A MONTAGE METHOD: THE OVERLAYING OF THE COMPUTER GENERATED IMAGES ONTO A BACKGROUND PHOTOGRAPH

Eihachiro Nakamae, Koichi Harada and Takao Ishizaki
Hiroshima University
Saijo, Higashi-Hiroshima
Hiroshima 724, Japan

Tomoyuki Nishita
Fukuyama University
Higashimura, Fukuyama
Hiroshima 729-02, Japan

ABSTRACT

A system of computer programs has been established to generate high quality montage image of considerable usefulness in architectural simulation which combine computer-generated images and photographed background pictures.

Traditionally, there are two methods of creating architectural montages: (1) an artist paints new buildings onto a background scene usually generated photographically, and (2) a three-dimensional scale model is created to simulate the whole landscape, and this model is then photographed. The montage method described here combines aspects of both traditional montage methods with significant improvement in accuracy and reduction of time and cost of preparation. Specifically, a digitized photograph is used as a background scene onto which is superimposed a 3D computer-generated image of a new building. The outstanding points of the new method are:

(i) The shading and shadows of each computer generated image are calculated with higher accuracy, (ii) the fog effect is taken into account, and (iii) a new anti-aliasing technique improves the quality of the final montage image.

CR Categories and Subject Descriptors: I.3.3[Computer Graphics]:Picture/Image Generation -- display algorithm; I.3.5[Computer Graphics]: Computational Geometry and Object Modelling -- curve, surface, solid and object representations; I.3.7[Computer Graphics]: Three-Dimensional Graphics and Realithm -- color, shading, shadowing and texture; I.4.3[Image Processing]: Enhancement -- sharpening and deblurring; I.4.6[Image Processing]: Segmentation -- region growing, partitioning; I.4.8[Image Processing]: Scene Analysis -- depth cues; J.4 [Social and Behavioral Sciences]: Sociology

General Terms: Algorithms, Measurement

1. Introduction

In order to assess the potential impact of a new construction on local environments, it is very important to understand what is going to happen to viewing angles and local lighting/shadows if new buildings are constructed. Assessment grows more important and more difficult as the number and scale of the new buildings increases.

Traditionally, two methods have been used for attacking this problem: (1) A landscape where new construction occurs is photographed, and an artist paints the new building(s) onto this picture. (2) A 3D scale model is made to gauge the influence of the new buildings on their surroundings.

The major difficulty of the second method is that it is expensive even though it offers three-dimensional investigation. And it should be noted that every part of the model is artificial. Though the first method is less costly, the montage is totally affected by the artist's personal preference.

Modern computer graphics techniques offer a viable solution to this problem. In 1979, Uno et al.[1] proposed a general purpose graphics system that was used to create simple image compositions. In 1980, Feibush et al.[2] overlaid an imaginary house onto a background using a complex database composed of many polygons and textures. In 1984, Porter et al.[3] presented an interesting approach (Alpha channel method) to the digitized composition problem that allows processing of complicated three-dimensional images.

Image compositions in these earlier studies have been investigated chiefly from the viewpoint of anti-aliasing. The montage system in this paper includes the following unique characteristics along with the anti-aliasing processing:

(1) The ratio between the maximum illuminance due to the sun beam [4] and the ambient illuminance of each computer generated image described in an early paper by the authors is calculated by using information contained in the background picture. This enables superior simulation of natural shading and shadows.

(2) Atmospheric moisture effects (fogginess) is added to simulate various weather conditions. In the new method, the fog effects are processed by using exponential functions[5]. These functions are determined with a sophisticated algorithm based on the illuminance information

contained in the background picture. This makes the montage more reliable and applicable for wider range of environment impact assessment.

2. Overlaying of Computer-Generated Images onto a Background Photograph

2.1 Schematic Description of Proposed Method

The method consists of the following five steps:

(1) The data for computer-generated building images are prepared.

(2) The land where the buildings will be constructed is photographed. This picture is processed with an A/D converter; geometric transformation (translation, rotation, etc.) are carried out to fit the photographic image on a CRT screen.

