

LOCALLY-ADAPTIVE IMAGE CONTRAST ENHANCEMENT WITHOUT NOISE AND RINGING ARTIFACTS

Sascha D. Cvetkovic^{**}, Johan Schirris^{*}, Peter H.N. de With^{*}

^{*} Bosch Security Systems/ ^{**} University of Technology Eindhoven/LogicaCMG
{sacha.cvetkovic, johan.schirris}@nl.bosch.com and P.H.N.de.With@tue.nl

ABSTRACT

For real-time imaging in surveillance applications, visibility of details is of primary importance to ensure customer confidence. Additional constraints are the absence of human interaction and low computational complexity. Usually, image quality is improved by enhancing contrast and sharpness. Many complex scenes require local contrast improvements that should bring details to the best possible visibility. However, local enhancement methods mainly suffer from ringing artifacts and noise over-enhancement. In this paper, we present a new multi-window real-time high-frequency enhancement scheme, in which gain is a non-linear function of the detail energy. Our algorithm controls perceived sharpness, ringing artifacts (contrast) and noise, resulting in a good balance between visibility of details and non-disturbance of artifacts. Its advantage is that gains can be set now much higher than usual and the algorithm will reduce them only at places where it is really needed.

Index Terms— Local contrast enhancement, ringing, noise

1. INTRODUCTION AND PROBLEM STATEMENT

For real-time imaging in surveillance applications, visibility of details is obtained amongst others via video signal enhancement that should accommodate for the widely varying light conditions and versatility of scenes with non-ideal luminance distribution. The primary task of these algorithms is to improve perception, sharpness and contrast and bring visibility of details in all parts of the scene to the highest possible level. In real-time surveillance applications, almost no human interaction occurs for the enhancement (autonomous processing), while the computation complexity must remain low, thereby posing a significant problem.

A possible solution is to apply advanced contrast control. In earlier work [1], we gave an overview of state-of-the-art dynamic range expansion and video signal enhancement techniques employed in security cameras and proposed a new method to improve both aspects. Almost all current algorithms used in surveillance systems including our previous solution belong to the group of the global methods, where one transformation is applied to all pixels of the input image. 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 overall contrast is good but

local contrast is low. In these cases, locally-adaptive contrast enhancement will provide significant advantages. Kuroda [2] proposed an algorithm for real-time adaptive image enhancement, but its application and capabilities are limited. Dah-Chung Chang *et al.* [3] observed that image enhancement with contrast gain which is constant or inversely proportional to the Local Standard Deviation (LSD) produces either ringing artifacts or noise over-enhancement, due to the use of too large contrast gains in regions with high and low spatial activities. They developed a new method based on extending Hunt's image model in which gain is a non-linear function of LSD. Although promising, this approach has very high complexity making it inappropriate for real-time implementation. In addition, no general rule is given to determine the optimal window size, which anyhow varies from point to point. This problem led to multiscale methods. Boccignone [4] proposed to measure contrast at multiple resolutions generated through anisotropic diffusion. Once local contrast has been estimated across an optimal range of scales, its value is used to enhance the initial image. Again, computation time and complexity are obstacles for employing it in a real-time system. Narendra and Fitch [5] presented a real-time high-frequency enhancement scheme, in which they amplified medium and higher frequencies in the image by gain that is inversely proportional to the LSD. Schutte [6] introduced a multi-window extension of this technique and showed how the window sizes should be chosen. Prevention of excessive noise amplification is included, but the method is not sufficient in case a higher noise level is present or higher amplification factors are used. At the same time, a mechanism for the ringing-artifact suppression is not given.

The sequel of this paper is organized as follows. Section 2 describes system aspects and our improvements. Section 3 discusses noise improvement and Section 4 provides results.

