TEACH-IN OF A ROBOT BY SHOWING THE MOTION

Gernot Hoffmann

ICIP, Lausanne 1996

ABSTRACT

A computer controlled life size marionette, which is equipped with 12 mechanical degrees of freedom, can be taught to move by a human dancer. The person carries a rucksack with four flashing infrared light emitting diodes, which are strobed by radio transmission. Two CCD-cameras observe the motions. The generalized coordinates of the dancer are computed by use of the images of the light sources and filtered by several methods. When the same music is played, the marionette shows the choreographical motions of the dancer. The delays of the servo-systems are taken into account and the playback speed is synchronized by correlation control.

1. COMPUTER CONTROLLED MARIONETTE

The life size marionette Fig.1 consists of an aluminium skeleton with different textile costumes and heads. She moves vertically by 9 high speed string winches , can turn about the z-axis and move in the horizontal plane in an area of 1.5m x 1.5m.

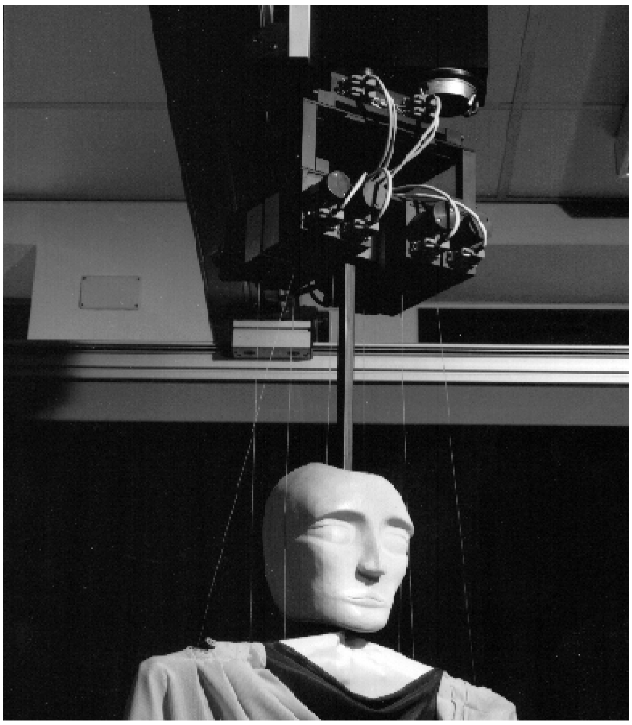

Fig.1 Computer controlled marionette

Two floodlights track the puppet automatically. The altogether 16 degrees of freedom are electrically driven. 5 personal computers, 1 signalprocessor and 2 microcomputers for the horizontal axes are combined in a realtime network.

The dynamics of the servomechanisms were identified and described by behaviouristic models which include the main features like time constants, time delays, small signal bandwidths, limited acceleration and speed [1] .

The marionette dances also fully automatically to any kind of music in real time with no prior knowledge. This project is finished and the results were published [2], [3]. The automatic motions consist of bundles of oszillations, which are synchronized with respect to a twelve-dimensional rhythm pattern, computed by the Fourier-Time-Transformation FTT [2], [4], which is quite different to the standard FFT. Thus they include no intellectual or arbitrary elements and it is some need to teach the machinery directly the choreographical part.

Obviously motion to music cannot be taught using a step by step method. The marionette is a trainer for experiments with man machine interfaces which accept human habits. Probably in some years robots will learn smooth and fast trajectories by looking at their master, e.g. for painting, avoiding obstacles or handling heavy loads by multiwire cranes where the operator shows the angles and directions by a platform in his hands.

2. VIDEO MOTOGRAPHY

The term "video motography" is used here for the threedimensional on line measurement of single light sources by use of two video CCD-cameras. The cameras deliver 50 half images per second each, odd and even rows alternating, where the actual half image includes the information of the preceeding half image in an adjacent row. Because of the very short video frame time of 20ms it is not reasonable to extract the coordinates from the natural video image. Instead of this attempt, the moving person is equipped with four light emitting infrared diodes (LEDs). In each video frame only one LED flashes for a very short time. The EX-TRAC system by *Wente/Thiedig* [5] computes the column and row numbers of the brightest pixel by hardware for both cameras. Result is a sequence of two pairs of pixel coordinates for LED1, LED2 and so on. Gathering four LEDs needs 80ms. The object space coordinates are computed by methods of photogrammetry, developed by the author and students [6],[7]. The application of EX-TRAC was initiated by a suggestion of *Baum* [8], who had also contributed to other projects of - as he says - On-line-Motografie [9]. The EX-TRAC station was originally connected to the LEDs by cables. To enable free motions the wireless "Motography Rucksack" has been developed..

