Intra-and-Inter-Constraint-Based Video Enhancement Based on Piecewise Tone Mapping

Yuanzhe Chen, Weiyao Lin, *Member, IEEE,* Chongyang Zhang, *Member, IEEE,* Zhenzhong Chen, *Member, IEEE,* Ning Xu, and Jun Xie

***Abstract*—Video enhancement plays an important role in various video applications. In this paper, we propose a new intra-and-inter-constraint-based video enhancement approach aiming to: 1) achieve high intraframe quality of the entire picture where multiple regions-of-interest (ROIs) can be adaptively and simultaneously enhanced, and 2) guarantee the interframe quality consistencies among video frames. We first analyze features from different ROIs and create a piecewise tone mapping curve for the entire frame such that the intraframe quality can be enhanced. We further introduce new interframe constraints to improve the temporal quality consistency. Experimental results show that the proposed algorithm obviously outperforms the state-of-the-art algorithms.**

I. Introduction

VIDEO services have become increasingly important in many areas including communications, entertainment, healthcare, and surveillance [1], [4], [8]–[10]. However, in many applications, the quality of video service is still hindered by several technical limitations such as poor lightening conditions, bad exposure level, and unpleasant skin color tone [8], [9]. Thus, it is crucial to enhance the perceptual quality of videos. In this paper, we focus on two issues for video enhancement.

A. Intraframe Quality Enhancement With Multiple Regions of Interest (ROIs)

Since a frame may often contain multiple ROIs, it is desirable for the enhancement algorithm to achieve high intraframe quality of the entire picture where multiple ROIs can be adaptively and simultaneously enhanced. However, many existing works [3], [4], [11]–[14] focus on enhancing the intraframe quality of a picture based on some predefined global metrics. Since these methods do not consider region differences within an image, they cannot guarantee all the important regions inside the image will be enhanced properly.

Other approaches [7], [8], [12] identify and improve the perceptual quality of some specific regions in an image. Shi *et al*. [7] and Battiato *et al*. [12] designed exposure correction methods [3], [4] to perform the exposure correction based on the features of some relevant regions. Liu *et al*. [8] proposed a learning-based color tone mapping method to conduct global color transfer by turning the color statistic of the face region according to a pretrained set. Although these methods can improve the quality of some specific regions, such as the face, the quality of other regions may be deteriorated. Thus, the overall perceptual quality may not be appealing.

There are some approaches that try to enhance multiple regions simultaneously. Some methods [11] adapt enhancement strategy according to the local neighborhood characters around each pixel. Although these methods suitably consider the image local characters, they are still not region-based methods since the ROIs cannot be specified. Other methods [9], [10] enhance multiple ROIs within an image by segmenting the image into regions and performing color tone mapping for each region, respectively. However, these tone mapping methods are highly dependent on the segmentation results and may result in "fake" edges between different regions.

B. Interframe Quality Enhancement Among Frames

Most of the existing enhancement algorithms only focus on improving the intraframe qualities within a single frame or an image. They are not suitable for enhancing videos since the interframe quality consistencies among frames are not considered. Some state-of-the-art algorithms can be extended for enhancing interframe qualities under some specific applications. For example, Liu *et al.* [8] proposed a learning-based method for video conferencing where frames share the same tone mapping function if their backgrounds do not change much. Although this method can achieve good interframe quality in video conferencing scenarios, it cannot be applied to other scenarios if the video backgrounds or contents change frequently. Toderici *et al.* [13] introduced a temporally coherent method by combining the frame feature and the shot feature to enhance a frame. Their method can effectively enhance both the shot-change frames and the regular frames. However, this method relies on the performance of shot detector and fails to suitably

Manuscript received September 20, 2011; revised December 16, 2011 and February 22, 2012; accepted March 27, 2012. Date of publication June 8, 2012; date of current version January 9, 2013. This work was supported in part by the National Science Foundation of China, under Grants 61001146, 61025005, and 61101147, in part by the Chinese National 973 Project, under Grant 2010CB731401, in part by the Open Project Program of the National Laboratory of Pattern Recognition (NLPR), and in part by the SMC Scholarship of Shanghai Jiao Tong University. This paper was recommended by Associate Editor S. Battiato.

Y. Chen, W. Lin, C. Zhang, N. Xu, and J. Xie are with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: yzchen0415@sjtu.edu.cn; wylin@sjtu.edu.cn; sunny_zhang@sjtu.edu.cn; xn8812@sjtu.edu.cn; jxie@sjtu.edu.cn).

Z. Chen is with MediaTek USA, Inc., San Jose, CA 95134 USA (e-mail: zzchen@ieee.org).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCSVT.2012.2203198

enhance the interframe quality within a shot. Furthermore, Sun *et al.* [1] used additional hardware to compensate for the lighting condition for keeping consistency. Since this method has specific system requirements, it cannot be easily applied in other applications. Therefore, it is desirable to develop a more generalized algorithm that can handle the interframe quality enhancement of various videos.

In this paper, a new intra-and-inter-constraint-based (A+ECB) algorithm is proposed. The proposed algorithm analyzes features from different ROIs and creates a "global" tone mapping curve for the entire frame such that different regions inside a frame can be suitably enhanced at the same time. Furthermore, new interframe constraints are introduced in the proposed algorithm to further improve the interframe qualities among frames. Experimental results demonstrate the effectiveness of our proposed algorithm.

The remainder of this paper is organized as follows. Section II describes the motivation as well as the detailed process of our A+ECB algorithm. Section III shows the experimental results. Finally, conclusions are drawn in Section IV.

II. A+ECB Video Enhancement

A. Motivations

As mentioned, most existing approaches have various limitations in enhancing videos. Fig. 1 shows an example. Fig. 1(c) and (e) shows the enhanced results by the modified global histogram equalization algorithm [4] and a region-based method [8], respectively.

From Fig. 1(c), we can see that since the image is enhanced based on a global contrast metric without considering the region difference, some of the important regions such as the face are not properly enhanced. Compared with Fig. 1(c), since the region-based method [8] identifies the face region and performs enhancement accordingly, the visual quality of Fig. 1(e) is much improved. However, since the tone mapping function trained from the face region will be applied to the entire image, the quality of some other regions such as the screen becomes poorer [e.g., the character on the screen in Fig. 1(e) becomes difficult to see]. Similarly, when extending the region-based method [8] to enhance the screen region [as in Fig. 1(g)], the quality of the face region becomes less appealing.

