Multi-scale contrast enhancement with applications to image fusion

Alexander Toet

This paper was published in

Optical Engineering, 31(5), pp. 1026-1031

and is made available as an electronic reprint with permission of SPIE.

Multiscale contrast enhancement with applications to image fusion

Alexander Toet, MEMBER SPIE
Institute for Perception TNO
Kampweg 5
Soesterberg NL-3769-DE, The Netherlands

Abstract. A method to merge images from different sensing modalities for visual display was introduced by Toet, van Ruyven, and Valeton in 1989, which produces a fused image by nonlinear recombination of the ratio of low-pass (RoLP) pyramidal decompositions of the original images. The appearance of merged images that are produced by this scheme is highly dependent on the contrast and mean gray level of the input images. That nonlinear multiplication of the successive layers of a ratio of low-pass pyramid results in a contrast-enhanced image representation that is highly invariant for changes in the global gray-level characteristics of the original image is shown. Application of this nonlinear multiplication procedure in the image fusion process results in composite images that appear highly independent of changes in lighting and gray-level gradients in the input images. The method is tested by merging different degraded versions of parallel registered thermal (FLIR) and visual (CCD) images.

Subject terms: contrast enhancement, low-pass pyramid; multiscale image representation; multisensor fusion.

Optical Engineering 31(5), 1026-1031 (May 1992).

1 Introduction

The workload of a human operator severely increases with the number of imaging systems that need simultaneous monitoring. Moreover, a human observer cannot reliably integrate visual information by viewing multiple images separately and consecutively. The integration of information across multiple human operators is nearly impossible. An imaging system that fuses signals from multiple imaging sensors into a single image is therefore of great practical value.

An image fusion method intended for human observation was recently introduced.[1] In this scheme, the input images are first decomposed into sets of light and dark blobs on different levels of resolution. This is done by computing a ratio of low-pass (RoLP) pyramid[2] for each of the input images. A RoLP pyramid for the fused image is then obtained by selecting nodes with maximum absolute gray-level contrast from the sets of corresponding nodes in the RoLP pyramids of the individual images. The fused image is reconstructed from the set of pyramid nodes or pattern primitives thus obtained. As a result, perceptually important details (i.e., details with a relatively high local gray-level contrast) of both images are preserved in the composite image. A serious shortcoming of this image fusion method is its inherent sensitivity to variations in mean image intensity and global gray-level gradients.

Paper 2866 received Feb. 1, 1990; revised manuscript received Oct. 17, 1991; accepted for publication Oct. 19, 1991.

This paper presents a scheme to enhance image contrast by nonlinear multiplication of successive layers of the RoLP image decomposition. In this multiscale contrast enhancement process, image contrast at finer scales is weighted by local contrast extrema at coarser scales. As a result, image contrast is enhanced at all levels of resolution and fine-scale texture becomes more visible. The recombination of a contrast-enhanced RoLP pyramid with a constant-valued top layer results in a reconstructed image that is, to a high degree, independent of global variations in the mean gray level of the original image. Hence, the result of the image fusion scheme can also become highly insensitive to global variations in the intensity of the input images when this multiscale contrast enhancement procedure is applied to each of the input images prior to the actual fusion process.

The organization of this paper is as follows: Section 2 introduces the nonlinear multiscale contrast enhancement scheme and shows some experimental results. Section 3 presents some results of the application of this contrast enhancement scheme in the image fusion process. Finally, some concluding remarks are given in Sec. 4.

2 Toward an Invariant Image Representation

2.1 *Multiscale Image Decomposition*

The RoLP (ratio of low-pass) pyramid was recently introduced by Toet.[2] The construction of the RoLP pyramid is very similar to that of the popular difference of low-pass[3] (DoLP) or difference of Gaussians[4] (DoG) pyramid structures. First a Gaussian or low-pass pyramid is generated for the input image. This is a family of low-pass-filtered copies of the input image, each with a band limit one octave lower than its predecessor.

Let array G_o contain the original image. This array becomes the bottom, or zero, level of the pyramid. Each node of pyramid level i ($1 < i < N$, where N is the index of the top level of the pyramid) is obtained as a (Gaussian) weighted average of the nodes at level $i-1$ that are positioned within a 5 x 5 window centered on that node. Because of the reduction in the spatial frequency content, each image in the sequence can be represented by an array that has half the dimensions of its predecessor.

