AN AUTOMATIC REAL-TIME HD DEFOGGING SYSTEM

J. P. Oakley[1] and Z. Marceta[2]

[1]Dmist Research Ltd. and the University of Manchester [2]RT-RK Computer Based Systems LLC, Serbia

ABSTRACT

This paper describes the automatic mitigation of airlight noise (defogging) in a new in-line image processing system. Solutions to two key technical challenges are presented. The first challenge is the design of an algorithm to produce ‘maps’ of airlight in RGB space using information derived from sample frames. The second challenge is to process the HD stream on a pixe-by-pixel basis with low latency by appropriate subtraction and re-scaling. Although the video processing is relatively simple, it is necessarily performed in RGB space and so colour conversions are required to translate from and to the YUV representation used in transmission, leading to a significant computational requirement. This requirement is met by an asynchronous dual-processor architecture that allows sample frames to be downloaded for airlight analysis with concurrent high-speed pixel processing. Test results show effective enhancement of degraded images with no distortion of clear images and no requirement for the user to adjust settings for different conditions. The latency for 1080i/50 streams is 71μs.

INTRODUCTION

Significant loss of image quality can arise in adverse atmospheric conditions such as rain, drizzle, smoke and fog. This is due to light scattering from particles between the camera and the subject, generating what is often called “airlight”.

Under clear conditions, the only light entering a camera is that directly reflected by objects in the field of view. If haze, fog, drizzle, rain or light smoke is present then some of the light originating from the primary light source (normally the sun, but could also be an artificial light source) is scattered so that it enters the camera. This is known as the “airlight” and the effect is illustrated in figure 1. The resultant image, I(x,y), produced by the camera is essentially a sum of two components: the scene component S(x,y) and an airlight component A(x,y).

$$I(x,y) = S(x,y) + A(x,y) \quad \ldots(1)$$

The intensity of the airlight is a function of the size and composition of scattering particles, the concentration of particles, the distance between subject and camera and the angle of illumination.

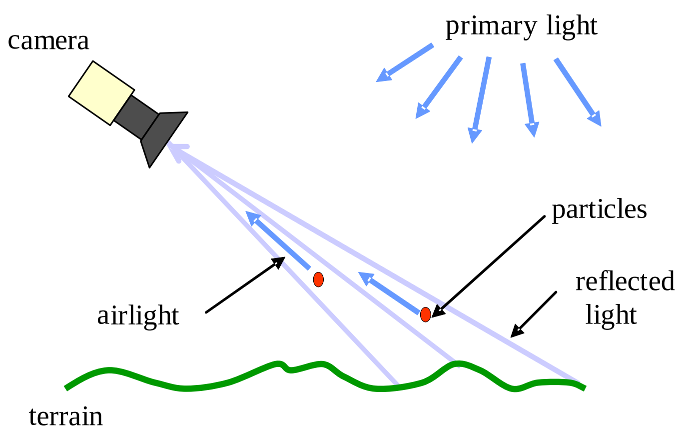


Figure 1 –Effect of haze

All these parameters are subject to change – some over the image area, and some over time. The amount of scattering is also dependent on the wavelength of the light (e.g., sometimes greater for blue than red wavelengths). The scene component S(x,y) is also attenuated by the atmosphere; a process known as extinction. The combined effect of these scattering phenomena is to reduce image contrast.

Airlight degradation is an important problem in outside broadcast, particularly with HD services as viewers pay a premium for high picture quality. The best current solution is manual adjustment of black level at the video editing desk using “Proc Amp” controls. However the colour can easily be distorted and different parts of the image can vary in brightness. It is also a demanding task when the camera is tracking a moving subject.

Figure 2 -Left part enhanced

This paper describes the automatic mitigation of airlight noise (defogging) in the new in-line image processing system called ClearVue. An example is shown in figure 2. Only the left-hand part of the image is processed to allow comparison.

