n_{example}$ is the number of pairs.

The input data are the posture in each frame of the given motion and the joint indices defined by the user. The system calculates two sets of weights, one is for the user defined joints and the other one is for the remaining joints. Then we blend the postures P_i^t using those weights, and finally make an output through the post-processing.

User defined correspondence between joints

The user can define the controlling part in the source character and the part to be controlled in the target character. Figure 2 illustrates the example of correspondence between joints. s_s and s_e denote the starting and the end point of the controlling part. t_s, t_e denote the starting and the end point to be controlled.

Figure 2. The example of user defined correspondence between joints

Those points are expressed by joint indices. Corresponding target joints can be controlled by the input motion. So the system generates both the whole postures and the controlled part of the target character at the same time.

3. Motion synthesis

Linearization of the Postures

A posture can be denoted by $P = (p, q_1, q_2, ..., q_{n_{joint}})^T$ where $p \in \mathbb{R}^3$ is the position and $q \in \mathbb{S}^3$ is the orientation of each joint. Because of the non-linearity of the unit quaternion space, it is hard to calculate the weights which are used to blend the examples. So we used posture linearization process [PSS02, PS04]. All the orientations are converted into their corresponding displacement vectors using the logarithm map and they can be recovered using the exponential map. So the final linearized posture can be represented by $\hat{P} = (0, 0, v_2, ..., v_{n_{joint}})^T$, $(1 \le j \le n_{joint})$ where j is the joint index and n_{joint} is the total number of joints.

Weight computation for the remaining joints

In this process, the system computes the weight for the remaining joints excluding the user defined joints. The posture of the remaining joints are determined by the weight w_i^g of P_i^s, so that P_i^s matches P^{input} as much as possible. In the following equation, $\hat{P}^{input}$ and $\hat{P}_i^s$ are linearized input posture and the example postures respectively.

$$\hat{P}^{input} = \sum_{i=1}^{n_{example}} w_i^g \hat{P}_i^s , \qquad (1)$$

We used damped least square method to solve this linear system for w_i^g.

Then the weights are applied to the corresponding example postures to generate the final target character posture.

Weight computation for the user defined joints

In this process, the system computes the weight w_i^p for the corresponded joints $t_s \le j \le t_e$. The motion cloning method [PS04] needs to be provided with the new example postures of the whole joints when the user wants to make the new correspondence between joints. But our system only needs the existing examples using direction vectors between the joints. The direction vector removes the restriction in the number of joints and simplifies the high degree of freedom [KMA05]. So we can make the postures without additional examples.

Figure 3. The weight computation for defined joints

We calculate the direction vector $\vec{s}$ toward the joint s_e from the joint s_s of the input posture P^{input}. The system uses the vector $\vec{t}_i$ toward the joint t_e from the joint t_s of the example postures P_i^t of the target character. The system restores the $\vec{s}$ direction with blending $\vec{t}_i$ using weights w_i^p. This process can be expressed as follows:

$$\vec{s} = \sum_{i=1}^{n_{example}} w_i^p \vec{t}_i \,. \tag{2}$$

The weights w_i^p are calculated by solving the linear system. The linear system for calculating weights is expressed as follows:

$$\begin{bmatrix} \vec{t}_{x,1} & \vec{t}_{x,2} & \cdots & \vec{t}_{x,n_{example}} \\ \vec{t}_{y,1} & \vec{t}_{y,2} & \cdots & \vec{t}_{y,n_{example}} \\ \vec{t}_{z,1} & \vec{t}_{z,2} & \cdots & \vec{t}_{z,n_{example}} \end{bmatrix} \begin{bmatrix} w_1^p \\ w_2^p \\ \vdots \\ w_{n_{example}}^p \end{bmatrix} = \begin{bmatrix} \vec{s}_x \\ \vec{s}_y \\ \vec{s}_z \end{bmatrix} . \tag{3}$$

The rows of the matrix on the left hand side represent x, y, z values of $\vec{t}_i$, and columns are examples. We adopt the damped least square method to solve this linear system. This method generates the shape of the target character according to the source character motions. With the examples and the directional vector, plausible postures can be generated even if the numbers of bones are different.

Posture blending and post-processing

The system finally generates the target character's posture P^{output} with applying w_i^g and w_i^p to P_i^t. The example posture P_i^t is linearized to $\hat{P}_i^t$ and the weights are applied as follows:

$$\hat{P}^{output} = \sum_{i=1}^{n_{example}} w_i \hat{P}_i^t \,. \tag{4}$$

In the above equation, w_i is determined by the following rule:

(a)

(b)

Figure 4. Elephant and bowling pin:

(a) The correspondence between joints (b) The results of our method

$$w_i = \begin{cases} w_i^p & if\ t_s \le j \le t_e, \\ w_i^g & otherwise \end{cases} \,. \tag{5}$$

 ISBN 978-80-86943-95-4

(a) (b)

Figure 5. Rabbit: (a) The correspondence between joints (b) The results of our method

In the post-processing step, root position of the target character is constructed by scaling the trajectory of the input motion and adapting to the target character.