The process that generates each image in the sequence from its predecessor is called a REDUCE operation because both the sample density and the resolution are decreased. Thus, for $1<=l<=N$, we have G_l=REDUCE (G_{l-1}), meaning

$$G_l(i,j)=\sum_{m,n=-2}^{2} w(m,n)G_{l-1}(2i+m,\ 2j+n) \ .$$

The weighting function $w(m,n)$ is separable: $w(m,n) = W(m)w'(n)$, with $w'(0) = a$, $w'(1) = w'(-1) = 0.5$, $w'(2) = w'(-2) = a/2$. A typical value of a is 0.4.

Because the levels in the low-pass pyramid differ in sample density, it is necessary to interpolate new values between the given values of a low-frequency image before it can divide a high-frequency image. Interpolation can be achieved simply by defining an EXPAND operation as the inverse of the REDUCE operation. Let $G_{l,k}$ be the image obtained by applying EXPAND to G_l k times. Then,

$$G_{l,0} = G_l ,$$

and

$$G_{l,k} = \text{EXPAND}\ (G_{l,k-1}) \ ,$$

meaning

$$G_{l,k}(i,j) = 4 \sum_{m,n=-2}^{2} w(m,n)G_{l,k-1}\left(\frac{i+m}{2}, \frac{j+n}{2}\right) \ ,$$

where only integer coordinates contribute to the sum.

The RoLP pyramid is then defined by

$$R_i = \frac{G_i}{\text{EXPAND}(G_{i+1})} \text{ for } 0 \le i \le N-1 \ ,$$

and

$$R_N = G_N \ .$$

Thus every level R; is the ratio of two successive levels in the low-pass pyramid.

2.2 *Multiscale Contrast Enhancement*

The RoLP pyramid is a complete representation of the original image. Therefore, Go can be recovered exactly by reversing the steps used in the construction of the pyramid:

$$G_N = R_N \ ,$$

and

$$G_i = R_i \cdot \text{EXPAND}(G_{i+1}) \text{ for } 0 \le i \le N-1 \ .$$

Different recombination rules can be used to display different aspects of the information encoded in this representations For instance, partial image representations, corresponding to locally filtered or structurally simplified versions of the original image, are obtained when only a subset of the nodes in the RoLP pyramid are used in the reconstruction process.

To enhance the contrast of the original family G_i, we define the following recombination rule:

$$\text{CE}G_N = L_N \ ,$$

and

$$\text{CE}G_i = R_i \cdot \text{CE-EXPAND}(\text{CE}G_{i+1}) \text{ for } 0 \le i \le N-1 \ ,$$

where L_N can be either G_N or a constant valued image, CEG_i denotes the contrast-enhanced version of G_i, and CE-EXPAND is defined by

$$\begin{aligned}&\text{CE-EXPAND}(\text{CE}G_{i+1})\ (i,j)\\ &= \min[\text{CE}G_{i+1}(2i+m,2j+n):m,n=-2\ldots2] \text{ if } R_i(i,j)<1\\ &= \max[\text{CE}G_{i+1}(2i+m,2j+n):m,n=-2\ldots2] \text{ if } R_i(i,j)>1\\ &= \text{EXPAND}(\text{CE}G_{i+1})(i,j) \text{ if } R_i(i,j)=1 \ .\end{aligned}$$

Luminance contrast is sometimes defined as $C = (L/L_b) - 1$, where L denotes the gray level at a certain location in the image plane and L_b represents the gray level of the local background. Thus, the value of nodes in the RoLP pyramid can be related to local contrast values directly.