Two key technical challenges addressed in the design of ClearVue are presented. The first challenge is the design of a reliable algorithm to produce ‘maps’ of airlight in RGB space using information derived from sample frames. The second challenge is to process the HD stream with low latency by appropriate subtraction and re-scaling. In this paper the theoretical background for the image enhancement is described. An overview of the implementation via an asynchronous dual-processor architecture is then given. Some illustrative results are presented followed by discussion.

BACKGROUND

Previous Approaches

Mitigation of this type of atmospheric degradation can be effected in various ways. The best-known image enhancement tools are based on histogram equalisation. Most of these programs will provide some improvement in image quality when applied to atmosphere-degraded images. However the time required for the enhancement computations introduces a delay, known as latency, between input and output. Latency is an important issue in outside broadcast. Also the previous enhancement algorithms distort clear images. More recently, specialised algorithms have been reported to mitigate atmospheric degradation (1-5). Such algorithms are *idempotent* in the sense that they correct a specific defect in the image. If no atmospheric degradation is present then they will introduce no changes in the image.

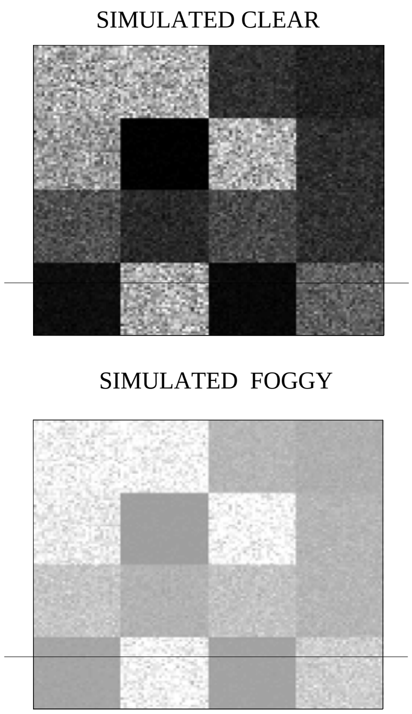
The basic idea is to invert equation (1) to recover S(x,y) from I(x,y). If the airlight distribution A(x,y) can be estimated by some means then the image may be recovered by simply subtracting A(x,y) from I(x,y), followed by appropriate rescaling.

The reported algorithms differ in how the airlight component A(x,y) is estimated. In (1) the fact that the airlight varies with range is exploited using nonlinear regression to produce an estimate for A(x,y). In cases where the range does not vary significantly across the image this approach runs into difficulties. Narasimhan and Nayar (2,4) describe a method for producing airlight estimates on the assumption that the range (and hence the airlight) is piecewise constant in the images. Again this method runs into difficulties in applications where the assumption is not valid. Oakley and Bu (3) describe a more general method based on minimising a cost function. This latter technique can work in the widest range of conditions and so is preferred here. An outline of this method is given below.

Cost Function Approach

The Coefficient of Variation (CV) is defined as the ratio of the local standard deviation of the pixel intensities to the local mean.

The assumption in the Oakley-Bu method is that the statistics of a clear image are, to a first approximation, stationary. An image typically contains some dark objects and some brighter regions. The assumption is that the CVs in the bright and dark regions should be similar. This is reasonable for natural scenes since differences in illumination generate image regions with different lightness but similar CVs. In foggy conditions the CVs differ considerably for light and dark regions. This is illustrated by the simulation shown in figure 3. Two synthetic images are shown; the first represents a clear image with a constant CV. An estimate for the CV is calculated from the equation displayed in which p_k is the value of the image at pixel position k and $\overline{p_k}$ is the output of a spatial low-pass filter at pixel position k. The second image is transformed using equation (1) to represent the foggy case. Plots of the CV for one selected line are shown for both clear and foggy cases. It can be seen that the clear image has a relatively uniform CV, although it is subject to statistical fluctuation. The foggy image shows greater variation in the CV, with darker regions showing lower values. This is a fundamental difference that can be detected by appropriate statistical analysis and this is the basis of the Oakley-Bu method.



