Procedure for Detecting Rain and Snow Particles on Digital Television Images

Vladimir Dmitriev [1] and Andrew Baukov [1]

[1] Ryazan State Radio Engineering University named after V.F. Utkin, Gagarina Street, 59/1, Ryazan, 390005, Russia

Abstract
The relevance of the task of detecting and eliminating rain and snow particles in television images taken with static cameras is shown. The procedure for detecting rain and snow particles on video images, which is supposed to be used as part of an algorithm for reducing the visibility of precipitation, has been developed. The approach developed is presented in the form of three-stage classification of frame pixels into zones with moving objects and areas of stationary background, distorted and undistorted by precipitation particles. The procedure proposed has been studied and optimal values of selected accumulated frames for its correct operation have been determined: 100 frames for video images with rain; and 140 frames for video with snow. The gain of the approach developed in comparison with known by levels of errors of the first and second kind is 0.8 % and 3.4...9.1 %. The described procedure can be used in TV systems where the delay of 3...5 seconds relative to video source is not critical.

Keywords
television image, precipitation detection, distribution bimodality coefficient, sample size, classification.

1. Introduction

Television images taken with static cameras can be obtained under various weather conditions. For example, footage of live reports from the scene, sport TV broadcasts and other video sequences often contain "traces" of certain weather phenomena, such as rain, snow, hail. Falling particles of such atmospheric precipitation cause the effect of dynamic noise on video images. Such interference, especially during high-intensity precipitation, often makes it difficult to see and distinguish certain objects in a video, reduce observation range and thus prevent viewers from correct perceiving of what is happening on the screen.

In order to improve the quality of video images with this type of interference, special algorithms to eliminate the visibility of precipitation particles have been developed [1–5]. As a rule, such algorithms contain the stage of detecting precipitation particles, at which pixel classification procedure takes place. The correlation of each pixel with a certain group (pixel class of stationary background, pixel class of moving objects, pixel class of precipitation particles) allows us to avoid unwanted distortion of useful information of a frame and allow only pixels of raindrops or snowflakes to be subsequently processed.

One of the common methods underlying the described classification of pixels is threshold comparison of successive frames [4, 5]. The disadvantage of this approach may be the difficulty of choosing a global (i.e., the same for entire frame) threshold value that determines the difference between the intensity of pixels covered by precipitation particles and the intensity of undamaged pixels. Thus, the intensity of raindrop or snowflake edge pixels can be higher than background pixel intensity by only 5–8 units (in 8-bit monochrome image), which often coincides with temporal fluctuations in the intensity of background points associated with small changes in illumination or fluctuations, e.g., leaves on trees.

GraphiCon 2022: 32nd International Conference on Computer Graphics and Vision, September 19-22, 2022, Ryazan State Radio Engineering University named after V.F. Utkin, Ryazan, Russia
EMAIL: vol77@rambler.ru (V. Dmitriev); baukov.andrej@yandex.ru (A. Baukov)
ORCID: 0000-0001-5521-6886 (V. Dmitriev); 0000-0002-0003-058X (A. Baukov)

Time clustering method presented in [2] provides for the possibility of automatically determining local threshold values and demonstrates good results when processing video sequences with intense precipitation throughout a frame. However, in this case, it does not take into account the fact that some significant areas of frame pixels (or individual pixels) may not be exposed to precipitation at all for a long time, approximately several seconds (comparable to the time necessary for the algorithm to learn how to automatically determine threshold) or more. This can lead to incorrect choice of threshold value for a large number of pixels, and, consequently, to the loss of useful information or to undesirable effect of "ripples" for viewers or observers in the areas of a frame that do not require elimination of precipitation particle visibility.

In summary, the above facts lead to the conclusion that the task of developing a procedure to detect the particles of intense precipitation in television images taken by static cameras is relevant. The procedure developed should provide the ability to automatically determine passage or non-passage of precipitation in certain areas of a frame (ideally, in each specific pixel), as well as the ability to automatically select local threshold values.