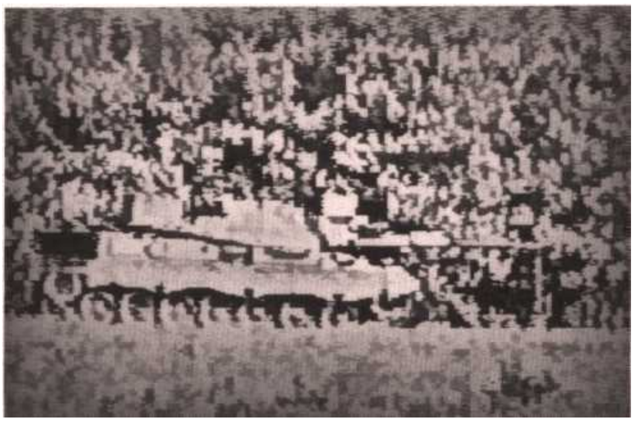
The nonlinear recombination algorithm proceeds from coarse scales to fine scales. At each step, fine-scale ratios are modulated by large-scale ratios, which are regional extrema. The process is such that ratios larger than 1 will generally increase and ratios smaller than 1 will generally decrease. Thus, fine-scale image contrast is enhanced by nonlinear modulation with the regional gray-level extrema at coarser scales. This process is repeated until the finest scale is reached.

This method effectively combines Peli and Lim's[6] and Fahnestock and Schowengerdt's[7] adaptive contrast stretching methods because it involves (1) separate processing of a range of spatial frequency bands and (2) the use of local maximal gray-level contrast to determine the local contrast stretching factor. In contrast to previous methods, the technique presented here is, in principle, parameter free.

2.3 *Eliminating Global Gray-Level Variations*

Information about the overall image intensity is only retained when the nonlinear pyramid recombination starts at some level $M \le N$ and $L_M = G_M$. When a constant-valued image is substituted for L_M, the reconstructed image can be independent of the original mean gray level. In this case, smooth shading and the relative contrast of regions will no longer be represented in the recombination. Hence, the reconstructed image becomes independent of the mean intensity of the original image and is even independent of gray-level gradients across the original image.

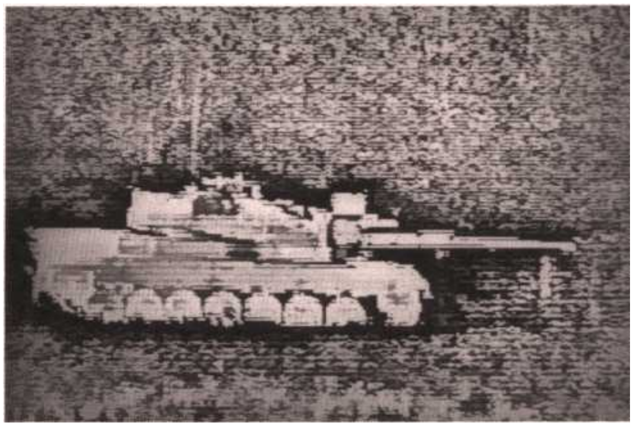
(a)



(b)

Fig. 1 (a) Original 257 x 257 CCD tank image. (b) The result of the nonlinear recombination of the RoLP pyramid representation of Fia. 1(a).

(a)



(b)

Fig. 2 (a) Original 257 x 257 FLIR tank image. (b) The result of the nonlinear recombination of the RoLP pyramid representation of Fig. 2(a).

To illustrate its effect, the multiscale contrast enhancement algorithm was applied to original 8-bit 257 x 257 visual [CCD, Fig. 1(a)] and thermal [FLIR, Fig. 2(a)] images of a tank, as well as to several degraded versions of these images. In all cases, L_N was set equal to 128.

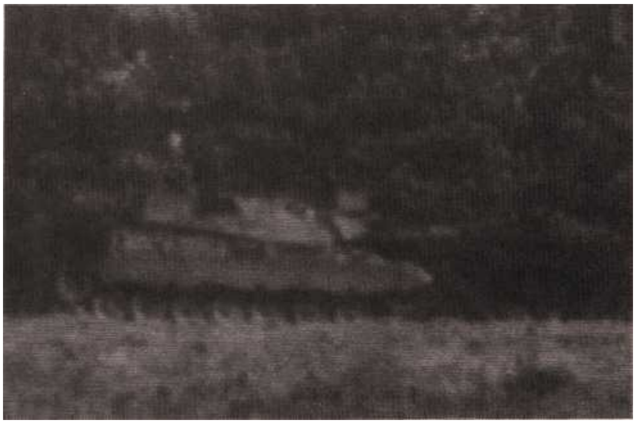
Figure 1(b) shows the result of the nonlinear recombination of the 8-level RoLP pyramid representation of the original CCD image of Fig. 1(a). The contrast of this image is clearly enhanced compared to the original image. Close inspection of this image reveals Mach bands at all resolution levels. This is a direct result of the nonlinear local contrast stretching procedure, which increases the gray-level contrast of pixels on the brighter side of a gray-level edge and decreases the gray-level contrast of pixels on the darker side of an edge.

