New fast search algorithm for base layer of H.264 scalable video coding extension

Livio Lima † Daniele Alfonso * Luca Pezzoni * Riccardo Leonardi †
† *University of Brescia, Brescia, Italy*
* *Advanced System Technology (AST) - STMicroelectronics, Agrate Brianza, Italy*
{*livio.lima,riccardo.leonardi*}*@ing.unibs.it* {*daniele.alfonso,luca.pezzoni*}*@st.com*

In this contribution, a fast search motion estimation algorithm for H.264/AVC SVC (scalable video coding) [2] base layer with hierarchical B-frame structure for temporal decomposition is presented and compared with fast search motion estimation algorithm in JSVM software [1], that is the reference software for H.264/AVC SVC.

The proposed technique is a block-matching based motion estimation algorithm working in two steps, called Coarse search and Fine search. The Coarse search is performed for each frame in display order, and for each 16x16 macroblock chooses the best motion vector at half pel accuracy. Fine search is performed for each frame in encoding order and finds the best prediction for each block type, reference frame and direction, choosing the best motion vector at quarter pel accuracy using R-D optimization. Both Coarse and Fine Search test 3 spatial and 3 temporal predictors, and add to the best one a set of updates. The spatial predictors for the fine search are the result of the Fine search already performed for the previous blocks, while the temporal predictors are the results of Coarse Search scaled by an appropriate coefficient. This scaling is performed since in the Coarse search each picture is always estimated with respect to the previous one, while in the Fine Search the temporal distance between the current picture and its references depend on the temporal decomposition level. Moreover in Fine search the number and the value of the updates tested depend on the distance between the current picture and its references. These sets of updates are the result of a huge number of simulations on test sequences with different motion features.

The proposed algorithm has been tested on the set of test sequences proposed by JVT group, using different resolutions and temporal decomposition structures. The proposed method can reduce the average coding complexity in terms of motion vector tested from 70 to 90 percent with respect to the Fast-ME JVT method, while the quality loss depends on the GOP dimension, that is the most critical parameter for the performance of the algorithm. In fact for small GOP dimensions (4 or 8) the algorithm has the same quality at equal bit-rate respect to the Fast-ME JVT method for almost all the sequences and better quality for some sequences. For medium and long GOP dimensions (16-32) the algorithm has a quality loss lower than 0.5 dB for all the tested sequences.

References

[1] Z. Chen, P. Zhou, and Y. He. Fast Motion Estimation for JVT. JVT input document, JVT-G016, march 2003.

[2] J. Reichel, H. Schwarz, and M. Wien. Joint Scalable Video Model JSVM-6. 2006.