From the above discussions, we have the following observations:

1) in order to achieve suitable enhancement results, features from ROIs need to be considered;
2) it is desirable to enhance the entire frame "globally" but with the consideration of different ROIs at the same time.

Therefore, we propose a new intra-frame-constraint-based (ACB) step for enhancing the intraframe quality of a frame, as will be described in detail in Section II-B.

Furthermore, as mentioned, most existing works [1]–[10] cannot effectively handle the interframe consistencies in a video. Although these methods may achieve proper visual qualities in each frame, the qualities among different frames

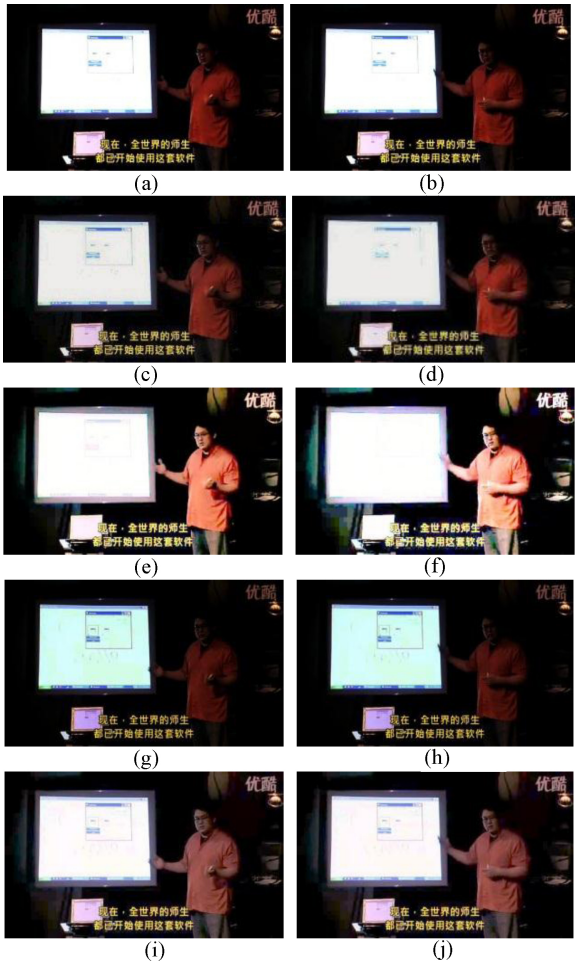


Fig. 1. (a) and (b) Original image. (c) and (d) Image enhanced by HEM [4]. (e) and (f) Image enhanced by [8]. (g) and (h) Enhanced image focusing on the screen region by [8]. (i) and (j) Enhanced image by our A+ECB method. Best viewed in color.

(i.e., interframe qualities) may vary. This inconsistency may become severe when the algorithm adaptively adopts different enhancement parameters for different frames or when the color histograms of the original videos are changing quickly. Therefore, new algorithms that can handle interframe consistency are needed.

For the ease of description, we will discuss the idea of our algorithm based on the histogram equalization modification-based (HEM) method [4]. However, it should be noted that the idea of our algorithm is general and it can be extended to other enhancement algorithms [1], [7]–[10]. Furthermore, in order to make the description clear, in the remainder of this paper, we will use boldface to represent vectors (i.e., color histograms) and use lightface to represent scalars and functions.

The HEM method can be described as (1). Instead of using the histogram distribution to construct the tone mapping function directly, the method formulated a weighted sum of two objectives as

$$\mathbf{h} = \arg\min_{\mathbf{h}^*}\left(\left|\mathbf{h}^* - \mathbf{e}\right|^2 + \lambda \cdot \left|\mathbf{h}^* - \mathbf{u}\right|^2\right) \tag{1}$$

where $\mathbf{u}$ is the uniform distribution, $\mathbf{h}$ is the desired color histogram, $\mathbf{h}^*$ is the possible candidate of $\mathbf{h}$, and $\mathbf{e}$ is the

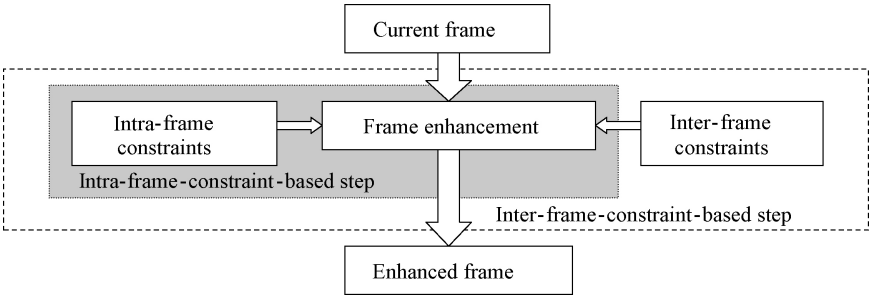


Fig. 2. Framework of the proposed A+ECB algorithm.

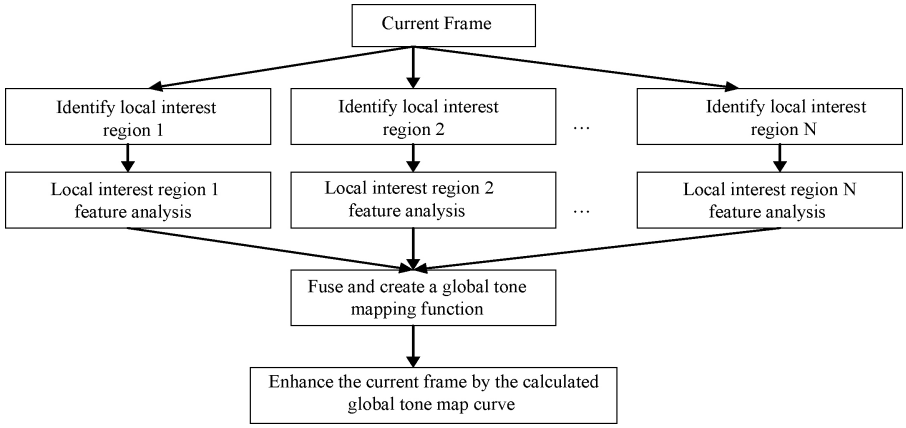


Fig. 3. ACB step process.

color histogram by traditional HE. λ is a parameter balancing the importance between **u** and **e** [4]. From (1), the desired **h** can be achieved by

$$\mathbf{h} = \left(\frac{1}{1+\lambda}\right) \cdot \mathbf{e} + \left(\frac{\lambda}{1+\lambda}\right) \cdot \mathbf{u}. \quad (2)$$

Based on (2), a tone mapping function can be calculated that enhances the original image to the desired color histogram **h** [4].

