NOTE

A Practical Approach to Image Restoration for Computer Vision

FEDERICO BUMBACA

Laboratory for Intelligent Systems, Division of Electrical Engineering, National Research Council, Montreal Road, Ottawa, Canada K1A 0R6

AND

KENNETH C. SMITH

Department of Electrical Engineering, University of Toronto, Toronto, Canada M5S 1A4

Received August 5, 1986; accepted November 5, 1987

The objective of this paper is to present a practical approach to image restoration for computer vision applications. Emphasis is placed on the practical application of restoration processes to images obtained from a commonly employed camera (vidicon) possessing typical degradation sources. A brief review of applicable standard image processing techniques is given in conjunction with the description of a typical distortion process. Three types of preprocessing are found to be effective: noise filtering, radiometric correction, and inverse filtering. A case study outlining a practical approach to extract the original input image from its distorted counterpart is then presented.

1. INTRODUCTION

In the literature, there exists much work concerning all aspects of image processing theory, whether it be linear filter theory or nonlinear correction algorithms [1]. However, very few researchers have presented a more practical approach to dealing with the entire restoration process, linear and nonlinear methods alike. In addition, even less work seems to exist in the area of image restoration directed at applications for computer vision. A complete model of the distortion process exhibited by the commonly employed vidicon camera is thereby presented, followed by a practical procedure to extract the original input image from its given distorted counterpart.

Emphasis is placed on preparing the image for subsequent computer vision processing as opposed to human vision processing. This point cannot be overstressed, for the two processes are quite unique and distinct. Computer vision processes which include edge finding and segmentation algorithms [29] are clearly not as robust as those exhibited by the human visual system. They may be easily fooled and led astray by local perturbations in the imagery. The human system, however, easily overlooks minor distortions and further unambiguously interprets most images which would seem to be grossly distorted and buried in noise. Therefore, in order to quantitatively analyze the results of an image restoration process, as a computer vision algorithm does, graphical plots not grey-scale images are required. Grey-scale images which seem acceptable in quality to humans may, in fact, be quite unacceptable to a computer. Only through plots can we accurately...
observe any resulting noise, blurring, and radiometric distortion which may confuse a computer vision algorithm. To this end, the images presented are in the form of graphical plots of the grey-scale images.

In some computer vision domains where geometric precision is critical, it may be necessary to account for nonlinear geometric distortions of the camera. The following restoration process does not account for such effects and the interested reader is urged to refer to [30] which describes a general methodology for geometric correction in three-dimensional space. Nevertheless, many computer vision tasks still require calibration of the camera to determine the correspondence between the image geometry and the actual 3D geometry. Provided that significant geometric correction is not necessary, the calibration stage should occur after preprocessing. This is necessary to account for any further geometric distortion introduced by the preprocessing stages. That is, both radiometric correction and inverse filtering may adversely affect the relative geometric position of the imagery which will undoubtedly affect any prior geometric calibration.

All imaging systems have non-ideal image formation processes due to several degradation sources. Three types of degradation [2-3] are considered with respect to a vidicon camera:

- intensity nonlinearity (radiometric distortion)
- bandwidth reduction
- noise distortion.

Nonlinearity occurs in every recording process. This nonlinearity, to be referred to as radiometric distortion, may involve intensity nonlinearity as well as spatial nonuniformity, i.e., radiometric distortion may be viewed as a nonlinear shift variant (NLSV) process. All physical recording systems are also band limited, causing a loss of resolution. This blurring process is assumed to be linear shift invariant (LSI). The measurement process introduces noise to the image as well. This noise must be filtered out without causing a further blurring effect on the image.

The overall recording process for the case of a vidicon camera can be described by (1):

\[ g(x, y) = S \left( \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x - r, y - s) f(r, s) \, dr \, ds, x, y \right) + n(x, y); \]  

(1)

\( S(\cdot, \cdot, \cdot) \) represents the radiometric distortion which is a function of the intensity and position \((x, y)\) of the particular pixel, \( h(\cdot, \cdot) \) is a linear spatial operator that takes into account the bandwidth limitation, \( n(\cdot, \cdot) \) is the random noise, which is assumed to be additive, \( f(\cdot, \cdot) \) is the ideal image, and \( g(\cdot, \cdot) \) is the measured image.