2. AN IMPROVED METHOD FOR IMAGE/VIDEO CONTRAST ENHANCEMENT

The enhanced output image resulting from the multi-scale approach in [6], can be described as:

$$o = i + \sum_{j=1}^k M \frac{C_j}{LSD_j} \cdot (i - m_j), \quad (1)$$

where symbols used are: o , i – output and input image, respectively; k – number of scales used; M – global mean of

the image; m_j -local mean of the kernel, C_j - gain factors that control the enhancement level per kernel and LSD_j - local standard deviation per kernel. The global mean of the image is used here as a multiplication factor to compensate for the well-known effect of increasing JNDs for the higher luminance levels. It will automatically increase the local gain in the high-luminance images and enable details to surpass the JND threshold and become visible. Use of the global luminance mean can help if overall scene illumination is high, but it might be better to use the local luminance mean m_j instead and make this scheme content sensitive. This approach will allow higher gains only locally, where they are needed. Unfortunately, a problem of these approaches can occur in bright areas: enhanced output high-frequencies are added to the already high luminance signal, which can easily lead to clipping of the total signal, and therefore to the loss of details. Very advanced content-dependent parameter tuning is needed to prevent this effect with still questionable results. Anyhow, we will use contrast gains delivering an output high-frequency signal at least a factor of 10 larger than the corresponding JNDs and there will be no need for additional luminance-dependent gain.

Let us further explain some interesting properties of this frequency enhancement scheme. Let $H_{j=i-m_j}$ represent an input high-pass signal, F is a square Box filter and \otimes is 2D convolution: $m_j=i\otimes F_j$. If the LSD is a measure of the energy calculated as the sum of the abs. differences between the input image and the low-pass image: $LSD=|i-m_j|\otimes F$, then it can be proven that some interesting properties hold:

- Ho (enhanced high-frequency output) is invariant to changes in the gain and offset of the input image, so:

$$Ho(a \cdot i + b) = Ho(i) \cdot \quad (2)$$

- Ho is proportional to the parameter “C”, that is:

$$Ho(s \cdot C) = s \cdot Ho(C) \cdot \quad (3)$$

The first property insures that variations in lighting of the input image are in fact compensated. The second property indicates that “C” (contrast gain) is the parameter that can be used to determine the required amount of contrast in the amplified high-pass image, regardless of the amount of contrast in the input image. In practice, the gain limits will decrease the validity of the above properties. It should be noted that the above properties hold only for the amplified high-pass image Ho and not for the output image o , because the latter also includes the low-pass image.

2.1 Frequency bands instead of high-frequency regions and change of filter types

Intrinsically, the algorithm is trying to equalize amplitudes of all output frequencies by assigning higher gains to frequencies with smaller energy (amplitude) and vice versa. This is valid for each scale employed, although the range of spatial frequencies that can be enhanced by different

windows varies. This enhancement scheme has a pyramidal form in the frequency domain in which bigger windows cover smaller ones and also add certain new, lower frequencies. Basically, overlap of the frequency ranges occurs: instead of being responsible only for a certain scale (object size), a bigger window enhances all smaller scales as well (Fig. 1, top). This is one of the reasons why controllability and performance of any common multi-scale contrast enhancement scheme is not satisfactory. We propose windows in which frequency bands are separately controlled and enhanced (Fig. 1, bottom). This is achieved by subtracting windows of two subsequent scales: the difference between windows represents a supplementary frequency band which is added to a certain pixel neighborhood and which is additionally enhanced by the use of the bigger kernel:

$$o = i + \sum_{j=1}^k \frac{C_j}{LSD_j} \cdot (m_j - m_{j+1}) \cdot \quad (4)$$

We already mentioned in Eqs. (2) and (3) that we can easily control local output contrast. These properties are valid when previously employed Box kernel windows are used. Unfortunately, these filters yield rather poor frequency characteristics. They have a very poor suppression of the side-lobes in the stop-band range of the filter. The resulting effects can be well observed on the prominent edges (ringing artifacts). Instead, we employ raised-cosine types of windows, with much better frequency characteristics and visual results, particularly on edges. Consequently, the properties (2) and (3) only hold partly. We designed “soft clippers” to cope with this problem of controlling the output contrast: they compress amplitudes of high-frequency pixel outputs if they exceed a certain threshold that is directly coupled to the level of image enhancement, pixel luminance and sign of the overshoot (undershoot). It should be noted that a JND- guided gain approach as explained in Ji *et al.* [7] suffices for the image enhancement in terms of perception of details. If local details are above the visibility threshold level, perception of details of that region is acceptable. Often we want even higher edge overshoots, since this gives the impression of an extremely high-contrast image. At the same time there are certain edge types where such large signal overshoots can destroy the perceptual impression of the image. There are also certain cases where applying even small contrast gain would cause problems. Bigger windows are most critical regarding this effect, which is particularly