3. MOTOGRAPHY RUCKSACK

The Rucksack Fig.2 is equipped with four high power infrared LEDs, type OD-100 [10], mounted on a square on adjustable rods. The power is typically 600mW (3A) and max. 1300mW (10A) for short pulses. The half intensity beam angle is ±55° from the radiation axis, which is here vertical.

Fig.2 Motography Rucksack, four LEDs on a square

Every 20ms during the vertical retrace of the video signal (1.3ms), the shutters of the cameras are open for 0.5ms. In this interval the one actual LED must be strobed. This is done in the sequence 1,2,3,4 by a quartz controlled counter in the rucksack. The next sequence cannot be triggered directly by the strobe No.1 from the ground station, because the delay of the radio frequency transmisssion, including the modulator and the receiver filters, is not neglectable: $\Delta T \approx 100\mu s$. Therefore the ground station strobe signal No.4 triggers a digital delay generator that produces a delay $T_d=(20ms - \Delta T_1)$, $\Delta T_1 \geq \Delta T$, which switches the transmitter modulator in advance while the shutter is open. This necessity results also from the fact, that between the ground strobe and the shutter closing a gap of only 90µs is available. If the shutter is opened longer, then he *starts earlier*, but the ground strobe is allways fixed to the same point of time.
No reliable informations about the timing of the whole system were available, therefore this part has been highly experimental.
Other methods of synchronization with the ground station, like optical or acustical transmission, did not look promising.
A direct *acustical* distance measurement, sometimes suggested, is quite hopeless because of the long wavelength of sound.
The Rucksack electronics are supplied by a reliable lead battery and voltage regulators. The LEDs are flashed by constant current controllers with a settling time of only 2µs and the current 3A is delivered by the battery without voltage controller.
The LEDs are mounted vertically, thus on large distances the camera views because of unfavovourable radiation angles rather low intensities; nevertheless the signals are detected over more than 5m. The person can move in an area of 4m x 4m.
The printed card boards are small, but the Rucksack is on purpose rather bulky: the dancing person should be reminded that fuzzy motions are not welcome - the sample rate for each set of four LEDs is only 12.5 Hz.
The Rucksack electronics were developed by S*cherge* [11].

4. PHOTOGRAMMETRY

Fig.3 shows the scenery of photogrammetry: Compute the coordinates $\mathbf{x}=(x,y,z)^T$ of a point in the object space by using coordinates $\mathbf{r}=(r,0,t)^T$ in the *image plane*, delivered by two cameras. Objects near to the image plane are mapped with scalefactor one; the center is fixed to the *viewpoint* in the object space. The image plane is a part of the object space.

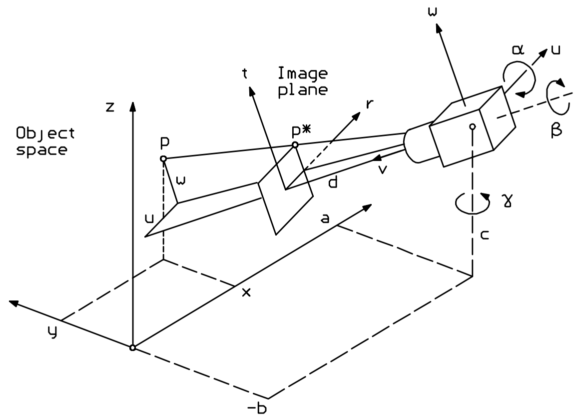

Fig.3 Photogrammetry, one of two cameras

The *image* is projected by the lens to the *picture* in the pixel plane and delivers pixel coordinates $\mathbf{R}=(R,0,T)^T$ for two cameras. The camera transformation is given by Eq.(4.1). The camera constant K has the unit pixel/m. R_0,T_0 are the pixel numbers near to the optical axis, a question which is discussed later. Dividing by the image distance d and introducing a new camera constant L = K d results in Eq.(4.2), where d is merely a scalefactor which must *not* be known. L is in pixel only. No focal length or similar parameters must be determined.