$$CV \simeq \left| \frac{p_k - \overline{p_k}}{\overline{p_k}} \right|$$

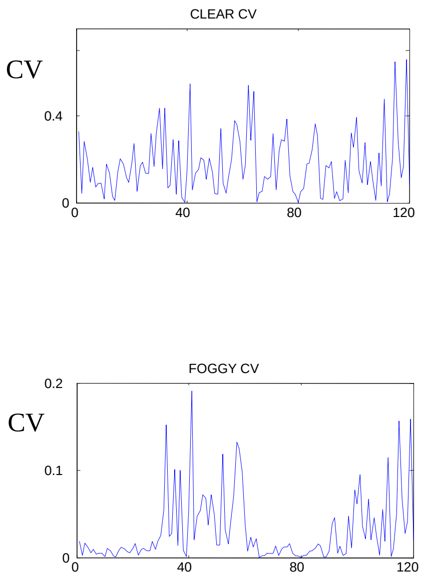


Figure 3. CV Variation in Clear and Foggy Images

The Oakley-Bu cost function is:

$$S(A)=\frac{1}{k}\sum_{k=0}^{k<K}\left(\frac{p_k-\overline{p_k}}{\overline{\overline{p_k}}-A_k}\right)^2\cdot\exp\frac{1}{k}\sum_{k=0}^{k<K}\left(\ln\left(\overline{p_k}-A_k\right)^2\right) \qquad \ldots(2)$$

The airlight values $\{A_k\}$ are chosen to minimise the value of this function. Even with a good estimate for $\{A_k\}$ the CV will still vary significantly in different parts of the images. However the spread of values of the CV is minimised in a specific sense[1]. A can be represented as a parametric function, as in (1), in a global minimisation or as a smooth non-parametric function, in which case some kind of iterative local minimisation is required. The latter approach is used in the ClearVue system in order to give the greatest possible flexibility in application. The airlight estimation algorithm is coded in C++ and implemented on a conventional IA32 processor.

IMPLEMENTATION

Once airlight has been estimated the required enhancement computation is a pixel-by-pixel subtraction and scaling. The level of airlight in general varies with wavelength and hence is different for the red, green and blue channels. For this reason the processing is performed in RGB colour space. If the input pixel is (x_r, x_g, x_b) and the output pixel (y_r, y_g, y_b) , then the required transformation is:

$$\begin{aligned} y_r &= m_r\,(x_r - A_r) \\ y_g &= m_g(x_g - A_g) \\ y_b &= m_b\,(x_b - A_b) \end{aligned} \qquad \ldots(3)$$

where m_r, m_g, m_b A_r, A_g and A_b represent scaling and airlight (offset) parameters. The required gain and offset parameters may vary for different parts of the image since the extent of the degradation will depend on range. Since the actual video processing is very simple, i.e. subtraction and scaling according to equation (3), it is advantageous to separate the relatively complicated statistical analysis algorithm from the video processing pipeline. In this way such enhancement can be applied to a high-definition video stream in real-time (6) whilst achieving low latency (in the order of microseconds).

Figure 4. ClearVue HD Product

[1] It can be shown that the Oakley-Bu cost function is equivalent to minimising the Theil index T0, a well-known metric for variability used in the analysis of economic inequality.

For medium-volume applications the video pipeline could be implemented either using a Field Programmable Gate Array (FPGA) or a Digital Signal Processor (DSP). The DSP route was chosen for ClearVue, mainly on the grounds of lower production cost. The DSP device selected is the DM642 from Texas Instruments. The assembly is mounted in a 1U enclosure as shown in figure 4 above. The processing architecture is shown in figure 5 below. The two dotted boxes show functionality implemented on a bespoke DSP printed circuit board and functionality implemented in software on a general-purpose IA32 board. The two boards are linked via a PCI connector.