4. EXPERIMENTAL RESULTS

The examples in this paper were computed on a 1.86GHz Intel Core2Duo CPU, with 2GB RAM and nVIDIA geForce 8800GTS graphics card. We used the source character which has human-like structures with 26 bones. Figure 4 shows the results after generating the elephant and the bowling pin's locomotion. The elephant has 36 bones and different articulated structure from the source character. In the result (a), the user corresponds the right arm to the elephant's nose. With 11 pairs of examples, the elephant locomotion and nose poses were generated. Bowling pin has 14 bones and also different articulated structure from the source character. The user corresponds the left arm from elbow to finger to the bowling pin from neck to head. It used 15 pairs of examples. Figure 5 shows the result of rabbit example with two corresponding parts.

5. CONCLUSION

In this article, we have presented a motion generation technique for arbitrary skeletal structures with the user defined correspondences between joints. By simplifying the controlled part to the direction vector, the system can generate the motion without additional examples and does not restrict the number of bones. So it can reduce the animator's manual efforts. Exploiting this correspondence between joints, our system can generate the part of the posture according to the intention of the animator, which is hard to generate previously.

If the existing examples don't have enough movements, the generated part may not meet the user's intention. We are planning to solve this problem as future research. The interactivity of our system and no limitation in the structure of target character could be applied to many other applications such as games or interactive art works.

6. ACKNOWLEDGMENTS

This work was supported by grant No. R01-2006-000-11214-0 from the Basic Research Program of the Korea Science & Engineering Foundation.

7. REFERENCES

[BB04] Baerlocher, P., and Boulie, R. An Inverse Kinematic Architecture Enforcing an Arbitrary Number of Strict Priority Levels. The Visual Computer, 20(6), pp.402-417, 2004.

[Gle98] Gleicher, M. Retargetting Motion to New Character. 25th Computer graphics & Interactive techniques SIGGRAPH'98, pp.33-42, 1998.

[HRE+08] Hecker, C., Raabe, B., Enslow, R.W., DeWeese, J., Maynard. J., and Prooijen. K.v. Real-time Motion Retargeting to Highly Varied User-Created Morphologies. ACM Transaction on Graphics 2008, 27(3), pp.27:1-27:11, 2008.

[KMA05] Kulpa, R., Multon, F., and Arnaldi, B. Morphology independent representation of motions for interactive human-like animation. Computer Graphics Forum, Eurographics 2005 special issue 24(33), pp.343-352, 2005.

[LLC+07] Lee, T.Y., Lin, C.H., Chu, H.K., Wang, Y.S., Yen, S.W., and Tsai, C.R. Mesh pose-editing using examples. International Conference on Computer Animation and Social Agents, 18(4-5), pp.235-245, 2007.

[PS04] Park, M.J., and Shin, S.Y. Example-based motion cloning. Computer Animation and Virtual World, 15, pp.245-257, 2004.

[PSS02] Park, S.I., Shin, H.J., and Shin, S.Y. On-line Locomotion Generation Based on Motion Blending. In Proceedings of the 2002 ACM SIGGRAPH/Eurographics symposium on Computer animation, pp.105-111, 2002.

[SLSG01] Shin, H.J., Lee, J., Shin, S.Y., and Gleicher, M. Computer puppetry: an importance-based approach. ACM Transactions on Graphics 2001, 20(2), pp.67-94, 2001.

POSTER: A Visualization and Management System for Chat Dialogues on 3D Virtual Avatars

Seung-Hyun Ji, Soo-Hyun Park, Dong-Sung Ryu and Hwan-Gue Cho
Dept. of Computer Engineering
Pusan National University, Republic of Korea
{shji,shpark07,dsryu99,hgcho}@pusan.ac.kr

ABSTRACT

Chat is one of basic features in current virtual world communication. So tracking of chat dialogue is an important issue. However most chat programs only provide textual history transcripts based on the temporal sequence, without any explicit relationship tags. In this poster, we propose a novel data structure, Chatting Flow Graph (CFG)[6] which keeps track of the all chatting dialogues among virtual agents. On this CFG we newly propose one visualization method for managing Chat Dialog efficiently and hierarchically.

Keywords: Virtual World, Chat, Communication, Graph Visualization.

1 ISSUES IN VIRTUAL CHAT

Internet chatting is one of common amusements for people in Web Space. Due to rapidly evolving communication technologies, the framework for Internet chatting is evolving. Traditional chatting with plain texts is evolving into systems supporting multimedia. Especially current Internet technologies such as 'Second Life' and 'IMVU' enable us to chat in virtual space. Chatting among avatar agents in virtual space is quite different from those who chat on the plain text, since avatars can move around the virtual space and talk to different agents with different time. So, it is not easy to reconstruct the corresponding pairs of question/answer from a set of all dialogue texts appeared in the virtual space. In the previous chatting system, they should prepare another communicating session layers to make a pairwise chatting. or additional manual intervention such as clicking to designate talking objects is required as was done in 'Yahoo' and Microsoft MSN instant messenger. Since the current chat systems provide unstructured dialog sets and the distribution of chat history data are so widespread, there are many tasks for user to browse their chat history.

