Self-detection of optical contamination or occlusion in vehicle vision systems

Yi Zhang
Jie Yang
Kun Liu
Xiang Zhang
Shanghai Jiao Tong University
Institute of Image Processing and Pattern Recognition
P.O. Box A0603221
800 Dongchuan Road
Shanghai 200240, China
E-mail: zypk@sohu.com

Abstract. We present a novel and practical algorithm for the self-detection problem of contamination or occlusions on the lens of a camera mounted on a vehicle. First, we analyze the intrinsic characteristics of such contamination on the video image. Based on this, cumulative differences are used to segment the static region in the image. A blurred edge detection algorithm based on wavelet decomposition is introduced to confirm if the static region belongs to contamination or an occlusion. Through the combination of these algorithms, contamination or occlusions can be detected. Experimental data are analyzed to show the detection performance of our algorithm and the effect of different contamination or occlusion material. © 2008 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.2947578]

Subject terms: lens contamination detection; lens occlusion detection; cumulative difference; blurred edge detection.

Paper 070975R received Dec. 16, 2007; revised manuscript received Apr. 8, 2008; accepted for publication Apr. 10, 2008; published online Jun. 25, 2008.

1 Introduction

Vision-based recognition for use in intelligent vehicles is an active research topic in automotive electronics. It attempts to use image processing and computer vision techniques to comprehend road conditions and features such as road boundaries, lane markings, other vehicles and pedestrians, etc. The algorithm then makes decisions based on this data to ensure the safety of the vehicle with appropriate braking distances, speeds, and correct steering. In abnormal situations such as accidents, it is essential for the automatic driver to instigate an appropriate response in the shortest possible time. These tasks all depend on the correct information provided by the vehicle camera. If the camera is contaminated or occluded by something like dust, oil, or leaves, the decisions made by the algorithm can be useless or even dangerous.

The automatic determination of whether a camera lens is occluded or clean is a key problem. Robins and Nelson1 used a separate viewfinder mounted on the lens, and a light source located between the lens cover and the digital image sensor. If the lens is contaminated, a portion of the light from the light source will be scattered and directed onto the digital image sensor. A contamination detector for receiving the digital image signals caused by the scattered light can determine the presence or absence of contamination on the lens.

We detect contamination or occlusions on the lens of the camera from our actual test video sequences. When the lens is contaminated or occluded, these areas remain static with the movement of the vehicle. Cumulative difference is used to segment the static region from the moving background. Cumulative difference is an accumulation of a series of temporal differences. Temporal differencing is a comparison of pixel-wise differences between two or three consecutive frames to distinguish between static and dynamic regions. Temporal differencing is very adaptive and robust to dynamic environments, and is often used in moving-target detecting and tracking.2

The contaminated or occluded regions form a corresponding blurred region in the digital image (Fig. 1). This is because the contamination is adhering to the lens; their images are an out-of-focus blur while other regions are clear. This is a dramatic feature that can be used to distinguish contamination or occlusions and unobscured areas of the image. It also accords with our subjective judgment. A blurred edges, detection algorithm with wavelet transform is used to define and detect blurred regions based on the cumulative difference result. If a region is both static and has blurred edges we can be sure that it is a contamination or occlusion region. Through the combination of the two algorithms, we can identify the contamination and occlusion. The system structure is shown in Fig. 2.

Our work is structured in the following way. Section 1 introduces the research background and analysis intrinsic characteristics of contamination and occlusion. Section 2 presents the cumulative difference method to segment static regions in the video. Section 3 presents a blur edge detection algorithm based on wavelet transform. Section 4 shows the combination process of the two algorithms. Section 5
analyzes contamination and occlusion material and shows the experimental result. Section 6 concludes the work.

2 Static Region Segmentation with Cumulative Difference

Static region segmentation is based on cumulative difference of video sequences. The difference matrix \( D_n(m,n) \) of each pixel \((m,n)\) between the current frame \(f_n\) and the last frame \(f_{n-1}\) is defined as:

\[
D_n(m,n) = \begin{cases} 
1, & \text{if } [f_n(m,n) - f_{n-1}(m,n)] > T_1 \\
0, & \text{if } [f_n(m,n) - f_{n-1}(m,n)] \leq T_1 
\end{cases}
\]

(1)

\(T_1\) is a threshold defined according to the sensor type. Points that have value equal to 1 in matrix \(D_n(m,n)\) are moving points and 0 are static points. Then difference matrix \(D_n(m,n)\) carries out an “or” operation. That is:

\[
A_n(m,n) = D_0(m,n) | D_1(m,n) | \ldots | D_n(m,n).
\]

(2)

Thus \(A_n(m,n)\) comprise the dynamic information in the scene from time 1 to \(n\). Let \(S_n\) represent the 0 value point set in \(A_n(m,n)\), which represents the static region.

In this system, the cumulative difference operation will stop when it deems the static region is segmented, or there is no static region and executes the next step or restarts for the next detection period. The question then posed is when should the cumulating operation stop and restart? Since the contamination and occlusions are correspondingly steady during certain times while the scene is changing, we have designed a self-adaptive stop method with the following details.