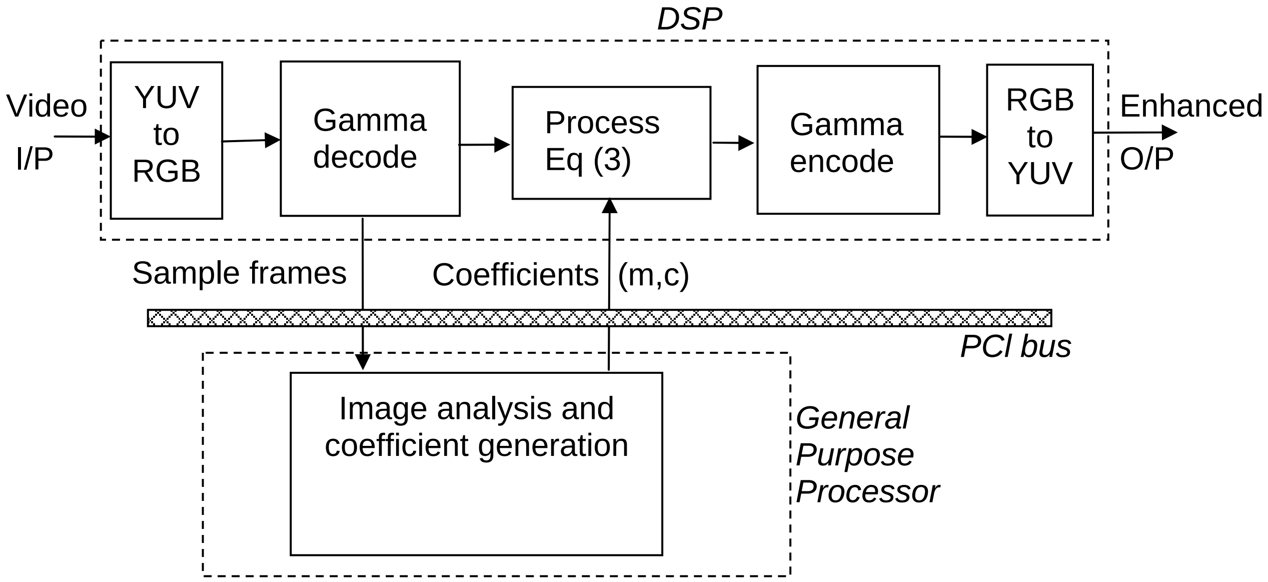


Figure 5. Implementation architecture

Although the pixel processing is simple, for HD streams the computational requirement is such that carefully optimised DSP code is required. As the processing must be carried out in RGB space, colour conversions, both from and to YUV colour space, are required. Conventional video sources are gamma-encoded and this non-linear transformation must also be reversed prior to processing and re-applied after processing. The central task of the video processor is to implement the transformation specified by equation (3) using stored values of m_r, m_g, m_b A_r, A_g and A_b. These enhancement coefficients are held in high-speed memory within the DSP system. In principle the video process can be achieved on a pixel-by-pixel basis. In practice, in a DSP implementation, it is advantageous to process up to four lines of the image at a time. This increases the latency but the values are considered acceptable for outside broadcast application. Table 1 shows the latency values for different video formats. The HD values are lower because the coding of the DSP program so highly optimised for these cases. An FPGA-based video processor could be used to provide lower values of latency if required.

Operation

In operation the image analysis and video processing run as asynchronous tasks, communicating via a PCI bus. The image analysis process signals the video process when it is ready to analyse an image. An image is then

INPUT	RESOLUTION	RATE (Hz)	LATENCY (μs)
SDTV	720 x 576i	50	256
	720 x 480i	60	256
HDTV	1280 x 720p	50/60	53.33/44.44
	1920 x 1080i	50/60	71.11/59.26

Table 1. Latency of Enhancement Process

transferred without affecting the ongoing video process. When the analysis is complete, new enhancement coefficients are sent to the video processor and loaded during the blanking interval. The image analysis task then requests a new frame, and so on. Although the video process operated at full video rates (50Hz for PAL and 60 Hz for NTSC), only a subset of frames, typically one in four, are used for airlight analysis. The reason for this is that the pattern of atmospheric degradation changes relatively slowly.