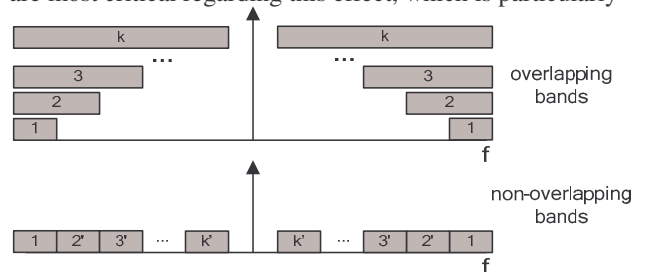


Figure 1 Multiscale enhancement scheme: $band_i = band_i - band_{i-1}$.

visible on the transitions between two very distinct objects and becomes very annoying if one of the objects is very bright. This is why for the two aforementioned cases, additional ringing-suppression mechanisms were developed.

2.2 Additional ringing-suppression mechanisms

First, for the large overshoots, we introduce a formula modification for the pixel-based gain control in a certain band which concentrates on the gain behavior for the often problematic large values of the LSD . We propose an offset for the gain calculation leading to $G = C / LSD + a$. If $a < 0$, less gain is given to pixels with larger LSD values

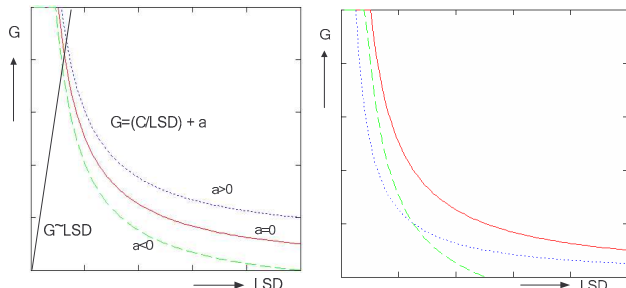


Figure 2 (left): Gain/ LSD forming transfer functions.

Figure 3 (right): different gain/ LSD curves with variables C and a ; Colors: $C_{red} = C_{green} = 2 * C_{blue}$ · $a_{red} = a_{blue} = 0$; $a_{green} < 0$; (gain is limited between certain minimum and maximum value).

and vice versa (see hyperbolic transfer functions in Fig. 2). The choice of the parameter a is important: it directly controls perceived contrast and can prevent annoying perceptual artifacts. As observed in Fig. 3 (hyphenated transfer function), it is possible to keep the contrast gain larger without introducing an excessive gain for the higher LSD values with a negative a . To find the most optimal value of the parameter a , we need to have the knowledge of the image type which is often not possible. To mitigate this problem but also give this system higher flexibility, we can additionally reshape the gain function by calculating edge amplitudes over smaller window support and for instance decrease gain for large edge amplitudes (Fig. 4, *edge sum*). Like this, we will also couple gain reduction to the edge size and not only to the LSD calculated over a large window.

Second, it is observed that the most critical cases of ringing at which even small overshoots create problems are at the very sharp transitions between objects (isolated edges), particularly when objects have a low spatial activity. A detail-anisotropy detector (Fig. 4) is used to detect these cases and reduce the gain symmetrically around such edges. In order to detect an isolated edge, it checks if energy around edges is distributed in a non-symmetrical fashion (*edge dif*), for instance along one of the axes, and reduces the gain along that axis. An extra counting circuit of neighboring edges checks for occurrence of isolated edges.

