Efficient Operator for Local Energy Measurements to Enhance an Image Signal

Sascha Cvetkovic, Member, IEEE, Johan Schirris, and Peter H.N de With, Fellow, IEEE

Abstract—For real-time imaging with digital video cameras and high-quality display with TV systems, the obtained picture quality including visibility of details, local contrast and absence of artifacts, is very important to ensure user quality acceptance. We present a multi-window real-time high-frequency enhancement scheme, in which gain is a non-linear function of the detail energy. Then we discuss the computation of commonly used local energy measurements and show that a selection of those measurements can be calculated efficiently. In our first contribution, we propose a new local energy measurement APS that can be calculated more efficiently than the existing metrics in a 2D-separable fashion. In addition, we also show that the new APS measurement gives better performance than standard energy measurements. The second contribution is the use of local contrast and a modified contrast gain formula that can substantially improve the overall algorithm performance, especially when a high-level contrast enhancement is desired. Our algorithm trades off between added contrast and “halo” artifacts, resulting in a good balance between visibility of details and an acceptable level of artifacts. The new scheme can be successfully applied to cameras and TV systems to improve their visual quality.

Index Terms—Local contrast enhancement, local energy measurement, non-linear sharpness enhancement, TV displays.

I. INTRODUCTION

For digital video cameras and high-quality TV displays, a high visibility of details is obtained amongst other ways via video signal enhancement processing. Many complex scenes as well as low-quality video footage require local sharpness and contrast improvements that should bring image details to the best possible visibility. Image enhancement techniques often depend on measuring the local signal energy for their processing. However, the computational complexity of the involved calculations is very high for large filter kernels. On the other hand, large filter kernels are needed to achieve the desired level of quality improvement due to intrinsic high image resolutions. In real-time video applications, the computation complexity should be low, which poses a significant design problem and has a clear impact on the possible solution for the image quality optimization: whatever the adopted solution is, simplicity and efficiency are a prerequisite. Global contrast control algorithms can be applied [1]-[4], however, there are often more complex situations, where contrast can be poor in some parts of the image, but adequate in other parts, or when the overall contrast is good but local contrast is low. Besides this, low-quality video footage originating from low-quality capturing devices poses significant challenges if an acceptable output video quality is required. In all these cases, locally-adaptive contrast enhancement will provide significant advantages for the perceived output image quality and the visibility of details [5]-[8], [11]-[16].

Fig. 1. (a) Multi-scale sliding window in the spatial domain: overlapping bands 1, 2, …, k, and non-overlapping bands Bk = Fk − Fk+1 marked as 1, 2’, …, k’. (b) Multi-scale enhancement by contrast gain functions [5].

In our previous work [5], [6], we proposed an image enhancement method for processing an input image I(m, n) where the enhanced output image O(m, n) resulting from the multi-scale algorithm, is described by:

$$O = \sum_{k=1}^{K} I + G_k \cdot (F_k - F_{k+1}).$$

Here, K is the number of scales used, F_k is the local mean of the kernel k and G_k = G_k(C_k, LD_k) represents the contrast gain of the kth kernel. Parameters C_k are gain factors that control the enhancement level per kernel and LD_k is the local energy deviation of kernel k (see Figs. 1(a) and 1(b)). The size of the filter kernels is in the range from (2a_1 + 1)(2b_1 + 1) till (2a_k + 1)(2b_k + 1), often increasing with two or three times at each step when going to a subsequent scale (i.e., a_k = 2a_{k-1} - 1 or a_k = 3a_{k-1} - 1). This enhancement scheme is employed for the luminance channel, while chrominance is adjusted to preserve saturation. Image improvement is obtained by amplifying medium and higher frequencies in the image by a gain that is inversely proportional to the local deviation LD, with limits. Likewise, areas that have a low level of detail will be amplified more to increase their visibility. We also pointed out that using non-overlapping frequency bands improves performance of this method (non-overlapping bands are created by subtraction of low-pass signals from subsequent scales: B_k = F_k − F_{k+1}, Fig. 1(a)), and that the shape of the rejection band of the low-pass filters and the choice of LD function play a vital role in the output video quality. We also discussed ways for improving the noise performance of the algorithm and the minimization of the “halo” artifacts [5], [6]. The principal remaining problem related to image enhancement without artifacts is that the LD calculations are very expensive in terms of HW/SW computations, because they are often 2D-inseparable. In addition, we want to have a better control of the enhancement process of the image and be
able to completely prevent “halo” artifacts. Finally, clipping of the output signal after enhancement can sometimes occur and then negatively influences the image quality.

Our approach to solve the complexity problem is to review the existing energy metrics and design a new alternative that has a good performance and at the same time can be very efficiently computed. To this end, we first describe several choices of the energy measurement LD and show their involved costs and complexity. Afterwards, we propose our new energy measurement APProximated Sum of absolute differences (APS) that is 2D-separable and thus suitable for any HW/SW implementation. APS approximates the well-known Sum of the Absolute Differences (SAD), but at the same time has better properties than SAD, especially for “halo”-artifacts suppression. APS can be used to replace the calculation of the SAD for the local energy calculation used in the process of the local contrast enhancement. However, we claim that it can also be used in all applications for arbitrary local energy measurements. We also show that it is possible to down-sample the APS computation and even further reduce the costs of HW/SW, thereby further improving the efficiency of this method.

Then, we propose a new contrast gain function that eliminates “halo” artifacts for middle-to-large edges by designing it in a more natural way. In addition, by considering the local contrast measurement, we are able to perform image enhancement based on the perception of local contrast by human visual system. Namely, we want to exercise a better control of the contrast gain function, since the amount of desired local contrast depends not only on the existing amount of contrast, but also on the signal intensity (DC level) of the local feature. Finally, we also show how to reduce clipping of the high-frequency enhancement signal when it is added to the original image.

The sequel of this paper is organized as follows. Section II describes system aspects and implementation details of several LD measurements. Section III discusses separable approximate calculations of APS measurements and its down-sampling simplification. In Section IV we propose a new contrast gain function, explain ways to include a local contrast measurement in the enhancement process and achieve the reduction of clipping of the enhanced output signal. Section V provides experimental results and Section VI concludes the paper.