(3) The computer-generated images and the photographed background scene are matched from the viewpoint of geometry and adjusted by calculating the control parameters (the coordinates of the camera position, the view angle, the ambient illuminance and the fogginess).

(4) The computer generated images are overlaid on the background scene and displayed on the CRT. Hidden-surface removal, anti-aliasing and fog calculations are implemented at this stage.

(5) The foreground scene in front of the superimposed computer-generated image is re-displayed.

Further explanation of Steps (3), (4) and (5) is given below.

2.2 Geometric Matching Between Computer-Generated Images and Background Scene

The view reference point (a point existing on the object at the center of the photograph) and the viewpoint (the camera location) must be designated to carry out projection mapping. The object point is easily known by tracing the position of this point on a geophysical map.

However, the viewpoint is not easily determined if the camera is set on an airplane for instance. The least square method (see Appendix) may be used to estimate the viewpoint comparing several points on the photograph with their actual position coordinates on a map.

The size of the perspective plane is aligned with that of the photograph by selecting approximate horizontal and vertical angles; these angles are calculated based on the focal length of the camera lens and the size of the photograph. They can also be determined by tracing the objects existing at the edge of the photograph on a map. The photographs are assumed to have been taken with a standard (55mm) or telephoto lens to avoid the distortion caused by wide angle lenses.

2.3 Shading

The position of the sun must be determined from the photograph's time and date as well as the longitude and latitude of the camera position, in order to correctly calculate the shading of buildings and the shadows they cast in the simulated montage image. Let the angle between the x-axis (that defines the computer-generated image) and the southern direction be β and the hour angle be τ (τ=0 deg. corresponds to the noon and τ =15 deg. indicates one hour later). Furthermore, let ϕ and δ denote the longitude and the solar declination, respectively. The direction toward the sun is determined by:

$$\begin{bmatrix} x_\tau \\ y_\tau \\ z_\tau \end{bmatrix} = \begin{bmatrix} \cos\beta & -\sin\beta & 0 \\ \sin\beta & \cos\beta & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} -\cos\delta\sin\phi & 0 & \sin\delta\cos\phi \\ 0 & \cos\delta & 0 \\ \cos\delta\cos\phi & 0 & \sin\delta\sin\phi \end{bmatrix} \begin{bmatrix} \cos\tau \\ \sin\tau \\ 1 \end{bmatrix} \quad (1)$$

Once the direction of the sun is determined, shadows and shading are calculated by traditional methods. The shading on each object is described by:

$$E = I\cos\alpha + E_h \quad (2)$$

where I, cos α and E_h stand for the intensity of sun light, the cosine of the angle between the surface normal and the direction vector toward the sun and the ambient illuminance, respectively. For simplicity, specular reflection is neglected.

For montage generation, the ratio between E_h and I, called H, is important because the ultimate aim is to match computer-generated images to the background scene. First, H is calculated for the background scene by using the method explained below. Then, the computer generated images are designed so that their H values are equal to that of the background image.

Two walls of a building in the background scene are chosen (one is preferred to face the shadow side), and their intensity level is detected (this quantity is proportional to the illuminance and the reflectance). H is then calculated as:

$$H = (D_2\cos\theta_1 - D_1\cos\theta_2)/(D_1 - D_2), \quad \cos\theta_i\,(i=1,2) \gtrsim 0 \quad (3)$$

where D_1, D_2 are illuminance on each wall. θ_1 and θ_2 denote the angle between the direction vector toward the sun and the surface normal of each wall surface, respectively.

If there is no such building, a pilot box (a white box is useful for this purpose) may be set in the vicinity of the camera to serve as a reference for ambient illuminance.

Since the ambient illuminance is affected by the surrounding objects, H varies depending on the object position. H should be calculated for several points and the average value of these calculation results is used.

2.4 Fog Effects

In general an object far from the viewpoint seems to lose its color and tone because of atomospheric moisture effects. If these effects are not taken into account, the new buildings, computer-generated images, do not match the photographed background pictures.

The reduction rate of color and tone depends on the various factors such as seasons, weather conditions and time. This fact means that the informations of the reduction rate should be obtained from the background picture, even though it is already known that the color and tone of each object changes exponentially with respect to its distance from the viewpoint [5].