2. Description of the procedure developed

At the first stage of pixel classification, it is advisable to select the areas of a frame containing a significant number of moving objects other than background, and then to detect rain and snow particles in such areas using a special algorithm [5]. The pixels belonging to these areas of a frame can be selected over a larger (compared to the pixels affected only by drops) range D of change in time over n frames of intensity I in such pixels, and then combined together with the surroundings into certain areas, forming a binary image $G_1(x,y)$. The first (preliminary) classification stage described may be performed once per n frame. Further, the processing of selected areas corresponding to logical values 1 in $G_1(x,y)$ is performed in accordance with the algorithm [5] and pixels of other zones are classified according to the developed approach.

In order to develop further steps of classification procedure, experimental studies of video sequences with a frequency of 25 frames/s with fixed background, taken in intense rain and snow, in which pixel brightness changes are caused only by precipitation, as well as small differences in illumination and camera noise, were carried out. In the course of these studies, time distributions $p(I)$ in intensity values of the pixels exposed to precipitation particles were obtained. The examples of such distributions for three different pixels are shown in Figure 1 (a, b, c).

From the analysis of Figure 1(a, b, c) it can be seen that temporal distributions of the pixels studied appear to be asymmetric bimodal or asymmetric unimodal. The large mode in bimodal distribution and the only mode in unimodal one correspond to intensity values of a pixel at the time moments in which it is not distorted by precipitation, since the "hit" of a particle into a pixel leads to increase in its intensity [4]. Accordingly, a smaller mode in bimodal distribution and a flat part in unimodal distribution correspond to pixel values in the frames in which it is covered by a rain particle. Positive skewness of such distributions is characteristic for vast majority of video sequences exposed to rain or snow, since the frames in which precipitation occurs in a pixel are usually smaller than the frames in which a pixel is clean. For the image points that are not covered by precipitation particles for several seconds, distributions $p(I)$ of pixel intensity values in most cases are symmetrical unimodal and are similar in shape to a normal distribution (Figure 1(d)).

Thus, the criterion that a pixel is exposed to precipitation for a certain period of time, or, on the contrary, this point belongs to an undistorted area of a frame, can be the form of distribution of pixel intensity values over time. The article [6] describes the Sarlet coefficient which is used to determine bimodality degree of random variable distribution. For a finite sample of size n, the Sarlet coefficient is calculated using formula:

$$s = \frac{A^2+1}{E+\dfrac{3(n-1)^2}{(n-2)(n-3)}}, \tag{1}$$

where A is a sample skewness, E is a sample kurtosis [6]. If the value is $s < 5/9 \approx 0.555$, then the distribution is considered unimodal, if $s > 0.555$, then it is bimodal or multimodal, and if $s \approx 0.555$, then it is uniform [6].