1. When the cumulative difference process is running, calculate \((S_n - S_{n-1})/S_n\).

2. If \((S_n - S_{n-1})/S_n < T_2\), add 1 to counter \(COU\). Else, \(COU=0\). (The initial value of \(COU\) is 0. \(T_2\) is called the stability threshold.)

3. If \(COU > T_3\), stop the cumulative difference process. \(T_3\) is called the convergence threshold.

It means that when the changing part of \(A_n(m,n)\) has increased a small amount for the period of time, we believe it can be stopped. In practice, \(T_2=0.002\) and \(T_3=50\) are approximate values and very suitable for our application.

The last step is to use a morphologic method to “erase” some noise. In practice, some materials such as oil or gray matter can produce noise or tiny gaps in the image \(A_n(m,n)\). We remove such noise through “dilation,” followed by a “corruption” operation. Fig. 3 and 4 show such a result image \(A_n(m,n)\) following these steps.

Using this method, we can observe static regions within the video sequence comprising contamination and occlusion regions, as analyzed previously. Not all these regions are necessarily contamination or occlusions. The possibility exists that there are some unchangeable gray areas within the scene. For example, Fig. 4 shows the cumulative difference result; the sky in the top-left region is identified as a
3 Blurred Edge Detection with Wavelet Transform

The existing methods for blurred image detection cannot discriminate if a single image is blurred or not. The method detects the clearest within a group of images in the same scene. However, there is a direct method that examines a discriminative feature such as an edge. When blurring occurs, both the edge type and its sharpness will change. By examining the absence of the alternating component (AC) coefficients, indicating edge sharpness, the authors of Ref. 6 proposed using discrete cosine transform (DCT) information to qualitatively characterize the extent of the blur. Tong et al. have proposed that such a scheme takes advantage of the ability of the Harr wavelet transform to discriminate different edge types while recovering the sharpness from the blurred version. This is effective for both out-of-focus blur and linear-motion blur. Direct methods are not so precise and are not used for direct blur discrimination. We have achieved a good out-of-focus blur detection result in the frame image by modifying and improving the Tong et al. method. The following are the details of our improved method.

Decompose the current frame \( f(x,y) \) with the Harr wavelet transform to level three.

\[
W_m(m,n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \psi_{j,m,n}(x,y),
\]

\[
W_{\phi}(j,m,n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \psi_{j,m,n}(x,y) \quad i = \{H,V,D\},
\]

\[
\psi_{j,m,n}(x,y) \text{ is a 2-D Harr basic wavelet. To compute the edge map } W_j(m,n) \text{ in each scale:}
\]

\[
W_j(m,n) = (LH_j^2 + HL_j^2 + HH_j^2)^{1/2} \quad (i = 1,2,3).
\]

\( LH_j, HL_j, \) and \( HH_j \) are horizontal low-pass/vertical high-pass, horizontal high-pass/vertical low-pass, and horizontal high-pass/vertical high-pass, respectively. We find the maximum value in appointed windows of each \( W_j(m,n) \).

The window in \( W_1(m,n) \) is \( 4 \times 4 \), in \( W_2(m,n) \) is \( 2 \times 2 \), and in \( W_3(m,n) \) is \( 1 \times 1 \). The result is saved in each \( W_{\text{max}}(m,n) \). \( W_{\text{max}}(m,n) \) represents the intensity of the edge. The larger \( W_{\text{max}}(m,n) \) is, the more intense the edge is. Different edge types have different intensity characteristics in each \( W_{\text{max}}(m,n) \). For example, a sharp edge is distinct in \( W_{\text{max}} \), but weaker in \( W_{\text{max}} \), and little in \( W_{\text{max}} \). On the contrary, a blurred edge is indistinct in \( W_{\text{max}} \), and more and more clear in the next. We design three detailed rules consulting the method of Tong et al.

1. If any point of \( W_{\text{max}}(m,n) > T_A \), point \((m,n)\) is an edge point.
2. If \( W_{\text{max}}(m,n) < W_{\text{max}}(m,n) < W_{\text{max}}(m,n) \) or \( W_{\text{max}}(m,n) < W_{\text{max}}(m,n) \) and \( W_{\text{max}}(m,n) < W_{\text{max}}(m,n) \), \((m,n)\) is more likely to be a blurred edge point.
3. For any point \((m,n)\) that satisfies rule 2, if \( W_{\text{max}}(m,n) < T_A \), then \((m,n)\) is a blurred edge point.

The three rules are used to identify whether \( W_{\text{max}}(m,n) \) belong to a blurred edge point. The result is saved in \( W_{\text{blur}}(m,n) \). White points represent blurred edges. Figures 5 and 6 show the processed result of one image. We find that the blurred edge points are extracted completely but with heavy noise. We reduce the noise with...
the following method: every 30 frames, we implement the wavelet image blur measure to one frame. After 300 frames have been processed, we calculate the average of the ten result images, giving us a final single image. See Fig. 7.

