PTZ Visibility Detection Based on Image Luminance Changing Tendency

Zhaozheng Chen
Jiangxi Expressway Networking Mangement Center
Honggutan New Area Hongjiaozhou district Wolong Road No.1 402Room
Jiangxi Nanchang, China
e-mail: zaogege@163.com

Abstract—**Due to the fact that the existing shortage of visibility detection methods, a novel algorithm is presented to measure atmospheric visibility base on image luminance changing tendency using PTZ (Pan/Tilt/Zoom) video cameras. First, the background of image was used to remove the interference from cars and passengers, and the region of interest was figured out to ensure the height constancy of the selected pixels. Then the homogeneous road surface, without roadbed or lane marks, to ensure the luminance constancy of the selected pixels. From the selected road region, combined with camera calibration algorithm, the luminance curve which reflect that road luminance vary with the distance, was extracted to search the feature point, and then the extinction coefficient is calculated, finally, the maximum visibility distance was obtained. The experimental results show that the algorithm is consistent with observation characteristic of human eyes, and the accuracy is up to 96% while the inspection error within 10 meters. Compared with other visibility detection methods, this algorithm is more effective, robust and precise.**

Keywords-visibility detection; height constancy; luminance constancy; luminance feature pixels

I. Introduction

Road visibility has great influence on the safety of transportation, especially on expressway. However, traditional visibility meter has lots of limitations on fog and haze detection: insufficient sample space, sensitive to dust and smog. Besides, it's too expensive to dispose with high density, which makes very difficult to reflect the aero conditions of the entire road.

Considering that visibility sensing is a psychophysics phenomenon of human visual system, which can be simulated by analyzing images from roadside camera, video-based visibility detection have great advantages than traditional method. It has advantage of low cost, easy operation, and full coverage. But the key problem is to figure out the relationship between degrading of image and traffic visibility.

T. M. Kwon [1] used multiple large targets in front of camera, the size, color, position of these targets are carefully designed, the degrade of each target in image is used to measuring visibility, however, this method costs a lot because each roadside camera should equip 6~8 large targets, and the camera can’t move because it is quite hard to locate multiple specific target from a complex image scene.

Robert G. Hallowell [2] proposed a method that comparing the images obtained from the video cameras with the images stored when weather is quite good, then the relative visibility can be calculated without auxiliary equipment. However, it is not compatible with the existing PTZ (Pan/ Tilt/Zoom) cameras, and quite sensitive to the moving objects in scene.

AerotechTelub and Dalarma University [3] has introduced low noise infrared cameras to deal with visibility. The edges of image are first figure out, then using neural network to classify the visibility level. But it is too expensive with complicated maintenance.

Chen [4] proposed a visibility detection algorithm using the wavelet transformation, first the wavelet transform is used to get the edge information of predetermined areas on road, and then a model of edge variation trends of these areas is established, finally with curve fitting to calculate visibility. This method is quite sensitive to the noise on road for curve fitting algorithm is used, and also the PTZ camera on road can’t be used. Another algorithm based on video-image contrast has been presented by Li [5]. By comparing the contrast of every pixel with its 4-neighbors, distinguishable pixels can be confirmed when the max contrast is greater than the preset threshold. Then, combined with camera calibration technology, the visibility value can be obtained. With camera calibration, this algorithm complies with PTZ camera. But the distinguishable pixels are obtained by a fixed threshold, which are easily disturbed by noise, including CCD imaging noise, quantization noise.

In this thesis, a novel algorithm base on luminance characteristic is presented. Begin with the luminance variations feature detection of the entire image, to find out the road region to ensure eight consistency and luminance consistency, then the luminance variation trend caused by atmospheric extinction is analyzed, to find out feature points of luminance curve, combined with camera calibration technology, the visibility is calculated. By means of analyzing the brightness variation trend, image noise is mostly removed. The experimental results show that the algorithm has good stability and accuracy.

II. PTZ Camera Calibration

A. *Traffic Camera Visual Model*

The traffic camera model is shown in figure 1, which drew from Li’s model [5] and made some improvement.

