MOTION VECTOR PROCESSING USING THE COLOR INFORMATION

Ai-Mei Huang and Truong Nguyen

Video Processing Lab
ECE Dept, UCSD, La Jolla, CA 92093
E-mail: aihuang@ucsd.edu, nguyent@ece.ucsd.edu

ABSTRACT

In this paper, we investigate color distribution around object edges and further exploit this information for motion vector reliability analysis. Since the motion vectors are often estimated only based on luminance component, the resulting motion vectors may become unreliable for the areas with similar intensity. By making use of the color information for the received residual energy calculation, these unreliable motion vectors can effectively be identified. Moreover, the motion boundaries can be easily detected due to stronger residual energy distribution on the object boundaries. This characteristic can also be used to assist the motion vector processing in motion-compensated frame interpolation so that the interpolated object structures can be well maintained. Experimental results show that using same motion vector processing approach but with additional consideration on color information outperforms other methods.

Index Terms— color information, motion compensated frame interpolation, frame rate up conversion, motion vector processing

1. INTRODUCTION

In order to reduce computational complexity, most of video encoders usually perform block-based motion estimation to obtain motion vectors (MVs) by minimizing the sum of absolute difference (SAD). Since these MVs are estimated to maximize the coding efficiency instead of finding true motion, the received MVs are often unreliable and cannot be directly used for applications that require correct motion vector field (MVF), such as motion-compensated frame interpolation (MCFI). To overcome the motion reliability problem for MCFI, a more robust block-based hierarchical motion estimation was presented in [1], in which three different search window sizes are employed to find both global and local motions. Haan *et al.* proposed using motion temporal and spatial correlation for motion estimation [2]. This method was further improved by adopting different block sizes near the object boundaries [3]. Instead of considering the motion correlation, the method in [4] considered the pixel correlation, boundary pixel differences between the surrounding well-established interpolated blocks and the current predicted block, to find more accurate motion. Conventional approaches exploited the concept of the smooth motion flow and the neighboring pixel information to obtain better motion quality for MCFI. However, most of them search the motion in luminance domain and it is not clear whether color information has been considered in the literature.

Color information has been shown to be effective in the object edges detection due to its insensitivity on specular reflection that can prevent false edge detection as compared to luminance-based methods [5]. An affine motion estimation model using image segmentation in color domain was proposed in [6]. To obtain better motion tracking results of human body, Lee *et al.* employed a edge refinement for labeled regions using the color information [7]. Since the color has sharper and more consistent variations between object boundaries, applications that need accurate object edge information, often take the color information to assist the image segmentation process. Comparing to other conventional luminance-based approaches, their experimental results also show that the visual quality can benefit from using the color information.

In our previous work [8], we proposed a MV processing approach to correct unreliable MVs from the bit-stream for MCFI. The unreliable MVs are identified by analyzing the received residual energy based on 8×8 block size. During the MV classification, the color information was found very useful for the unreliable MV detection especially in the areas where the luminance component tends to distribute uniformly. However, further analysis on the chrominance component has never been conducted. In this paper, we would like to examine the color information and analyze how the chrominance components can be used to assist the MV processing in MCFI.

The rest of this paper is organized as follows. We first analyze the color distribution around the object boundaries in Section 2. The unreliable MV classification using the color energy distribution is presented in Section 3.1. We describe the MV processing method using color information in Section 3.2. The simulation results are shown in Section 4. Finally, the conclusions are given in Section 5.

2. COLOR INFORMATION

The luminance components have stronger intensity distribution than the chrominance components as shown in Fig. 1(a), in which the Y, Cb, and Cr maps are demonstrated clockwise. As a result, the conventional motion estimation and motion vector processing approaches often ignore the color information due to the concern of the complexity. On the other hand, the color information has been widely exploited in image edge detection and image segmentation fields. This is because some of color characteristics are distinct from luminance, such as the insensitivity in highlight areas, which can be further used in preventing the false detection of object edges.

Fig. 1(b) and (c) show the results of the object edge detection for the decoded FOREMAN frame 13 based on Y, Cb, and Cr components, respectively. Comparing these two results, we can observe that the chrominance improves the edge identification for the static text, face features, the cap, and the shirts. With additional color information, not only the object edges become sharper but the areas that luminance cannot produce clear edges are effectively detected in Fig. 1(c). The reason we analyze the color gradient strength around

This work is supported by NVIDIA Inc.