II. LOCAL ENERGY MEASUREMENT METHODS FOR IMAGE/VIDEO CONTRAST ENHANCEMENT

One important aspect of the proposed detail enhancement scheme is the choice of the local signal energy measurement LD [5]. Let us now discuss some important possibilities and the related consequences for the implementation.

A. Local Standard Deviation (LSD)

LSD is an energy measurement which is very expensive in terms of hardware/calculation, but it gives very good results without almost any disturbing “halo” artifacts [5]. If $M = 2a_k + 1$ and $N = 2b_k + 1$ are vertical and horizontal kernel sizes in the $k$th signals $V^2$ and $V'D_k$ in the $k$th kernel [5].

\[
\text{LSD}_k^2(m, n) = \frac{1}{MN} \sum_{i=-a_k}^{a_k} \sum_{j=-b_k}^{b_k} (F_k(m, n) - I(m + i, n + j))^2.
\]

The 2D low-pass filtered signal $F_k$ can be pre-calculated by means of a 2D-separable convolution. To calculate the LSD value at each pixel position, we have to perform $MN = (2a_k + 1)(2b_k + 1)$ operations of memory fetching, multiplications (squaring operation) and additions, which is too expensive for real-time low-cost processing. Therefore, we propose a much cheaper solution by modifying the original equation and introducing an intermediate step of calculation.

We separate the horizontal and vertical component of the 2D calculation to reduce implementation costs. To simplify the explanation, we mark all pixel values in an image window as $X = I(m + i, n + j)$, while the average image luminance in the window centered about the current position $(m, n)$ is represented as $\bar{X} = F_k(m, n)$. The definition of energy variance $\text{LSD}_k^2$ is: $\text{LSD}_k^2 = (1/MN) \sum (X - \bar{X})^2$. Knowing that the number of elements in the summation equals $MN$, we rewrite this equation to:

\[
\text{LSD}_k^2 = \frac{1}{MN} \sum (X - \bar{X})^2 = \frac{1}{MN} \sum X^2 - \bar{X}^2 = \bar{X}^2 - \bar{X}^2.
\]

This expression can be calculated in a 2D-separable fashion. We first compute the vertical sum $V_k^2$ of $X^2$ of a certain column in the $k$th kernel (see Fig. 2), which for the $(n+j)$th column results in:

\[
V_k^2 = \frac{1}{M} \sum_{(n+j)\text{th column}} X^2 = \frac{1}{M} \sum_{i=-a_k}^{a_k} \sum_{j=\text{const}} I^2(m+i, n+j).
\]

Then we sum results of this calculation horizontally to calculate the total squared LSD:

\[
\text{LSD}_k^2(m, n) = \left[ \frac{1}{N} \sum_{j=-b_k}^{b_k} V_k^2(m, n+j) \right] - F_k^2 = \left[ \frac{1}{N} \sum_{j=-b_k}^{b_k} \left( \frac{1}{M} \sum_{i=-a_k}^{a_k} I^2(m+i, n) \right) \right] - F_k^2.
\]
sum of those values (all scaled with the number of elements). Finally, after subtracting the square of the signal average value \( F_k^2 \) we find the \( \text{LSD} \) as the square root of (5).

These operations are easy to implement, since the processing pipeline performs processing in the horizontal (pixel) direction and after finishing with one line it proceeds with the next line. However, although we can perform a 2D-separable LSD calculation, we need to calculate the sum of squares of all elements, which is expensive due to the squaring operations and large accumulators needed to preserve the intermediate results for \( V_k^2 \). We notice that in a non-separable LSD version, accumulators can be smaller, as we are summing differences between the signal and its low-pass version. Similar to the previous derivations, an unbiased version of LSD, called \( \text{LSD}_u \), is defined as
\[
\text{LSD}_u = \sqrt{\frac{1}{MN} \sum_{i=-a}^{b} \sum_{j=-b}^{b} \left| F_k(m,n) - I(m+i,n+j) \right|^2}
\]
(7)

SAD cannot be calculated in a 2D-separable fashion. The enhancement performance of SAD is somewhat better than the performance of LSD in the uniform texture areas [5], although the LSD distinguishes small and large details better than the SAD due to its quadratic function. As a result, LSD also gives less disturbing “halo” artifacts. To prove this difference in performance, we discuss the computation of various averaging metrics, such as the quadratic and arithmetic mean. Starting from the inequality of the quadratic and arithmetic mean, we will derive the validity of the above performance claims. First, for positive real numbers \( x_1, \ldots, x_p \), we define a generalized mean with exponent \( r \) (\( r \) is a non-zero real number) as in [9]:
\[
M_r(x_1, \ldots, x_p) = \left( \frac{1}{p} \sum_{i=1}^{p} x_i^r \right)^{1/r}
\]
(8)

Without the loss of generality, when \( r < 0 \), then the following inequality of means holds \( M_r(x_1, \ldots, x_p) \leq M_q(x_1, \ldots, x_p) \), where the equality sign is valid if and only if \( x_1 = x_2 = \ldots = x_p \). Some interesting cases are obtained for \( r=2 \) (quadratic mean Q), \( r=1 \) (arithmetic mean A), \( r=-1 \) (harmonic mean H) and \( r \to 0 \) (geometric mean G). For \( r = 1 \) and \( q = 2 \), we find that the inequality of quadratic and arithmetic means results in:
\[
\frac{(x_1 + x_2 + \ldots + x_p)^2}{p} \leq \left( \sum_{i=1}^{p} x_i^2 \right)/p.
\]
(9)

