NEW BIDIRECTIONAL MOTION ESTIMATION USING MESH-BASED FRAME INTERPOLATION FOR VIDEOCONFERENCING APPLICATIONS

V. Muñoz-Jiménez, A. Zergaïnoh-Mokraoui, J.-P. Astruc
L2TI, Institut Galilée, Université Paris 13
99, Avenue Jean Baptiste Clément, 93 430 Villetaneuse, France
e-mail: {munoz, zergainoh,astruc}@galilee.univ-paris13.fr

Over the last few years, the gradual integration of video communication services into different network communication infrastructures has received an increasing interest. The new service types include streaming internet video, videoconferencing, broadcasting television, interactive multimedia databases, video games, …

In videoconferencing application, the most important moving objects in the video sequence are limited to the heads and the different elements of the speaker face such as lips and eyes. Since these moving objects are explicitly identified, one can use them to improve the accuracy of motion estimation in the video sequence. However in the real world, the camera introduces some additional transformations such as zooms, pans and rotations deforming the object. In most traditional standards, only translational motion model is considered using block-matching algorithm for its computational simplicity. However at low bit rate, the visual quality of the decoded image may be degraded.

To overcome these drawbacks, a new bidirectional motion estimation algorithm for videoconferencing application at very low bit rate is proposed. The approach is based on spatio-temporal spline interpolation. The proposed algorithm processes as follows. Some frames in the original video sequence are purposely missed. Afterwards, these frames are predicted by the decoder side using only the transmitted frames with no additional information. The motion of each selected and decoded reference object in the scene is modeled by rectangular deformable mesh grid which is constructed according to a coarse to fine quad-tree decomposition tilling algorithm. An optimization step adapts the mesh nodes to the object gradient minimizing then the error of the reconstructed object in an acceptable computational time. From these nodes, a temporal cubic spline interpolation predicts the mesh nodes of the moving object in the missed frame reconstructing therefore the meshed object. While the current frame background is interpolated using temporal cubic spline.

The proposed approach is integrated in the H.264/AVC video coding standard. Simulation tests conducted on videoconferencing sequences, at very low bit rate, show that the method provides interesting results in terms of rate-distortion compared to H.264/AVC video coding standard since no additional information is sent to the decoder.

Index Terms – Video coding, Mesh, Deformation, Interpolation, Motion estimation, Quad-tree decomposition