The basic idea of the HEM method is that by introducing another constraint (i.e., $|\mathbf{h}^* - \mathbf{u}|^2$), the unnatural effects in the HE histogram can be effectively reduced. However, the HEM method is still an intraframe-based method that does not consider the temporal continuities among frames.

In order to handle interframe consistency, we can extend (1) by including an additional interframe constraint by

$$\mathbf{h} = \arg\min_{\mathbf{h}^*}\left(\left|\mathbf{h}^* - \mathbf{e}\right|^2 + \lambda \cdot \left|\mathbf{h}^* - \mathbf{u}\right|^2 + \gamma \cdot \left|\mathbf{h}^* - \mathbf{h_{t-1}}\right|^2\right) \quad (3)$$

where $\boldsymbol{h}_{t-1}$ is the desired color histogram of the previous frame $t-1$ and γ is another balancing parameter handling the importance of the interframe constraint. These parameter values would directly relate to the final enhancement results. In our experiment, λ and γ are set to be 2 and 3, respectively, based on the experimental statistics. Note that these parameters and the other parameters in (4)–(8) can also be selected in an automatic way by selecting a set of parameters that maximizes the objective quality measurements in a suitable image database [11].

Therefore, we propose a new inter-frame-constraint-based (ECB) step for handling interframe constraints. The proposed ECB step will be described in detail in the following.

B. Intra-and-Inter-Constraint-Combined Algorithm

The framework of our A+ECB algorithm can be described as in Fig. 2. In Fig. 2, an input frame is first enhanced by the proposed ACB step for improving the intraframe quality. Then, the resulting frame will be further enhanced by the proposed ECB step for handling the interframe constraints. The ACB step and the ECB step will be described in detail in the following.

1) *ACB Step:* The process of the ACB step can be further described by Fig. 3. In Fig. 3, multiple ROIs are first identified from the input video frame. In this paper, we use video conferencing or video surveillance as example application scenarios and identify ROIs (such as human faces, screens, cars, and whiteboards) based on an AdaBoost-based object detection method [6]. Other object detection and saliency detection algorithms can also be adopted to obtain the ROIs. Note that since our algorithm creates a global tone mapping curve for enhancement, these regions do not need to be perfectly extracted, as will be shown in the experimental results.

After extracting and analyzing the features from these ROIs, a global tone mapping curve is created by fusing these features from different regions. Finally, the enhanced frame by this global tone mapping curve can simultaneously provide appealing qualities for different ROIs. It should be noted that the creation of the global tone mapping curve is the key part

of our algorithm. Therefore, in the remainder of this section, we will focus on describing this part.

In order to create global tone mapping curves, features need to be first extracted and analyzed for each ROI. In this paper, we utilize a simple but effective method by extracting the mean $m_{Ri,j}$ and the standard deviation $\sigma_{Ri,j}$ for each ROI R_i and for each color channel j (we use R-G-B color channels and perform enhancement in each channel independently). However, note that our algorithm is general and the feature extraction as well as the color channel processing modules can also be implemented by more sophisticated ways. For example, if one ROI include multiple major colors, we can also view each major color region as a "sub-ROI" and prefuse these sub-ROI features before fusing with other ROIs. Furthermore, the correlation constraints among color channels can also be included when performing enhancement for each color channel [4], [5], [9]. For the ease of description, we focus on discussing the global tone-mapping-curve creation from only two ROIs (i.e., R_A and R_B) in this paper. The global curve creation from more ROIs can be easily extended in an iterative way (i.e., fuse two ROIs at each time and then view the fused ROIs as an entire ROI for later fusion). The ACB strategy for fusing multiple ROI features and for creating the global tone mapping curve can be described by

$$f_{g,j}(x)=\begin{cases} f_{g,j}^{PB}(x) \;\; if \;\; \left(m_{RA,j}+\rho\cdot\sigma_{RA,j}\right)-\left(m_{RB,j}-\rho\cdot\sigma_{RB,j}\right)<\mathrm{TH} \\ f_{g,j}^{FB}(x) \;\; if \;\; \left(m_{RA,j}+\rho\cdot\sigma_{RA,j}\right)-\left(m_{RB,j}-\rho\cdot\sigma_{RB,j}\right)\geq\mathrm{TH} \end{cases} \tag{4}$$

where $f_{g,j}(x)$ is the fused global tone mapping curve for intraframe enhancement in color channel j. x is the color level value. $f_{g,j}^{PB}(x)$ is the tone mapping curve by the piecewise-based strategy for color channel j and $f_{g,j}^{FB}(x)$ is the curve by the factor-based strategy. $m_{Ri,j}$ and $\sigma_{Ri,j}$ are the mean and the standard deviation for ROI R_i in color channel j. In this paper, we assume that $m_{RA,j}<m_{RB,j}$. TH is a threshold and ρ is a parameter for reflecting the separability of the two ROIs. In our experiment, TH is set to be 50 and ρ is set to be 1 based on the experimental statistics.

From (4), we can see that if the features of the two ROIs R_A and R_B are different from each other [i.e., the pixel values for regions R_A and R_B seldom overlap and can be roughly separated: $(m_{RA,j}+\rho\cdot\sigma_{RA,j})-((m_{RB,j}-\rho\cdot\sigma_{RB,j})<TH)$], the fused global curve $f_{g,j}(x)$ will take the piecewise-based form $f_{g,j}^{PB}(x)$. Otherwise, $f_{g,j}(x)$ will take the factor-based form $f_{g,j}^{FB}(x)$.

