Ensuring High Compression Ratio for Real-Time Video Streams

D. B. Mukhamedova

The transition of television broadcasting to the digital format allows to significantly increase the number of transmitted TV programs, reduce the cost of TV broadcasting, provide high-quality high-definition images, and transmit stereo images.

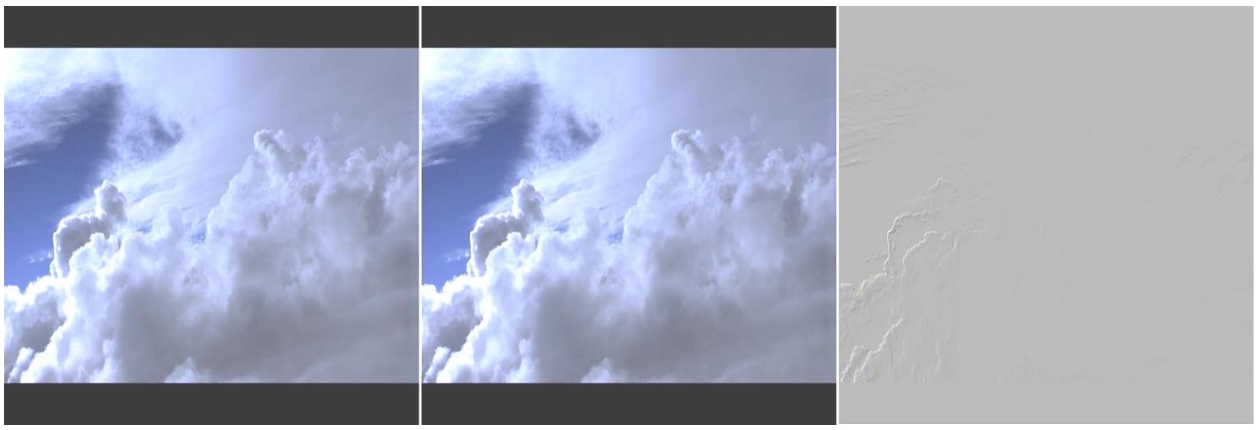
The conducted studies have shown that at high video stream data compression ratios, traditional compression algorithms based on video object motion compensation lose their effectiveness. This is due to the fact that at high compression ratios, the array of block movement vector metadata becomes comparable to the volume of compressed video data. So, to ensure an output stream speed of 2 Mbit/s, in standard definition television, the frame data volume should not exceed 9 KB, while the volume of small block coordinate metadata can reach 5 KB, which negates the advantage of this method of interframe image processing. And the time spent on finding the exact positioning of moved video object blocks is very significant. Therefore, image processing methods are required that do not create metadata or greatly minimize their volume.

Such methods include calculating the interframe difference (Fig. 1) and changing the image scale. Usually, interframe difference is not used in video stream compression due to large values of interframe information of moving images, but it does not create additional data and has high speed. Therefore, when compressed by 100 times or more, it can show greater efficiency compared to motion compensation. The fact is that in television, the main compression of a video stream is provided by eliminating temporal or interframe redundancy, since within one video story, information in adjacent frames usually changes little (Fig. 1). Thus, if only image changes relative to the reference or intermediate frame are transmitted, then fairly high video stream compression ratios can be obtained. However, due to the movement of video objects or changes in the shooting angle, the coordinates of the pixels of video objects in adjacent frames do not match. Therefore, to compensate for their mutual movement, complex interframe processing is used based on the search for image fragments of the first frame in the zones of their expected displacement in the next frame. If such fragments are found, their new coordinates are transmitted instead - displacement vectors, thus, for example, 255 bytes of a block of 16x16 pixels can be replaced by 1-2 bytes of its new coordinates. In this case, the structure of the video stream consists of a reference frame, where only intra-frame redundancy is eliminated, and one or more types of frames transmitting inter-frame differences and displacement vectors of blocks, which is implemented in the MPEG compression standards and other codecs.

Motion compensation can be performed on objects, blocks of either a fixed size or with subdivision, as well as on the entire frame. In this case, motion compensation is most often sought in the class of parallel shifts, although rotation and scaling operations can be used.

However, all methods of motion compensation of scene fragments form an additional array of metadata, which is attached to the compressed video data of the frame. This leads to the fact that the greater the compression of the video stream, the higher the accuracy of motion compensation and approximation of video objects, and accordingly the metadata array becomes larger. Moreover, if video data is usually compressed with the loss of information that is barely noticeable to the eye, then metadata does not allow any data distortions, since the integrity of the reconstructed images depends on it. Therefore, methods of noise-resistant coding are often used for metadata, which further increase their volume. Thus, to achieve video stream compression ratios of 100 times or more, it is proposed to use interframe difference in the interframe processing path, transmitting only interframe differences. And to increase its efficiency on images with a variable shooting angle, it is possible to use only global motion compensation, based on plane-parallel shifts of the entire frame and using only 4 bytes of metadata. In addition, preliminary rescaling of images additionally increases the amount of video stream compression

by 2-4 times. Thus, the integrated use of the listed image processing methods allows increasing the amount of video stream compression up to 150-200 times without significant deterioration in image quality.



Pic. 1. Images of the video stream "Clouds" and their interframe difference

The report materials provide more detailed research results, including the results of processing test images, tables and histograms.