We know Figure1 (a) is a typical snapshot representing virtual chat system and Figure1 (b) shows typical plain-textual chat system. In Figure1 (b) we can not easily match the corresponding pair of question and answer. So we need another ways of representing and managing chat dialogues, especially in 3D virtual space.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Figure 1: Internet Chat System. (a) 2.5-D Chat in Second Life. (b) Textual history transcripts(NATEON tool supporting Korean) with 8 chat agents.

2 CHAT DIALOGUE

As far as we know, the prevailing chatting tools only keep textual data according to the temporal sequence. In this case, if we do not find the corresponding answer for a query question. So we proposed Chatting Flow Graph (CFG)[6] for managing chat dialogues among multiple agents in virtual space. This helps us to explicitly designate the talking dialogue pairs with considering the geometric information of virtual space(the location of avatars and their facing direction and distance in-between.) Let us define Virtual Chat Bandwidth(VCB)[5] for creating CFG.

Definition 2.1 *For two chat agents A_a and A_b,* ***VCB(Virtual Chat Bandwidth)*** *is defined the communication capacity bandwidth as follows. If $(\overrightarrow{N_a} \cdot \overrightarrow{N_b} \geq 0)$, then $VCB(A_a, A_b) = 0$. Otherwise let,*

$$VCB(A_a, A_b) = C_1 \cdot \frac{(\sin\theta_a \cdot \sin\theta_b)^{k_1}}{(dist(A_a, A_b) + C_2)^{k_2}}$$

 ISBN 978-80-86943-95-4

Figure 2: Avatars A_a and A_b are chatting. The solid line segment denotes the width of word balloons W_a and W_b appearing on each head to show chat texts.

, where C_1, C_2,k_1 and k_2 are control constants. [6]

Figure3 shows our Chat Flow Graph. Each node denotes the message with timestamp and agent. Edges are determined through timestamp Δt and VCB by considering the spatial information.

Figure 3: Chat Flow Graph.

This directed acyclic graph,CFG, could be more flexible to the way of keeping chat linearly.

(a) (b)

Figure 4: (a) Our 3D-Chat System. (b) Chat Flow Graph for the conversation between two avatars.

3 VISUALIZATION OF CHAT DIALOGUE

In previous work, we have shown CFG could be a good solution for managing a set of chat dialogues [6]. But according to the features of the graph, if there are multiple nodes then CFG is difficult to understand. So, we need to make a reasonable visualization system that makes easier for people to understand the chat history. Suppose that there are 100 agents chatting in a virtual space in the form a small(5-6) talking group. If we want to cluster a set of all dialogues into a meaningful chat group(a talk in a specific topic), then subgraph clustering is needed to show a group of chat subgroup. Thus we give a new hierarchical graph decomposition method where a group of topic chat is represented into a rectangle. So whole CFG is successively partitioned into sub-rectangles recursively. Algorithm 1 explains how to apply graph-cut on Chatting Flow Graph(CFG).

Algorithm 1 Graph-cut Algorithm

Input: *CFG(Chat Flow Graph)*
Output: Set of divided Graph

procedure $CFGRAPHCUT(Gs, min-depth)$
  if $checkHcut(Gs) = TRUE$
    $HGs \leftarrow$ **do** $Hcut(G)$;
    $CFGRAPHCUT(HGs, min-depth)$;
  else if $checkVcut(G, min-depth) = TRUE$
    $VGs \leftarrow$ **do** $Vcut(Gs, depth)$;
    $CFGRAPHCUT(VGs, min-depth)$;
  else return
end procedure

The basic idea of CFG cut is the minimum edge cut to find a more dense sub-graph partitioned. And $min-depth$ is a control constant and $checkHcut()$ and $checkVcut()$ are responsible to find the vertical/horizontal cut. If we find the minimal graph cut, then we separate them by the function $Hcut()$(horizontal-cut) and $Vcut()$(vertical-cut). The graph decomposition works till the size of subgraph divided is greater than a threshold value which is determined empirically. In our procedure we prefer the vertical cut to the horizontal cut.

Now we explain how to map these graph partition into a nested rectangle structures. Whenever we get two partitioned subgraph, then we assign then into a rectangle. The size of rectangle is dependent on the number of nodes(= the number of dialogues of chat avatar). Next we perform this procedure recursively till we get a small enough graph cluster(a topic chat).

Figure 5 shows an example of Chat Flow Graph(CFG). And Figure 6 shows the visualization result from Figure 5. In Figure6, the color of the recursively divided rectangles help us to distinguish each other. Also we provide an nice interactive user interface to review the a set of highly related chat text our of all dialogues by a simple mouse click. If you point one of divided unit sub-rectangle, then it immediately represents the chat dialogues in the form of directed acyclic graph as was done in CFG. So we you can read the a set of related dialogues assigned in a rectangle. Also if you control the viewing scale of

 ISBN 978-80-86943-95-4

rectangle visualization, we can view the more wider chat dialogues.