$$r = (R - R_0)/K; \qquad t = (T - T_0)/K \tag{4.1}$$

$$r/d = (R - R_0)/L; \qquad t/d = (T - T_0)/L \tag{4.2}$$

The camera is rotated first about the yaw angle γ, then about the elevation angle α and finally about the roll angle β, according to the matrix **C**, where c and s mean cosine and sine.

$$\mathbf{C} = \begin{pmatrix} c\beta\, c\gamma - s\alpha\, s\beta\, s\gamma & c\beta\, s\gamma + s\alpha\, s\beta\, c\gamma & -c\alpha\, s\beta \\ -c\alpha\, s\gamma & c\alpha\, c\gamma & s\alpha \\ s\beta\, c\gamma + s\alpha\, c\beta\, s\gamma & s\beta\, s\gamma - s\alpha\, c\beta\, c\gamma & c\alpha\, c\beta \end{pmatrix} \tag{4.3}$$

The perspective projection is given in coordinates u,v,w:

$$r = d\,u/v; \qquad t = d\,w/v \tag{4.4}$$

By use of the matrix **C** and the camera position $\mathbf{a}=(a,b,c)^T$, this can be written down in a homogeneous formulation:

$$\mathbf{A} = \begin{pmatrix} -1 & r/d & 0 \\ 0 & t/d & -1 \end{pmatrix} \tag{4.5}$$

$$\mathbf{A}\mathbf{u} = \mathbf{0} \tag{4.6}$$

$$\mathbf{A}\mathbf{C}(\mathbf{x} - \mathbf{a}) = \mathbf{0} \tag{4.7}$$

This formulation is the base for the identification of camera parameters and the computation of object points as well. The author, not an expert in photogrammetry, did not find equivalent descriptions in recent books of *Meisel* [12] and *Sasse* [13].

The identification of camera parameters $\mathbf{p}=(a, b, c, \alpha, \beta, \gamma, L)^T$ is *here* explained for the method of individual calibration.

We use an arrangement of laboratory fixed n=5 LED-markers $\mathbf{x}_i$, and each matrix $\mathbf{A}_i$ is filled with the image of the marker $\mathbf{x}_i$. The alltogether 10 equations cannot be solved simultaneously, therefore the errors ε_i are introduced:

$$\mathbf{A}_i \mathbf{C}(\mathbf{x}_i - \mathbf{a}) = \varepsilon_i \,; \;\ldots\; \mathbf{A}_n \mathbf{C}(\mathbf{x}_n - \mathbf{a}) = \varepsilon_n \qquad (4.8)$$

We use then the *method of steepest descent* [14], which is slow but reliable, to minimize a scalar error function:

$$F = \sum_{i=1}^{n} \varepsilon_i^T \varepsilon_i \qquad (4.9)$$

A weighting matrix could be used, if information is available.

The computation of object points is based on two cameras, indicated by the upper index for the same point $\mathbf{x}$.

$$\mathbf{A}^1 \mathbf{C}^1 \mathbf{x} = \mathbf{A}^1 \mathbf{C}^1 \mathbf{a}^1; \quad \mathbf{A}^2 \mathbf{C}^2 \mathbf{x} = \mathbf{A}^2 \mathbf{C}^2 \mathbf{a}^2 \qquad (4.10)$$

This can be written down with block matrices, though it is not necessary for the practical computation.

$$\begin{pmatrix} \mathbf{A}^1\mathbf{C}^1 \\ \mathbf{A}^2\mathbf{C}^2 \end{pmatrix} \mathbf{x} = \begin{pmatrix} \mathbf{A}^1 & 0 \\ 0 & \mathbf{A}^2 \end{pmatrix} \begin{pmatrix} \mathbf{C}^1\mathbf{a}^1 \\ \mathbf{C}^2\mathbf{a}^2 \end{pmatrix} \qquad (4.11)$$

$$\mathbf{M} \quad \mathbf{x} = \quad \mathbf{N} \quad \mathbf{b} \qquad (4.12)$$

This are four equations and three unknown coordinates $\mathbf{x}$. The matrix $\mathbf{b}$ includes no image information. The best compromise is found by the well known *Gauss transformation* or *normal equation,* with the formal solution:

$$\mathbf{M}^T \mathbf{M}\, \mathbf{x} = \mathbf{M}^T \mathbf{N}\, \mathbf{b} \qquad (4.13)$$

$$\mathbf{x} = (\mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T \mathbf{N}\, \mathbf{b} \qquad (4.14)$$

The system matrix is symmetric positive definite, thus the equations are solved by the *method of Cholesky* [15] .