A discretized version of (1) follows:

\[ g(k, l) = S \left( \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} h(k - m, l - n) f(m, n), k, l \right) + n(k, l). \]  

(2)

The preprocessing to be described will attempt to recover \( f(\cdot, \cdot) \) from \( g(\cdot, \cdot) \). However, this restoration process is extremely difficult due to lack of exact informa-
tion concerning $S(\cdot, \cdot, \cdot)$, $h(\cdot, \cdot)$, and $n(\cdot, \cdot)$. These degrading processes are measured experimentally whenever possible in order to compute an estimate $\hat{f}(\cdot, \cdot)$ of $f(\cdot, \cdot)$ given the information $g(\cdot, \cdot)$. The mean square error is used as the criterion to determine this optimum estimate.

First of all, a filter is applied to the image $g(\cdot, \cdot)$ such that $n(\cdot, \cdot)$ is eliminated as much as possible without affecting the actual image $f(\cdot, \cdot)$ significantly. The filter to be used must have certain properties, such as noise reduction and edge preservation. The new image produced is

$$g'(k,l) = S\left\{ \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} h(k-m, l-n)f(m,n), k,l \right\} + n(k,l) - \hat{n}(k,l)$$

(3)

where $n(k,l) - \hat{n}(k,l)$ is the reduced noise in the image $g'(k,l)$, and it is assumed that $f(m,n)$ has been unaffected by the noise filtering process.

A radiometric correction NLSV process, $S^{-1}$, is applied to $g'(\cdot, \cdot)$ in order to compensate for this type of distortion. $S^{-1}$ is measured experimentally and is obviously not ideal. The new image produced has the general form of (4):

$$g''(k,l) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} h(k-m, l-n)f(m,n) + n_s(k,l).$$

(4)

It is assumed that $n_s(k,l)$ is negligible, resulting in (5):

$$g''(k,l) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} h(k-m, l-n)f(m,n).$$

(5)

It may be noted from (5) that the input and output images are related by a convolution operation. In order to extract the original image, $f(k,l)$, from the output image, $g''(k,l)$, a deconvolution is required. In the frequency domain, this may be expressed as an inverse filter. Rewriting (4) in the Fourier domain as (6),

$$G''(u,v) = H(u,v)F(u,v) + N(u,v)$$

(6)

an inverse or restoring filter, $R(u,v)$, may easily be found from $H(u,v)$;

$$R(u,v) = \frac{H^*(u,v)}{|H(u,v)|^2}$$

(7)

where $H^*(u,v)$ denotes the conjugate of $H(u,v)$. This quite trivial filter assumes that $h(k,l)$ is known exactly and that no noise is present. It is therefore very inefficient near the zeroes of $H(u,v)$, where $R(u,v)$ goes to infinity. If the noise is non-zero at these frequencies, the estimated image will be poor:

$$\hat{f}(u,v) = R(u,v)G''(u,v) + \frac{N(u,v)H^*(u,v)}{|H(u,v)|^2}$$

(8)
Additional noise is added due to the approximations of \( h(k, l) \) to the true impulse response and of \( r(k, l) \) to the true inverse of \( h(k, l) \). The former approximation is due to an inaccurate estimate of the true camera response and the latter is due to the actual filter implementation.

Rewriting (8) in the space domain and including the aforementioned additional noise, (9) is obtained,

\[
\hat{f}(k, l) = r(k, l) * \{ g''(k, l) + n_s(k, l) \} + n(k, l),
\]

which may be simplified to (10),

\[
\hat{f}(k, l) = r(k, l) * g''(k, l) + n(k, l),
\]

where \( * \) signifies convolution, and \( \hat{f}(k, l) \) represents the approximation to the original image \( f(k, l) \). The noise term, \( n(k, l) \), may become quite significant, as mentioned previously. This is dealt with by further processing with a filter having similar properties as the noise filter discussed earlier. Figure 1 summarizes the overall process discussed above.