4 EXPERIMENT

We show some experimental results obtained from a movie, 'Dead Poets Society'. In one movie scene, one group of five men is talking together in a room. We calculated the VCB virtually considering the the movie scene to construct and visualize the CFG corresponding. In this example we adopted the the actual dialogues obtained the movie script and checked the time stamps of each dialogue by manual work in realtime.

Table 1: Textual script for agents in Figure 7.

Time	Agent	Message
00	B	"*Keen. Meeks, door closed.*"
02	E	"*Yes sir.*"
04	A	"*Gentlemen, what are the four pillars?*"
07	every	"*Travesty, horror, decadence, excrement.*"
09	B	"*Okay, good study group. Meeks aced Latin. I didn't quite flunk English. So, if you want, we got our study group.*"
13	A	"*Sure, Cameron asked me, too. Anyone mind including him.*"
15	B	"*What's his specialty, bootlicking?*"
17	A	"*He's your roommate.*"
18	B	"*That's not my fault.*"
21	E	"*I'm sorry. My name is Stephen Meeks.*"
24	A	"*This is Todd Anderson.*"
25	E	"*Nice to meet you.*"
26	C	"*Nice to meet you too.*"
28	B	"*Charlie Dalton.*"
30	D	"*Knox Overstreet.*"
31	A	"*Todd's brother was Jeffrey Anderson.*"
32	B	"*Oh, yeah. Valedictorian. National Merit Scholar.*"
33	E	"*Oh, well. Welcome to Hell-Ton.*"
35	B	"*It's every bit as tough as they say, unless you're a genius like Meeks.*"
37	E	"*He flatters me. That's why I help him with his Latin.*"
38	B	"*And English, and Trig.*"
43	A	"*It's open. Father, I thought you'd gone.*"

Using the proposed 3D virtual system and CFG algorithm, we can construct the corresponding Chat Flow Graph(CFG) for a take in the movie scene(about 2.5 minutes). Figure 8 shows the corresponding CFG. Our visualization algorithm can depict this dialogue graph into a nested rectangle as was shown in Figure 9.

Figure 5: Chat Flow Graph(CFG) showing a conversation on the planning a short trip in an elementary school. Conversation includes one teacher(P1) and ten students(P2P11). In the conversation, first, a teacher is talking about trip plan(data and place) to all students. After this question, students are divided into a three small group to discuss this matter. Each group is talking about their picnic plan. Finally P_{1-14} node shows that all students agreed on the picnic plan.

 ISBN 978-80-86943-95-4

Figure 6: Visualization of CFG in Figure 5. We set threshold parameter min-depth = 3. The number in each sub-rectangle denotes the number of dialogue nodes included.

Figure 7: A scene taken from the movie 'Dead Poets Society'. Five men are discussing together.

Figure 8: The corresponding Chat Flow Graph for the movie scene of Figure 7.

Figure 9: The rectangle visualization result from the CFG in Figure 8. We set the threshold parameter min-depth = 2.

5 CONCLUSION

As Internet chatting in virtual environments becomes prevalent, managing the communication information as a textual script is becoming important. The straightforward method used in the common text based chat protocols(e.g., instant messengers provided by MSN, Yahoo and NateOn) cannot be applied directly to tje virtual reality environments such as the SecondLife model, since they do not keep no semantic relationships in a linear text script. Also they do not utilize the geometric information of virtual avatars. In this paper we proposed a nice algorithm for clustering a set of text dialogue by considering the geometric information about virtual chatting agents. Also we newly proposed a novel visualization method to show the hierarchical structure of chatting dialogues to help us to search a small, but highly related dialogue set(a topic chat). Some experiments show that our approach works properly and efficiently.

REFERENCES

[1] James Abello, Frank van Ham, and Neeraj Krishnan. Ask-graphview : A large scale graph visualization system. *IEEE TRANS. ON TVCG, 2006*, 2006.

[2] Renaud Blanch and Eric Lecolinet. Browsing zoomable treemaps: Structure-aware multi-scale navigation techniques. *IEEE TRANS. ON TVCG, 2007*, 2007.

[3] Soo-Hyun Park, Seung-Hyun Ji, Dong-Sung Ryu, and Hwan-Gue Cho. Avachat: A new comic-based chat system for virtual avatars. *15th ACM Symposium on VRST 2008*, pages 279–280, 2008.

[4] Soo-Hyun Park, Seung-Hyun Ji, Dong-Sung Ryu, and Hwan-Gue Cho. A new cognition-based chat system for avatar agents in virtual space. *The 5th ACM SIGGRAPH International Conference on VRACAI 2008, To appear*, 2008.

[5] Soo-Hyun Park, Seung-Hyun Ji, Dong-Sung Ryu, and Hwan-Gue Cho. A smart and realistic chatting interface for gaming agents in 3-d virtual space. *2008 International Conference on CyberGames 2008*, pages 32–38, 2008.

