A Study on The Application of Color Transfer Technique for Video Compression

K. Madhusudhana Rao Suman K. Mitra
Dhirubhai Ambani Institute of Information and Communication Technology
Gandhinagar, Gujarat-382009, India.
{msrao_kallu, suman_mitra}@da-iict.org

Abstract

In this paper, we explore the technique of colorization for video compression. The chrominance of the intermediate frames is discarded during encoding. At the time of decoding, chrominance values, which are dropped, are approximated from colorizing grayscale frames by transferring color from a source color frame to a targeted grayscale frame. So far human intervention is an integral part of colorization and pseudo colorization. But the present work can be looked upon as a novel application of the color transformation methodology without any human intervention for video compression. The videos regenerated at the decoding end demonstrate the potential and the utility of color transferring techniques in video compression domain.

1. Introduction

Color can be added to grayscale images in order to increase the visual appeal of images such as old black and white photos, classic movies or scientific illustrations. In addition, the information content of some scientific images can be perceptually enhanced with color by exploiting variations in chromaticity as well as luminance. The task of colorizing a grayscale image involves assigning three-dimensional (RGB) pixel values to an image, which varies along only one dimension (luminance or intensity). Mathematically the process of colorizing could be stated as mapping from g to RGB ($g \rightarrow \{R,G,B\}$) where g is the luminance or intensity of gray scale image and RGB is 3D vector. Since different colors may have the same luminance value but vary in hue or saturation, the problem of colorizing grayscale images has no inherently correct solution, meaning no exact mapping from $g \rightarrow \{R, G, B\}$. But there exists nearly correct solutions for this mapping [1,2]. Though there exists the process of pseudocoloring, [4] where the mapping of luminance values to color values is automatic, yet the choice of the colormap is commonly determined by human decision. On the other hand the process of colorizing is expected to have less human intervention possible. Since most colorization software used in the movie industry is proprietary, detailed technical documents describing the process are generally not publicly available. In the real-time colorization scenario a, traditional techniques where segmented image regions are hand colored one by one are grossly inadequate. An important step towards minimizing the amount of human intervention needed to color grayscale images was the recent work of Welsh [1]. The techniques to transfer the chromatic information from a source color image to a target grayscale image have been proposed in [1]. Unfortunately, empirical evidence suggests that the degree of similarity between these source and target images has a strong influence on the quality of the resultant colorized Image. Thus, obtention of reasonable coloring with the techniques developed in [1], in principle, still strongly dependent on human selection of an appropriate source color image for each given grayscale image.

Possibly less human intervention in selecting the source color image is required in case of video image sequences. In this work, we explore the idea of coloring gray scale frames of a video sequences from the previous color frames leading towards compression of the entire video sequence. We designed, implemented and experimentally assessed a video encoder that discards the chrominance part and encodes only the luminance component. At the decoder end the gray scale frames are colorized with intermediate color frames as the source frames. The videos thus regenerated are subjectively compared with the ones constructed with original chrominance values and the compression results are promising too. Though integrating sophisticated compression algorithms can easily enhance the performance of the present scheme, yet we have not explored this idea in the present paper. The compression algorithm based on DCT [5,6], wavelet [7,8] could be incorporated.

This paper is organized as follows. Section 2 discusses the various steps of the proposed compression algorithm and reports on some experimental results. Section 3 draws the conclusions.

2. Colorization

In this section we are going to present the process of colorization suggested in [1]. The basic assumption for colorization is that the reference color image should posses all the colors that are expected to be present in the targeted grayscale image. Missing of one or more colors in the reference image leads to unrealistic colorization. The general procedure for transferring color requires a few steps. First, color image is converted into the $l\alpha\beta$ color space. Next, for each pixel in the grayscale image in scan-line order, select the best matching sample in the color image using neighborhood statistics. The best match is determined by using a weighted average of pixel luminance and the neighborhood statistics. Variance could be one of the measures of neighborhood statistics [1]. The chromaticity values (α, β) of the best matching pixel are then transferred to the grayscale image to form the final image. Color transfer using swatches involves the same global image matching procedure but only between the source and target swatches. The colorized pixels in the target swatch regions are then used as the source pixels for the color transfer to the remaining non-colorized pixels using a texture synthesis approach [1]. This automatic, global matching procedure works reasonably well on images when corresponding color regions between the two images also correspond in luminance values. However, regions in the target image that do not have a close luminance value to an appropriate structure in the source image will not appear correct.

In order to account for global differences in luminance between the two images we perform luminance remapping to linearly shift and scale the luminance histogram of the source image to fit the histogram of the target image. This helps create a better correspondence in the luminance range between the two images but does not alter the luminance values of the target image.

3. The Methodology

The methodology consists of two things, 1) designing encoder and 2) designing decoder. The procedure at the encoder is quite simple as there is no additional computation. The video is divided into groups of frames. The number of frames in a group is chosen to be eight. For each of the group, the chrominance part of the first frame is down sampled in both the directions whereas for the remaining frames the chrominance part is discarded. The block diagram at the encoder side is shown in Figure 1. The third component of the bock diagram (Figure 1) includes source and entropy coding. This is a standard procedure of simple video compression. It consists of motion estimation, transformation, quantization followed by run length coding. Note that source and entropy coding has not been carried out in the present experiment. Our intention here is to evaluate the performance of colorization for video compression. So only the first two components of Figure1 have been implemented and the compression is computed on this basis.