A new method has been established to determine this exponential function based on the background picture: The color and tone become gray as the distance becomes large. Let red-, green- or blue-component of the illuminance at a point in the picture be C_e (e=r,g,b, respectively). If the point moves from the viewpoint to the infinite point, the apparent color C'_e of it approaches exponentially to a color which depends on the background picture. This color is called the standard color F_e. The apparent color C'_e is therefore given by:

$$C'_e = (1-\exp(-x/R_e))(F_e - C_e) + C_e, \qquad e = r,g,b \tag{4}$$

R_e (standard distance) and F_e are determined as follows: At least, three sampled points are designated on the photograph; more sampled points help refine the final result. These points should be selected on objects which are considered to share a color (trees, for example) but stand at different distances from the viewpoint. The illuminance on these points, which is defined as the average value in the vicinity of each point, is detected on the photograph. The distance to each of the points should be obtained by using the geophysical map. Data points are plotted on a illuminance-distance plane (see Fig.1). The least mean square measure is used to give the most appropriate exponential function from these data.

A simpler method is to use a pair of objects. These objects are assumed to possess different colors and locate at a near position (x=0 in Eq.(4)) and at the distance x. Let A'_e and B'_e be the colors of the pair of objects at the distance x. Then, these colors are given with their original colors A_e and B_e (the color at a near position of each object) as:

$$\begin{aligned} A'_e &= K_e(F_e - A_e) + A_e, \\ B'_e &= K_e(F_e - B_e) + B_e, \\ K_e &= 1 - \exp(-x/R_e), \\ e &= r,g,b \end{aligned} \tag{5}$$

By solving a simple system of equations, the standard distance R_e (which dominates the change of each color with respect to the distance) and the standard color F_e is determined by:

$$\begin{aligned} F_e &= (A_e B'_e - B_e A'_e)/(A_e - A'_e - B_e + B'_e), \\ R_e &= -x/\log_e[(F_e - A'_e)/(F_e - A_e)] \end{aligned} \tag{6}$$

The fogginess of an image heavily depends on the weather conditions when the photographs are taken. Therefore, the requirement to add or even remove the fog in the photographed scenes might occur in order to observe the montage image under the different conditions.

Strictly speaking, a background scene should be processed in a three-dimensional space to know the exact fog effect. Since this is almost impossible, the next best approximation is useful: Each object on the background scene is described with a combination of planes. The distance between the objects and the viewpoint is calculated by using this approximated image. In the case of Fig.2, mountains (A through D) and a grove of trees (E) are approximated with vertical planes. F and G are processed with horizontal planes. The road (H) and two banks of earth (I and J) are treated as convex polyhedra.

A cursor or a tablet is used to interactively divide the background scene into several parts, and each part is checked whether it should be approximated with vertical planes, horizontal planes or polyhedra. The distance from the viewpoint is also specified at this step.

Because each part consists of planes, Eq.(4) may be applied to calculate the fog effects at each point. It should be noted that the longer the standard distance, the smaller the fog effect, and vice versa.

2.5 Anti-Aliasing

It is well known that the diagonal edges of objects displayed on a raster-scan type CRT are jagged. In some cases, moire patterns appear; e.g. the same shaped objects are set periodically. These undesireable phenomena called 'aliasing' make the computer-generated images unnatural against the background scene.

The anti-aliasing for montage methods should be different from other usual ones because: (i) the background scene includes pixels which in general change pixel by pixel, and (ii) image overlay operation is performed several times (overlay process initiates from the further scene, and ends at the closest scene, or foreground).

This first characteristic ((i) above) excludes the use of filters that work based on more than one pixel. If these filters were used, the resulting image would be very vague. This second characteristic ((ii) above) increases calculation cost, so a simpler filter is preferred over more complicated filters, such as the conical filter or Gaussian filter[2].