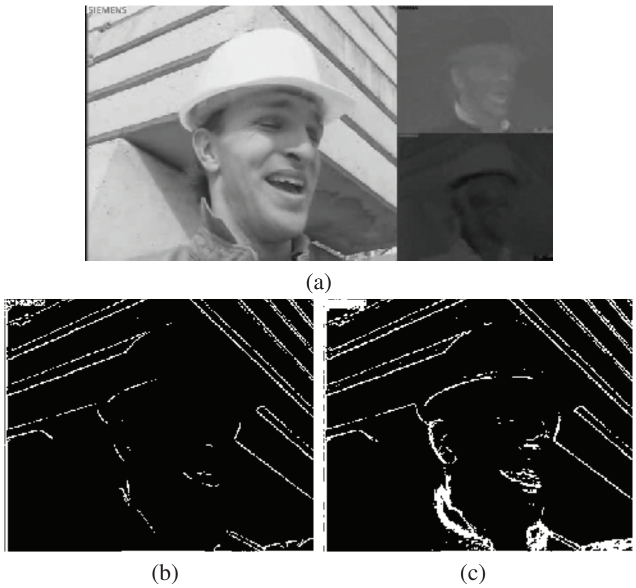

Fig. 1. (a) The Y, Cb, and Cr components for the frame 13 of FOREMAN sequence. (b) The object edge detection map using the luminance information. (c) The object edge detection map using both the luminance information and the chrominance information.

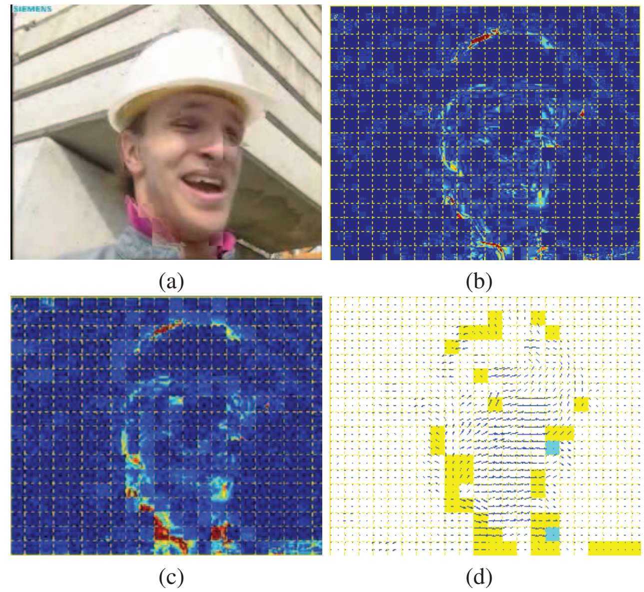

Fig. 2. (a) Interpolation result of the frame 14 of FOREMAN sequence using direct MCFI from reconstructed frames 13 and 15. (b) The luminance residual energy map. (c) The combination map of the luminance residual energy and the chrominance residual energy. (d) The MV reliability classification map for the reconstructed frame 15 based on the combination residual energy map (yellow MBs), and intra-MBs (cyan MBs).

the object boundaries is that if the moving objects have sharp edges, the ambiguous motions seem more unlikely to appear, except the areas with similar patterns. According to this assumption and the edge result in Fig. 1(c), the color seems to provide useful information to differentiate object boundaries, which we can further use to refine the motion on the motion boundaries.

The block-based motion estimation usually has the drawback that the estimated motion can only represent the major movement of the block, and the areas with different motion will be compensated using high residual values. Therefore, we classified the unreliable MVs using their associated residual energy from the bit-stream in [8]. From Fig. 1(b), we can observe that the luminance has very smooth variations around the face and shirts areas. Since the bit-stream motion is mainly determined using the luminance difference, the motion can be easily wrong in these areas. To resolve this, we should take advantage of the color information around the motion boundaries to identify the MVs with small luminance difference but actually pointing to neighboring objects with similar intensity.