[6] Soo-Hyun Park, Seung-Hyun Ji, Dong-Sung Ryu, and Hwan-Gue Cho. A textual dialogue management system for internet chatting over virtual agents communication. *4th International Conference on NCM 2008*, pages 165–171, 2008.

[7] Ying Tu and Han-Wei Shen. Visualizing changes of hierarchical data using treemaps. *IEEE TRANS. ON TVCG, 2007*, 2007.

POSTER: Information Transformation for Point Cloud

Koji Nishio
Osaka Institute of Technology
1-79-1, Kitayama
Hirakata, Osaka
Japan(573-0196)
nishio@is.oit.ac.jp

Yusuke Takebayashi
Osaka Institute of Technology
1-79-1, Kitayama
Hirakata, Osaka
Japan(573-0196)
takebayasi@is.oit.ac.jp

Yuji Teshima
Shizuoka Institute of Science and Technology
2200-2 Toyosawa
Fukuroi, Shizuoka
Japan(437-8555)
teshima@cs.sist.ac.jp

Takayuki Kanaya
Hiroshima International Univ.
555-36, Kurose Gakuendai
Higashi Hiroshima, Hiroshima
Japan (739-2695)
t-kanaya@hw.hirokoku-u.ac.jp

Ken-ichi Kobori
Osaka Institute of Technology
1-79-1, Kitayama
Hirakata, Osaka
Japan(573-0196)
kobori@is.oit.ac.jp

ABSTRACT

We propose a method of Information transformation for point cloud, which reduces information amount of three dimensional point data. Our method generates clusters and spiral chain lists. These chain lists are consists of three dimensional points. After that, we adopt a predictive encoding to compress these chain lists. In addition, we will show effectiveness of our method with same experimental results. In these experiments, our method reduces the information amount of point cloud coordinate to 31.7% of original model.

Keywords

Computer graphics, CAD, data compression, information transformation, and point cloud.

1. INTRODUCTION

Many types of 3d scanner have been developed in recent years, and we can use point cloud representation for three dimensional shape design, sampled by those scanners from real shape. This is an easy way to modeling 3d shapes from real object. So it is thought that in the near future, point cloud data will be distributed in CAD and computer graphics areas. Fig.1 shows an example of point cloud data. Each point has three-dimensional coordinate and its normal vector.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Figure 1. An example of point cloud

However, point cloud data is constructed by enormous numbers of points. As a result, it causes enlarging data size and increase of network traffic, in case of transfer via networks. So it is required to compress those data by transforming them into smaller representations. [Was04a] proposed a method which compresses coordinates and normals progressively by using binary tree. However it is for

both coordinates and normals, so the process needs both of them and does not work separately.

In this paper, we propose an information transformation method for point cloud captured by range scanner such as 3d laser scanner, reducing its information amount. With this method, we can convert it into smaller file using conventional compression techniques such as zip and so on.

Fig.2 shows outline of data compression process. The process consists of two procedures, information transformation and entropy encoding. First, information transformation technique is applied to point cloud data in order to reduce information amount with redundancy elimination. After that, entropy coding is applied to compress transformed data, for example, Huffman coding, arithmetic coding and so on.

2. Information transformation

First, our method generates clusters to manage points structuring

In this section, we propose a method which is one of information transformation techniques using chain list and differential encoding.

2.1 Point clustering

Our process generates clusters of point cloud to get more efficient result. Each cluster consists of points within a defined distance R_{th} form center of cluster, as shown fig.2. s_i is center of cluster and cluster C_i holds points and each distance from s_i is smaller than defined radius R_{th}. If we use larger R_{th}, the number of clusters becomes smaller. On the other hand, if we use smaller R_{th}, the number of clusters becomes larger. It is prefer to get more effective result, properties of points which belong to same cluster are similar, Where, properties are three-dimensional coordinate, normal vector, color and so on. So we should have to appropriate radius R_{th}. In general, smaller radius is preferable for the point cloud which has many flat or low curvature features, and lager radius is preferable for complex one.

Figure 2. Clustering of point cloud and its overlapping.

Figure 3. Reduction of overlapping.

If cluster circles overlap each other, the number of points in each cluster gets to be smaller and the information amount of each cluster gets to be smaller too. So our method chooses centers of cluster in order to place apart from each other. With this strategy, the process can generate large clusters efficiently and get effective result.

In the latter of the process, remaining point which does not already belong to any cluster are scattered and the size of cluster becomes small as process advances. However, sizes of clusters become larger as a whole result.

2.2 Spiral chain list

This process applies differential encoding to points in each cluster. Differential encoding converts properties of points into difference value. So the process generates point lists like a scroll print as shown fig.4. In this paper, we call it spiral chain list.