Second, suppose we now substitute \( x_1 \) with \( |F_k(m,n) - I(m+i,n+j)| \) for \( l = 1, \ldots, p \), covering all combinations of indices \( i \) and \( j \), we would find that \( \text{SAD}_k(m,n) \leq \text{LSD}_k(m,n) \) for all \( m, n \) and \( k \). The previous inequality becomes equality \( \text{SAD}_k(m,n) = \text{LSD}_k(m,n) \) if and only if \( x_1 = x_2 = \ldots = x_p \). This would mean that all terms \( |F_k(m,n) - I(m+i,n+j)| \) are equal for \( i = -a_k, \ldots, a_k, j = -b_k, \ldots, b_k \). In other words, if local texture is uniform, the SAD is equal to the LSD. For edges, these conditions are not satisfied, as they are actually distinguishable by large signal changes, giving rise to a non-uniform texture. The LSD would therefore give a higher output than the SAD, providing a smaller local signal gain and less “halo” artifacts. However, the SAD will provide a higher enhancement level in the other areas. This property can also be illustrated with a simple model in Fig. 3, where various mean values are presented for two positive numbers \( a \) and \( b \).
We can notice that the quadratic mean Q (representing LSD) gives the highest output, followed by the arithmetic mean A (representing SAD calculation). Similar reasoning holds for the p-dimensional case with variables $x_1, \ldots, x_p$. We can also observe that the larger the discrepancy of values $a$ and $b$ (for $a + b$ constant), the larger $Q$ will be, while $A$ will remain the same. Summarized, the LSD provides a higher penalty on edges (distinguished by large signal energy variations) than SAD. Hence, the larger value of LSD at edges gives a lower contrast gain and correspondingly leads to less “halo” artifacts.

C. Vertical SAD and Horizontal Convolution (VSHC)

Finally, we discuss a third metric called VSHC [5], which is a modification of the SAD metric, such that it can be calculated in a 2D-separable fashion. VSHC gives measurements and perceptual results which are somewhat worse than SAD, but it is simpler to implement. We separate the horizontal and vertical components of the 2D calculation to reduce computational costs. We first compute the vertical sum of absolute differences $V D_k$ between a low-pass filtered value at a certain position and all pixels in the vertical column to which that $F_k$ value belongs (see Fig 2):

$$V D_k(m, n) = \frac{1}{N} \sum_{i=0}^{N} |F_k(m, n) - I(m + i, n)|.$$  

(10)

In the second pass, we perform a horizontal convolution of the pre-calculated vertical differences $V D_k$, so that

$$V SHC_k(m, n) = \frac{1}{N} \sum_{j=0}^{N} V D_k(m, n + j).$$  

(11)

Instead of always subtracting the low-pass value of the kernel’s central pixel (as in the SAD approach), we now subtract low-pass values placed at the center of the corresponding column (and likewise, effectively increase the horizontal kernel size two times). This action decreases the energy value at the edges, providing for the increased contrast gain (and likewise more “halo”) at the edges. However, this operation is less expensive than the original SAD formula due to its considerably lower memory access.

III. APS : NEW ENERGY MEASUREMENT BASED ON THE SAD CALCULATION

A. Approximation of SAD (APS)

In this section, our new approach APS estimates the 2D non-separable SAD measurement, such that it can be computed in a 2D-separable way. However, the approximation algorithm is defined in such a way that a new metric is created, which outperforms all the other energy measurements, both on edges and on non-edge areas. The algorithm is constructed as an addition to the VSHC measurement to repair the poor performance of VSHC on strong edges. The design constraint is to make a good approximation of the SAD in a separable way, so that computational complexity remains limited. The first step is to perform the VSHC measurement. The difference of VSHC to SAD is that, instead of always subtracting a low-pass signal value from the center of the kernel, we are using a shifted version of the low-pass signal that is positioned at the center of the vertical column that we are currently processing. To prepare ourselves for the additional step that we are going to make for VSHC, we write this operation in a 2D non-separable manner and start the summation column-wise, giving

$$V SHC_k(m, n) = 1/MN:$$

$$\sum_{i=-a_k}^{a_k} \sum_{j=-b_k}^{b_k} |F_k(m, n + j) - I(m + i, n + j)|.$$  

(12)

However, we would actually like to perform a calculation with a non-shifted low-pass signal $F_k$, which equals

$$SAD_k(m, n) = 1/MN:$$

$$\sum_{i=-a_k}^{a_k} \sum_{j=-b_k}^{b_k} |F_k(m, n) - I(m + i, n + j)|.$$  

(13)

Observing each member of the SAD and VSHC sums, we will name the contribution to the SAD and VSHC calculation from a single summation term at the position $(m + i, n + j)$ as $sad_k(m + i, n + j)$ and $vshck(m + i, n + j)$, respectively. Hence, from (12) and (13), it follows that

$$sad_k(m + i, n + j) = |F_k(m, n) - I(m + i, n + j)|$$

and

$$vshck(m + i, n + j) = |F_k(m, n + j) - I(m + i, n + j)|.$$  

Furthermore, when we use the triangle inequality $|a + b| \leq |a| + |b|$, which holds for all real numbers $a$ and $b$, we obtain a new inequality (see also Fig 2):

$$sad_k(m + i, n + j) = |F_k(m, n) - I(m + i, n + j)| =$$

$$|F_k(m, n) - F_k(m, n + j) + F_k(m, n + j) - I(m + i, n + j)| \leq$$

$$|F_k(m, n) - F_k(m, n + j) + F_k(m, n + j) - I(m + i, n + j)| =$$

$$|F_k(m, n) - F_k(m, n + j)| + vshck(m + i, n + j).$$  

(14)

In other words, by adding extra elements $|F_k(m, n) - F_k(m, n + j)|$ around each pixel to the known VSHC calculation, we over-estimate the SAD. Thus, $sad_k(m + i, n + j) \leq vshck(m + i, n + j) + |\Delta F_k(j)|$. The term $\Delta F_k(j)$ represents the first order difference of the low-pass signal $F_k$ on a distance $j$. The final result for the calculation of the APS at a position $(m, n)$ becomes now:

$$SAD_k(m, n) \leq$$

$$V SHC_k(m, n) + \frac{k}{N} \sum_{j=-b_k}^{b_k} |F_k(m, n) - F_k(m, n + j)| =$$

$$V SHC_k(m, n) + \frac{k}{N} \sum_{j=-b_k}^{b_k} |\Delta F_k(j)| = APS_k(m, n).$$  

(15)