We use Fig. 2 and Fig. 3 to further illustrate how the MV reliability classification can benefit from using color information. The *direct* interpolated result as shown in Fig. 2(a), has many artifacts around the nose and the shirts due the received unreliable MVs. Comparing the received residual energy in Fig. 2(b) and (c), the unreliable MVs around the shirts can only be detected using both luminance and chrominance information. In shirts region, the luminance seems to have uniformly distribution so that the encoder always chooses the face motion for the shirts. From the edge detection result in Fig. 1(c), we can also observe that the color has stronger gradients than luminance around the shirts. The corresponding MV reliability map using combined residual energy is presented in Fig. 2(d) where macroblocks (MBs) having at least one unreliable MV are denoted by yellow color. According to this map, most of artifacts in Fig. 2(a) can be identified by analyzing the high residual energy values. Another comparison is given in Fig. 3. The pavement and lawn have very similar intensity, so incorrect MVs cannot be detected merely by using the luminance residual energy as shown in Fig. 3(b). However, if we consider both luminance and chrominance information during the MV reliability classification, these unreliable MVs can be effectively identified as shown in Fig. 3(c) and (d). Generally, the chrominance residual distribution is similar to the luminance components but with smoother variations. However, once the encoder favors the motion belonging to different objects due to the plain luminance changes, the color difference will become relatively large as compared to the luminance components. By imposing the color information in the residual energy, the accuracy of MV reliability classification can be greatly improved due to stronger gradient variations of the residual distribution.

3. MV PROCESSING FOR MCFI USING THE COLOR INFORMATION

In MCFI, the skipped frame can be bidirectionally interpolated based on the received MVF between two consecutive reconstructed frames, f_{t-1} and f_{t+1}, and the interpolation scheme can be represented as follows:

$$\begin{aligned} f_t(i,j) &= \frac{1}{2}\cdot f_{t-1}(i+\frac{1}{2}v_x, j+\frac{1}{2}v_y) + \\ & \quad \frac{1}{2}\cdot f_{t+1}(i-\frac{1}{2}v_x, j-\frac{1}{2}v_y) \qquad (1) \end{aligned}$$

where $\mathbf{v} = (v_x, v_y)$ is the received MVF in the bit-stream between f_{t-1} and f_{t+1}. Directly using the received MVF for MCFI often results unpleasant visual artifacts, so further MV analysis and MV correction should be considered for *direct* MCFI.

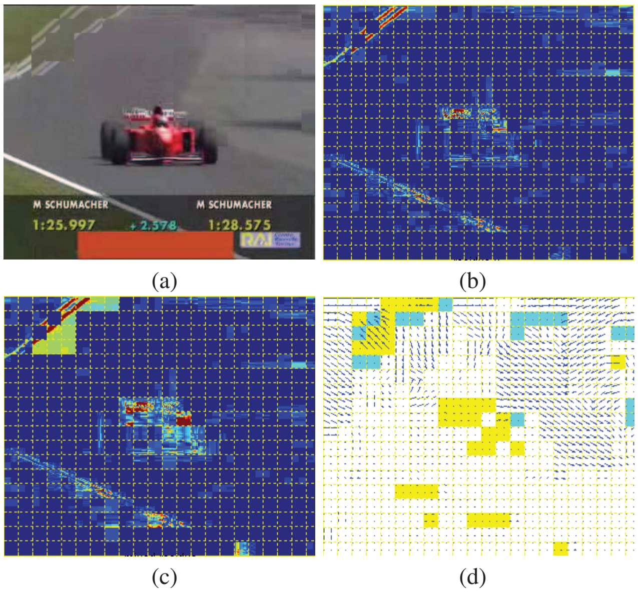

Fig. 3. ((a) Interpolation result of the frame 56 of FORMULA 1 sequence using direct MCFI from reconstructed frames 55 and 57. (b) The luminance residual energy map. (c) The combination map of the luminance residual energy and the chrominance residual energy. (d) The MV reliability classification map for the reconstructed frame 56 based on the combination residual energy map (yellow MBs), and intra-MBs (cyan MBs).