2. NOISE FILTERING

An image may be subject to noise and interference from several sources including electrical sensor noise, scanner noise, noise introduced by an optical system, and noise due to analog-to-digital conversion. These noise effects can be minimized by classical statistical filtering techniques (restoration filters) or spatial ad hoc processing techniques [1]. The success of each of these filters is dependent on the type of noise to be filtered and on the tolerable additional noise introduced to the image as a side effect.

In many computer vision systems, edges are of prime importance since they define object boundaries from which features and other information are extracted. Any noise filtering process, then, should remove unwanted noise such as random impulses without further blurring. Two filters have been investigated in this regard [4]: (1) a mean filter and (2) a median filter. The mean or low pass filter simply replaces the center pixel of a sliding two-dimensional window with the average or mean of the pixels within the window. The median value of a discrete sequence of length \( N \) is defined to be that member for which \( (N - 1)/2 \) elements are smaller or equal in value, and \( (N - 1)/2 \) elements are larger or equal in value.

Median filtering has two main advantages [5]: (1) it preserves sharp edges and (2) it is very efficient for smoothing spiky noise. These advantages led to their adoption as pre-filters for smoothing images with spiky noise distributions and post-filters for smoothing amplified noise or spurious oscillations due to high-emphasis or restoration filtering [6]. However, in cases where the noise is not spiky, the mean filter has
better noise suppression characteristics [5]. Figure 2 represents a degraded image which exhibits spiky noise [4], thereby warranting the use of a median filter.

The real-time hardware implementation of a two-dimensional median filter can be quite costly since \( n \) pixels must be sorted before their median is actually known. A separable median filter has a much simpler implementation at the cost of a slight decrease in performance. It consists of successive applications of a one-dimensional median filter on the rows and columns of an image. Table 1 compares the actual performance obtained for the two types of median filters for observed levels of input intensity and noise. A separable five-by-five filter seems to present the best trade-off between noise suppression, edge preservation, and cost [4]. Figure 3 represents the image in Fig 2 after processing with such a filter. Various software and hardware implementations may be found in [6–12].

### 3. RADIOMETRIC DISTORTION

The vidicon camera introduces radiometric distortion into the imagery that it produces. There are two primary causes for this distortion. The first is due to the nonlinear response of the camera to light intensity in object space. The second is

---

**TABLE 1**

<table>
<thead>
<tr>
<th>Input intensity ( g(\bar{u}) )</th>
<th>Input noise ( \sigma^2 )</th>
<th>Output noise ((n = 9))</th>
<th>Output noise ((n = 25))</th>
</tr>
</thead>
<tbody>
<tr>
<td>205.3</td>
<td>12.917</td>
<td>1.8153</td>
<td>2.1693</td>
</tr>
<tr>
<td>102.9</td>
<td>2.0199</td>
<td>0.3700</td>
<td>0.4162</td>
</tr>
<tr>
<td>64.92</td>
<td>0.9578</td>
<td>0.2014</td>
<td>0.2327</td>
</tr>
<tr>
<td>40.96</td>
<td>1.8119</td>
<td>0.7955</td>
<td>0.8349</td>
</tr>
<tr>
<td>32.54</td>
<td>1.9974</td>
<td>0.9483</td>
<td>0.9821</td>
</tr>
</tbody>
</table>
due to the lack of spatial uniformity in the system’s response to light. That is, the camera’s nonlinear response varies with respect to its field of view [13].

Computer vision algorithms are also adversely affected by radiometric distortion when grey level intensities are required. This distortion becomes even more significant in the case of colour vision systems [4, 31]. For example, thresholds [29] are commonly utilized in computer vision algorithms; however, they are not dependent on either the particular area of the image in which they are applied nor the intensity level of that area. Although two different regions of the output image from a camera may have similar intensities, their true intensities may vary significantly in the original image, thereby potentially causing a vision algorithm to make an incorrect decision based on a given threshold level on the output image. In other words, radiometric distortion which varies with the particular area of an image, may quite easily lead a vision algorithm astray. It is imperative that a uniformly illuminated shade of grey (or red, green, blue) produce a uniform grey level image in the camera output. It is also vital that the input–output intensity relationship be linear.