The basic idea of our piecewise-based curve-fusion strategy $f_{g,j}^{PB}(x)$ can be described as follows; if the features of the two ROIs R_A and R_B are different from each other (i.e., the pixel values for R_A and R_B seldom overlap and can be roughly separated), then the fused global tone mapping curve $f_{g,j}(x)$ can be divided into two parts, where each part can be tuned to suitably enhance the quality of its corresponding ROI. More specifically, $f_{g,j}^{PB}(x)$ can be described by

$$f_{g,j}^{PB}(x)=\begin{cases} f_{RA,j}(x), & \text{if } x\in[0,\ P_c] \\ f_{RB,j}(x), & \text{if } x\in[P_c,\ 255] \end{cases} \tag{5}$$

where $f_{RA,j}(x)$ and $f_{RB,j}(x)$ are the piecewise tone mapping parts corresponding to the local ROIs R_A and R_B, respectively. P_c is the conjunctive point for the two parts. In this paper, $f_{RA,j}(x)$ and $f_{RB,j}(x)$ are modeled as piecewise cubic spline functions. The piecewise part $f_{RA,j}(x)$ can be derived from the constraints in (6) and (7). The second piecewise part $f_{RB,j}(x)$ can be calculated in a similar way as follows:

$$\begin{cases} f_{RA,j}(0)=0 \\ f_{RA,j}(m_{RA,j}-\alpha\cdot\sigma_{RA,j})=k_1\cdot f_{A,j}(m_{RA,j}-\alpha\cdot\sigma_{RA,j}) \\ \quad +(1-k_1)\cdot f_{B,j}(m_{RA,j}-\alpha\cdot\sigma_{RA,j}) \\ f_{RA,j}(P_c)=k_2\cdot f_{A,j}(P_c)+(1-k_2)\cdot f_{B,j}(P_c) \end{cases} \tag{6}$$

where $f_{A,j}(x)$ is the local tone mapping curve for enhancing the ROI R_A that can be easily calculated by some learning-based methods according to the extracted features in each ROI (i.e., $m_{RA,j}$ and $\sigma_{RA,j}$) [8]. k_1 $(0\leq k_1\leq 1)$ is the parameter to determine the level of the similarity between $f_{RA,j}(x)$ and $f_{A,j}(x)$. $f_{RA,j}(x)$ will be closer to $f_{A,j}(x)$ around ROI R_A's major color region (i.e., around $m_{RA,j}$) for larger k_1 values. k_2 $(0\leq k_2\leq 1)$ is a parameter to determine the pixel values around the intersection region between the two piecewise parts $f_{RA,j}(x)$ and $f_{RB,j}(x)$. Similar to k_1, $f_{RA,j}(x)$ will be close to $f_{A,j}(x)$ around the intersection region for large k_2 values, and will be close to $f_{B,j}(x)$ for small k_2 values. In our experiment, $k_1$is set to be 0.9 since we would like $f_{RA,j}(x)$ to be close to $f_{A,j}(x)$ around the R_A's major color region. k_2 is set to be 0.5 such that the color information of R_A and R_B can be equally enhanced at $f_{RA,j}(P_c)$.

Furthermore, in order to smooth the shape of the fused global curve $f_{g,j}^{PB}(x)$ and to prevent quick saturation at two ends, we also add the following derivative constraints on $f_{RA,j}(x)$:

$$\begin{cases} f'_{RA,j}(0)=0.5\cdot\dfrac{f_{RA,j}(m_{RA,j}-\alpha\cdot\sigma_{RA,j})}{m_{RA,j}-\alpha\cdot\sigma_{RA,j}} \\ f'_{RA,j}(m_{RA,j}-\alpha\cdot\sigma_{RA,j})=k_1\cdot f_{A,j}(m_{RA,j}-\alpha\cdot\sigma_{RA,j}) \\ \quad +(1-k_1)\cdot f'_{B,j}(m_{RA,j}-\alpha\cdot\sigma_{RA,j}) \\ f'_{RA,j}(P_c)=k_3\cdot\dfrac{f_{RA,j}(P_c)-f_{RA,j}(m_{RA,j}-\alpha\cdot\sigma_{RA,j})}{P_c-(m_{RA,j}-\alpha\cdot\sigma_{RA,j})} \end{cases} \tag{7}$$

where k_3 is a parameter to determine the shape of $f_{RA,j}(x)$ in higher intensities. By introducing the constraints in (7), the smoothness and the nondecreasing properties for $f_{RA,j}(x)$ can be guaranteed around the starting region (i.e., 0), the major color region of ROI R_A (i.e., $m_{RA,j}-\alpha\cdot\sigma_{RA,j}$), and the intersection region between $f_{RA,j}(x)$ and $f_{RB,j}(x)$ (i.e., P_c). The value of k_3 is determined by the values of $f_{A,j}(P_c)$ and $f_{B,j}(P_c)$. If $f_{A,j}(P_c)>f_{B,j}(P_c)$, k_3 is a real number between 0 and 1. Otherwise, k_3 is a real number between 1 and +8. In our experiment, k_3 is set to be 0.5 or 1.5 in these two cases, respectively, to make the gradient at $f_{RA,j}(P_c)$ smooth with both the $f_{RB,j}(x)$ curve part and the early part of $f_{RA,j}(x)$ around $m_{RA,j}$.

The conjunctive point P_c in (5)–(7) can be obtained by

$$P_c=\begin{cases} \dfrac{(m_{RA,j}+\alpha\cdot\sigma_{RA,j})+[m_{RB,j}-\alpha\cdot\sigma_{RB,j}+(n_{A,j}-n_{B,j})\cdot\sigma_{RA,j}]}{2} & n_{A,j}\geq n_{B,j} \\ \dfrac{[m_{RA,j}+\alpha\cdot\sigma_{RA,j}-(n_{B,j}-n_{A,j})\cdot\sigma_{RB,j}]+(m_{RB,j}-\alpha\cdot\sigma_{RB,j})}{2} & n_{A,j}<n_{B,j} \end{cases} \tag{8}$$

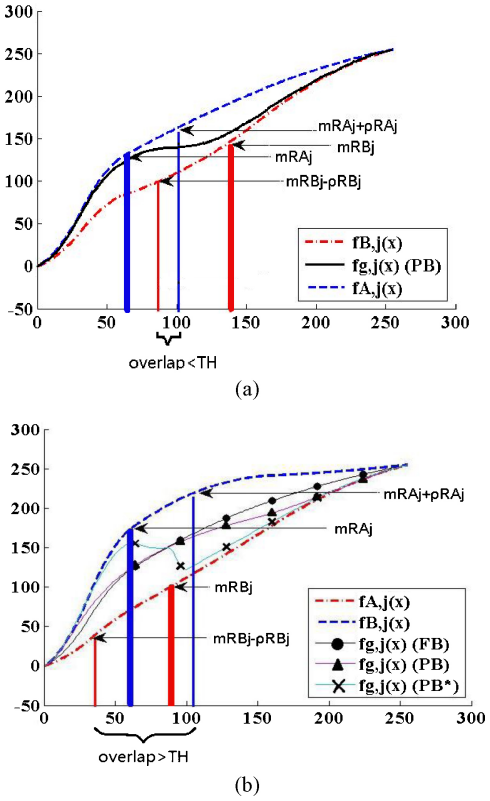


Fig. 4. (a) Fusion for different ROI features for the red color channel, including the local tone mapping curves for ROI R_A ($f_{A,j(x)}$, blue dotted curve), ROI R_B ($f_{B,j(x)}$, red dash-dot curve), and the fused global mapping curve obtained by the piecewise strategy ($f_{g,j(x)}$, black solid curve). (b) Fusion for similar ROI features for the red color channel, including the local tone mapping curves for ROI R_A (blue dotted curve), ROI R_B (red dash-dot curve), the fused global mapping curve obtained by the piecewise strategy without constraint (blue cross-mark curve), the fused global mapping curve obtained by the piecewise strategy with constraint (pink triangle-mark curve), and the fused global mapping curve obtained by the factor-based strategy (black diamond curve).