There are four coordinate systems: the first is the road surface coordinate system described by Xw－Yw－Zw with its origin at the point where the optic axis of the camera intersects with the road surface. Assuming that the road surface is flat, so Zw=0 for every point on road. The second is the camera coordinate system described by Xc－Yc－Zc with its origin Oc at the optic center of the lens, the axis Zc coincident with the optic axis, and the plane Xc－Yc parallel with the CCD image plane. The third is the image coordinate system described by u-v.

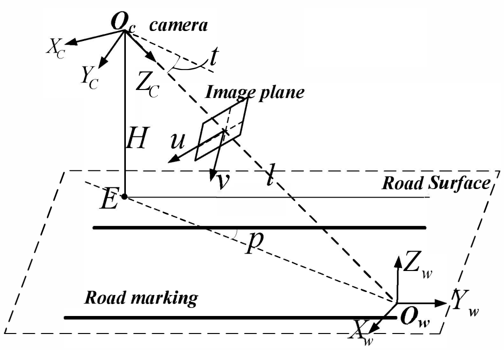


Figure 1. The visual model of the traffic camera.

The point E is the projection of camera coordinate system origin Oc on road, which makes line section Oc-E is perpendicular to the road surface with length H, approximately the height of the camera. The length of the line section Oc-Ow is described by l. The pitch angle is t between the optic axis and the road surface. The yaw angle is p between the road marking line and the projection line of the optic axis on the road surface.

$$u = \frac{-f\,[\cos(p)\,X_w - \sin(p)\,Y_w + \sin(t)\sin(p)\,X_w + \sin(t)\cos(p)Y_w]}{(\cos(t)\sin(p)\,X_w + \cos(t)\cos(p)\,Y_w + l)}$$
$$v = \frac{-f[\cos(p)\,X_w - \sin(p)\,Y_w - \sin(t)\sin(p)\,X_w - \sin(t)\cos(p)Y_w]}{(\cos(t)\sin(p)\,X_w + \cos(t)\cos(p)\,Y_w + l)} \tag{1}$$

$$X_w = \frac{l[-u\cos(p)\ \sin(t) + v\ \sin(p)]}{-v\ \cos(t) + f\ \sin(t)}$$
$$Y_w = \frac{l\,[u\sin(p)\sin(t) + v\cos(p)\,]}{-v\ \cos(t) + f\ \sin(t)} \tag{2}$$

B. Camera Model Solution

According to perspective projection, the images of parallel lines intersect at the horizon, as shown in figure 2.（u0, v0）are so-called vanishing points.

Back to the eq. (2), assuming at the horizon line where Yw=+∞, the simple relationship between focus f and pitch angle t can write in eq.(3).

$$f = \frac{v_0}{\tan(t)} \tag{3}$$

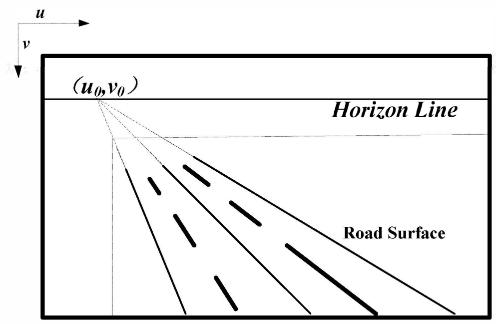


Figure 2. Images of parallel lines.

And the eq.(3) can rewrite as flows, presenting the relationship between a point with coordinate v in image to vanishing point v0.

$$X_w = \frac{l[-u\cos(p)\ \sin(t) + v\ \sin(p)]}{(v_0 - v)\cos(t)}$$
$$Y_w = \frac{l\,[u\sin(p)\sin(t) + v\cos(p)\,]}{(v_0 - v)\cos(t)} \tag{4}$$

Cause the height of the camera is usually known, the distance of two coordinate system l can written in eq.(5).

$$l = \frac{H}{\sin(t)} \tag{5}$$

According to the parallel and corresponding relationship among the corner points, where Yw(A) =Yw(B),Xw (A)=Xw(D). Integrating (3) and (5) the parameters p and t can be solved.