One very simple correction technique is to use a template for all images (Fig 4). By recording a uniformly illuminated shade of grey at an average level of intensity, this template may be used to remove spatially varying shading effects [13]. An implementation may be found in [14]. More complex correction techniques require the recording of several uniformly illuminated shades of grey at varying levels of intensity within the camera’s dynamic range. Interpolation techniques are then applied to remove the distortion effects. The result in an image whose intensity varies linearly with the ideal image intensity.

Due to the nature of the vidicon camera’s radiometric distortion, a best-fit cubic polynomial was chosen as the interpolating function. This provides a more precise correction exhibiting a greater degree of precision throughout the entire dynamic range of the camera.

If the response is slowly varying enough, as it is in this case (Fig 5), the same correction factor may be applied for all pixels within a given region of the field of view. It was decided to compute the cubic coefficients for each two-by-two pixel
area, thereby effectively reducing the computational and/or memory cost by a factor of four, without significantly affecting the precision.

The most general input/output relation for system intensity nonlinearities is given by (11) [15],

$$g(k, l) = S\{ f(k, l), k, l \} = S_{k,l}\{ f(k, l) \}, \quad (11)$$

where $f(k, l)$ is the input intensity and $g(k, l)$ is the output signal. Perfect intensity correction would be possible if we could return to the ideal image $f(k, l)$ with the
knowledge of $g(k, l)$ at any point $(k, l)$. This operation can be written as

$$\hat{f}(k, l) = S^{-1}\{g(k, l), k, l\} = S_{k, l}^{-1}\{g(k, l)\},$$

where $S^{-1}$ is the inverse distortion operation and $\hat{f}(k, l)$ is the ideal image estimate. Since the nature of the sensor response is approximately third order and it is not desired to limit its dynamic range, the approximation to $S^{-1}$ is made,

$$\hat{f}(\tilde{u}) = d_0(\tilde{u}) + d_1(\tilde{u})g(\tilde{u}) + d_2(\tilde{u})g^2(\tilde{u}) + d_3(\tilde{u})g^3(\tilde{u}).$$

where $\tilde{u}$ is the position of the pixel $(k, l)$, and the $d_i(\tilde{u})$ are the position dependent coefficients.

Using five plates [16] representing uniformly illuminated shades of grey of known intensity, as inputs, the coefficients, $d_i(\tilde{u})$, are determined for each two-by-two pixel area. Linear least squares estimation theory is used to fit a cubic polynomial such that the mean squared error is minimized. A typical polynomial is plotted in Fig 4. Using Fig 5 as the input, $g(\tilde{u})$, to (13), Fig 6, $\hat{f}(\tilde{u})$, results. $\hat{f}(\tilde{u})$ is the intensity estimate of the original plate having a uniformly illuminated shade of grey. Figure 7 is the resulting image after radiometrically correcting the image of Fig 3.

4 INVERSE FILTERING

As mentioned previously, edges are of prime importance to computer vision algorithms. It is therefore necessary to enhance the blurred edges produced by the camera. However, the transfer function or impulse response of the camera used must be determined before an inverse filter can be designed. Due to the extreme difficulty and impracticality of experimentally measuring its response it seems that the best approach is to estimate it.
Several papers [17–19] discuss the modulation transfer function of electro-optical devices, which may be expressed in the form of (14),

\[
H(\omega_x, \omega_y) = \exp \left[ - \left( \frac{\sqrt{\omega_x^2 + \omega_y^2}}{\omega_c} \right)^n \right],
\]

(14)

where \( n \) is the modulation transfer index, \( \omega_c \) is the frequency constant in cycles per millimeter, \( \omega_x \) is the spatial frequency, \( u \), in the \( x \) direction, and \( \omega_y \) is the spatial frequency, \( v \), in the \( y \) direction. From the manufacturer's specifications [20] and the papers previously mentioned, the values of \( n \) and \( \omega_c \) have been determined to be approximately equal to 1.6 and 26 cycles per millimeter, respectively, for the particular vidicon tube used. Figures 8 and 9 illustrate the transfer function and impulse response, respectively, using these values.
The inverse restoration filter, \( R(u, v) \), is now computed using (15):

\[
R(u, v) = \frac{H^*(u, v)}{|H(u, v)|^2}.
\]  