The Alpha channel method[3] is useful for anti-aliasing when making extremely complex montages requiring many overlay loggers. However, anti-aliasing is only really required along the edges of each computer-generated image. Hence, for simple montages such as discussed here, the following easy anti-aliasing method is suggested:

The color at each pixel on the CRT screen is first read out. The color of the background scene is determined pixel by pixel based on this information. Then, a modified two-dimensional Fourier window (box window) is activated. The detailed explanation is:

A pixel is approximated with a unit square shown in Fig. 3. This square is divided into n parts. S_0 is an actual scan line. Each of other lines within the pixel is called a virtual scan line (VSL). The color of this pixel is determined according to the product of the length of the image on each VSL S_ℓ (ℓ=1,2,...,n) and the illuminance of this image. Each product is added with a proper weight to decide the pixel color: Let m objects exist within the pixel (i,j). The background scene color at this pixel is supposed to be $B_{i,j}$ (this vector consists of three components that correspond to red, green and blue). Furthermore, let $L_{f,\ell}$ and $b_{i,j,\ell}$ stand for the visible interval on the VSL of the object O_f and the background scene, respectively. The color $C_{i,j}$ of the pixel (i,j) is calculated according to the trapezoidal integration rule as:

$$C_{i,j} = \sum_{\ell=1}^{n} d_\ell (b_{i,j,\ell} B_{i,j} + \sum_{f=1}^{m} L_{f,\ell} C_f), \qquad (7)$$
$$d_\ell = \begin{cases} 1/n; & 0 < \ell < n \\ 1/2n; & \ell=0,n \end{cases}$$

where d_ℓ indicates the weight factor. C_f is the color of object O_f.

2.6 Removal of Computer-Generated Images Hidden by Foregrounds

A basic problem of this overlay method is that the inserted computer-generated images might hide the scene which exists in front of them. To avoid this drawback, the scene which is supposed to exist in front of the computer-generated images is interactively designated with a cursor or a tablet, and the images inside the designated area are overlaid again. The boundaries of each designated area are depicted with straight lines or splines[6] -- the former is useful for buildings, and the latter to mountains or trees.

The aliasing problem mentioned above also occurs when the foregrounds are superimposed onto the computer-generated images. In this case, the boundaries of the foregrounds become jagged. This problem is serious especially when the foregrounds cross the objects, which consist of straight boundaries (buildings, for example). The anti-aliasing method is also applicable here: If a pixel includes the boundaries of a foreground, its color is determined based on (i) the area of the foreground, (ii) the background color and (iii) the color of the computer-generated image within this pixel. The color of the pixel is determined considering the intersection of the boundary (straight line or spline) of the image and each VSL.

3. Implementation

Fig. 4 indicates R, G and B intensity variation with respect to the distance. Trees at various distance from the viewpoint are selected from the original background scene. Actual R, G and B intensity calculation was carried out by averaging each small region, about ten pixels, within a group of trees. The exponential function described in Section 2.4 was determined in this way. The exponential functions in Fig. 3 is used for the second example.

Fig. 5 shows the outputs of the proposed method. (a) is the background scene. (b) depicts the overlay of the computer-generated images onto the background scene, and the foreground scene in front of the superimposed computer-generated image has been re-displayed. However, anti-aliasing and the distance from the viewpoint are not taken into account. Therefore, the windows of the buildings seem uneven, and the contrast of the building far from the viewpoint looks too strong because the inserted computer-generated images do not match their surroundings.

In (c), anti-aliasing and distance are considered, but there are no shadows from the new buildings. As a result, the building near the viewpoint seems to stand above the earth level. In (d), the shadows are cast properly on the ground up to a grove of trees and a house near the right side.

In (e), the fog effect is activated. The buildings seem to be natural due to this effect. In (f), the fog effect is removed. The mountains look as if they are too close. The computer-generated images match their surroundings in cases (d), (e) and (f).

Figures 5(g), (h) and (i) demonstrate another set of examples. (g) correponds to (d) in the previous example (anti-aliasing, the distance from the viewpoint and the shadows are all considered). (h) and (i) show the cases when the fog effect is activated and when it is removed, respectively. Note that the difference between (g) and (i) is not so clear, while (f) is very different from (d). This is simply because the weather condition in (d) was better (fog was thinner) than that in (f).