3.1. Motion Vector Analysis

In [8], we have proved that the residual energy distribution is highly correlated with the reliability of the received MVF. Based on the observation in the previous section, the residual energy for MV reliability classification with color consideration can therefore be represented as follows:

$$E_{m,n} = \sum_{(i,j)\in \mathbf{b}^{Y}_{m,n}} |r_Y(i,j)| + \alpha \cdot \Big(\sum_{(i,j)\in \mathbf{b}^{Cb}_{m,n}} |r_{Cb}(i,j)| + \sum_{(i,j)\in \mathbf{b}^{Cr}_{m,n}} |r_{Cr}(i,j)| \Big) \tag{2}$$

where $r_Y(i,j)$, $r_{Cb}(i,j)$, and $r_{Cr}(i,j)$ are the reconstructed residual signals of Y, Cb and Cr components of the 8×8 block, $\mathbf{b}_{m,n}$, respectively. α is the weight used to emphasize the degree of color difference.

For selecting the α value, we only need to be careful not to overemphasize the color since the luminance is still the fundamental elements of the image. By observing what is the minimal required degree for the color factor so that the weighted color energy can effectively detect the unreliable MVs having small luminance difference, we empirically set α to be eight for 420 YUV planar clips in our experiment. The residues are embedded in the reconstructed signals during the decoding process, so there is no additional calculation required other than Eqn. (2).

The MV classification process is the same as the method in [8]. We compare $E_{m,n}$ to a predefined threshold, ε_1, based on the combined residual information and assign the MVs to three different reliability levels according to $E_{m,n}$ scale, coding types, and the reliability levels of neighboring MVs, which can be written as follows:

$$MVCM(m,n) = \begin{cases} L_1, & \text{if } E_{m,n} \geq \varepsilon_1, \\ L_2, & \text{if MVs adjacent to MBs that have MVs} \in \mathrm{L}_1, \\ L_3, & \text{otherwise.} \end{cases}$$

After creating the MV reliability map, the MB merging process and MV selection process will be adopted accordingly.

3.2. Motion Vector Correction using the Color Information

From [8], the adjacent MBs will be merged as a group using the residual energy distribution for the MV correction process. For each merged group, the best MV, $\mathbf{v}^*_b$, is selected from the MV candidate set, which is obtained from the neighboring MVs and the MVs within each merged group, by minimizing the absolute bidirectional prediction difference ($ABPD$) between forward and backward predictions. This process can be represented as follows:

$$\mathbf{v}^*_b = \arg\min_{\mathbf{v}\in S}(ABPD(\mathbf{v})),$$

where S denotes the set of the MV candidates and $ABPD$ is described in Eqn. (3). In Eqn. (3), G and G' denote the merged groups based on luminance and chrominance components, respectively. It is noted that we consider both luminance and chrominance information in $ABPD$ calculation in Eqn. (3) and use the same weighting factor as in Eqn. (2). The selected MV will be regarded as the major motion and be assigned to all MBs within the merged group. The areas with different motion will be examined using the bidirectional prediction difference energy (BPD) for the subsequent MV refinement process. The bidirectional energy distribution, $BPD(m,n)$, can be obtained in a similar way:

$$BPD_{m,n} = BPD^{Y}_{m,n} + \alpha \cdot (BPD^{Cb}_{m,n} + BPD^{Cr}_{m,n}),$$

where $BPD^{Y}(m,n)$, $BPD^{Cb}(m,n)$, and $BPD^{Cr}(m,n)$ are the sum of bidirectional prediction difference for Y, Cb, and Cr components of block $\mathbf{b}_{m,n}$ using the updated MV $\mathbf{v}^*_{m,n}$, respectively.

4. SIMULATIONS

In this section, we present simulation results to evaluate the MCFI performance with chrominance processing. Three different methods are used for visual comparisons, direct interpolation, and the methods in [8] without color consideration and with color consideration, respectively. Two video sequences, FOREMAN and FORMULA 1, of CIF frame resolution are used, in which the motion estimation easily produces ambiguous results since the luminance have much smoother distribution than other clips. They are all encoded using H.263 with even frames skipped and the skipped frames are interpolated to evaluate different interpolation schemes.