Now we consider an improved camera error model:

In Fig.4 the pixel plane has coordinates $\mathbf{F}=(F,0,H)^T$ instead of $\mathbf{R} = (R,0,T)^T$ and F_o, H_o correspond to previous values R_o, T_o .
Because the viewline is not identical with the optical axis, the projection is oblique or tilt. The tilt angles δ, ε are correlated to offsets Δf, Δh or $\Delta F = L\, \Delta f / d$, $\Delta H = L\, \Delta h / d$. The pixel plane is assumed mechanically orthogonal to the optical axis.

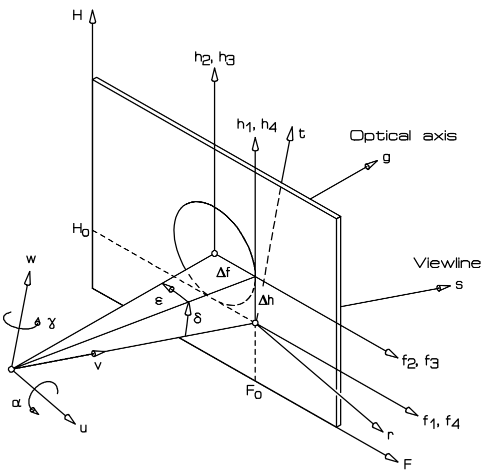

Fig. 4 Projection on a tilt pixel plane

The complete transformation, including the radial rectification by the orthogonal *Legendre polynoms* Q_i [16], is as follows:

$$f_1 / d = (F - F_0) / L \,; \qquad h_1 / d = (H - H_0) / L \qquad (4.15)$$

$$f_2 / d = f_1 / d - \Delta f / d = f_1 / d - \varepsilon \,; \; h_2 / d = h_1 / d - \Delta h / d = h_1 / d + \delta$$

$$\rho = \sqrt{(f_2 / d)^2 + (h_2 / d)^2} \,/\, \rho_{max} \qquad (4.17)$$

$$Q_3 = (1/2)(5\rho^3 - 3\rho) \,; \; Q_5 = (1/8)(63\rho^5 - 70\rho^3 + 15\rho) \qquad (4.18)$$

$$K_\rho = 1 + q_3 Q_3 / \rho + q_5 Q_5 / \rho \qquad (4.19)$$

$$f_3 / d = K_\rho f_2 / d \,; \qquad h_3 / d = K_\rho h_2 / d \qquad (4.20)$$

$$f_4 / d = f_3 / d + \varepsilon \,; \qquad h_4 / d = K_\rho h_3 / d - \delta \qquad (4.21)$$

$$r / d = (f_4 / d) / (\, 1 + (f_4 / d)\varepsilon - (h_4 / d)\delta \,) \qquad (4.22)$$

$$t / d = (h_4 / d) / (\, 1 + (f_4 / d)\varepsilon - (h_4 / d)\delta \,) \qquad (4.23)$$

The background about the physics of radial distortion is based on [11] and *Regensburger* [17],[18], but *Legendre polynoms* are more suitable than non orthogonal polynoms or sine functions.
For many cases Q_5 alone would be sufficient. Then the parameters of the camera are $\mathbf{p}=(a, b, c, \alpha, \beta, \gamma, L, \delta, \varepsilon, q_5)^T$.
Eq.(4.22) is the linearized version of a projection on a tilt pixel plane. The complete transformation is available from the author.
By this attempt we achieve a minimal number of significant parameters. Correlated parameters cannot be identified.

5. GENERALIZED COORDINATES

The coordinates of the moving person are represented by the position and the angles of the Rucksack: $x_o, y_o, z_o, \psi, \theta, \phi$. From $\mathbf{x}$ to $\mathbf{u}$ we rotate about the yaw angle ψ, then about the pitch angle θ and finally about the roll angle ϕ (similar to an aircraft).

The LEDs $L_1,..,L_4$ have body fixed coordinates $\mathbf{u}=(\pm b, \pm b, 0)^T$, but measured values in $\mathbf{x}$ will not fit to a square. We introduce a set of generalized coordinates $\mathbf{q}=(x_o, y_o, z_o, \phi, \theta, \cos\Psi, \sin\Psi)^T$.