The Fletcher-Powell nonlinear minimization algorithm is used to design a second-order infinite impulse response (IIR) two-dimensional filter by minimizing the mean squared error, (16), between the ideal filter \( R(u, v) \), and the designed filter, \( R'(u, v) \) [21].

\[
\text{mean squared error} = \sum_{k=0}^{N} \sum_{l=0}^{N} |R(k, l) - R'(k, l)|^2,
\]  

where \( N \) represents the maximum spatial frequency in the \( x \) and \( y \) directions. That is, (16) represents the mean squared error over the entire frequency space. Figure 9 illustrates the transfer function of the ideal filter, \( R(u, v) \). The designed filter takes the following form in the \( Z \) domain:

\[
R'(z_1, z_2) = \frac{\sum_{l_1=0}^{2} \sum_{l_2=0}^{2} a(l_1, l_2) z_1^{-l_1} z_2^{-l_2}}{\sum_{k_1=0}^{2} \sum_{k_2=0}^{2} b(k_1, k_2) z_1^{-k_1} z_2^{-k_2}},
\]  

where

\[
\begin{align*}
(a_{l_1,l_2}) &= a(l_1, l_2) \quad \text{and} \quad (b_{k_1,k_2}) = b(k_1, k_2) \\
R_{\text{ideal}}(u, v) &= \frac{1}{1 + H(u, v)} \\
R_{\text{designed}}(u, v) &= \frac{1}{1 + H'(u, v)}
\end{align*}
\]
TABLE 2
IIR Filter Coefficients

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-9.135 \times 10^{-2}</td>
<td>-0.1982</td>
<td>-8.546 \times 10^{-2}</td>
</tr>
<tr>
<td>2</td>
<td>-0.1982</td>
<td>3.472</td>
<td>5.3666</td>
</tr>
<tr>
<td>3</td>
<td>-8.546 \times 10^{-2}</td>
<td>5.3666</td>
<td>7.796</td>
</tr>
</tbody>
</table>

where $a(l_1, l_2)$ and $b(k_1, k_2)$ are computed from (16) and tabulated in Table 2. These values define a filter which seems indistinguishable from the ideal [4] and will therefore not be reproduced here. Had the filter specifications been more stringent, a fourth or higher order implementation may have been required. Figure 11 represents the image of Fig 7 after inverse filtering by convolution with the filter specified by (17). Although somewhat difficult to observe, the restored image does indeed have pronounced edges. Various hardware implementations of inverse filters may be found in [21–28].

It is also noted that the matrices represented by the filter coefficients in Table 2 are symmetric due to the symmetry of the inverse filter transfer function (Fig 10). $R'(u, v)$ is not isotropic as predicted from (14) since it has been shown [32] that the imposition of exact circular symmetry on the class of causal recursive filters results in a denominator polynomial of unity. Therefore, to achieve a good approximation to a circularly symmetric filter, octagonal symmetry was imposed. This leads to numerator and denominator matrices which are symmetric (Table 2).

Due to noise and spurious oscillations which may be introduced into the digital imagery by the restoration filter, a separable three by three median filter is further

![Fig. 10. Inverse transfer function, $R(\omega_x, \omega_y)$.](image-url)
utilized to reduce these effects. Figure 12 gives the image of Fig 11 after further noise filtering. Unfortunately, due to the nature of the degradation problem the true ideal image, $f(k, l)$, is not available for comparison with its approximation, $f'(k, l)$, in Fig 12.

5. CONCLUSIONS

A practical approach to image restoration for applications in computer vision has been presented. It has been shown that four preprocessing stages are necessary for acceptable image restoration. A separable median filter is required to eliminate noise as well as preserve edges at a relatively low cost for real-time applications. A radiometric correction procedure employing a best-fit cubic polynomial is necessary to achieve an acceptable level of quality. Finally, a second order IIR inverse filter is required to reduce any blurring effects followed by noise filtering to further
eliminate noise and spurious oscillations. This practical approach has been applied as a preprocessing stage for the authors’ work in colour computer vision [4, 31].

ACKNOWLEDGMENTS

This research has been supported by a Natural Sciences and Engineering Research Council postgraduate scholarship and operating grant A1753.

REFERENCES

20. RCA Electro Optics and Devices, 4875/U Series, RCA, Lancaster, PA.