The scaling factor $k = N/(N - 1)$ is used to compensate for the fact that we do not use the central column (elements on a position $(m + i, n), i = -a_k, \ldots, +a_k$ in the calculation–hence, we have $N - 1$ elements $|\Delta F_k(j)|$ to normalize with. Usually, factor $k < 1$, but if we use values of parameter $k$ that are larger than one, we can boost the APS measurement on edges and thus even further minimize the “halo” artifact and/or enhancement effect of the algorithm. The resulting new energy measurement APS, which stands for APproximation of Sum of absolute differences, is actually an upper bound for the SAD. The equality sign is obtained when all elements $A = (F_k(m, n) - F_k(m, n + j))$ and $B = (F_k(m, n + j) - I(m + i, n + j))$ have the same sign. In the low-detail areas where $|\Delta F_k(j)| \approx 0$, we ensure the usage of the high detail gain and the same performance as the SAD energy measurement. In order to analyze the cases where the equality sign is obtained, we observe the sign of
the product \(AB\). In Fig. 4, one can see that in case of a single horizontal edge, where \(I(m, n+j) = I(m+i, n+j)\), \(AB \geq 0\) always holds: terms have equal signs. Likewise, using the combination of VSHC measurement and \(|\Delta F_k(j)|\), we exactly estimate the SAD on such an edge, which benefits the performance. For the other edge types, we over-estimate the exact energy signal terms in (12), since it is by default a separable energy measurement APS with a performance better than the SAD metric, and better than the LSD on textures, because it then resembles SAD. At the same time, we have reduced the computational cost needed to achieve that performance. Prior to giving experimental results, we first propose a new contrast gain function \(G_k\) that eliminates “halo” artifacts for large edges by design. Then, we improve the obtained algorithm thus far to achieve a perceptually uniform enhancement performance and prevent signal clipping.

IV. FURTHER IMPROVEMENTS ON “HALO” AND CLIPPING ARTIFACTS

A. New proposal for the non-linear gain functions \(G_k\)

To further improve the performance of our contrast enhancement algorithm, we can modify the original equations of the local contrast enhancement by making the contrast gain a non-linear function of the signal energy and some other signal metrics. Therefore, a basic gain formula \(G_k = C_k/LD_k\) can be substituted with \(G_k = f(LD_k, p_1, \ldots, p_n)\), where \(p_1, \ldots, p_n\) is a set of parameters that represent additional signal features/preferences, like a local/global noise measurement, local contrast, signal gradient measurement, local energy measurements derived from other kernels, characteristics of the human visual system and contrast sensitivity functions, characteristics of the display medium (such as CRT, LCD, paper, etc.), user preferences (contrast gain, application-specific parameters), etc. Function \(f\) is non-linear and can have any form that combines previously mentioned parameters in a beneficial manner. For example, we have already discussed a way to improve the noise performance of our algorithm [4] by limiting the original contrast gain by introducing the Maximum gain limit and a linear response of a gain function to the LD, \(G \propto 1/LD\), (see Fig. 1(b)). We also use the amplitude of the high-pass signals to discriminate noise from the relevant image features and prevent noise boosting. To this end, we propose to further improve the general shape of the non-linear contrast gain function \(G_k\) to \(G_k = C_k/LD_k^p\), to gain a better control of the “halo” artifacts for medium-to-large edges. Let us now elaborate further on this approach.

In [6], we have described various methods that are used to avoid “halo” artifacts. These methods [12], [14], [15], [16] attempt to split the original image into low- and band-pass components (texture, surface reflectances). As a second step, only the low-pass image (illumination) is compressed, so that texture details are given more dynamic range for display. For
example, Tumblin and Turk [12] use a form of anisotropic diffusion in an attempt to split edges into two types, according to their size. Then, large edges (boundaries between the High dynamic Range (HDR) regions) can be compressed independently of the small-to-medium edges (image texture), as long as their split is perfect. Since this cannot be fully achieved in practice, smooth compression across the different resolution bands is still necessary, which is achieved by adjusting the compression coefficients manually. This algorithm reduces the effects of “halos”, but it is computationally expensive. Similarly, a robust filter [14], or bilateral filter [15], is used to reduce “halo” artifacts. These edge-preserving filters include a second weight that depends on the intensity difference between the current pixel and its spatial neighbor, and therefore avoid “strong” filtering across the edges. A nearby pixel whose intensity is very different from the current pixel intensity is considered an outlier and its contribution to the local mean is reduced. Likewise, by preserving the sharpness at large transitions in the image (that correspond to HDR edges), the low-pass filtered version of the image at such places is very similar to the original image, yielding negligible high-pass $F_k$ and band-pass $B_k$ image signals. Hence, the creation of “halos” is minimized or completely reduced. However, the implementation costs of edge-preserving filters are much higher than the costs of Gaussian filters. In addition, they can also create double edges and false contours because they make concave/convex edges steeper than the original edges in the input. However, the original goal of using the edge-preserving filter for HDR image tone-mapping is attractive, but we will use it in a different form that does not introduce these artifacts and achieve a much lower computational complexity.