where $n_{i,j}$ ($i = A$ or B) denotes the feature difference between ROI R_i and the other ROI for channel j. They are calculated by the ratio between the pixel value overlapping area of the two regions and the variance $\sigma_{Ri,j}$ of ROI R_i. In this paper, the pixel value overlap-ping area is calculated based on the 3σ areas [2] of the two ROIs, i.e., $\alpha = 3$ in (6)–(8). Note that $n_{i,j}$ can also be viewed as the measure of the pixel-value separability between the ROIs, and it will be equal to 0 if there is no overlapping area and the pixel value of the two ROIs can be completely separated.

Fig. 4(a) shows an example of a fused global tone mapping curve by the piecewise-based strategy. From Fig. 4(a), we can see that with the piecewise-based strategy, the final tone mapping curve (i.e., the black solid curve) is composed of two parts, where the first part is close to the local curve of R_A around its 3σ area and the second part is close to the local curve of R_B around R_B's 3σ area. This way, both ROIs can be properly enhanced by the created global tone mapping function.

Although the proposed piecewise-based strategy is effective in handling the difficult problem of simultaneously enhancing multiple ROIs that have large differences in pixel value statistics (i.e., features), it is not suitable in cases when the ROIs have similar features, as it may create uneven tone mapping curves. In this case, $f_{g,j}(x)$ will take the factor-based form $f_{g,j}^{FB}(x)$ by

$$P_c = \begin{cases} \frac{(m_{RA,j}+\alpha\cdot\sigma_{RA,j})+[m_{RB,j}-\alpha\cdot\sigma_{RB,j}+(n_{A,j}-n_{B,j})\cdot\sigma_{RA,j}]}{2} \\ \frac{[m_{RA,j}+\alpha\cdot\sigma_{RA,j}-(n_{B,j}-n_{A,j})\cdot\sigma_{RB,j}]+(m_{RB,j}-\alpha\cdot\sigma_{RB,j})}{2} \end{cases} \quad (9)$$

where x is the color level value, and λ_j ($0 \le \lambda_j \le 1$) can be obtained by solving a optimization problem as

$$\lambda_j = \arg\min_{\lambda_j^*}\{||f_{g,j}^{FB}(x) - f_{A,j}(x)||_{x\in[m_{RA,j}-\alpha\cdot\sigma_{RA,j},m_{RA,j}+\alpha\cdot\sigma_{RA,j}]} + ||f_{g,j}^{FB}(x) - f_{B,j}(x)||_{x\in[m_{RB,j}-\alpha\cdot\sigma_{RB,j},m_{RB,j}+\alpha\cdot\sigma_{RB,j}]}\}. \quad (10)$$

From (9), we can see that the factor-based strategy creates the global tone mapping function by embedding the fusion information into the weighting factor λ_j. By this way, both of the ROIs R_A and R_B can be suitably enhanced.

Figs. 4(b) and 5 are examples to show the necessity of using the factor-based strategy when ROI features are similar. Fig. 5(a) is the original image, where we identify two ROIs for enhancement: the left car and the road. Fig. 5(b) is the enhancement results by the piecewise-based strategy without derivative constraints [i.e., creating the piecewise curve only based on (7) and without (8)]. Fig. 5(c) is the enhancement results by the regular piecewise-based strategy [i.e., creating the curve based on both (7) and (8)]. Fig. 5(d) is the result by the factor-based strategy. Fig. 4(b) shows the corresponding tone mapping curves for Fig. 5(b)–(d), respectively. Furthermore, the local tone mapping curves of the two ROIs in Fig. 5(a) are also plotted in Fig. 4(b) (the two dashed curves).

From Figs. 4(b) and 5, we can see that since the color features of the two ROIs are similar (i.e., their mean values are very close to each other), the pixel values for the two ROIs have large overlapping areas. If we directly use the piecewise strategy without the derivative constraints in (7), the created curve (the blue cross-mark curve) becomes uneven or even not monotonically increasing. Thus, the enhancement result [Fig. 5(b)] includes unnatural colors. If we include the derivative constraints of (7) during piecewise strategy, the created curve (the pink triangle-mark curve) becomes smoother and more reasonable. Its corresponding enhancement result [Fig. 5(c)] is much improved from Fig. 5(b). However, since the features of the two ROIs are close to each other, the tone mapping curve by the piecewise strategy is still less effective in creating a satisfactory curve balancing both ROIs. In Fig. 5(c), the street regions are tuned to be green since the piecewise curve around these color values is unsuitably tuned to favor the car while neglecting the effect of the street. Compared to the piecewise curve, the tone mapping curve created by the factor-based method can suitably balance both the ROIs and creates the most appealing result [Fig. 5(d)].

It should be noted that although the piecewise strategy is less effective for similar ROI feature cases, it will create better results than the factor-based strategy when ROI features are different. This is because when ROI features are different, the colors from different ROIs seldom overlap, and thus the tone mapping for the color values in one ROI will become less effective to the other. In this case, the piecewise strategy

Fig. 5. (a) Original image. (b) Image enhanced by the piecewise strategy without constraint factor-based strategy. (c) Image enhanced by the piecewise-based strategy with constraint. (d) Image enhanced by the factor-based strategy. Best viewed in color.

has more flexibility to make full use of the color resource to enhance multiple ROIs satisfactorily. This will be further demonstrated in the experimental results.