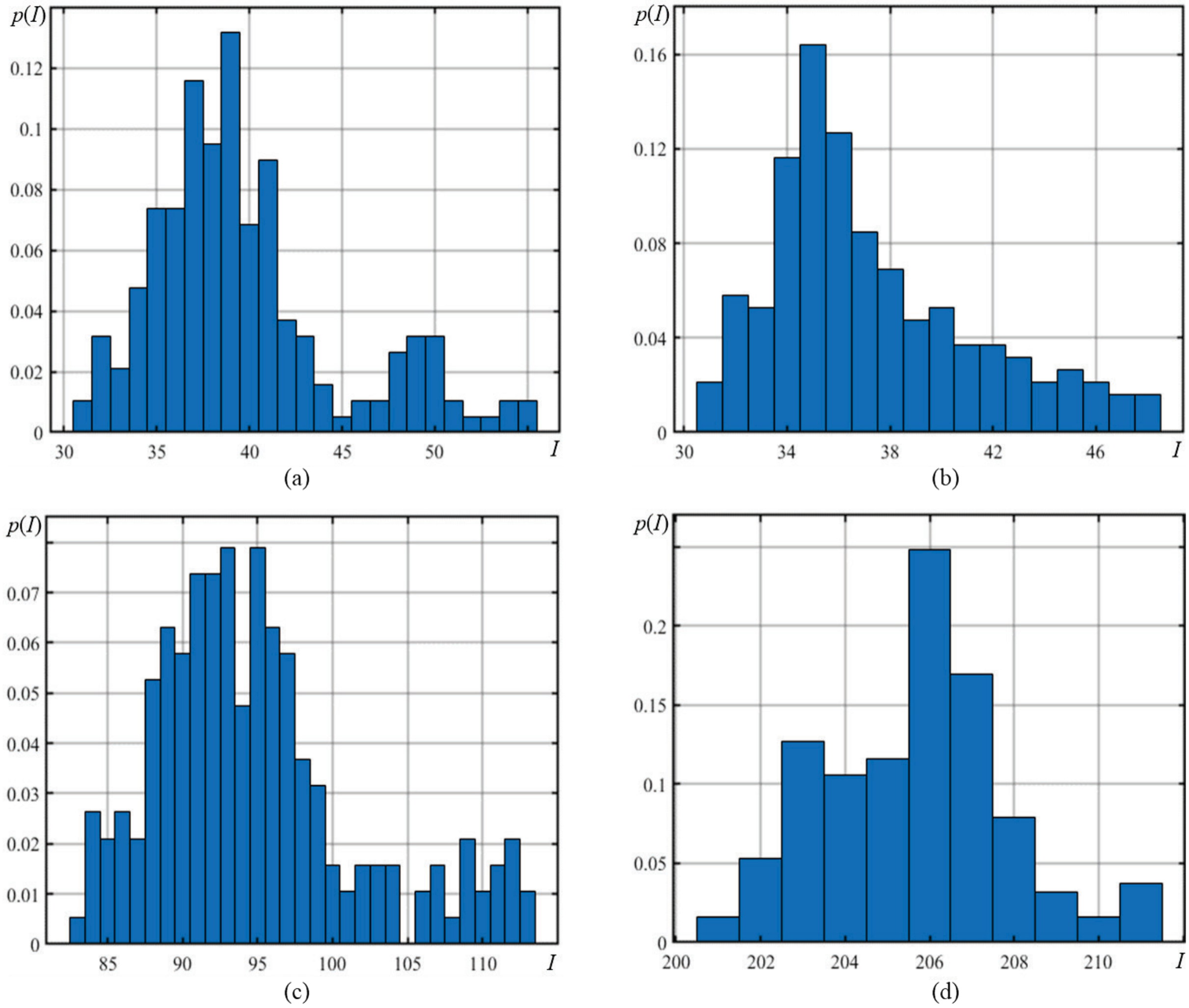

Figure 1: Histograms of distributions $p(I)$ in time of intensity values I of a pixel taken from: (a, b) frame area with rain; (c) areas of frame with snow; (d) area of the frame not distorted by precipitation

Also the article [6] indicates that for unimodal distributions with a large asymmetry coefficient, the Sarlet coefficient also takes on values $s > 0.555$. This fallacy of the criterion is beneficial from the viewpoint of developing an algorithm for detecting precipitation particles, since the distributions in intensity values of the pixels exposed to rain or snow described above can be not only bimodal, but also highly asymmetric unimodal.

In order to apply the Sarlet coefficient in the procedure developed, its modification was performed. First, the sign of A is taken into account in a modified coefficient, since only positive asymmetry is typical for intensity distribution of the distorted pixel. Secondly, range D of pixel intensity changes during the observation of n frames is taken into account. The introduction of correction factor $D/8$ made it possible to increase the accuracy of pixel classification at boundary value $s \approx 0.5...0.6$ (8 is an empirically obtained average value of minimum pixel intensity difference that is noticeable to a viewer when a particle falls).