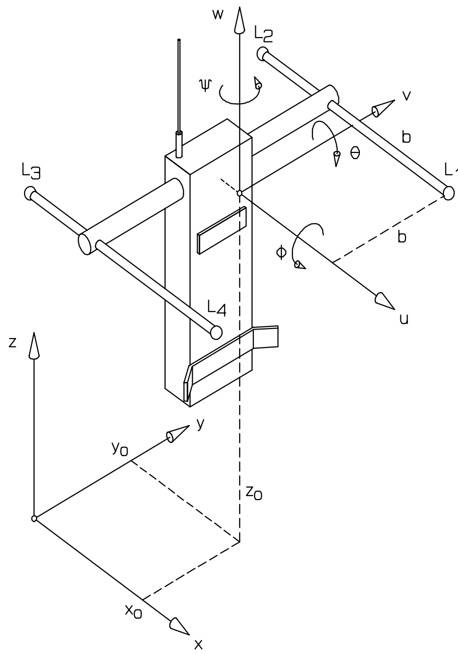

Fig. 5 Coordinates of the Rucksack

The angles ϕ and θ are assumed to be small, but the yaw angle is fully turning and represented by cosine and sine. By use of the *linearized* rotation matrix **B** we get 4 sets of 3 equations. These are reassembled to 12 *linear* equations for 7 variables **q**:

$$\mathbf{x}_i - \mathbf{x}_0 = \mathbf{B}^T \mathbf{u}_i \qquad (5.1)$$

$$\mathbf{M}\ \mathbf{q} = \mathbf{m} \qquad (5.2)$$

The measured values $\mathbf{x}_i$ are stored in $\mathbf{m}=(x_1,y_1,z_1,...,x_4,y_4,z_4)^T$. Then the *Gauss transformation* is applied and the resulting system can be solved *analytically*:

$$\mathbf{x}_0 = \sum_{i=1}^{4} \mathbf{x}_i / 4 \qquad (5.3)$$

$$\Phi = (z_1 - z_4 + z_2 - z_3)/(4b) \qquad (5.4)$$

$$\theta = (z_2 - z_1 + z_3 - z_4)/(4b) \qquad (5.5)$$

$$\cos\psi = (x_1 - x_2 + x_4 - x_3 + y_2 - y_3 + y_1 - y_4)/(8b) \qquad (5.6)$$

$$\sin\psi = (x_3 - x_2 + x_4 - x_1 + y_1 - y_2 + y_4 - y_3)/(8b) \qquad (5.7)$$

The computed LED-square is shown on a screen in *realtime*.

6. FILTERING OF THE TRAJECTORY DATA

Sometimes one LED is shaded or several LEDs are invisible. One missing LED can be complemented. This is called *syntactic filtering*. Fig.6 shows that other manipulations on the data are necessary to prepare them off line for sending to the robot.
The trajectory must be smooth, this is a *semantic* information. Graph(B) shows linear bridging. Graph(C) shows lowpass filtering by a 21-stage FIR-filter [19] and a nonlinear compression of the working area.