C. ECB Step

The ECB step can be implemented by the HEM-based framework as mentioned in (3). In our paper, besides (3), we also propose another ECB step described by (11). It should be noted that both (3) and (11) are based on the same interframe enhancement idea discussed in Section II as follows:

$$f_{g,j}^{A+E}(x) = (1 - \lambda_j^{EA}) \cdot f_{g,j}(x) + \lambda_j^{EA} \cdot f_{\text{preg},j}^{A+E}(x) \tag{11}$$

where $f_{g,j}(x)$ is the intraframe global tone mapping curve from (4). $f_{\text{preg},j}^{A+E}(x)$ is the tone mapping curve by the A+ECB method in the previous frame. λ^{EA} is the balancing parameter with the interframe constraints embedded, calculated by

$$\lambda_{\text{EA}} = \max\left(\arg\min_{\lambda_{\text{EA}}^*} |E(t) - E(t-1)|, \text{LB}\right) \tag{12}$$

where $E(t)$ is the entropy of frame t and it can be calculated by

$$E = \sum_k -p(k) \cdot \log p(k) \tag{13}$$

where $p(k)$ is the histogram value at bin k. Note that a lower-bound LB is defined in (12) to ensure that the interframe constraint can be effective in controlling the interframe consistencies. LB is set to 0.5 in our experiment. From (11)–(13), we can see that this ECB strategy embeds the interframe constraints in the balancing parameter λ_{EA} such that results of the ACB step can be shifted to a relatively stable intensity level. Thus, the visual qualities in each frame can be kept coherent and the interframe discontinuity can be effectively reduced.

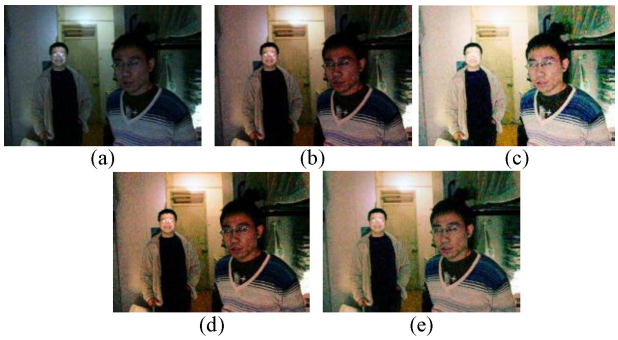


Fig. 6. (a) Original image. (b) and (c) Image enhanced by [8] based on the left and the right face region, respectively. (d) Image enhanced by the factor-based strategy. (e) Image enhanced by our proposed ACB step with the piecewise strategy. Best viewed in color.

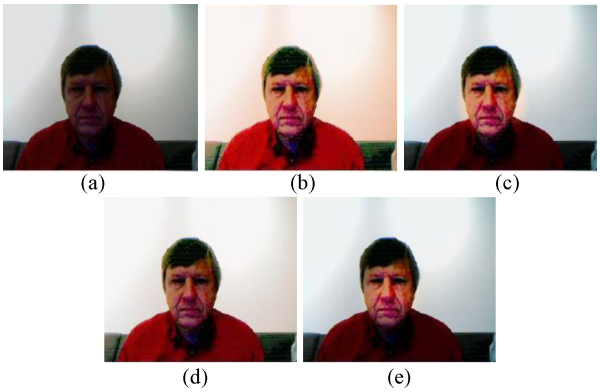


Fig. 7. (a) Original frame. (b) Frame enhanced by [8]. (c) Image enhanced by directly putting different locally enhanced regions together. (d) Image enhanced by the factor-based strategy. (e) Image enhanced by the proposed ACB step with the piecewise strategy. Best viewed in color.

III. Experimental Results

A. Results for the ACB Step

Fig. 6 compares the enhancement results for different intraframe enhancement methods. From Fig. 6, we can see that since the color of the two people are far different to each other, the learning-based method cannot properly enhance both faces simultaneously. As in Fig. 6(b) and (c), when it enhances the face of one person, the quality of another person's face becomes unsatisfactory. Although by using the factor-based strategy in Fig. 6(d), the trade-off between the two faces can be improved, it is still less effective in creating a tone mapping curve for enhancing both ROIs. We can see that the face of the right person is still dark in Fig. 6(d). Comparatively, our ACB algorithm will select the piecewise strategy that calculates a fused piecewise global tone mapping function based on both regions. By this way, we can achieve satisfactory qualities in both faces, as shown in Fig. 6(e). Moreover, although the original video from each party may have large difference in illumination conditions, the enhancement results of different users are more coherent by our algorithm. This leads to better visual experience to users.

Furthermore, Fig. 7 shows another example. From Fig. 7, we can see that since the learning-based method performs enhancement only based on the color statistics in the face region, the color of the white background is not properly

Fig. 8. (a) Original video sequence (correspond to the third, fourth, and seventh frames of the original video). (b) Enhanced video sequence obtained by HEM [4]. (c) Enhanced video sequence obtained by applying our interframe constraints (ECB).

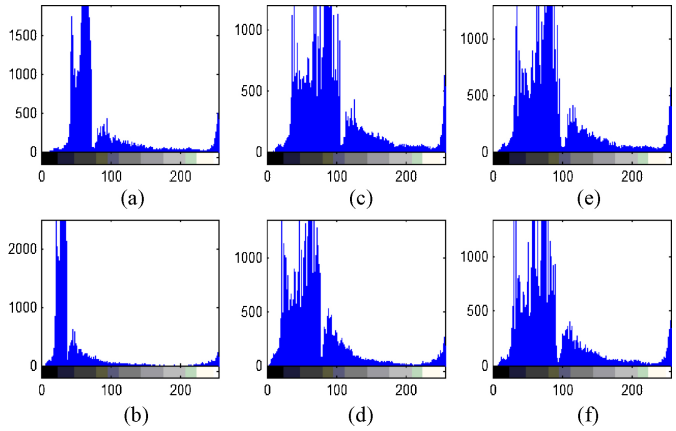


Fig. 9. (a), (c), and (e) Histograms of the first row of images in Fig. 8. (b), (d), and (f) Histograms of the second row of images in Fig. 8.

enhanced in Fig. 7(b) (i.e., it becomes red). Furthermore, although Fig. 7(c) properly enhances the background and the head, it creates fake edges around the person's head due to its local enhancement process (i.e., the red part around the head). Similar to Fig. 6, since the features of ROIs are different, the factor-based strategy in Fig. 7(d) still creates less appealing result as the background still turns a little bit red. Comparatively, the result in Fig. 7(e) is more appealing in both the face and the background regions while avoiding the fake edge effect at the same time.