Spiral chain list connects points sequentially which begins at center of cluster s_i and glow out to the boundary of cluster C_i. Differential values, such as differential coordinate and differential normal are calculated along this spiral chain list. In addition our method has adopted a predictive encoding. In this process, points are connected sequentially as shown in fig.5. $p_i(i=1\ldots n)$ is a point on spiral chain and this process calculates predictive point q. Then there is difference between next point p_{i+1} and q. So the differential Vector v_{diff} is stored as differential encoding result and in most cases, v_{diff}

Figure 4. Spiral chain list.

 ISBN 978-80-86943-95-4

is smaller than v_{i+1} . With this prediction, the information amount of point cloud becomes smaller than original data.

Figure 5. Prediction vector.

3. Experiments

We have evaluated effectiveness of our method and measured the difference between original point cloud data and transformed data about information amounts. In this experiment, we used 2 models, Bunny and Buddha. Each model consists of 34,834 and 543,652 points with three dimensional coordinate and normal.

Experimental results are shown in table 1 and table 2. The information amounts of original Bunny are 72.8KB for coordinate and 88.6KB for normal. The results of our method are 23.1KB and 86.1KB.Total compression ratios are 67.7%. However, the compression ratio for coordinate is 31.7%. The results for Buddha are 195.4KB, 852.7KB and 1046.1KB respectively. Total compression ratio is 44.1%, and Coordinate compression ratio is 19.7%.

Fig.6 and 7 show frequencies of original and transformed Bunny respectively using 8-bit quantization,. Fig 8 shows the path of spiral chain list generated by proposed method and its close up image is shown in fig.9.

The results of Buddha are shown in fig 10,11 and 12.

From this experiment, our method is effective just for coordinate to reduce information amount of point cloud. It is thought that in our method predictive encoding is applied only to coordinate, so compression ratio for normal is not effective. We think this may be improved by adopting prediction to normal.

Table 1. Information amounts for Bunny.

	Original	Transformed	Ratio
Coordinate	72.8KB	23.1KB	31.7%
Normal	88.6KB	86.1KB	97.2%
Total	161.4KB	109.2KB	67.7%

Table 2. Information amounts for Buddha.

	Original	Transformed	Ratio
Coordinate	989.7KB	195.4KB	19.7%
Normal	1388.2KB	852.7KB	61.4%
Total	2377.8KB	1048.1KB	44.1%

4. Conclusions

From this experiment, our method is effective to reduce information amount of point cloud and provide more compact representation for point cloud. It is thought that our method is effective for coordinate to reduce file size and reduction of network traffic when point cloud data is transferred over network.

(a) Coordinates

(b) Normals

Figure 6. Histograms of original Bunny.

 ISBN 978-80-86943-95-4

(a) Coordinates

(b) Normals

Figure 7. Histograms of transformed Bunny.

Figure 8. Path of spiral chain list for Bunny.

Figure 9. Detail of spiral chain list.

(a) Coordinates

(b) Normals

Figure 10. Histograms of original Buddha.

 ISBN 978-80-86943-95-4

(a) Coordinates

(b) Normals

Figure 11. Histograms of transformed Buddha.

Figure 12. Path of chain list for Buddha.

5. REFERENCES

[Was04a] Waschbüsch, M.. Progressive Compression of Point-Sampled Models. Proceedings of the Eurographics Symposium on Point-Based Graphics 2004, pp. 95-102.

POSTER: A Composition-Based Image Retrieval Using Line Feature

Jooyeon Hwang
School of Media
Soongsil Univ. 511
Sangdo 5 Dong, Dongjak Gu
156-743, Seoul, Rep. of Korea
zooyouny@gmail.com

Dongsup Lim
School of Media
Soongsil Univ. 511
Sangdo 5 Dong, Dongjak Gu
156-743, Seoul, Rep. of Korea
picmuse@gmail.com

Doowon Paik*
School of Media
Soongsil Univ. 511
Sangdo 5 Dong, Dongjak Gu
156-743, Seoul, Rep. of Korea
ssudlab@gmail.com

ABSTRACT

In this paper, we propose an image retrieval scheme that retrieves compositionally similar images with the query image from the line-used image set. To develop a new effective similarity measurement, we use line feature to discriminate in similarity, by investigating the relation between the value of the property and the degree of the compositional similarity. We have developed a similarity measure and compared effectiveness with a simple similarity measure in the sense of recall and precision. Our result proves that proposed scheme provides higher effectiveness.

Keywords

image retrieval, line feature, compositional similarity

1. INTRODUCTION

Image retrieval has been a favorite topic in computer vision research area. Text-based method and visual-based method, for example content-based image retrieval (CBIR), are the example of approaches that used for image retrieval. In content-based image retrieval, color is a common feature that used for retrieval. [Swa91] and [Smi95] used color histogram and color set as a form to represent the color feature. In sophisticated retrieval, spatial features have been considered in as a visual feature. [Lu94], [Smith96], and [Ko02] used the spatial feature by segmented sub-regions and the spatial relations between the sub-regions are considered as visual feature. In the aspect of the photographic composition, a line is basic and important feature which gives viewer specific feelings [Pra06]. Few studies, however, used line feature in image retrieval. [Fra00] proposed matching scheme using line feature but it is restricted in line-drawing images.