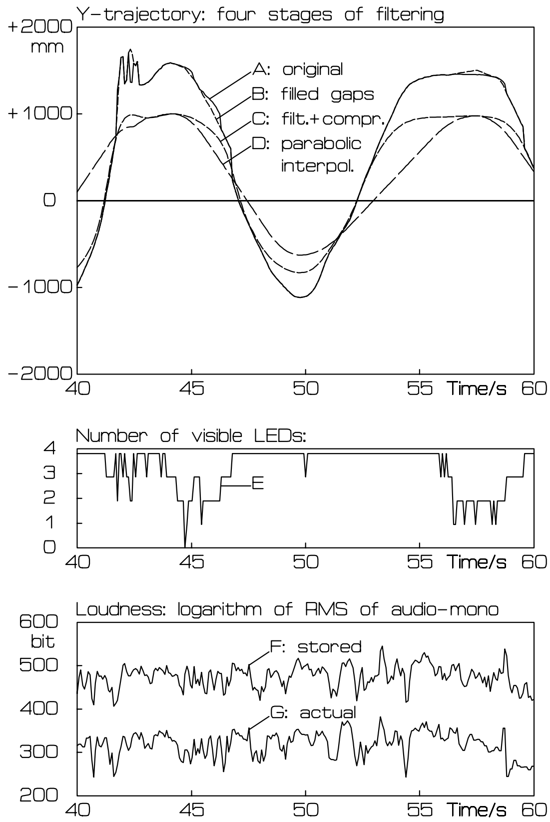

Fig.6 Several stages of digital filtering

For the application of teaching choreography there is a pragmatic aspect too: the dancing motion contains points of rest at musical climaxes, which must be reproduced during the playback by the robot exactly at the same time, even though the machinery is slower than the moving person. This problem is solved by *pragmatic filtering*: Minima and maxima of the trajectory are connected by parabolic interpolation, Graph(D). The ends are parabolic and in the middle the slope is constant with the maximal speed of the servoystem, which must be known. The mean time delay of each servosystem is taken into account by sending the data in advance.
Fig.6 (bottom) shows the speed control by correlation of stored and actual loudness of the music, sampled by T=20ms. Depending on the correlation, stored data are either doubled or skipped. Thus any speed variation of recorders is compensated.

CONCLUSIONS

The method of *Teach-In of a Robot by Showing the Motion* is not very acccurate, mainly because of small elevation angles of the two cameras. More cameras would be helpful.
In further applications the LEDs must be fixed on irregular patterns on the moving object and the appropriate set of generalized coordinates must be treated fully *nonlinear*.

REFERENCES

[1] Hoffmann,G.: Regelung und Identifizierung von Servosystemen. Antriebstechnik 34 (1/1995), p.49-54.
[2] Hoffmann,G.: Musikgesteuerte Marionette. Automatisierungstechnik 42 (8,1994), p.356-365.
[31 v.Randow,G: Techno in Ostfriesland. DIE ZEIT, Nr.4, 20.01.1995, p.38.
[4] Terhardt,E.: Fourier Transformation of Time Signals: Conceptual Review. Acustica 57 (1985), p.242-256.
[5] EX-TRAC S100. Ing.-Büro Wente/Thiedig , Rebenring 33, D-38106 Braunschweig, 1993.
[6] Meyeran,F.: Programmentwicklung und Installation für ein Motografie-System, Diploma, FHO Emden, 1994.
[7] Habbe,M.: Automatisches Einmessen eines fotogrammetrischen Motografie-Systems, Diploma, FHO Emden, 1995.
[8] Baum,E.: Motografie III. Wirtschaftsverlag NW, Bremerhaven, 1986
[9] König,D.H., Kirchner,J.-H., Fischer,D.: Freiraum für Beinbewegungen an handgeschobenenWagen. Wirtschaftsverlag NW, Bremerhaven, 1993.
[10] Opto Diode Corp.. 750 Mitchell Road, Newbury Park, California 91320.
[11] Scherge,R.:Entwicklung eines drahtlosen Motografie-Systems. Diploma, FHO Emden, 1994.
[12] Meisel,A.: 3D-Bildverarbeitung für feste und bewegte Kameras. Vieweg Verlag, Braunschweig, 1994.
[13] Sasse,R.: Bestimmung von Entfernungsbildern durch aktive stereoskopische Verfahren. Vieweg Verlag, Braunschweig, 1994.
[14] Leitman,G.(Ed.): Optimization Techniques. Academic Press, New York London, 1962.
[15] Schwarz,H.R.: Numerische Mathematik. B.G.Teubner Verlag, Stuttgart, 1993.
[16] Bronstein,I.N., Semendjajew,K.A.: Taschenbuch der Mathematik. Verlag Harri Deutsch, Zürich u.Frankfurt/Main, 1973.
[17] Regensburger,K.: Photogrammetrie. VEB Verlag für Bauwesen, Berlin, 1990.
[18] Rüger,W., Pietschner,J., Regensburger,K.: Photogrammetrie. VEB Verlag für Bauwesen, Berlin, 1987.
[19] Schrüfer,E.: Signalverarbeitung. Carl Hanser Verlag, München Wien, 1990.

May 1996. Graphics by Gernot Hoffmann. Photos by Wilhelm Kettwig. The author thanks him and Ralph Scherge for their assistance.

Old links:

http://www.fho-emden.de/~hoffmann/

http://www.fho-emden.de/~hoffmann/filename.pdf

Please use since February 2013

New links:

http://docs-hoffmann.de/

http://docs-hoffmann.de/filename.pdf