LOW BIT-RATE FACIAL VIDEO COMPRESSION

by

SHOMA CHATTERJEE

Department of Computer Science and Engineering

Submitted
in fulfillment of the requirements of the degree of Doctor of Philosophy
to the

Indian Institute of Technology, Delhi

June 2001

deo compression

CERTIFICATE

This is to certify that the thesis entitled "Low Bit-Rate Facial Video Compression" submitted by Miss. Shoma Chatterjee in the Department of Computer Science and Engineering, Indian Institute of Technology, New Delhi for the award of Doctor of Philosophy is a record of the original bonafide research work carried out by her under my guidance and supervision.

The thesis and any part thereof has not been submitted to any other university / institution for award of any degree or diploma.

KBiswas

Prof. K.K. Biswas

Deptt. of Computer Sc. & Engg.

Indian Institute of Technology,

New Delhi - 110 016.

ACKNOWLEDGEMENTS

I wish to express my heartfelt gratitude to my supervisor Prof K.K. Biswas who has been my constant source of encouragement for my work. I am equally thankful to Prof Subhashis Banerjee who was not my formal supervisor, but with whom I have shared many of ideas, and who has been kind enough to spend his valuable time to formulate ideas, critically analyzed the results and offered suggestions to improve my work. I am thankful to the members of my research committee Prof H. M. Gupta, Prof Santanu Chaudhury, Prof P.K. Kalra for spending time to evaluate my work periodically and to offer advice at various stages of my work.

I am also very thankful to Shri K.R. Kaushik, senior technical assistant of Computer Science and Engineering Department, who has guided me in all my lab setups.

I am thankful to the CSIR for providing my financial support for the last year of my Phd work.

S.Chatterjee

June 2001 Shoma Chatterjee

LOW BIT-RATE FACIAL VIDEO COMPRESSION

ABSTRACT

This thesis focuses on compression of facial video sequences to achieve low bit rates. The need for facial video compression arises due to increasing electronic transmission of digital images and facial video sequences over the internet and the growing demand for multimedia applications like on-line classroom lectures and video telephones. Video compression maximizes the amount of information to be transmitted over a given channel bandwidth. Compression of video sequences is also essential for digital transmission in the telephone network, storage of video on digital storage media including conventional storage devices CD-ROM, DAT, tape drives, writable optical drives, as well as telecommunication channels such as ISDNs and local area networks.

Traditional methods of video compression, e.g. MPEG 1-2, H.261, H.263 etc. address the issue of video compression solely in the image domain. In contrast, in the thesis we investigate methods to achieve video compression using techniques of computer vision to estimate 3 dimensional structure of the moving object.

We present an affine structure based global reconstruction scheme in which we utilize the invariant property of affine structure and send the $3D$ affine structure to the decoder at the bootstrapping stage. After that we need to send only the bases parameters which are few in number in order to reconstruct the frame at the decoder.

In our projective view morphing scheme to reconstruct global motion we rectify the image planes to parallel planes before interpolation to minimize the projective distortion in case of non-affine motion.

To reconstruct independent local motion we develop three separate schemes. The first one utilizes $2D$ affine local transformations, the second scheme is snake contour tracker based which develops an on-line database of distinct templates and recognizes current expression and encodes as an index into the database and the third scheme is based on optical flow parameters. Finally we conclude the thesis.

Contents