Permission to make digital or hard copies of all or part ofthis work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

In this paper we propose effective image retrieval from the line-used image set. We organize our paper as follows: In section 2, we describe which property of the line is related with compositional similarity and then analysis the relation between the value of the property and the degree of the compositional similarity. Then, we develop a new measure for compositional similarity in section 3. In section 4, we evaluate the performance of the proposed scheme. The experiment shows that proposed scheme gives higher performance in the sense of recall and precision than the basic similarity scheme.

2. COMPOSITIONAL SIMILARITY

In this section, we describe how to find the properties of the line that affect the compositional similarity and then analysis the relation between the value of the property and the degree of the compositional similarity.

We investigated human's subjective retrieval results to define the properties of the line that affect the compositional similarity. Description is as follows:

1. Some query images are given by a user. If the query is compositionally similar with the image in the DB, a vote will be given. The process repeated for all of queries.
2. From the voting result, we compute the degree of subjective similarity for all queries with all images in the DB individually.
3. For each image in the DB, we extract the

*Corresponding author

representative line from the image and compute the values of the property (angle, position, and magnitude) of a line.

4. Then we select discriminative property in subjective similarity from the result of 3rd process.

In next subsection, we will explain the process in detail.

2.1 Discriminative Property of the Line

Subject users consist of 13 people, and the number of query images is 8. The each query image has a representative line whose angle ranges separately from 0 to 180 degree (figure 1). The DB is made of images which have a line that is fairly distributed on the angle, position and magnitude.

Figure 1. The query images for finding a distinctive property of line

The 13 subject users were given the query images and voted for the DB as the first step of the above process. After voting, we measured the degree of subjective similarity between each query and each image in the DB, as follows:

$$Similarity_{q,t} = \frac{Score_{q,t}}{\sum_i Score_{i,t}} \quad (1)$$

q : index of the query

t : index of the image in the DB

$Score_{q,t}$: the number of voting of image t for image q

Then, we extract the representative line of a image by using Hough transform. Before conducting the Hough transform, we normalized the size of the image, and extracted the edges. Then we transformed the image by the Hough method, so we could get the representative line by finding the maximum value in the Hough transformed image. We modified the voting process in Hough transform so that the higher intensity of the edge, the more voted into accumulation array while the general ways do not consider the magnitude of edge.

After finishing the representative line extraction, we represented each image in the DB as a point in the graph which horizontal and vertical axis corresponds to subjective similarity and the value of one of three properties of line respectively in figure 2. In these three of graphs, we can find that the angle of the representative line reflects distinctiveness of two separate groups that one is similar to a given query and the other is not, while other two properties, position and magnitude do not so. As the results of other query images was also of the same, it would be better not to show all of those results.

Figure 2. Distribution of the property's values of line of each image in the database

2.2 Distributions of the Angle of Similar Images with a Query Image

The figure 3 illustrates the distribution of the angles of the representative lines included in the images of the group which is the set of images similar to the given query. It points out that if the angle of the representative line of a query image corresponds to horizontal or vertical line, the range of the distribution is narrow comparing to the case of diagonal.

Figure 3. The distribution of the angles of the representative lines included in the images of the similar images group to the given query

In the case that the representative line is near vertical, the range of the distribution in direction of diagonal side is wide while the range of the distribution in direction of vertical side is narrow. Also, the case of near horizontal line provides the same result. It means that the human recognize the differences sensitively in near vertical or horizontal angles while insensitively in diagonal angles.

3. SIMILARITY MEASURE

In this section, we describe a measure for compositional similarity, which is made based on the observation that the distribution range of the angle of the representative line included in the result pictures for a query is affected by the angle of representative line included in the query.

As mentioned previously in the section of 2.1, it is possible to use the difference of angles for measuring similarity between a query and each image of the DB. However, when using the difference directly, a problem would be occurred. The problem is that images which are not similar to the query image may be included in the retrieval results in the case of the query images that have a near vertical or horizontal line. To resolve this problem, the weight is applied to the difference of angles according to the angle of the representative line of a query image as shown in (2).

$$\text{Distance}(\theta_q, \theta_i) = \begin{cases} |\Delta(\theta_q, \theta_i)| / \mathrm{W}_{ccw}(\theta_q) & \text{, if } \Delta(\theta_q, \theta_i) \geq 0 \\ |\Delta(\theta_q, \theta_i)| / \mathrm{W}_{cw}(\theta_q) & \text{, if } \Delta(\theta_q, \theta_i) < 0 \end{cases} \quad (2)$$

The θ_q and θ_i is the angle of the representative line of a query image, *q,* and that of an image of the DB, *i*, respectively. $\Delta(\theta_q, \theta_i)$ is the acute angle from two possible differences of θ_q and θ_i. If θ_i is located toward more clockwise direction rather than θ_q, the sign of the distance is positive, otherwise negative. We designed the weight function, $\mathrm{W}_{cw}(\theta_q)$, that produces smaller value when the θ_q is closer to horizontal or vertical in clockwise direction. The $\mathrm{W}_{ccw}(\theta_q)$ is also same to the $\mathrm{W}_{cw}(\theta_q)$ except the direction.