$$t = \arcsin\left\{\frac{-v_0^{\,2}(-v_D + v_A)(v_A - v_B)}{[((u_A - u_D)v_0 - u_A v_D + u_D v_A) *((u_A - u_B)\,v_0 - u_A v_B + u_B v_A)]}\right\}^{1/2}$$
$$p = \arctan\left\{\frac{\sin(t)(u_A v_0 - u_A v_D - u_D v_0 + u_D v_A)}{v_0(-v_D + v_A)}\right\} \tag{6}$$

Thus far, we have got all the four parameters in the model. The two expressions (2) and (6) can be used directly to transform coordinates from road surface to image plane or the inverse. Notice that the origin of the road surface coordinate system is located at the center of the image which is an uncertain region on the road. Eq. (9) can be used to transfer the origin to the basement of the camera pole so that the position of an object on the road with reference to the camera and the distance d can be calculated.

$$d = \sqrt{X_{wd}^{\,2} + y_{wd}^{\,2}}$$
$$X_{wd} = X_w + l\cos(t)\sin(|\,p\,|)$$
$$Y_{wd} = Y_w + l\cos(t)\cos(|\,p\,|) \tag{7}$$

Because all interested points are on road, which have little difference on Xw coordinate, and the main part of distance is Yw, and in most case, the direction of camera is the same as road direction, which makes the yaw angle p≈0. Alliance eq.(7), the distance of road area represent by image point v to camera can be write as follows.

$$d = \begin{cases} \alpha \dfrac{v}{(v - v_h)} + \beta & v > v_h \\ \infty & v \le v_h \end{cases}$$
$$\alpha = \frac{l\cos(p)}{\cos(t)}, \tag{8}$$
$$\beta = l\cos(t) * \cos(|p|)$$

III. VISIBILITY DETECTION MODEL

After ROI is set and camera calibration are get, a visility detection model is set up based on Monocular Camera Calibration, and the detail of model is introduced as following.

A. Relationship between Visibility and the Extinction Coefficient of Atmosphere

Koschmieder [6] proposed that light is attenuated against background sky when it goes through atmospheric layer. A simple relationship between the distance d of an object with intrinsic luminance L0 and its apparent luminance L was established as follows:

Suppose Vmet is the farthest distance that human eyes can see, namely, as discussed before, atmospheric visibility is the farthest pixel to camera whose contrast is 0.05.

$$L = L_0 e^{-kd} + L_f(1 - e^{-kd})$$
$$V_{met} = -\frac{1}{k}\ln\left(\frac{C_d}{C_0}\right) = -\frac{1}{k}\ln(0.05) \approx \frac{3}{k} \tag{9}$$

Eq.(9) shows the relationship between atmospheric visibility and extinction coefficient, and then figure out the atmosphere visibility, object in scene with common luminance should first figure out, with camera calibration, the luminance change tendency varies pixel distance is calculated.

B. Luminance Inflection Point

From Eq.(9), both the luminance of object in image varies with the distance to camera, and values of the extinction coefficient k. When the density of fog increases, which means the extinction coefficient k goes up, the object tends to get obscured by the luminance emanating from the sky.

By virtue of both Eq.(9), and taking the derivative of L with respect to v, the luminance tendency in image varies to the position v of pixel can be written in Eq.(10).

$$\frac{dL}{dv} = k\frac{\alpha v_h (L_0 - L_f)}{(v - v_h)^2} e^{-k(\alpha\frac{v}{(v - v_h)} + \beta)}$$
$$\frac{d^2L}{dv^2} = k\frac{\alpha v_h (L_0 - L_f)}{(v - v_h)^3} e^{-k(\alpha\frac{v}{(v - v_h)} + \beta)}\left(\frac{k\alpha v_h}{(v - v_h)} - 2\right) \tag{10}$$

If make the Eq.(10) to be zero, which means to find out the luminance inflection point in image, one solution is k=0 which of no interest. Therefore the other useful solution is given in Eq.(11):

$$k = \frac{2(v_i - v_h)}{\alpha v_h} \tag{11}$$

where vi denotes the position of the inflection point and vh is the position of the vanishing point. Thus, the atmospheric visibility can be deduced as follows:

$$V_{met} \approx \frac{3}{k} = \frac{3\alpha v_h}{2(v_i - v_h)} \tag{12}$$

When vi is approximate the same as vh, Vmet is in the critical state, just the presence of fog or no fog appears. If vi is greater than vh, fog can be detected.