Thus, the Sarlet coefficient modified to solve the problem of classifying pixels (for drops or snowflakes per n frame being present in them), has the form:

$$s_m = \operatorname{sign}(A) sD/8 . \tag{2}$$

For values $s_m > 0.555$, the decision is made that a particular pixel is located in the region distorted by precipitation (i.e., particles fall into this pixel at least several times per n frame) and it should be subjected to subsequent research and processing, and for $s_m < 0.555$ we decide that the pixel does not belong to distorted area of a frame (perhaps there is no precipitation at all in a frame) and it is not advisable to process it.

Thus, the criterion proposed forms the second stage of pixel classification, as a result of which a binary image is created according to the rule:

$$G_2(x,y) = \begin{cases} 1, & \left(s_m(x,y) > 0.555\right) \wedge \left(\neg G_1(x,y)\right), \\ 0, & \left(s_m(x,y) < 0.555\right) \vee G_1(x,y), \end{cases} \tag{3}$$

where $\wedge$, $\vee$, $\neg$ are logical operators "AND", "OR", "NOT". The second stage, like the first one, is executed once per n frame.

Next, the pixels belonging to the distorted areas of video image (i.e. $G_2(x,y)=1$) are checked for presence or absence of precipitation particles in each current frame, and thus are subjected to the third stage of classification. At this stage, the authors propose to use thresholding while local threshold values are determined automatically for each pixel selected at the previous stage. Based on the analysis of intensity distributions in distorted and undistorted pixels, the following values are proposed to be used as local threshold values:

$$c(x,y) = 2\hat{I}(x,y) - \min\left(I(x,y)\right), \tag{4}$$

where $\min\left(I(x,y)\right)$ is minimum intensity value taken by (x,y)-th pixel as n frames, $\hat{I}(x,y)$ is distribution $p(I)$ mode for (x,y)-th pixel, which was obtained at the second stage of classification. The result of thresholding and the third stage of pixel classification of k-th frame is a binary image:

$$G_3(x,y,k) = \begin{cases} 1, & \left(I(x,y,k) \geq c(x,y)\right) \wedge G_2(x,y), \\ 0, & \left(I(x,y,k) < c(x,y)\right) \vee G_2(x,y). \end{cases} \tag{5}$$

Thus, the procedure proposed for detecting precipitation particles in video images is a three-stage classification of pixels (the first and second stages are executed once per n frame, and the third - in each frame):

1. Obtaining in accordance with $D(x,y)$ a binary image $G_1(x,y)$ that classifies pixels of video image frames into areas with moving objects, subsequently processed by algorithm [5] ($G_1(x,y)=1$), and areas without such objects, considered at subsequent stages of this classification ($G_1(x,y)=0$).
2. Calculation of the modified Sarlet coefficient $s_m(x,y)$ by formula (2) for pixels, for which $G_1(x,y)=0$, and obtaining a binary image $G_2(x,y)$ according to rule (3) there is no manifestation of precipitation ($G_2(x,y)=1$), i.e., on distorted and undistorted areas.
3. Calculation of local threshold values $c(x,y)$ by formula (4) and obtaining a binary image $G_3(x,y)$ according to formula (5), which classifies the pixels of k-th frame into those that are currently covered by raindrop or snowflake and require enhancement (for which $G_3(x,y)=1$), and the pixels that are not in a current frame affected by precipitation and require no treatment $G_3(x,y)=0$.

According to [7], sample size strongly affects the ability to correctly estimate shape and form of the desired distribution. Therefore, the most important parameter of the procedure proposed is the number of accumulated n frames, according to which, using the modified Sarlet coefficient, the shape of distribution $p(I)$ of pixel intensity values at the second classification stage is estimated, since if it has too small or too large n, the probability of erroneous skipping of rain pixels by an algorithm or unnecessary processing of undistorted points, or incorrect definition of local thresholds increases. Therefore, experimental studies in order to select the optimal value of the number of n frames were carried out.