B. Results for the ECB Step

In this section, we show results for our ECB step. Note that our ECB step is most effective in the following two application scenarios.

1) The original video is flicking or temporally inconsistent and the proposed algorithm can effectively reduce these interframe inconsistencies.
2) The original video sequence is temporally consistent. But after applying intraframe enhancement methods on each frame, the interframe consistency decreases.

Fig. 10. (a) Original video sequence. (b) Enhanced video sequence obtained by HEM [4]. (c) Enhanced video sequence obtained by applying our interframe constraints (ECB).

TABLE I

OBJECTIVE MEASUREMENT COMPARISON FOR DIFFERENT METHODS

	Original	HEM [4]	LB [8]	ECB	ACB	A+ECB
H	4.14	4.70	4.58	4.65	4.85	4.84
HIBTE	0.292	0.285	0.331	0.248	0.309	0.249
TAMBE (μ)	7.69	6.36	8.84	3.35	8.54	3.41
TAMBE (σ)	19.16	15.63	7.74	7.24	7.51	7.19

TABLE II

USER TEST RESULTS FOR DIFFERENT METHODS

	Orig	HEM [4]	LB [8]	ECB	ACB	A+ECB
Average Score	2.08	2.76	3.02	2.88	3.52	3.82

Note: The scores are averaged over all users and all video sequences.

The example results for cases 1 and 2 are shown in Figs. 1, 8, and 10, respectively. The videos for Figs. 8 and 10 have inconsistent interframe qualities since they are captured under a lightning weather or a frequent light-changing scenario. The ECB step in Figs. 8 and 10 is implemented by (3).

Comparing the sequences in Figs. 8 and 10, the effect of our ECB algorithm is apparent. In the left columns of Figs. 8 and 10, the illumination flickers obviously in different frames due to the unstable light condition. Although the HEM method can properly improve the intraframe quality in each frame (the middle columns in Figs. 8 and 10), the temporal inconsistency among frames still exists. Compared with the HEM method, the proposed ECB step can effectively improve both the intraframe and the interframe qualities in the video. We can see from the right columns of Figs. 8 and 10 that the interframe inconsistency is properly eliminated and the contrast in each frame is also properly tuned to become visually appealing.

Fig. 9 shows the corresponding histograms of the first and second row images in Fig. 8. From Fig. 9, we can see that due to the lightening weather condition, there is a clear distribution shift in the histograms of the original sequence [i.e., Fig. 9(a) and (b)]. When enhanced by the HEM algorithm, the distribu-

tion shift is still obvious in Fig. 9(c) and (d). However, when enhanced by our ECB algorithm, the histogram distributions of the two continuous frames are tuned to have similar shapes and distributions. This also demonstrates the ability of our method in handling the interframe consistency.

C. Results for Combining Intraframe and Interframe Steps

An example result of the proposed A+ECB algorithm is shown in Fig. 1. Similarly, when performing intraframe enhancement only based on the face region [i.e., Fig. 1(e) and (f)] or only based on the screen region [i.e., Fig. 1(g) and (h)], the other regions cannot be properly enhanced. However, by using our A+ECB algorithm, both ROIs can be enhanced properly. Furthermore, since the face region is not perfectly identified, we can see that the enhanced results in Fig. 1(e) and (f) are not temporally consistent. However, by including our ECB step in Fig. 1(i) and (j), the interframe quality of the video can be guaranteed.

Table I compares the objective measurements for different methods. Table I compares the discrete entropy (H), temporal absolute mean brightness error [TAMBE(μ)] [4], the standard deviance of the difference image between the neighboring frames [TAMBE(σ)] [4], and histogram-intersection-based temporal error (HIBTE) [14]. Normally, large H values reflect good intraframe qualities while small TAMBE(μ), TAMBE(σ), and HIBTE values imply good interframe qualities [4], [14].

From Table I, we can see that the intraframe quality (i.e., H) of the ECB method is a little degraded from that of HEM. This is because by including the interframe constraints in ECB, the enhancement extent for each frame is limited in order to keep coherent with neighboring ones. However, this intraframe quality degradation by ECB is small. Comparatively, the interframe quality improvements by ECB are obvious, which leads to an overall improvement of ECB from HEM. Similarly, although our ACB method can obviously improve the intraframe quality of the video, its interframe quality is still poor. By further combining with our ECB method, the proposed A+ECB method can achieve improvement in both the interframe and the intraframe video qualities.

Furthermore, Table II shows the results of a subjective user test experiment. In this experiment, 30 users are asked to view the enhanced videos with different methods and give evaluation scores with a range of 1–5 [8], with 1 for very poor quality and 5 for very good quality. The orders of the enhanced videos are randomly placed and unknown to the users. The scores are averaged over different users and over different video sequences, as shown in Table II. The subjective evaluation results in Table II further demonstrate the effectiveness of our proposed method.

IV. Conclusion

In this paper, we proposed a new A+ECB algorithm for video enhancement. The proposed method analyzed features from different ROIs and created a "global" tone mapping curve for the entire frame such that the intraframe quality of a frame can be properly enhanced. Furthermore, new interframe constraints were introduced in the proposed algorithm to further improve the interframe qualities among frames. Experimental results demonstrated the effectiveness of our algorithm.

References

[1] M. Sun, Z. Liu, J. Qiu, Z. Zhang, and M. Sinclair, "Active lighting for video conferencing," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 12, pp. 1819–1829, Dec. 2009.

[2] G. R. Shorack, *Probability for Statisticians*. New York: Springer, 2000.

[3] J. A. Stark, "Adaptive image contrast enhancement using generalizations of histogram equalization," *IEEE Trans. Image Process.*, vol. 9, no. 5, pp. 889–896, May 2000.

[4] T. Arici, S. Dikbas, and Y. Altunbasak, "A histogram modification framework and its application for image contrast enhancement," *IEEE Trans. Image Process.*, vol. 18, no. 9, pp. 1921–1935, Sep. 2009.

[5] E. Reinhard, M. Ashikhmin, B. Gooch, and P. Shirley, "Color transfer between images," *IEEE Comput. Graphics Applicat.*, vol. 21, no. 5, pp. 34–41, Sep.–Oct. 2001.

[6] P. Viola and M. Jones, "Robust real-time object detection," in *Proc. 2nd Int. Workshop Statist. Computat. Theories Vision*, 2001, pp. 137–154.

[7] C. Shi, K. Yu, J. Li, and S. Li, "Automatic image quality improvement for videoconferencing," in *Proc. ICASSP*, 2004, pp. 701–704.