An advantage of the overall method is that C_j gains can be set much higher and the algorithm will reduce them only at the places where it is really needed. The edge signal can be

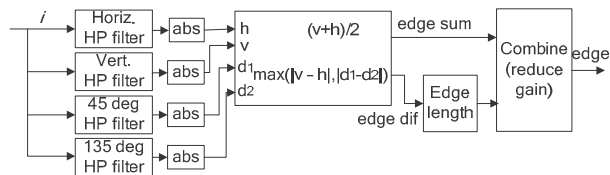


Figure 4: Detail anisotropy detector.

used in all windows to lift the local LSD values at such places and minimize ringing artifacts. It became evident that for complete handling of the ringing artifacts, all above-mentioned methods are necessary simultaneously.

3. NOISE BEHAVIOR IMPROVEMENT

When talking about noise reduction, adaptive dynamic noise reduction can be used to temporally/spatially filter the output image, depending on the locally-detected amount of motion. On the other hand, we are also interested in the visibility of noise, which can be rather high due to the high-frequency enhancement scheme previously described. Therefore, in the sequel, we aim at actually reducing noise visibility by controlling the noise-causing factors.

3.1 Global noise reduction

To improve noise behavior, the gain is limited to a certain maximum value and a first part of the gain/ LSD transfer function is made linear to suppress noise for the lower LSD values (see Fig. 2). If we temporally filter out the input image used for the LSD measurement, we ensure that low-level details receiving the highest gain are originating from the real structures in the image and not from the noise. Anyhow, it is clear that the presence of noise only increases the LSD , thereby automatically reducing the gain: $G \sim 1/LSD$). It is therefore usually not necessary to perform this.

We propose two additional noise-reduction mechanisms which concentrate on the noise reduction in “flat” (low-detailed) areas and overall noise reduction.

3.2 Local noise reduction

Noise is the most visible and annoying in the flat, low-detailed areas. An LSD measurement can identify sufficient amount of details in the flat regions, if a part of the window is lying over object edges or some detailed area. In those cases, the pixel-based LSD value is large enough not to be considered noise but still small, therefore receiving very high gain and amplifying overall present noise considerably. We propose the following solution: we check the local absolute values of the 3-tap high-pass signal (and not LSD only, which is gathered over a much bigger area), and we reduce the gain if they are below noise level $THR2$ (see Fig. 6 left, extra factor applied to the gain). This approach looks like a “coring” action, although we do not force all signal values to zero, since they can still be details. Instead, we assign them a gain that is more appropriate to the type of the

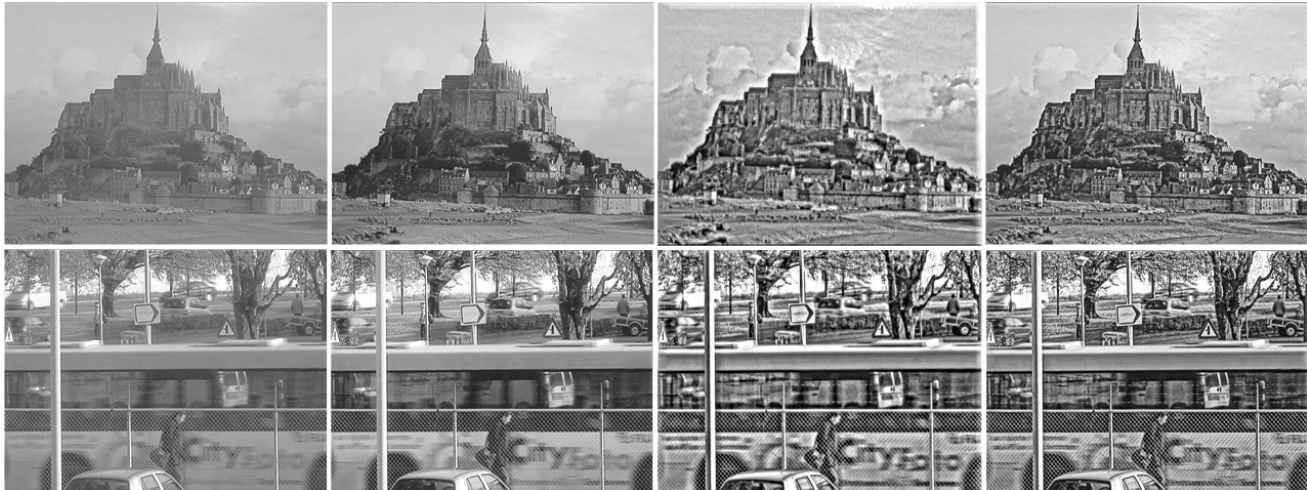


Figure 5: Results: from left to right: standard camera image, method from the [1], method from the [6], our method.

image or the highest acceptable noise level. Likewise, we potentially reveal some very fine, hidden details in the image by bringing them above the visibility threshold JND, while preserving an acceptable level of noise.