Figure 2(b) shows the result of the nonlinear recombination of the 8-level RoLP pyramid representation of the original FLIR image of Fig. 2(a). Again, the contrast of this image is clearly enhanced compared to the original image. Figure 2(b) reveals pronounced low-resolution Mach bands separating the image of the tank from its background. This is a direct result of the fact that the original FLIR image contains mainly details of low resolution.

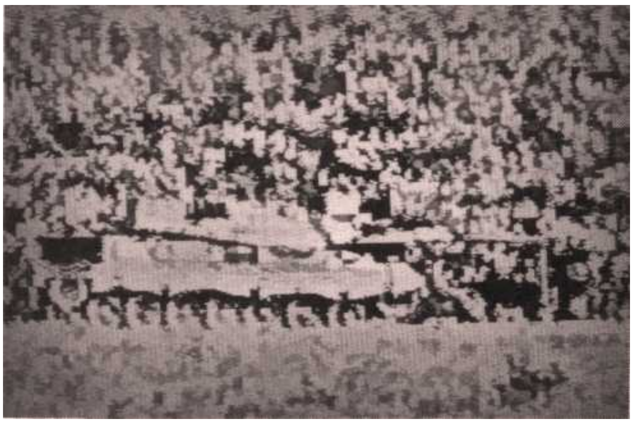
Figure 3(a) shows the image of Fig. 1(a) after the gray-level values have been multiplied by 0.1 to simulate a low-light condition. The result of the nonlinear pyramid recombination of this image is shown in Fig. 3(b). Note the perceptual equivalence of this recombination and Fig. 1(b) (which represents the recombination of the original image).

Figure 4(a) shows the image of Fig. 1(a) after the gray-level values have been multiplied by 0.1 and a constant overall background gray level (of value 128) has been added to simulate a fog condition. This image has low-gray-level contrast, making it very hard to distinguish the depicted scene. The result of the nonlinear pyramid recombination of Fig. 4(a) is shown in Fig. 4(b). This reconstructed image has good contrast and is clearly recognizable. Again, this recombination is similar to Fig. 1(b) [and to Fig. 3(b)]. The examples of Figs. 3 and 4 show that the reconstructed image is independent of changes in lighting.

Figure 5(a) shows the image of Fig. 1(a) after the application of a sinusoidal gray-level modulation to introduce global gray-level gradients. Average local brightness varies strongly in this image making its central part (the horizon and the tank itself) virtually unreadable. Figure 5(b) shows the result of the nonlinear pyramid recombination proce-

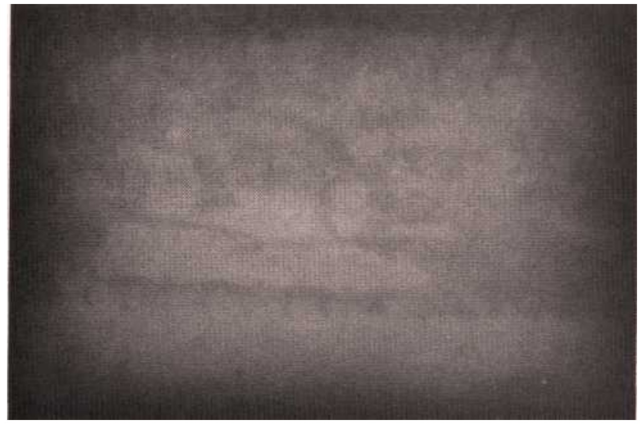


(a)