[8] Z. Liu, C. Zhang, and Z. Zhang, "Learning-based perceptual image quality improvement for video conferencing," in *Proc. ICME*, 2007, pp. 1035–1038.

[9] W.-C. Chiou and C.-T. Hsu, "Region-based color transfer from multi-reference with graph-theoretic region correspondence estimation," in *Proc. ICIP*, 2009, pp. 501–504.

[10] Y. W. Tai, J. Jia, and C. K. Tang, "Local color transfer via probabilistic segmentation by expectation-maximization," in *Proc. CVPR*, 2005, pp. 747–754.

[11] R. Schettini and F. Gasparini, "Contrast image correction method," *J. Electron. Imag.*, vol. 19, no. 2, pp. 5–16, 2010.

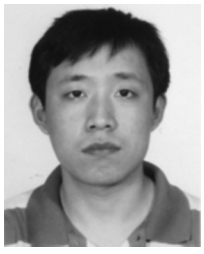
[12] S. Battiato and A. Bosco, "Automatic image enhancement by content dependent exposure correction," *EURASIP J. Appl. Signal Process.*, vol. 2004, no. 12, pp. 1849–1860, 2004.

[13] G. D. Toderici and J. Yagnik, "Automatic, efficient, temporally-coherent video enhancement for large scale applications," in *Proc. ACM Multimedia*, 2009, pp. 609–612.

[14] M. Swain and D. Ballard, "Color indexing," *Int. J. Comput. Vision*, vol. 7, no. 1, pp. 11–32, 1991.

Yuanzhe Chen received the B.S. degree in electrical engineering from Shanghai Jiao Tong University, Shanghai, China, in 2011. He is currently pursuing the M.E. degree in electrical engineering from Shanghai Jiao Tong University and Georgia Technical University, Atlanta.

His current research interests include image and video processing, machine learning, computer vision, and multimedia technologies.



Weiyao Lin (M'11) received the B.E. and M.E. degrees from Shanghai Jiao Tong University, Shanghai, China, in 2003 and 2005, respectively, and the Ph.D. degree from the University of Washington, Seattle, in 2010, all in electrical engineering.

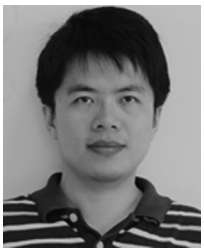
Since 2010, he has been an Assistant Professor with the Institute of Image Communication and Information Processing, Department of Electronic Engineering, Shanghai Jiao Tong University. He has authored or co-authored over 40 technical papers. His current research interests include video processing, machine learning, computer vision, and video coding and compression.

Dr. Lin is a Technical Committee Member of IEEE Visual Signal Processing and Communication, IEEE Multimedia Systems and Applications, and IEEE Multimedia Communications. He was a Guest Editor for special issues of the *Journal of Visual Communication and Image Representation* and the *Journal of Electrical and Computer Engineering*. He was a Lead Organizer of the special sessions at the IEEE VCIP'2011 Conference. He was a Technical Program Committee Member of the IEEE ISCAS, ICME, VCIP, the CVPR HAU3D Workshop, and the ACM MM SRED Workshop. He is a member of ACM.

Chongyang Zhang (M'12) received the Ph.D. degree from Shanghai Jiao Tong University, Shanghai, China, in 2008.

He is currently an Associate Professor with the Department of Electronic Engineering, Shanghai Jiao Tong University. As a Principal Investigator, he has owned or has joined over 10 projects or grants from various agencies in China, such as the National Natural Science Foundation of China Project and the National Key Technology Research and Development Program of the 12th Five-Year-Plan of China. His current research interests include video processing theory and application, in particular, video coding, video communications, intelligent video surveillance, and embedded multimedia systems. He has published over 20 international journal or conference papers on these topics.

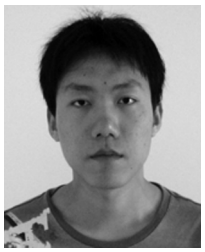
Dr. Zhang was a reviewer of various IEEE conferences and other journals. He is a member of IEICE.



Zhenzhong Chen (S'02–M'07) received the B.E. degree from the Huazhong University of Science and Technology, Wuhan, China, and the Ph.D. degree from the Chinese University of Hong Kong (CUHK), Shatin, Hong Kong, both in electrical engineering.

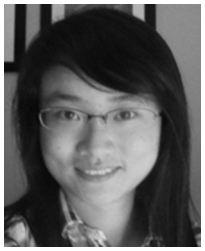
He is currently a Technical Staff Member with MediaTek USA, Inc., San Jose, CA. He was a Lee Kuan Yew Research Fellow and a Principal Investigator with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, and the European Research Consortium for Informatics and Mathematics (ERCIM) Fellow with the National Institute for Research in Computer Science and Control, Rennes, France. He held visiting positions with Polytech'Nantes, Nantes, France, Université Catholique de Louvain, Louvain-la-Neuve, Belgium, and Microsoft Research Asia, Beijing, China. His current research interests include visual perception, visual signal processing, and multimedia communications.

Dr. Chen was the recipient of the CUHK Faculty Outstanding Ph.D. Thesis Award, the Microsoft Fellowship, and the ERCIM Alain Bensoussan Fellowship. He is a Voting Member of the IEEE Multimedia Communications Technical Committee (MMTC) and an invited member of the IEEE MMTC Interest Group of Quality of Experience for Multimedia Communications from 2010 to 2012. He was a Guest Editor for special issues of the IEEE MMTC E-Letter and the *Journal of Visual Communication and Image Representation*. He is the Co-Chair of the International Workshop on Emerging Multimedia Systems and Applications in 2012. He has co-organized several special sessions at international conferences, including IEEE ICIP 2010, ICME 2010, and Packet Video 2010, and has served as a Technical Program Committee Member of the IEEE ICC, GLOBECOM, CCNC, ICME, and others. He is a member of SPIE.



Ning Xu received the B.E. degree in electrical engineering from Shanghai Jiao Tong University, Shanghai, China, in 2011. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana.

His current research interests include computer vision, machine learning, and multimedia.



Jun Xie received the B.S. degree in electrical and computer engineering from Shanghai Jiao Tong University, Shanghai, China, in 2011. She is currently pursuing the Ph.D. degree in electrical engineering from the University of Washington, Seattle.

Her current research interests include image and video processing, computer vision, and multimedia technologies.