The visual comparisons are presented in Fig. 4 and Fig. 5. In Fig. 4(c), although the artifacts around the nose and the eye are reduced using the MCFI method in [8], a lot of unreliable MVs still cannot be identified if we do not consider the color information during the MV processing. These artifacts are removed in the interpolated result as shown in Fig. 4(d), since the chrominance information sharpens the residual energy and BPD distribution and the unreliable MVs around the shirts and face areas can be identified and be corrected accordingly. In Fig. 5, the intensity between grass and pavement is very similar, so the motion estimation easily fails on

$$
\begin{aligned}
ABPD(\mathbf{v}) = \quad & \sum_{i,j \in G} |f_{t-1,Y}(i + \tfrac{1}{2}v_x, j + \tfrac{1}{2}v_y) - f_{t+1,Y}(i - \tfrac{1}{2}v_x, j - \tfrac{1}{2}v_y)| + \\
& \alpha \cdot (\sum_{i,j \in G'} |f_{t-1,Cb}(i + \tfrac{1}{4}v_x, j + \tfrac{1}{4}v_y) - f_{t+1,Cb}(i - \tfrac{1}{4}v_x, j - \tfrac{1}{4}v_y)| + \\
& \sum_{i,j \in G'} |f_{t-1,Cr}(i + \tfrac{1}{4}v_x, j + \tfrac{1}{4}v_y) - f_{t+1,Cr}(i - \tfrac{1}{4}v_x, j - \tfrac{1}{4}v_y)|).
\end{aligned} \tag{3}
$$

(a) (b) (c) (d)

Fig. 4. The interpolated results of frame 14 of FOREMAN sequence using (a) the original frame 14, (b) direct MCFI (PSNR: 32.11dB), (c) the method in [8] without color consideration (PSNR: 30.10dB), and (d) the method in [8] with color consideration (PSNR: 30.43dB), respectively.

(a) (b) (c) (d)

Fig. 5. The interpolated results of frame 56 of FORMULA 1 sequence using a) the original frame 56, (b) direct MCFI (PSNR: 29.68dB), (c) the method in [8] without color consideration (PSNR: 29.58dB), and (d) the method in [8] with color consideration (PSNR: 31.12dB), respectively.

the white lines areas as the camera pans to catch up with the race car. Comparing to the interpolated results in Fig. 5 (b) and (c), the color components can effectively detect the unreliable MVs and find more suitable motion during the bidirectional MV processing, so the artifacts do not appear in Fig. 5 (d).

5. CONCLUSIONS

In this paper, we present using color information for the MV reliability classification and MV processing in MCFI as well. Those unreliable MVs with small luminance difference can be effectively detected and corrected if the color information is explicitly considered in the residual energy analysis and the motion refinement process. From our simulation results, the MV processing approach using both luminance and chrominance outperforms other methods in both visual quality.

6. REFERENCES

[1] R. Thoma and M. Bierling, "Motion compensating interpolation considering covered and uncovered background," *Signal Processing: Image Communication*, vol. 1, pp. 191–212, 1989.

[2] G. Haan, P. W. A. C. Biezen, H. Huijgen, and O. A. Ojo, "True-motion estimation with 3-D recursive search block matching," *IEEE Trans. Circuit and System for Video Technology*, vol. 3, no. 5, pp. 368–379, Oct. 1993.

[3] R. Braspenning and Gerard de H., "Efficient motion estimation with content-adaptive resolution," *Proc. ISCE*, pp. E29–E34, Sept. 2002.

[4] C. Yang, P. Tao, and S. Yang, "An adaptive frame interpolation algorithm using satistic analysis of motions and residual energy," *In Proc. MMSP*, pp. 235–240, Sept. 2008.

[5] T. Gevers and H. Stokman, "Classifying color edges in video into shadow-geometry, highlight, or material transitions," *IEEE transactions onmultimedia*, vol. 5, no. 2, pp. 237–243, June 2003.

[6] P. E. Eren, Y. Altunbasak, and A. M. Tekalp, "Region-based affine motion segmentation using color information," pp. 3005–3008, 1997.

[7] N.S. Lee, T. Yamasaki, and K. Aizawa, "Hierarchical mesh decomposition and motion tracking for time-varing-meshes," *in Proc. ICME*, pp. 1565–1568, 2008.

[8] A.-M. Huang and T. Nguyen, "A novel multi-stage motion vector processing method for motion compensated frame interpolation," *in Proc. ICIP'07*, pp. 389–392, 2007.