As shown in the (2), the angles close to horizontal or vertical have their difference divided by relatively small number to react sensitively. On the other hand, the angles close to diagonal have their difference divided by relatively big number to react insensitively. By applying this method for measuring the compositional similarity between images, the results of the search are improved comparing to by using only difference value.

4. Retrieval Effectiveness

For verifying the measure suggested, the simple search system was implemented, and precisions and recalls computed from the results for several query images are compared.

In the search system, a series of the output images for a given query are retrieved sequentially in increasing order on the amount of the similarity. The DB consists of about 100 line-used images each of which are assigned the number for identifying and the angle of the representative line. The following is the process for the searching.

1. Input a query image
2. Extract the representative line from the query
3. Compute all distances of angle of representative line between the query image and every image in the DB
4. Output the identity number of each image of the DB retrieved for the query in increasing order on the distance

The performance of the search system are evaluated by comparing two different ways of computing distance that one is weighted related to the angle of the representative line, and the other is not. The followings are the formulas for computing the precision and recall used for evaluating.

$$\begin{aligned} A_k^j &= \sum_{i=0}^{k-1} V_i \text{(detections)} \\ B_k^j &= \sum_{i=0}^{k-1} 1.0 - V_i \text{(falses)} \\ C_k^j &= \sum_{i=k}^{N-1} V_i \text{(misses)} \\ \text{Recall}: R_k^j &= \frac{A_k^j}{A_k^j + C_k^j} \\ \text{Precision}: P_k^j &= \frac{A_k^j}{A_k^j + B_k^j} \end{aligned} \quad (3)$$

j : the index of query image

k : the rank of the retrieval result

V_i : the subjective similarity of ith image

The V_i is the subjective similarity between the query image and the ith image of the results. This value is computed by dividing the number of voting for the image *i* with given query image by the number of subject people. Thus, the value ranges from 0 to 1.0.

The Figure 4 shows the results of the experiments with each of the some query images in which the representative line's angle ranges variously. In the case that the angle of the representative line of a query is near vertical or horizontal, the result with the weight is better than the result without the weight. For the leftmost 10 samples, the result without the weight has 0.54 of recall and 0.30 of precision, and

the result with the weight has 0.64 of recall and 0.37 of precision.

In the case that the angle of the representative line of a query is near diagonal, although both results are similar for some of leftmost part of the samples, the result with weight is getting better than the result without weight as the number of samples increase In the case that the angle of the representative line of a query is vertical or horizontal, both results are almost same for the whole samples.

(a) The result about the query image which has the near vertical or horizontal angle of the representative line

(b) The result about the query image which has the near diagonal angle of the representative line

(c) The result about the query image which has the vertical or horizontal angle of the representative line

Figure 4. Average retrieval effectveness. The "o" marks the average precision P_k and recall R_k for the top k^{th} results (k = 10 if red, and k = 20 if black).

5. CONCLUSION

In this paper, we found the property of line that affects compositional similarity and the relation between the angle of the representative line of a query image and that of its retrieved images in the DB by analyzing the results of the experiments. In addition, we suggested the method for measuring compositional similarity, and designed and implemented the search system in which the method is applied. In order to verify the search system, we compared proposed scheme to simple scheme in the sense of precision and recall. Although the result is slightly changed according to the angle of the representative line of a query image, it shows reasonable performance in the most of case.

6. ACKNOWLEDGMENTS

This work was supported by the Soongsil University Research Fund.

7. REFERENCES

[Fra00] P. Franti, A. Mednonogov, V. Kyrki and H. Kalviainen, Content-based matching of line-drawing images using the Hough transform, Int. J. Doc. Anal. Recognition, Springer-Verlag, Vol. 3, pp.117-124, 2000

[Ko02] B. Ko and H. Byun, Integrated region-based image retrieval using region's spatial relationships, in IEEE int. Conf. on Pattern Recognition, 2002

[Lu94] H. Lu, B. Ooi, and K. Tan, Efficient image retrieval by color contents, in Proc. Of the 1994 Int. Conf. on Applications of Databases, 1994

[Pra06] D. Prakel, Basics Photography Composition, AVA Publishing SA, Switzerland, 2006

[Swa91] M. Swain and D. Ballard, Color indexing, International Journal of Computer Vision, Vol.7, No. 1, 1991.

[Smi95] J. R. Smith and S. F. Chang, Tools and techniques for image retrieval, in IS & T/SPIE Proceedings, Vol, 2670, Strorage & Retrieval for Image and Video Databases IV, 1995

[Smi96] J. R. Smith and S. F. Chang, Querying by color regions using the VisualSEEk content-based visual query system (M. T. Maybury, Ed.), in Intelligent Multimedia Information Retrieval, 1996.