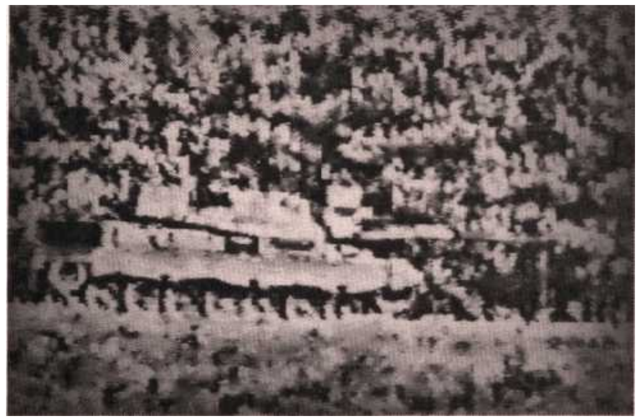


(b)

Fig. 3 (a) The image of Fig. 1(a) after the gray level has been reduced by a factor 10 to simulate a low light condition. (b) The result of the nonlinear recombination of the RoLP pyramid representation of Fig. 3(a).



(a)



(b)

Fig. 4 (a) The image of Fig. 1(a) after the gray level has been reduced by a factor 10 and a constant overall background gray level (of value 128) has been added to simulate a fog condition. (b) The result of the nonlinear recombination of the RoLP pyramid representation of Fig. 4(a).

dure. This image has uniform contrast, is readable in all of its parts, and is again perceptually equivalent to Fig. 1(b). This example shows that the algorithm removes global gray-level modulations and reproduces both the bright and dark parts of the image with excellent contrast.

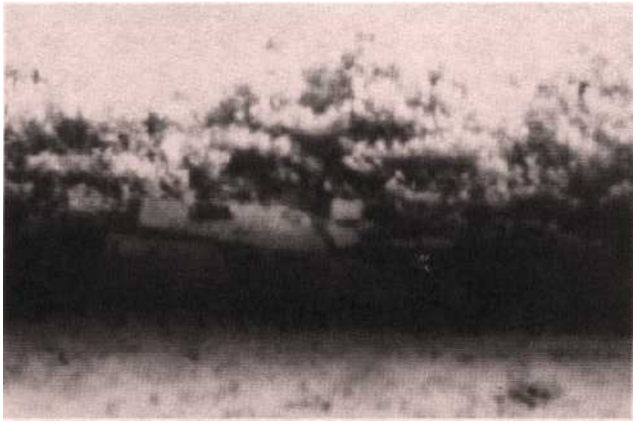
Figure 6(a) shows the FLIR image of Fig. 2(a) after the gray-level values have been multiplied by 0.1 to simulate a low-thermal-contrast condition. The result of the nonlinear pyramid recombination of this image is shown in Fig. 6(b). Note the perceptual equivalence of the recombination of the degraded FLIR image [Fig. 6(b)] to the recombination of the original FLIR image [Fig. 2(b)].

Summarizing, the results show that the recombined images clearly display almost all useful information that is present in the input images. A reconstruction differs from its original in two ways. First, information about the overall intensity of the image is lost when the pyramid layer at which the recombination starts is set to a prescribed gray-level value. As a result, the reconstructed image always looks the same, independent of the overall gray level of the input image. Second, intensity gradients across the image can be reduced and may even disappear completely.

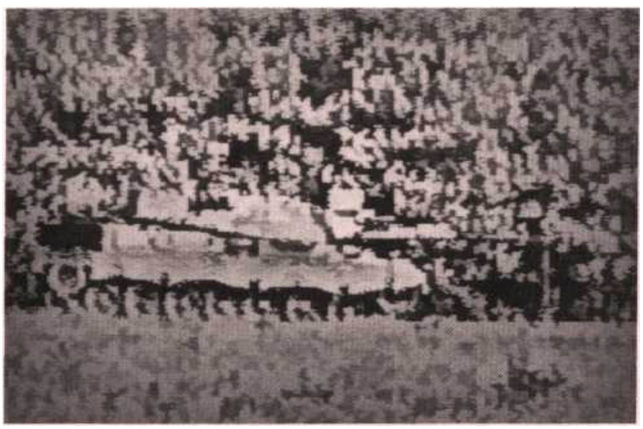
3 Invariant Image Merging

The RoLP image merging scheme can be cast into a three-step procedure.[2,8,1] First, an RoLP pyramid is constructed for each of the source images. The different source images are registered and have the same dimensions. The latter restriction is not very serious, because it can be shown that the method can also be applied to images with different definition regions as long as the intersection of these regions is not empty.[4] Second, an RoLP pyramid is constructed for the composite image by selecting values from corresponding nodes in the component RoLP pyramids. The actual selection rule depends on the application and can be based on individual node values or on masks or confidence estimates. For example, in the case of the fusion of two input images A and B into a single output image C, and maximum absolute local gray-level contrast as a selection criterion, we have $(\forall i,j,l)$:

$$R(C)_l(i,j) = R(A)_l(i,j) \text{ if } \|R(A)_l(i,j)-1\| > \|R(B)_l(i,j)-1\|$$
$$= R(B)_l(i,j) \quad \text{otherwise} ,$$