In essence, when we perform a low-pass filtering across the HDR edge, we would preferably like to have a much smaller high(band)-pass output than when filtering standard low- and medium-contrast edges. By doing so, we will completely remove the large overshoots and undershoots (“halos”) that are very typical artifact of many local-contrast rendering algorithms (see also Fig. 5). We can observe that the sum of the original band-pass signals already manifests large over/undershoot behavior (Fig. 5(b), signal labeled as “original”). Any kind of linear gain control like unsharp masking [17] would completely distort the image appearance on such an edge. The non-linear contrast gain often used in the literature, i.e. $G_k = C_k/LD_k$ compresses and limits this large over/undershoot, but still creates distortions (see Fig. 5, $G \propto 1/LD$). Therefore, we propose an alternative gain function in the form of $G_k \propto 1/LD_k^n$, which completely removes over and undershoots at large edges. Enhanced band-pass output of our multi-scale scheme $\sum_{k=1}^{K} C_k \cdot B_k/LD_k^2$ is much smaller for large edges than for the small, texture edges. This is visible in Fig. 5 ($G \propto 1/LD^2$), where enhanced output of the currently proposed gain function shows no “halo” artifacts. We have found this by noticing that the scaled envelope of the shape of the $|B_k|$ signal very well corresponds to the shape of the local energy signal $LD_k$, like in $LD_k \approx s \cdot |B_k|$, where scale $s$ varies between 2 and 4 depending on the band and steepness of the edge. This implies that the enhancement signal of a band $k$, for the original gain formula equals $C_k \cdot B_k/LD_k \approx C_k \cdot B_k/(s \cdot |B_k|) = \pm C_k/s = const.$ Thus, the resulting enhanced band-pass signal for a certain band is approximately constant, regardless of the input image. The contrast gain limits (Maximum gain, linear gain $G \propto LD$, Minimum gain) as presented in Fig. 1(b), will modify the contrast gain function, so that this amplitude equalization behavior is only approximately achieved. When this amplitude-equalized enhanced band-pass signal is added to the original image (we perform our pyramid reconstruction starting from the original image and not from the largest low-pass filter $F_K$), we add relatively more to small- and medium-size edges than to large-size (HDR) edges. This is valid because the signal of the same amplitude is added to both small and large edges, and its relative contribution to large edges is much smaller than the contribution to smaller edges. This operation is beneficial in many cases, however, we would preferably like not to add any signal to upper-medium and large edges, because these edges are already well visible and any enhancement would probably not be appreciated. This is why we propose the modified contrast gain formula $G_k \propto 1/LD_k^n$, which, for example, for $n = 2$ yields the following enhancement signal of a band $k$: \[
G_k \propto \frac{1}{LD_k^2} \sum_{k=1}^{K} C_k \cdot B_k
\]
outliers can also alter the real LC metric to some extent. To do this, a smooth function calculated by averaging a band-pass kernel (such as Gaussian) can be employed. The LD metric is also be calculated in a 2D-separable fashion. The LD metric is an approximation to consider both paradigms is the Local Contrast (LC) from the local average intensity level, the higher the local contrast depends on the local average intensity level: the higher contrast gain, we can eliminate “halos” at high-contrast edges, but it is still possible to create “halos” at medium-size edges when very high contrast gain is used. At the same time and secondly, following the Weber-Fechner law, the perception of local contrast depends on the local average intensity level: the higher the local average intensity level, the higher the local contrast needed to create an adequate visual perception. A good metric to consider both paradigms is the Local Contrast (LC) from [18], which can be defined as a relative difference between the maximum and minimum intensity values in a certain kernel $LC_k(m,n) = \max(I(m+i,n+j)) - \min(I(m+i,n+j))$]

$$C_k \cdot B_k/LD_k^2 \approx C_k \cdot B_k/(s \cdot |B_k|)^2 = \pm C_k/(s^2 \cdot B_k) = \text{const}/B_k.$$ Likewise, the larger the input signal amplitude of a band, the smaller the enhanced output: we can observe this effect well in Fig. 5(b). We can notice that for a large edge, the band-pass output is much smaller than for a small edge, which results in an almost unmodified total signal for the large edge, and a properly enhanced signal for the small edge, as in Fig. 5(a). In essence, by using a special form of the gain contrast function, we can mimic the same effect as techniques employing edge-preserving filters, but having much lower cost and complexity. Evidently, the proposed contrast gain formula is just a feasible example of the explained principle, other non-linear formulas with the similar property can also be used.

B. Perceptually uniform local contrast enhancement

As a main measurement input, we have used the local energy measurement LD to distinguish whether a certain region has a good or a poor visibility of local details, and likewise we have improved this visibility when necessary. However, the local signal energy is by nature an average measurement and cannot distinguish whether the measured local energy results from a lot of small-to-medium energy contributions (texture region), or perhaps from a single, large contrast edge (isolated edge). “Halo” artifacts mainly originate from isolated edges and we preferably need a metric that can distinguish the two previously mentioned cases. By using our modified contrast gain, we can eliminate “halos” at high-contrast edges, but it is still possible to create “halos” at medium-size edges when very high-contrast gains are used. At the same time and secondly, following the Weber-Fechner law, the perception of local contrast depends on the local average intensity level: the higher the local average intensity level, the higher the local contrast needed to create an adequate visual perception. A good metric to consider both paradigms is the Local Contrast (LC) from [18], which can be defined as a relative difference between the maximum and minimum intensity values in a certain kernel $LC_k(m,n) = \max(I(m+i,n+j)) - \min(I(m+i,n+j))$]

$$C_k \cdot B_k/LD_k^2 \approx C_k \cdot B_k/(s \cdot |B_k|)^2 = \pm C_k/(s^2 \cdot B_k) = \text{const}/B_k.$$ Likewise, the larger the input signal amplitude of a band, the smaller the enhanced output: we can observe this effect well in Fig. 5(b). We can notice that for a large edge, the band-pass output is much smaller than for a small edge, which results in an almost unmodified total signal for the large edge, and a properly enhanced signal for the small edge, as in Fig. 5(a). In essence, by using a special form of the gain contrast function, we can mimic the same effect as techniques employing edge-preserving filters, but having much lower cost and complexity. Evidently, the proposed contrast gain formula is just a feasible example of the explained principle, other non-linear formulas with the similar property can also be used.

B. Perceptually uniform local contrast enhancement

As a main measurement input, we have used the local energy measurement LD to distinguish whether a certain region has a good or a poor visibility of local details, and likewise we have improved this visibility when necessary. However, the local signal energy is by nature an average measurement and cannot distinguish whether the measured local energy results from a lot of small-to-medium energy contributions (texture region), or perhaps from a single, large contrast edge (isolated edge). “Halo” artifacts mainly originate from isolated edges and we preferably need a metric that can distinguish the two previously mentioned cases. By using our modified contrast gain, we can eliminate “halos” at high-contrast edges, but it is still possible to create “halos” at medium-size edges when very high-contrast gains are used. At the same time and secondly, following the Weber-Fechner law, the perception of local contrast depends on the local average intensity level: the higher the local average intensity level, the higher the local contrast needed to create an adequate visual perception. A good metric to consider both paradigms is the Local Contrast (LC) from [18], which can be defined as a relative difference between the maximum and minimum intensity values in a certain kernel $LC_k(m,n) = \max(I(m+i,n+j)) - \min(I(m+i,n+j))$]

$$C_k \cdot B_k/LD_k^2 \approx C_k \cdot B_k/(s \cdot |B_k|)^2 = \pm C_k/(s^2 \cdot B_k) = \text{const}/B_k.$$ Likewise, the larger the input signal amplitude of a band, the smaller the enhanced output: we can observe this effect well in Fig. 5(b). We can notice that for a large edge, the band-pass output is much smaller than for a small edge, which results in an almost unmodified total signal for the large edge, and a properly enhanced signal for the small edge, as in Fig. 5(a). In essence, by using a special form of the gain contrast function, we can mimic the same effect as techniques employing edge-preserving filters, but having much lower cost and complexity. Evidently, the proposed contrast gain formula is just a feasible example of the explained principle, other non-linear formulas with the similar property can also be used.