3. Experimental study

Main quality indicators (QI) of detection algorithms are normalized error levels of the first K_1 and second K_2 kind [8]. Since incorrect choice of parameter n of the algorithm under study can lead to errors both at second and third stages of classification, QIs K_1 and K_2 are calculated separately for the second (K_{12} and K_{22}) and third (K_{13} and K_{23}) steps.

As a result of experimental studies it was found that the choice of the number of n accumulated frames which minimizes the levels of errors of the first kind, is simultaneously accompanied by an increase in the levels of errors of the second kind, and vice versa. In such cases, a multicriterion approach is often used [9], which is based on minimizing or maximizing objective function J represented as a weighted sum of quality criteria. To more significantly reduce the number of errors of the first kind, i.e., erroneous correction of pixels that are not affected by precipitation, we choose weight coefficients of criteria K_{12} and K_{13} slightly exceeding coefficients K_{22} and K_{23}. Then, to solve the problem, objective function is:

$$J = 0.3K_{12} + 0.2K_{22} + 0.3K_{13} + 0.2K_{23}. \quad (6)$$

Separate studies for video sequences with rain and snow were carried out. In the course of experimental study, 10000 pixel sequences (10000 pixels "with a duration" of n frames each, i.e., $10000 \times n$ array) belonging to distorted frame areas of various video images, and 10000 pixel sequences belonging to the areas with no precipitation during frames were used. Then, 500 pixels affected and unaffected by precipitation particles in current frames were used for calculation. A smaller data sample size in the study of the third stage of classification is explained by the complexity of pixels preliminary selection in a "visual" way. Specified QIs and objective function values were calculated using the number of frames, varying from 40 to 210. The video images used in these studies were recorded at a frequency of 25…30 frames/s.

According to the results of experimental studies, the optimal values for the number of accumulated frames used to calculate the modified Sarlet coefficient and determine local threshold values are 100 frames for video sequences with rain and 140 frames for video images with snow. The difference in the obtained optimal values can be explained by the fact that rain drop falling rate is higher compared to a snow particle, therefore, in video images with rain, a pixel is distorted more often at the same time, and the shape of distribution $p(I)$ is estimated more accurately by the modified Sarlet coefficient. Hence, the probability of pixel correct classification for the videos with rain is higher than for snow videos.

It has been established that, at optimal values for the number of accumulated frames n, the approach developed makes it possible to detect pixels of rain regions on video images with rain with a normalized level of errors: first kind $K_{12} = 1.92$ %, second kind $K_{22} = 1.27$ %. For the videos shot in snow conditions, these indicators are $K_{12} = 3.85$ % and $K_{22} = 1.24$ %. Further, the procedure makes it possible to detect pixels of rain regions affected by a drop in current frame, with error levels: first kind $K_{13} = 0.8$ %, second kind $K_{23} = 1.8$ %. With precipitation in the form of snow: $K_{13} = 2.6$ %, $K_{23} = 8.4$ %.

Then, using total probability formula [10], it becomes possible to calculate total error levels of the first and second kind for detecting pixels of precipitation particles, i.e., when it is not known which region a pixel belongs to - distorted or undistorted by precipitation:

$$\begin{aligned} K_1 &= K_{12} + \left(1 - K_{12}\right) K_{13}, \\ K_2 &= K_{22} + \left(1 - K_{22}\right) K_{23}. \end{aligned} \quad (7)$$

Results calculated by these formulas: $K_1 = 2.7$ % and $K_2 = 3.05$ % for video sequences with rain; $K_1 = 6.35$ % and $K_2 = 9.54$ % for video with snow. Also, a comparison between synthesized and known algorithms to detect precipitation particles on video images was made. The results of the comparison are presented in Table 1. The QI K_2 gain of the developed algorithm compared to known approaches [2-5] is 3.4…9.1%. At the same time, according to the criterion K_1, the approach proposed is comparable with algorithm [5] and outperforms methods [2-4] up to 0.8%.