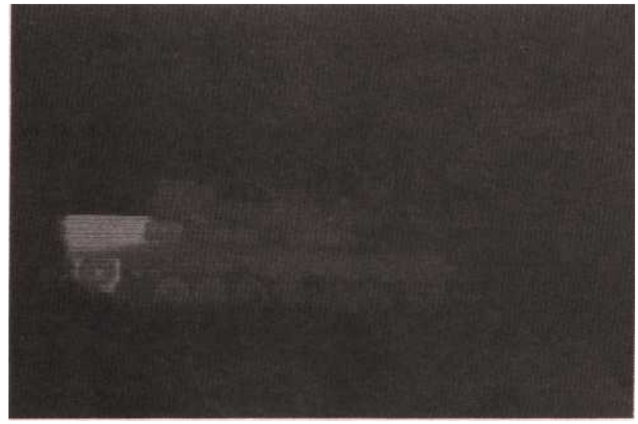


(a)



(b)

Fig. 5 (a) The image of Figure is after the introduction of extra global grey level gradients. (b) The result of the nonlinear recombination of the RoLP pyramid representation of Figure 4(a).



(a)

(b)

Fig. 6 (a) The image of Fig. 2(a) after the gray level has been reduced by a factor 10 to simulate a low-light condition. (b) The result of the nonlinear recombination of the **RoLP** pyramid representation of Fig. 6(a).

where $R(A)$ and $R(B)$ represent the RoLP pyramids for the two source images and R(C) represents the RoLP pyramid for the fused output image. Finally, the composite image is recovered from its RoLP pyramid representation through the corresponding reconstruction procedure.

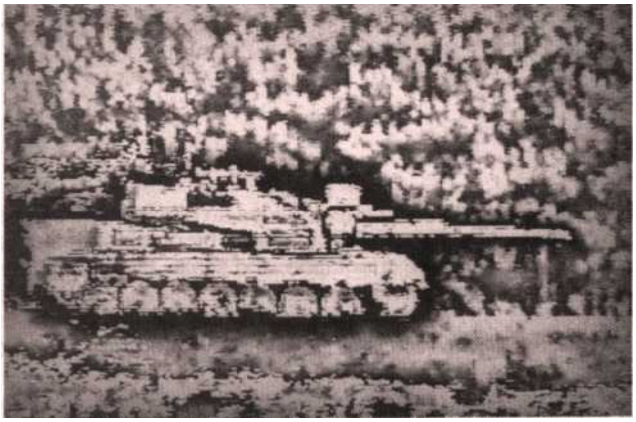
Figure 7(a) shows the result of the application of the image fusion scheme to the contrast-enhanced CCD image of Fig. 1(b) and the contrast enhanced FUR image of Fig. 2(b). In the CCD image, it is hard to distinguish the front part of the gun barrel, the back part of the vehicle (containing the engine room), and the man hood located on top of the tank. All these parts of the target have very low contrast in the visual image. However, in the thermal image these parts are clearly visible. The background, which is clearly visible in the CCD image, is nearly indistinguishable in the FLIR image. This result convincingly demonstrates that the fused images contain those details from both input images that have maximum local contrast. Notice that all the (aforementioned) details that can be obtained from both image modalities are clearly represented in the single fused image.

Figure 7(b) shows the result of the application of the image fusion scheme to merge the enhanced CCD image of Fig. 4(b) with the enhanced FLIR image of Fig. 6(b). Note the striking similarity between this image, which resulted from the fusion of the degraded FLIR and CCD images, and Fig. 7(a), which was obtained by merging the original FLIR and CCD images. Figure 7 clearly demonstrates that the application of the multiscale contrast enhancement algorithm in the image fusion process results in fused images that are highly invariant to global gray-level modulations of the input images.