IV. LIP BASED LUMINANCE CONSISTENCY

From the Eq.(13), the luminance of pixels lies in adjacent lines which is from the homogeneous road surface are quite close, and the homogeneous road surface can be separated out from image by scanning from the bottom to horizon line, the neighbor pixel whose luminance difference lies in a reasonable region is regarded as homogeneous area.

Considering that visibility sensing is a psychophysics phenomenon of human visual system, the difference of image pixels can mapping into a gray tone range, by using the logarithmic image processing (LIP) model as developed by Jourlin and Pinoh [7], provided a robust mathematical framework with highly desirable properties, such as, the result of the operations is bounded to a specific range.

This framework follows many laws of the human vision system, both physical and psychophysical (among others, Fechner's law, Weber's law, psychophysical contrast notion, brightness scale inversion, etc.). According to [8], LIP has been demonstrated to be a very appropriated model for many different tasks, because it is at the same time mathematically well-justified, physically consistent, psychophysically coherent with higher primates visual system, and computationally affordable.

The set of gray tone functions defined on the spatial support S and valued in the bounded real number interval [0, M], the Criteria for basic operations of two gray tone functions f(x, y) and g(x, y) are defined as follows:

$$\begin{aligned} &\forall f, g \in S && f \oplus g = f + g - \frac{fg}{M} \\ &\forall f, g \in S && f \ominus g = M\frac{f - g}{M - g} \\ &\forall f \in S, \forall \lambda \in R && \lambda \otimes f = M - M\left(1 - \frac{f}{M}\right)^{\lambda} \\ &\forall f \in S && \ominus f = \frac{-Mf}{M - f} \end{aligned} \tag{13}$$

Considering that the pixels Global luminance balance is evaluated by computing the difference of any given pixel at line l with gray level P(i， j)and the start line median luminance value Pg，such as Eq.(14).

$$P_l(i,j)\ominus P_g = M\frac{P_l(i,j)-P_g}{M-P_g} \le \rho n_r \min G_{\max}^{k} \qquad (k=-1,0,1) \tag{14}$$

where ρ is a positive value between [0, 1], nr designates the line number between Pl（i， j） and the initial seed point Pg. Gmaxk designates the constraining thresholds between one pixel and its 3 neighbor pixels on top, as shown in Fig.3.

$$G_{\max}^{-1} = G_{\max}^{1} < G_{\max}^{0} \tag{15}$$

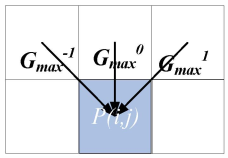


Figure 3. One pixel and its top 3 neighbor pixels.

Global luminance balance avoids that pixels stray too far from the seed line. Suppose the level of gray in the image is range from 0 to 255, if the vertical gradient tolerance is equal to 8, black and white pixels could then be in the target road region in fewer than 32 iterations without this condition.

The pixels Q(i, j) which meets the global luminance balance will be evaluated through image Neighborhood luminance model to keep the luminance of neighborhood pixels are close.

$$\exists m \in \{i-1, i, i+1\},$$
$$Q(i,j)\ominus Q(m,j+1) = M\frac{Q(i,j)-Q(m,j+1)}{M-Q(m,j+1)} < G_{\max}^{i-m} \tag{16}$$

This constraint can enhance noise robustness, and also effectively prevent the level of gray to be discontinuous. The pixels both meeting the luminance balance and the consistency, can be taken as homogenous road surface area.

V. VISIBILITY CALCULATION BASED ON LUMINANCE FEATURE PIXELS

Using the parameters from camera calibration, the relationship reflecting that luminance varies with distance, and the relationship between the inflection point and the extinction coefficient, as diagrammed in Fig.4, the visibility can be calculated as following.