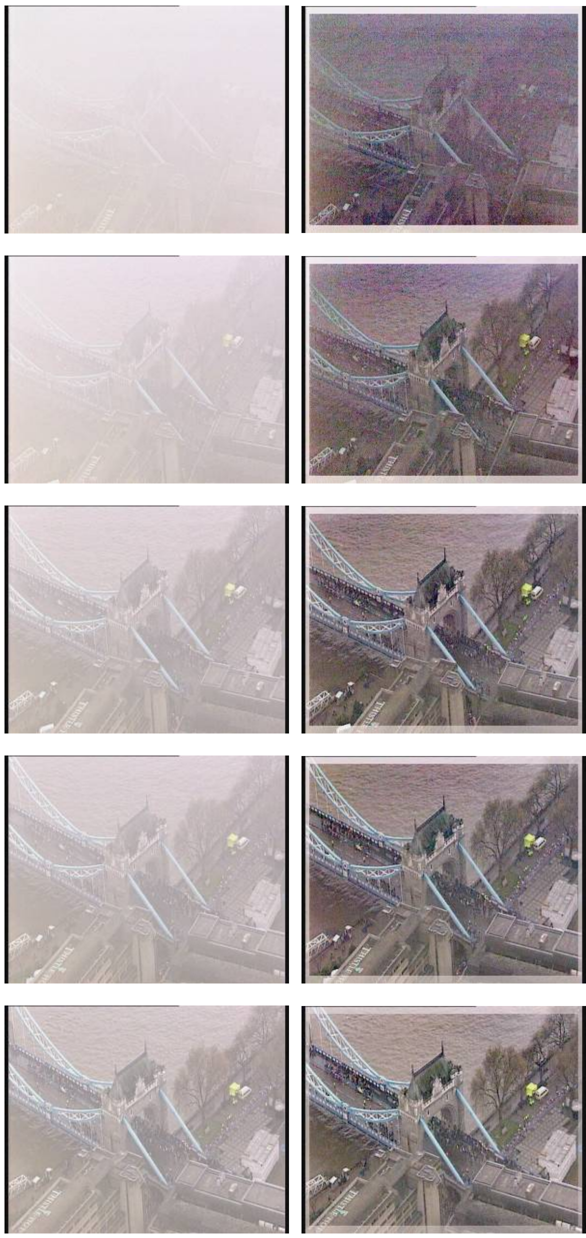
RESULTS

The main testing methodology used with Clearvue is the processing of many hours of archive footage of different subject matter acquired under a wide variety of atmospheric conditions, followed by painstaking subjective analysis. This shows consistently high output quality with no visible distortion. Testing with live camera feeds is also used. The aim of this testing is to establish:

1. **Safety**. If there is no adverse atmospheric condition, the image should ideally not be changed at all. At worse any change should not adversely affect subjective image quality and
2. **Effectiveness**. When adverse conditions are present the enhancement process image should be fully effective.

When the visibility is very poor the scaling effect of the transformation described by equation (3) increases any noise present in the image. This puts a fundamental limitation on the enhancement process since some noise is always present. The two main sources of noise are sensor noise and particle noise. Sensor noise arises mainly from shot noise caused by the discrete nature of the light detection process. Sensor noise is always present. Particle noise is caused by relatively large particles close to the lens and the level of particle noise varies greatly according to the atmospheric conditions. An extreme example would be snow. In general the ClearVue process is most effective in moderate visibility conditions where the scaling effect introduced by the enhancement does not raise noise levels to unacceptably high levels.

In some situations the atmospheric conditions can change quickly and corresponding sections of archive footage are particularly



1 2 3 4 5

Figure 6 –Sample results

(left is original and right is enhanced) useful in testing. Figure 6 shows image frames extracted at 0.5s intervals under conditions of light rain which improve rapidly. The contrast for each of the five unprocessed images, defined as $(I_{max} - I_{min})/I_{mean}$, where I is image intensity or lightness, ranges from around 0.3 in image 1 to around 1.0 in image 6. The contrast for each of the unprocessed images is shown by the front bars in figure 7. The processed images are shown to the right in figure 5 and the corresponding contrast is shown by the rear bars in figure 7. The processed contrast is relatively stable at around 2.2. Although the subjective quality of images 1 and 2 is improved by enhancement, the enhanced images show a high degree of noise and would not be suitable for production purposes. Images 3-5, with an unprocessed contrast ratio of between 0.6 and 1.0, represent situations in which the enhancement renders the video stream usable for production. Without some kind of processing this stream could not be used. The subjective effect of this processing is that the atmospheric problem is not noticeable by the viewer. More examples of ClearVue processing can be found in (7).