B. Perceptually uniform local contrast enhancement

As a main measurement input, we have used the local energy measurement LD to distinguish whether a certain region has a good or a poor visibility of local details, and likewise we have improved this visibility when necessary. However, the local signal energy is by nature an average measurement and cannot distinguish whether the measured local energy results from a lot of small-to-medium energy contributions (texture region), or perhaps from a single, large contrast edge (isolated edge). “Halo” artifacts mainly originate from isolated edges and we preferably need a metric that can distinguish the two previously mentioned cases. By using our modified contrast gain, we can eliminate “halos” at high-contrast edges, but it is still possible to create “halos” at medium-size edges when very high-contrast gains are used. At the same time and secondly, following the Weber-Fechner law, the perception of local contrast depends on the local average intensity level: the higher the local average intensity level, the higher the local contrast needed to create an adequate visual perception. A good metric to consider both paradigms is the Local Contrast (LC) from [18], which can be defined as a relative difference between the maximum and minimum intensity values in a certain kernel $LC_k(m,n) = \max(I(m+i,n+j)) - \min(I(m+i,n+j))$]

$$C_k \cdot B_k/LD_k^2 \approx C_k \cdot B_k/(s \cdot |B_k|)^2 = \pm C_k/(s^2 \cdot B_k) = \text{const}/B_k.$$ Likewise, the larger the input signal amplitude of a band, the smaller the enhanced output: we can observe this effect well in Fig. 5(b). We can notice that for a large edge, the band-pass output is much smaller than for a small edge, which results in an almost unmodified total signal for the large edge, and a properly enhanced signal for the small edge, as in Fig. 5(a). In essence, by using a special form of the gain contrast function, we can mimic the same effect as techniques employing edge-preserving filters, but having much lower cost and complexity. Evidently, the proposed contrast gain formula is just a feasible example of the explained principle, other non-linear formulas with the similar property can also be used.

B. Perceptually uniform local contrast enhancement

As a main measurement input, we have used the local energy measurement LD to distinguish whether a certain region has a good or a poor visibility of local details, and likewise we have improved this visibility when necessary. However, the local signal energy is by nature an average measurement and cannot distinguish whether the measured local energy results from a lot of small-to-medium energy contributions (texture region), or perhaps from a single, large contrast edge (isolated edge). “Halo” artifacts mainly originate from isolated edges and we preferably need a metric that can distinguish the two previously mentioned cases. By using our modified contrast gain, we can eliminate “halos” at high-contrast edges, but it is still possible to create “halos” at medium-size edges when very high-contrast gains are used. At the same time and secondly, following the Weber-Fechner law, the perception of local contrast depends on the local average intensity level: the higher the local average intensity level, the higher the local contrast needed to create an adequate visual perception. A good metric to consider both paradigms is the Local Contrast (LC) from [18], which can be defined as a relative difference between the maximum and minimum intensity values in a certain kernel $LC_k(m,n) = \max(I(m+i,n+j)) - \min(I(m+i,n+j))$]

$$C_k \cdot B_k/LD_k^2 \approx C_k \cdot B_k/(s \cdot |B_k|)^2 = \pm C_k/(s^2 \cdot B_k) = \text{const}/B_k.$$ Likewise, the larger the input signal amplitude of a band, the smaller the enhanced output: we can observe this effect well in Fig. 5(b). We can notice that for a large edge, the band-pass output is much smaller than for a small edge, which results in an almost unmodified total signal for the large edge, and a properly enhanced signal for the small edge, as in Fig. 5(a). In essence, by using a special form of the gain contrast function, we can mimic the same effect as techniques employing edge-preserving filters, but having much lower cost and complexity. Evidently, the proposed contrast gain formula is just a feasible example of the explained principle, other non-linear formulas with the similar property can also be used.
in lowering the contrast gain gradually from its nominal to its minimum level (sudden gain change would introduce perceptual artifacts). This concept is somewhat similar to the one taken in the design of the photographic tone-mapping operator [11] which mimics the dodge-and-burn technique known from photography practice. The author tries to find the largest surrounding area of a pixel that does not contain any large contrast edges. The low-pass value of this (largest) area will be used to stretch and therefore preserve the local contrast. The larger this area can be, the better the preservation of low-frequency local contrast. The low-pass value of the largest kernel can also be viewed as the local adaptation level of that area of the HDR image [11]. Photographic tone mapping is similar to our multi-band approach in that we add to the input signal the enhanced band-pass signal created by kernels of increasing sizes. Our contrast gain function \( G_{k_{\text{total}}} \) should stop that process at the moment when a sufficiently large contrast edge is crossed (defined by \( TH_k \)). For example (see Fig. 1(a)), if we detect a large-edge transition in the 4th kernel, the contrast gain of the 4th kernel will be almost zero and we add only enhanced bands 1, 2', and 3' to the output. Effectively, this is similar to using a high-pass band 3 only (largest band not containing large edges), however, we still have the flexibility to shape contributions of different frequencies of band 3 by means of the different contrast gains in its sub-bands 1, 2', and 3'. This is why we achieve favorable enhancement results [6].