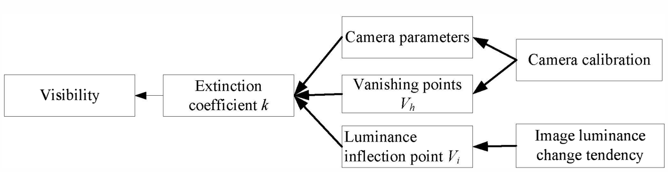


Figure 4. Visibility calculation based on luminance feature pixels.

Because of the road marks and green belts, the pixels in one line are always discontinuous in the obtained road region. So it may be erroneous to get the median luminance directly. To solve this problem, the maximum continuous point sets Pix（j）of the jth line of the obtained region is found out, and take the middle of Pix(j) as the center point of the measurement belt with maximum width LEN to keep the pixel number in a reasonable range to reduce computational cost. The measurement belt is gained with the condition of Eq.(11), and Fig.4 shows the schematic of measurement belt with middle points, and the median luminance of the road surface measurement belt is calculated by utilizing Eq.(17).

$$len(Pix(j)) = \min(LEN, length(j)) \tag{17}$$

$$L_j = median(Pix(j)) \tag{18}$$

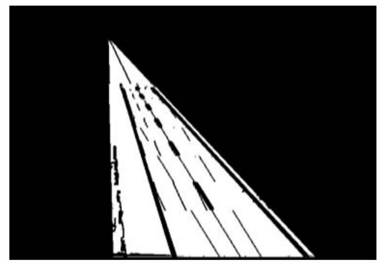


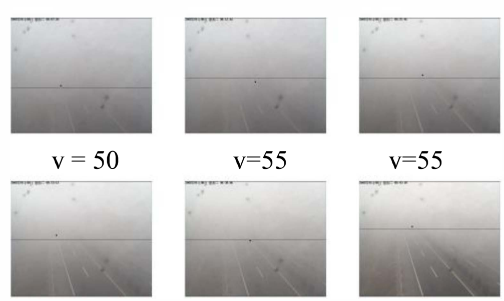
Figure 5. Schematic of road surface measurement belt.

By virtue of Eq.(12), the brightness-distance curve L can be get. For the second derivative of L to determine the mutation point vi, calculate extinction coefficient combined with camera calibration, and then get the maximum visibility distance. Brightness L gradually changes with the distance. However, because the brightness value is the discrete integer value between 0~255, the phenomenon of the same brightness values between the adjacent rows is easy to see. Or due to the impact of the noise points, a lot of confused second derivative zero exists. Therefore, to avoid the false detection, before we seek for the mutation point, interpolating and filtering are used to eliminate the confused second derivative zero.

VI. RESULT AND CONCLUSION

The visibility detection system is tested on the platform, which has Ubuntu Linux OS. Videos encoded to H.264 with the resolution 704×576.

The images, as indicated in Fig. 6, show the variation of visibility every 30 minutes. Within the images, "•" denotes the visibility by the contrast algorithm, and the line on the image is the visibility by the algorithm of this article. Below images, "v" denotes the visibility by the algorithm of this article.



v = 50 v=55 v=55

v = 60 v = 60 v = 90

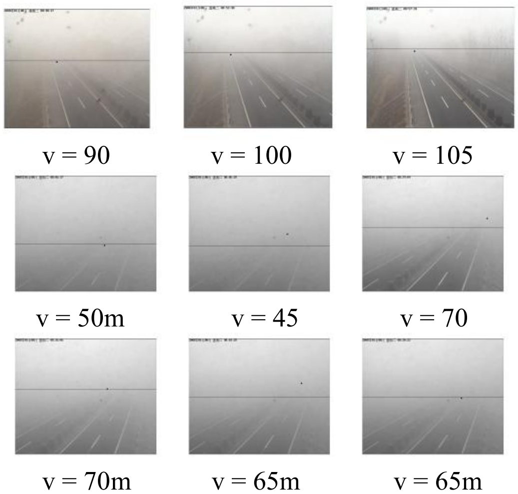


Figure 6. Results of visibility detection.