The fog effect is more apparent in (j). Upper-left is the original background scene. Other pictures demonstrate the fog effect. Four buildings in upper-right are set on different distance planes; the greater the distance of a building from the viewpoint, the stronger the fog effect. The fog effect is removed in lower right.

In (k), four different colors are assigned to the bridge. This is another application of the proposed method to assess the visual impact [8] of new construction on its surroundings.

Calculations were carried out on a TOSBAC DS600 computer and GRAPHICA M508 color graphic display (512x512 pixels, 8-bit each for R, G and B).

4. Conclusion

The proposed method produces useful montages because:

(1) The fog effect increases the sense of perspective in the montages, and it permits simulation of various weather conditions.

(2) The shading and shadows of the computer generated images help a montage match to the background scene.

(3) The new anti-aliasing technique enables cost-effective high quality display of objects which might include narrow vertical parts; e.g. electric light poles or chimneys.

Further improvement should enable:

(1) The correction of the distortion which appears on the photograph due to the use of a wide-angle lens.

(2) The computer image generation of roadside trees and/or flower beds.

(3) Processing for the light that comes from buildings.

Of these improvements, (1) requires complicated geometric transformation. The clue to achieve (2) is found in Ref.[7]. (3) is important when a night view is considered.

5. Acknowledgments

The background scenes were prepared with image scanners installed at GRAPHICA Inc. and Government Industrial Research Institute, Chugoku. The computer-generated images were created according to data supplied by Kajima Corporation. We are grateful to Laurin Herr for his discussions. Special thanks to Tsuyoshi Tokoro for his skillful computer operation.

APPENDIX Location of Viewpoint

The ultimate aim is to estimate the coordinate of the viewpoint. This point is defined by R(the distance between the viewpoint and the view reference point), θ (the azimuth angle) and ϕ (the elevation angle). n sampled points are selected on the picture for calculating the viewpoint. 3-D coordinate of these points, (x_i, y_i, z_i), $i=1,2,\ldots,n$, may be obtained on the geographical map.

The coordinate of these points on the picture, (M_{xi}, M_{yi}), along with the above 3-D coordinates are used for calculating the location of the viewpoint based on the following equation:

$$\begin{aligned} e_{xi} &= f_i(R,\theta,\phi) - M_{xi} , \\ e_{yi} &= g_i(R,\theta,\phi) - M_{yi} , \\ f_i(R,\theta,\phi) &\triangleq f(R,\theta,\phi,x_i,y_i,z_i) , \\ g_i(R,\theta,\phi) &\triangleq g(R,\theta,\phi,x_i,y_i,z_i) , \\ i &= 1,2,\cdots,n \ (n \geqq 3) \end{aligned} \tag{A1}$$

In Eq. (A1), f(g) and $e_{xi}(e_{yi})$ mean the relation function which maps the points in the 3-D space onto the 2-D picture coordinates and the difference between the mapped coordinates and its measured coordinates on the picture, respectively.

The best viewpoint estimation is described as the minimization problem of:

$$Q = \sum_{i=1}^{n} (e_{xi}^2 + e_{yi}^2) \tag{A2}$$

Partially differentiating this value with respect to R, θ and ϕ, the following system of equations is given:

$$\begin{bmatrix} A_{RR} & A_{R\theta} & A_{R\phi} \\ A_{\theta R} & A_{\theta\theta} & A_{\theta\phi} \\ A_{\phi R} & A_{\phi\theta} & A_{\phi\phi} \end{bmatrix} \begin{bmatrix} -dR \\ -d\theta \\ -d\phi \end{bmatrix} = \begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix}$$

$$\begin{aligned} A_{jk} &= \sum_{i=1}^{n} (a_j^i a_k^i + b_j^i b_k^i) , \\ B_j &= \sum_{i=1}^{n} (a_j^i E_{xi} + b_j^i E_{yi}) , \\ a_e^i &= (\partial f_i / \partial e)_{R=R_0, \theta=\theta_0, \phi=\phi_0} , \\ b_e^i &= (\partial g_i / \partial e)_{R=R_0, \theta=\theta_0, \phi=\phi_0} , \\ e &= R, \theta, \phi , \\ E_{xi} &= f_i(R_0, \theta_0, \phi_0) - M_{xi} , \\ E_{yi} &= g_i(R_0, \theta_0, \phi_0) - M_{yi} \end{aligned} \tag{A3}$$