C. Intensity-sensitive soft clipper

Starting from (1), we observe that the output image is created from the input image by adding amplified band-pass information. Adding a high-frequency signal can cause clipping of the total signal, especially in cases when the original input signal is close to its minimum or maximum value. The clipping leads to a very annoying waveform distortion, which should be prevented in any case. Since we control the gain function, it is best to avoid clipping at this position than instead at a later processing stage, after the addition of the enhancement signal. We distinguish two cases: clipping in black (at the minimum level \( Min \)) and clipping in white (at the maximum level \( Max \)). We have to ensure that the amplitude of the added enhancement band-pass signal \( G \cdot B \) does not surpass the space left until the clipping level, which we marked as Remaining distance \( RD \) in Fig. 7(a)). The total output signal is bounded by \( Min < I + G \cdot B < Max \). For a negative sign of the band-pass signal \( B \), at each position we can apply the gain limit formula: \( G < (I - Min)/|B| \) and for the positive sign of \( B \), the formula is \( G < (Max - I)/|B| \). To integrate this gain limitation better with the current gain formula and to increase its smoothness, we propose the following gain limits: \( G < s \cdot (I - Min)/LD \) for the clipping in black and \( G < s \cdot (Max - I)/LD \) for the clipping in white. We simplified this limitation by noticing that the shape of the local energy signal \( LD \) nicely emulates the shape of the \( |B| \) for \( s \approx 2 \), i.e. \( LD \approx s \cdot |B| \). Our basic gain formula for band-pass undershoots \( B < 0 \) becomes now \( G_k = min(C_k \cdot s \cdot LD_k \cdot (I - Min))/LD_k^2 \) and for the band-pass overshoots \( B \geq 0 \): \( G_k = min(C_k \cdot s \cdot LD_k \cdot (Max - I))/LD_k^2 \).

In our previous work [6], we showed that it is beneficial to perform contrast enhancement in parallel with the Global Tone-Mapping Function (GTMF), and to add the enhanced high (band)-pass signal both before and after this function. Whatever this function is, if we add this enhancement signal after the GTFM, the Remaining Distance \( RD \) from the signal clipping borders is altered (see new/old \( RD \), Fig. 7(b)). For example, with our employed GTFM function, it is decreased for higher values of the input signal \( I \) and high-pass overshoots and increased for the lower values of input \( I \) and high-pass undershoots. Therefore, the gain modifications have to fit with the characteristics of the GTFM function. This can be achieved in a relatively simple fashion: instead of taking the input signal \( I \) in the gain limit formula, we take a transformed input signal GTMF(\( I \)). To perform this operation less costly, a piecewise linear approximation of the GTFM can be performed. Our gain formula for the band-pass undershoots \( B < 0 \) becomes now \( G_k = min(C_k \cdot s \cdot LD_k \cdot (\text{GTMF}(I) - Min))/LD_k^2 \) and for the band-pass overshoots \( B \geq 0 \) it is \( G_k = min(C_k \cdot s \cdot LD_k \cdot (Max - \text{GTMF}(I)))/LD_k^2 \).

V. EXPERIMENTAL RESULTS

We have tested our algorithm for various input signals, both in simulations as well as in real-time applications (via FPGA), using sequences that contain both low-, medium-, and high dynamic-range video signals and we have compared the performance of various energy metrics. We have performed a subjective comparison, since it is the most relevant test for the local contrast evaluation and there are no well-established measurements for its evaluation. Results are very good in most of the scenarios and the new metric outperforms conventional techniques like LSD and SAD, both in terms of performance.
and complexity. We presented a comparison of our method to other state-of-the-art methods in [6]. In this paper we will focus on the contributions we have currently presented. An example is given in Fig. 8 where we show a standard test image "Japanese Lady" and several enhanced output images using various metrics like LSD, SAD, VSHC, APS and APS & local contrast (LC) check. Our algorithm improves a poor local contrast of this image regardless of the applied metric. However, it is interesting to observe the different metric’s performances in various areas of this image. Let us consider the performance in the texture areas (persons hair, face, sweater, rose) and at large edges (boundaries between a person and a flat background). The VSHC metric provides the highest enhancement level in the texture areas, boosting the local contrast significantly, but it also gives the highest amount of disturbing “halo” artifacts at the isolated edges (see for instance the hair area and contours of the person’s face). The application of the LSD and the SAD metric results in a mutually similar performance, however, the SAD metric gives a bit higher output contrast in the texture areas and slightly more “halo” artifacts in the isolated-edge areas. The APS metric still enhances the texture areas in a similar fashion like the SAD metric and its performance approaches the LSD metric at the large edges. We can observe that the proposed solution (Fig. 8(f), the APS metric & Local Contrast (LC) check) maintains a high level of detail in the texture regions, but controls well the disturbing “halo” artifacts, providing for the best overall performance.

In Fig. 9, we present a part of a video line from Fig. 8 (abscissa represents x axis of a video line and in ordinate we present the intensity of the original and enhanced video signals). We can observe very large over/undershoots that the VSHC metric creates (“halo” artifacts). We can also notice that
the proposed solution based on the APS metric & LC check prevents the appearance of disturbing “halo” artifacts and it maintains a high level of detail in the texture regions where enhancement is desired (remainder of the video line, where small to medium edges are present). Our proposal provides for the smallest overshoots/undershoots at large contrast edges, and the size of these over/undershoots, controlled by parameters $TH_{a}^{k}$, $TH_{b}^{k}$ and $minG_{k}$, can be chosen depending on the application, user preferences, image resolution and viewing distance, among others. Another example result is presented in Fig. 10. This image belongs to a class of a so-called High-Dynamic Range (HDR) images [11], since it has a very high global contrast ratio (87,962:1). For a correct rendering, it requires a global tone mapping [1], [4] in addition to the local contrast enhancement [5]-[8], [11]-[17]. In this particular case, the required level of global tone mapping is extremely high,
to allow visibility of details in the dark parts of the image. This operation significantly reduces the local contrast in bright image parts. Local contrast enhancement rectifies this loss and we show several output images enhanced using various local energy metrics. Let us observe the performance in the texture areas (doll’s clothing, white toy bear, . . . ) and at large edges (boundaries between a doll and a flat background, separation line between a wall in the background and a floor (line is passing trough the middle of the image)). We present now only results obtained using the LSD and APS metrics, since we want to prove that obtained results of APS metric are better than results using the LSD metric. Results and conclusions when using other metrics are similar to Fig. 8. The APS metric gives better enhancement performance in the texture areas than the LSD metric, while its performance is very similar to the one of the LSD metric at the large edges. The proposed solution (Fig. 10(c), the APS metric & Local Contrast (LC) check) maintains a high level of detail in the texture regions, but controls well the disturbing “halo” artifacts, providing for the best overall performance. For example, one can observe that “halo” artifact on a separation line between a wall in the background and a floor is the smallest compared to all the other results. At the same time, performance in the texture areas did not reduce: it remained visually similar to the performance using the APS metric only.