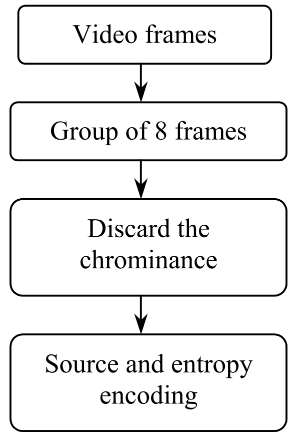

Fig. 1. Block Diagram of Encoder

The block diagram of the decoder is shown in Figure 2. The first step in the process is the inverse process of entropy encoding done at the encoder end. It typically consists of variable length decoding, inverse transform, and grayscale frame reconstruction. In the present experiment this has been discarded as no encoding has been done in the encoder. The chrominance part of the first frame is retained while the subsequent frames of a group will be gray scale. The color transformation is required for these frames. The procedure for the color transformation adopted here is as follows. For the second frame, the first frame is acting as a reference color frame to transfer color.

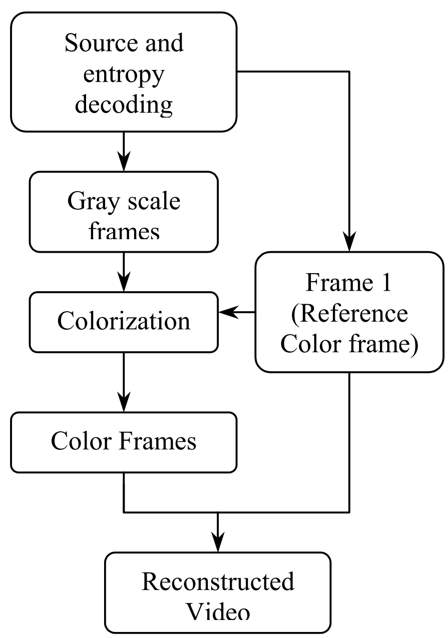

Fig. 2. Block Diagram of Decoder

The colored second frame now acts as a reference color frame for the third frame and so on. In order to transfer chromaticity values from the source to the target, each pixel in the grayscale frame must be matched to a pixel in the color frame. The comparison is based on the luminance value and neighborhood statistics of that pixel. The l values in $l\alpha\beta$ space determine the luminance value. The neighborhood statistics are computed over the frame and consists of the standard deviation of the luminance values of the pixel neighborhood. As the variation is not too high between consecutive frames, the neighborhood statistics are not necessary in case of video image sequences. A simple luminance matching works in almost all the cases. The search window of the matching process depends on the optical flow between the frames. The more the motion between the frames, the bigger should be the search window. The regenerated video frames are very close with the original color frames.

4. Results

The luminance frames can be coded using standard video compression techniques or can make use of 3D DCT. The main idea is to discard the chrominance part at the encoder end and colorize the grayscale frames at the decoder end. This will reduce the chrominance-coding overhead from the encoder. If a group of frames is taken as eight then for eight frames, chrominance information of only one frame is transmitted. If N*M is the size of the frame, and chrominance values of the reference color frame is down sampled by 2, then the total number of pixels that has to be processed is 1.5*N*M. For 1 sec of video with 24 fps, this turns to be 36*N*M. Using our proposed technique the number of pixels will be reduced to 25.5*N*M with almost 29% reduction in the number of pixels for processing. This reduction is independent of the encoding technique used at the encoder. The colorization methodology discussed in earlier section has been implemented for three video streams viz., foreman, news and Stefan. In each case streams of 6 seconds are encoded (keeping chrominance of the first frame of each group) and decoded (colorization). The results are shown in figure 3. For each video stream the original frames, first color frame along with subsequent grayscale frames and the reconstructed color frames are presented.

5. Conclusions

We have presented a simple scheme of video compression by discarding the chrominance information for most of the frames at the time of encoding and transferring the color from the color frames during reconstruction. The proposed technique reduces the number of pixels that are going to be processed for compression. The search window for luminance matching at the time of color transfer depends on the motion or optical flow between the frames. If the motion is high then the search window should be sufficiently large. This will increase the computational time at the decoder end. More compression can be achieved if one uses standard video compression standards or other compression algorithms. We would like to carry further research towards this direction.

References

[1] Tomihisa Welsh, Michael Ashikhmin and Klaus Mueller "Transferring Color to Greyscale Images", In Proceedings of ACM SIGGRAPH 2003.

[2] Luiz Filipe, M. Vieira Rafael, and D. Vilela Erickson R. do Nascimento "Automatically choosing source color images for coloring grayscale images", Proceedings of the XVI Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI 2003).

[3] Efros, Freeman, "Image Quilting for Texture Synthesis and Transfer", In Proceedings of ACM SIGGRAPH 2001.