The blurred edge of the contamination and occlusion does not change greatly, but the other blurred points change with time. So by averaging the results, the noise can be reduced. By defining a threshold, we can optimize a result. Figure 8 is the binary image of Fig. 7 and is obviously better than Fig. 6.

4 Combination Process

When the cumulative difference process has ended and each independent static region is labeled, blurred edge detection should end as well. The combination process is to identify regions that are both static and have a blurred edge, then decide if they are contamination or occlusions. Our method is detailed here.

1. Find all independent connected regions in $A_n(m,n)$ and name it $block_n$. Calculate the centroid of each $block_n$.
2. Search from the centroid of each $block_n$ to the edge in 16 directions to find if there is a blurred edge point calculated from Sec. 3. If the edge of a region is also the edge of an image, discard it.
3. Calculate the proportion of the directions that have a blurred edge in 16 directions. $block_i$ is not regarded as a contamination or occlusion region, if the proportion is less than $T_5$. This is the end of the contamination and occlusion decision process. However, in experiments we have found that large regions can be regarded as static but are not contamination regions. However, it may comprise contamination regions (Fig. 4). If such regions are detected, we need to continue the process.
4. If a region is not regarded as a contamination region but its area is greater than a certain value, divide it into small blocks.
5. Search for a blurred edge for each block not exceeding the overall region. The search process is similar to step 2.
6. Combine each block that has a blurred edge and define it as a contamination region.
7. Give the final discrimination result such as its area and position.

The detection process is then finished. In our experiment work this has been shown to work well. Its performance is analyzed in the next section.

5 Experimental Result

The experimental video samples were recorded by a camera mounted on a car. 43 video sequences of 640 × 480 resolution at 30 frames per second (FPS) containing different contamination and occlusion materials were collected. We divided the contamination and occlusion types into three categories according to different materials, dark, gray, and water-like contaminations.

Detection performance was evaluated and parameters are optimized by means of receiver operating characteristics (ROC) curves. Figure 9 has the solid line representing the detection performance of dark contamination; the dashed line represents gray contamination. Detection and false positive rates are calculated with reference to good detection within an experimental series. The curves are drawn with different thresholds $T_1$ (see Sec. 2). The experi-

![Fig. 7 Average result.](image1)

![Fig. 8 Binary result.](image2)

![Fig. 9 ROC curves of different contamination materials.](image3)
mental results show that both dark contamination and gray contamination can be detected with a performance up to 95%. Gray contamination detections are more sensitive to the change of threshold $T_1$ compared to dark contamination detection. So a suitable threshold is more important to gray contamination. The ROC curves can help us choose a suitable threshold for the detection process.

Some water-like contamination may not be detected, as light can pass through it and change as the scene changes. It is difficult to solve this problem. Fortunately, such contamination will dry and vanish or become dark contamination; in these cases, the effect is temporary. However, water contamination detection is an area for further research.

The computational complexity of cumulative difference is $O(N)$ ($N$ being the size of one image), and the computational complexity of discrete wavelet transform (DWT) is also $O(N)$. Thus the processing speed of our system is ideal. In our experiment, the system can be run in real time in our PC (P4 2.8G, 1-G DDR RAM). The detection process of the algorithm is relative to the vehicle speed and the changing of scenes. In our experiment, all video terminated within 1500 frames. If the video is streamed at 30 FPS, the algorithm can give a contamination or occlusion detection result within 50 sec.

We developed the system with VC2005 and tested our algorithms. The system interface is shown in Fig. 10. To the left are parameter settings. The right has a video window. The four windows in Fig. 11 are the cumulative difference operation results, blurred edge detection results, region labeling results, and the final detection result respectively. Figures 12 and 13 show gray contamination examples with detection.

6 Conclusion

We present a simple but practical framework to solve the lens contamination detection problem. Such contaminations are hard to detect with imaging alone. Other methods use additional equipment mounted on the lens to detect contamination.

As our aim is to detect contamination on a lens mounted on a vehicle, it is possible to use the motion information and the out-of-focus blur feature in videos to detect contamination and occlusions. It has proven to be an effective
and practical method. Although it is not suitable for all water contamination, it is adequate in many applications and situations. We feel that water contamination detection will require a new algorithm for an adequate solution and will be a future area of research.

References


Yi Zhang received the BS in automatic control from Southeast University in 2002, and the MS degree in cybernetics and control engineering from Donghua University in 2006. Currently, he is a PhD candidate in the Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University. His research interests include image processing and artificial intelligence.

Jie Yang received a PhD in computer science from the University of Hamburg, Germany. He is now a professor and vice-director of the Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University. He has performed more than 20 national and ministry scientific research projects in image processing, pattern recognition, data amalgamation, data mining, and artificial intelligence.

Kun Liu is working for his MS degree in the Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University. His current research interests include computer vision and human behavior analysis.

Xiang Zhang is a doctoral candidate in the Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University. His research interests include image processing and computer vision.