We finally present results based on the image “buildings” (Fig. 11). We can notice that parts of the image are in the shadow (lower part of the building, bridge), while other parts of the image are well visible. We can also observe that the right building is partially out of focus and it gives a less sharp impression, as the photographer’s focus was on the left building. In many texture areas all LD metrics give considerable improvements and all parts of the image seem in focus now (increased depth of field). This is both valid for shadow parts of the image as well as for the well illuminated parts. The difference in performance with respect to the texture areas and the large edges is also appearing here, despite the completely different nature of the image. Our proposal (Fig. 11(d)) presents very good, balanced results without “halo” artifacts, proving that it works well and provides a natural impression for various image types. We also have to take care of the noise performance, since the original image has a rather low SNR. Without additional measures [5], all the results would suffer from the noise increase, which is especially visible in the flat areas like the sky.

VI. CONCLUSIONS

In this paper, we have discussed the complexity and performance of various metrics for local energy measurements, which are used for the local contrast enhancement in a multi-window, real-time high-frequency enhancement scheme. In the enhancement, the gain is a non-linear function of the detail energy and local contrast. We have presented two major contributions: (1) an efficient way of computing an LD metric having beneficial properties, (2) a novel local contrast gain which combines an LC metric with LD, eliminating “halo” effects at edges and giving perceptually uniform enhancement.

With respect to the first contribution, we have shown that the LSD metric contrary to the SAD metric can be calculated in a 2D-separable fashion to reduce the complexity. Furthermore, it was found that the performance of the SAD metric is better than the contrast enhancement performance of the LSD metric for the texture regions, but worse for the large-edge regions, resulting in somewhat larger “halo” artifacts. Additionally, the
SAD can be modified into a VSHC metric, which provides a much higher contrast enhancement than all other metrics, but also gives the most disturbing “halo” artifacts. However, the complexity of the VSHC metric is the lowest, which makes it attractive for real-time enhancement applications. Our new proposal, called the APproximation of the Sum of absolute differences (APS), builds on the low complexity of the VSHC metric (also 2D-separable). We have shown that the APS measurement is optimal regarding the performance and computational costs and that it gives computational and performance benefits compared to the SAD metric, particularly for large kernels. The key to the APS algorithm is that we pre-calculate vertical differences as in a modified SAD calculation and add a correction term to this.

In the second contribution, we have improved the performance of the algorithm to avoid “halo” and achieve the perceptually uniform enhancement by introducing a new gain control definition and using a Local Contrast measurement (LC). In addition, we prevent clipping artifacts at the signal boundaries. The advantage of the proposed technique is that contrast gains can be set much higher than usual and our non-linear contrast gain, together with the LC measurement will reduce contrast gains only at image locations where this is really needed. Subjective evaluation of several images show that the proposed solution maintains a high level of detail in the texture regions, but controls well the disturbing “halo” artifacts, providing for the best overall performance compared to all other discussed metrics.

Finally, we have found that the enhancement parameters of all the local contrast enhancement algorithms must be carefully chosen, depending on the source material and the intended application. For this reason, it might be good to pre-classify the quality and type of images and perform their enhancement application. For this reason, it might be good to pre-classify chosen, depending on the source material and the intended application. For this reason, it might be good to pre-classify the quality and type of images and perform their enhancement application. For this reason, it might be good to pre-classify the quality and type of images and perform their enhancement application.

REFERENCES


Sascha D. Cvetkovic received his University degree in electrical engineering in 2000 from the Fac. of Electrical Engineering, Belgrade, Serbia. Since 2001 he is a senior research engineer at Bosch Security Systems. He develops various algorithms in the field of image and video signal processing, digital loop control and video content analysis. Since 2003, he is working part-time towards a PhD degree in the EE Dept. at the Technical Univ. of Eindhoven, the Netherlands. He is the recipient of the Nikola Tesla award, Chester W. Sell award for the second best paper in IEEE Trans. on CE for the year 2008 and as a part of the team he won 2005 NSCA Innovations in Technology Award for contributing to the design of the LTC 0495 DinionXF Day/Night Camera.

Johan L. Schirris completed his study in electrical engineering at the Hogere Technische School Breda, The Netherlands, in 1974. He started working at Philips TV Lab in 1974 in Eindhoven. He contributed to the development of digitally controlled TV and Teletext. He was involved in the development of the EISA awarded 100 Hz TV with "Natural Motion. Since 1997 he is employed at Philips VCM, after 2001 part of Bosch Security Systems. His work as an IC architect in the digital imaging area has resulted in a number of patents and he has contributed to the design of DSPs that are successfully applied in a range of cameras.

Peter H.N. de With Fellow of the IEEE, obtained the MSc and PhD degree from University of Technology Eindhoven and Delft, Netherlands. He has worked on video coding for recording at Philips Research Labs Eindhoven from 1984 till 1993. Between 1994 and 1997 he headed programmable TV architecture research at the same Lab. From 1997-2000 he was full professor at the University of Mannheim, Germany and between 2000-2007 distinguished/principal consultant at LogicaCMG. Since 2008, he is VP Video Technology at CycloMedia. Since 2000, he is also professor at the University of Technology Eindhoven. He has written and co-authored over 200 papers and received paper awards at ICCE, SPIE and ISCE conferences. He serves as committee member in ICIP, VCIP, ICCE conferences and chaired multiple working groups. He is regularly technical advisor to Philips, Logica, CycloMedia, and various other companies.