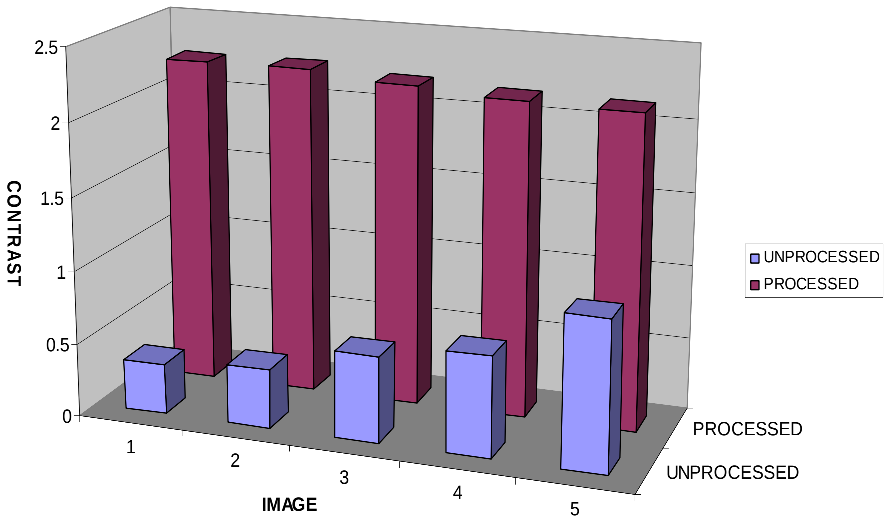


Figure 7. Image Contrast

DISCUSSION

ClearVue is the first commercially-available defogging system specifically designed for outside broadcast applications. It is effective in processing images in moderately poor visibility and restoring correct contrast and colour. The ClearVue process is completely automatic and there are no parameters to set. The system can be regarded as a kind of “fog filter”; when no airlight is present the video stream is not altered. Viewers are generally unaware that any processing of the video has taken place unless side-by-side presentation of unprocessed video is offered.

For very poor visibility the system will improve the clarity but the output will not, in general, meet retransmission standard. The main limiting factor is the noise present in the image. In general this is due to a combination of sensor noise (mainly shot noise) and noise introduced by particles close to the sensor.

Awareness and acceptance by the OB community will be an important milestone for ClearVue. ClearVue is currently available as a standalone add-on unit but the technology could potentially be incorporated within camera systems. The best approach will be to offer the algorithm as a set of integrated circuits and significant investment will be required to achieve this.

REFERENCES

1. J P Oakley and B Satherley, 1998, Improving image quality in poor visibility conditions using a physical model for contrast degradation, IEEE Trans. on Image Processing, 7 (2), pp 167-179.

2. S G Narasimhan and S K Nayar, 2003, Contrast restoration of weather degraded images, IEEE Trans. on Pattern Analysis and Machine Intelligence, 25 (6), pp. 713 – 724.

3. J P Oakley and H Bu, 2007, Correction of simple contrast loss in color images, IEEE Trans. on Image Processing, 16 (2), pp 511-522.

4. S K Nayar and S G Narasimhan, 1999, Vision in bad weather, Proc. of IEEE Int. Conference on Computer Vision, vol. 2, pp 820-827.

5. K K Tan and J P Oakley, 2001, A Physics-based approach to color image enhancement in poor visibility conditions, J. Opt. Soc. Am. A, 18 (10), pp 2460-2467.

6. M J Robinson, D W Armitage and J P Oakley, 2002, Seeing in the mist: real time video enhancement, Sensor Review, 22 (2), pp 157-161.

7. Dmist Research Ltd.,Manchester, UK.. http://www.dmist.com