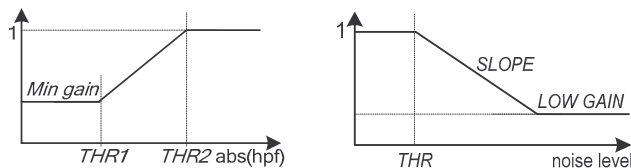


Figure 6(left):Coring-like action;(right):Noise based gain reduction

In order to reduce overall noise visibility, including the visibility in detailed areas, another approach is needed. We construct a noise-measurement detector by comparing the input pixel value to its (noise-poor) temporally recursive-filtered version. For this purpose, we will reduce the local gain for pixels with a significant amount of noise (noise level larger than THR, see extra gain factor in Fig. 6 right).

All gain reductions are performed in a smooth way, thereby circumventing inter-frame (field) differences that could lead to the temporal perceptual artifacts and noise break-through. All these techniques significantly improve the noise performance.

4. EXPERIMENTS AND RESULTS

We have tested our scheme and techniques on various inputs, both in simulations as well as in a real-time surveillance application and compared it to other approaches. We have performed a subjective comparison, since it is the most relevant test for a surveillance application. Sequences which contain both low and high dynamic-range video signals were chosen, where clipping of the signal was avoided if possible. Results are very good in most of the scenarios and the system outperforms conventional techniques (an example performance is shown in Fig. 5). As

already stated, *global* image enhancement (as in [1]) is beneficial but *local* image enhancement is revealing much more details. In the “Castle” image, one can observe disturbing ringing artifacts and noise over-enhancement resulting from the method in [6], as well as the much better result that we have obtained. Also, although looking sharper at first sight, “Traffic” scene of [6] exhibits loss of details in white and numerous signal disturbances giving it unnatural look. It can be concluded that careful filter selection and use of the band-pass image enhancement in combination with a further artifact reduction is necessary to exploit the perceptual benefits and reveal safely all the image details (as in our balanced approach). Combining global and local image enhancement techniques and further improvement of the presented approach is subject of our future interest.

5. REFERENCES

- [1] Sascha D. Cvetkovic and Peter H. N. de With, "Image enhancement circuit using non-linear processing curve and constrained histogram range equalization," Proc. of SPIE-IS&T Electronic Imaging, Vol. 5308, pp. 1106-1116 (2004).
- [2] I.Kuroda, "Algorithm and architecture for real time adaptive image enhancement," SiPS 2000. Pages:171 – 180 Oct. 2000.
- [3] D.C.- Chang and W.-R. Wu, "Image contrast enhancement based on a histogram transformation of local standard deviation," IEEE Trans. MI, vol. 17, no. 4, pp. 518-531, Aug. 1998.
- [4] G. Boccignone and M. Ferraro, "Multiscale contrast enhancement," *Electron. Lett.*, vol. 37, no. 12, pp. 751–752, 2001.
- [5] P.M. Narendra and R.C. Fitch, "Real-time adaptive contrast enhancement," IEEE Trans. PAMI, Vol.3, no. 6, pp.655-661, 1981.
- [6] K. Schutte, "Multi-Scale Adaptive Gain Control of IR Images," Infrared Technology and Applications XXIII, Proceedings of SPIE Vol. 3061, pp.906-914 (1997).
- [7] T.-L. Ji, M.K. Sundareshan and H. Roehrig, "Adaptive image contrast enhancement based on human visual properties", IEEE Trans. MI, vol. 13, no. 4, December 1994.