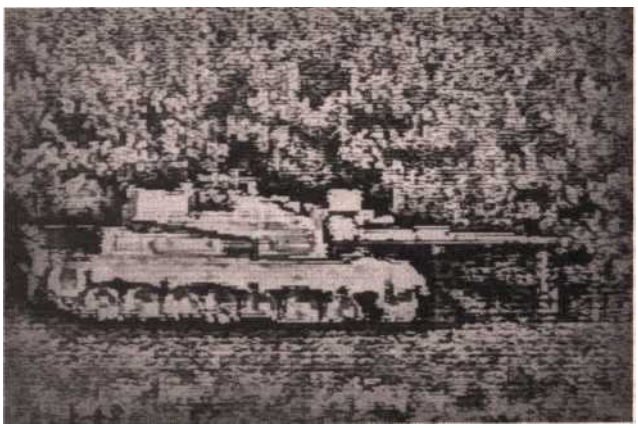
Summarizing, fused images produced by the combined application of the multiscale contrast enhancement algorithm and the RoLP image fusion process always look the same, independent of large variations in the global gray-level characteristics of the individual input images.

4 Discussion

The adaptive multiscale contrast enhancement method introduced in this paper is an effective and computationally efficient contrast enhancement technique that can be used parameter free. Extra parameters can easily be incorporated into the recombination scheme. For instance, gain factors, activation thresholds, and mutual node inhibition can be



(a)



(b)

Fig. 7 (a) Result of the RoLP fusion of Fig. 1(b) and Fig. 2(b). (b) Result of the RoLP fusion of Fig. 4(b) and Fig. 6(b).

introduced to regulate local contrast stretching. The recombination rule adopted in this paper was arbitrarily chosen. In principle, one can think of a large number of ways to reconstruct an image from its multiscale decomposition. The actual choice of the recombination rule depends on the type of information that is to be displayed. For instance, small-scale detail (noise) can easily be supressed by setting one or more of the lower RoLP pyramid layers equal to 1 before performing the recombination.

Fused images produced by the combined application of the multiscale contrast enhancement algorithm and the RoLP image fusion process present a more detailed representation of the depicted scene than either of the input images alone. Moreover, they are perceptually highly invariant, meaning that they appear independent of large variations in the global gray-level characteristics of the individual input images. As a result, the fused images are highly insensitive to the effects of sensor degradation and changing weather conditions. Detection, recognition, and search tasks can therefore benefit considerably from this new image representation.

References

1. A. Toet, L. J. van Ruyven, and J. M. Valeton, "Merging thermal and visual images by a contrast pyramid," Opt. Eng. 28, 789-792 (1989).
2. A. Toet, "Image fusion by a ratio of low-pass pyramid," Pattern Rec. Lett. 9, 245-253 (1989).
3. J. L. Crowley and A. C. Parker, "A representation for shape based on peaks and ridges in the difference of low-pass transform," *IEEE* Trans. Pattern Anal. Mach. Intell. 6, 156-170 (1984).
4. P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE* Trans. Commun. COM-31(4), 532-540 (1983).
5. M. M. Fleck, "Representing space for practical reasoning," Image Vision Computing 6, 75-86 (1988); M. M. Fleck, "Boundaries and topological algorithms," Report AI-TR 1065, MIT Art. Int. Lab., Cambridge, Mass. (1988).
6. T. Peli and J. S. Lim, "Adaptive filtering for image enhancement," Opt. Eng. 21, 108-112 (1982).
7. J. D. Fahnestock and R. A. Schowengerdt, "Spatially variant contrast enhancement using local range modification," Opt. Eng. 22, 378-381 (1983).
8. A. Toet, "Image fusion through multiresolution contrast decomposition," TNO Report IZF 1989-19 (1989).

Alexander Toet received the BS and MS degrees in physics from Utrecht University in the Netherlands in 1976 and 1982. He received the PhD degree in physics from the same university in 1987. The subject of his thesis was human visual perception of spatially ordered structures. His current research interests are in the fields of human vision and image analysis. In the area of human vision, he is studying the subjective geometry of visual space. In the area of image analysis, he is currently working on the construction of a hierarchical morphological graph description of shapes in images.