[4] Gonzalez, and Wintz, "Digital Image Processing", Addison-Wesley Publishing, 1987.

[5] Abdelhalim, Salama, "Implementation of 3D-DCT based video encoder/decoder system", Signals, Circuits and Systems, 2003. SCS 2003. International Symposium on, Volume: 2, pp. 389 - 392, 2002.

[6] Implementation of 3D-DCT based video encoder/decoder system Bakr, M.; Salama, A.E.; Circuits and Systems, 2002. MWSCAS-2002. The 2002 45th Midwest Symposium on, Volume: 2, pp. 13-16, 2002

[7] J.R. Ohm, "Three-Dimensional Sub band Coding with Motion Compensation", IEEE Trans. Imaging Processing, vol. 3, no. 5, pp. 559-571, September 1994.

[8] A. Seeker, D. Taubman, "Motion compensated highly scalable video compression using and adaptive 3D wavelet transform based on lifting", Proc. IEEE Intl. Conf. on Image Processing, pp. 1029-1032, October 2001.

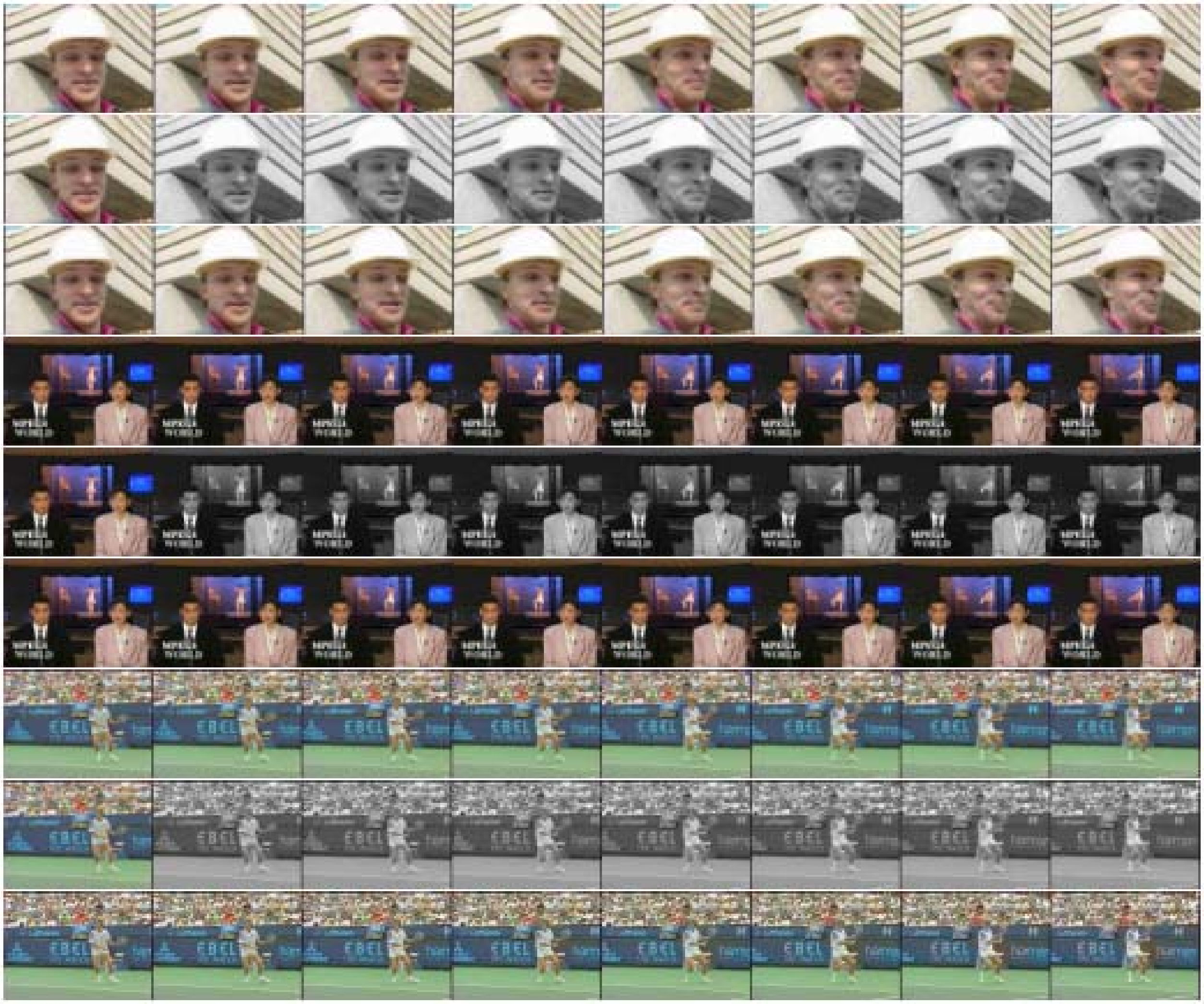

Fig 3. Rows 1,4 and 7 shows the original color frames of the foreman, news and Stefan sequences respectively. Rows 2,5 and 8 shows the corresponding grayscale frames. Rows 3,6 and 9 shows the colorized frames after color transfer.