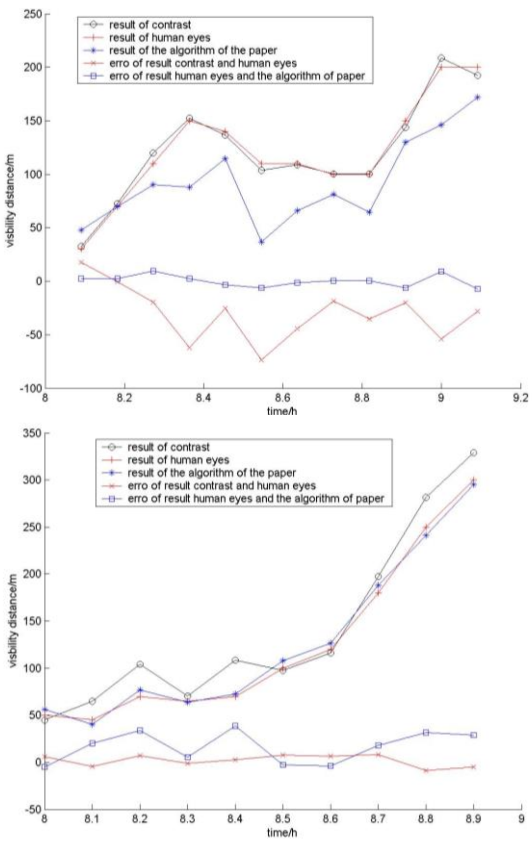


Figure 7. Comparison of results about contrast, the paper's algorithm and human eyes.

The detection results by different algorithms are shown in Fig. 7, including algorithm introduced in this article, the algorithm based on contrast, and detecting by human eyes.

The visibility detected by this algorithm get much closer to the top of the image along with the improvement of weather conditions; in majority case, it goes well with the observed value by human eyes, at the same time, it is in line with the algorithm result based on contrast. Of course it exist slight fluctuation, but the limit of error is permissible. Nevertheless, the result from the algorithm based on contrast can easily interfered, especially when it has loud noise or dense fog. So the algorithm in the article can get the visibility much better.

1)It can fully use the existing PTZ video cameras without setting any artificial markers, so it is of low cost;

This detection algorithm is highly stable and accurately. And it can adapt to the images better in every weather condition.

REFERENCE

[1] T. M. Kwon. Atmospheric visibility measurements using video cameras: Relative visibility. Technical report[R], America: University of Minnesota Duluth, July 2004.

[2] Robert G. Hallowell, Michael P. Matthews, and Paul A. Pisano, Automated Extraction of Weather Variables from Camera Imagery[C]. Proceedings of the 2005 Mid-Continent Transportation Research Symposium, Ames, IA, 2005.8

[3] Final Report on Signal and Image Processing for Road Condition Classfication[R], AerotechTelub and Dalarma University under the7 Swedish National Road Agency. Report #2002-02-06. 2002,6.

[4] Chen Zhaozheng, Zhou Qingkui, Chen Qimei. Video visibility detection algorithm based on wavelet transformation [J]. Chinese Journal of Scientific Instrument 2010,31(1):92~98

[5] Li Bo, Dong Rong, Chen Qimei. Visibility Detection Based on Traffic Video Contrast Analysis without Artificial Markers [J]. JOURNAL OF COMPU TER2AIDED DESIGN & COMPU TER GRAPHICS，2009,11(21):1575-1982

[6] Nicolas Hautiere, Raphael Labayrade, Didier Aubert, Real-Time Disparity Contrast Combination for Onboard Estimation of the Visibility Distance[J], in IEEE Transactions On Intelligent Transportation Systems, 2006,(2): 201-212

[7] Jourlin M, Pinoli JC. Logarithmic image processing: The mathematical and physical framework for the representation and processing of transmitted images. Advances in Imaging and Electron Physics,2001 115:129-94.

[8] J. C. Pinoli, A general comparative study of the multiplicative homomorphic, log-ratio and logarithmic image processing approaches, Signal Process., vol. 58, no. 1, pp. 11-45, Apr. 1997.