Table 1
Algorithm comparison results

Algorithm	K_1, %	K_2, %
Algorithm in [2]	3.49	7.53
Algorithm in [3]	2.8	10.31
Algorithm in [4]	3.27	12.15
Algorithm in [5]	2.64	6.45
Developed algorithm	2.7	3.05

4. Conclusion

A procedure to detect precipitation particles in video images which differs from many known methods by the possibility of automatic selection of local threshold values as well as a three-stage pixel classification, has been developed. At the first stage, the areas of the frame containing moving objects, e.g., roads, pedestrian zones, etc., which are supposed to be processed using the algorithm described in [5] are selected. The second stage is designed to detect areas where there is no rainfall, using intensity distributions of each pixel formed as a result of frames accumulation. At the third stage of classification, according to the obtained distributions, automatic determination of local threshold values is performed in order to detect pixels affected by drops or snowflakes in a current frame. According to the results of experimental study, optimal values of main parameter of the procedure developed, the number of accumulated frames, were established: 100 frames for video images with rain; and 140 frames for video with snow. The values of error levels of the first and second kind in the detection of precipitation particles by the proposed procedure are calculated and its advantage in comparison with known approaches before 0.8 % and on 3.4…9.1 % according to these quality criteria is confirmed. The developed procedure assumes the delay of 3...5 seconds relative to video source and can be used in TV systems where this delay is not critical.

5. References

[1] R. Gonzalez and R. Woods, Cifrovaja obrabotka izobrazhenij [Digital Image Processing], Moscow: Tehnosfera, p. 1104, 2012. (in Russian).

[2] Z. Jia, H. Wang, R. E. Caballero, Z. Xiong, J. Zhao and A. Finn, "A two-step approach to see-through bad weather for surveillance video quality enhancement", Machine Vision and Applications, vol. 23, no. 6, pp. 1059-1082, 2012.

[3] N. Brewer and N. Liu, "Using the Shape Characteristics of Rain to Identify and Remove Rain from Video", Structural, Syntactic, and Statistical Pattern Recognition, pp. 451-458, 2008.

[4] K. Garg and S. K. Nayar, "Vision and rain", International Journal of Computer Vision, vol. 75, no. 1, pp. 3-27, 2007.

[5] S. N. Kirillov, A. Lisnichuk, A. V. Batishchev, L. V. Aronov and A. A. Baukov, "Adaptive Software-Configurable System Development to Receive, Transmit and Process Information", 2020 International Scientific and Technical Conference Modern Computer Network Technologies (MoNeTeC), pp. 1-6, 2020.

[6] R. Pfister, K. A. Schwarz, M. Janczyk, R. Dale and J. Freeman, "Good things peak in pairs: a note on the bimodality coefficient", Frontiers in psychology, vol. 4, 2013.

[7] R. N. Vadzinskij, Spravochnik po verojatnostnym raspredelenijam [Handbook of Probability Distributions], Saint-Petersburg: Nauka [Science], p. 295, 2001. (in Russian).

[8] A. A. Baukov, "Razrabotka reshajushhego pravila obnaruzhenija chastic dozhdja v celjah uluchshenija kachestva videoizobrazhenij [Development of a decisive rule for detecting rain particles in order to improve the quality of video images]", Sovremennye tehnologii v nauke i obrazovanii-STNO-2020 [Modern technologies in science and education-2020], vol. 1, pp. 157-163, 2020. (in Russian).

[9] A. A. Lisnichuk and A. V. Batishhev, "Dvuhkriterial'nyj sintez OFDM-signalov dlja povyshenija jenergeticheskoj jeffektivnosti i pomehoustojchivosti [Two-criterial synthesis of OFDM signals to

increase energy efficiency and noise immunity]", Vestnik Rjazanskogo gosudarstvennogo radiotehnicheskogo universiteta [Vestnik of Ryazan State Radio Engineering University], no. 76, pp. 3-16, 2021. iIn Russian).

[10] V. E. Gmurman, Teorija verojatnostej i matematicheskaja statistika [Probability Theory and Mathematical Statistics], Moscow: Vysshaja shkola [Higher School], p. 479, 2003. (in Russian).