Appropriate values for R, θ and ϕ are set for initiating the iteration. On each iteration, the estimated value of the viewpoint (R_0, θ_0, ϕ_0) is replaced to the value (R, θ, ϕ) obtained by the following equation.

$$\begin{aligned} R &= R_0 - dR , \\ \theta &= \theta_0 - d\theta , \\ \phi &= \phi_0 - d\phi \end{aligned} \tag{A4}$$

The iteration continues until dR, $d\theta$ and $d\phi$ become small enough.

REFERENCES

1 S.Uno and H.Matsuka,"A General Purpose Graphic System for Computer Aided Design," Computer Graphics, vol.13, no.2,1979, pp.25-32

2 E.A.Feibush, M.Leroy and R.L.Cook,"Synthetic Texturing Using Digital Filters," Computer Graphics, vol.14, no.3, 1980, pp.294-301

3 T.Porter and T.Duff,"Composing Digital Images," Computer Graphics, vol.18, no.3,1984, pp.253-259

4 T.Nishita and E.Nakamae,"A Perspective Depiction of Shaded Time," Proc. of 4th Int. Symp. on the Use of Comp. for Env. Eng. Related to Build.,1983, pp.565

5 B.J.Schachter,Computer Image Generation, John Wiley & Sons.,New York, 1983

6 C.de Boor,Practical Guide to Splines, Springer-Verlag, New York, 1978

7 R.Marshall, R.Wilson and W.Carlson,"Procedure Models for Generating Three-Dimensional Terrain," Computer Graphics, vol.14, no.3,1980, pp.154-162

8 T.W.Maver, C.Purdie and D.Stearn," Visual Impact Analysis -- Modelling and Viewing the Natural and Built Environment," Comput. & Graphics, vol.9, no.2, 1985, pp.117-124

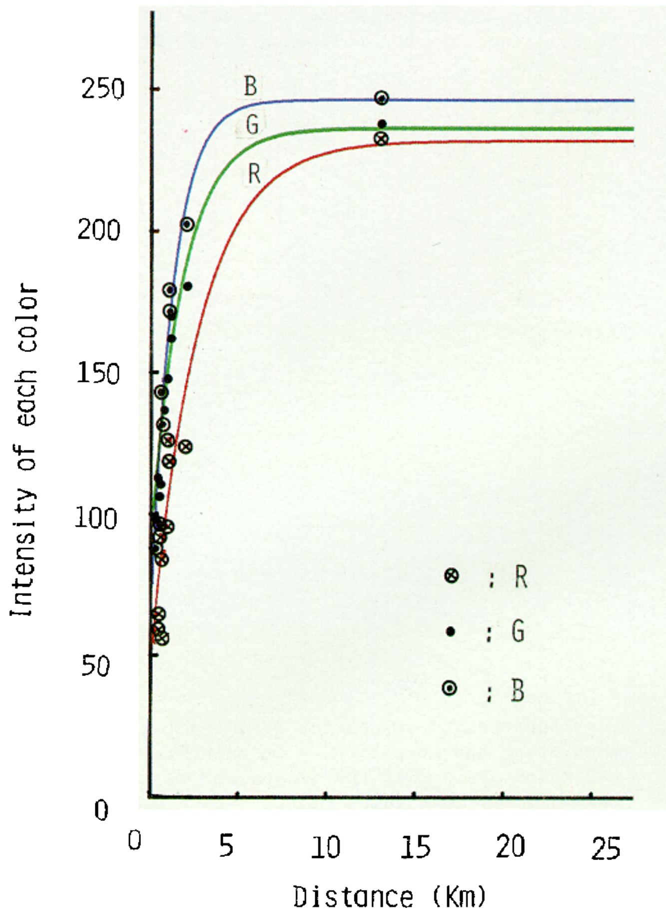


Fig. 1 Intensity variation with respect to the distance.
This Figure is referred along with Fig.4 (see Section 4).

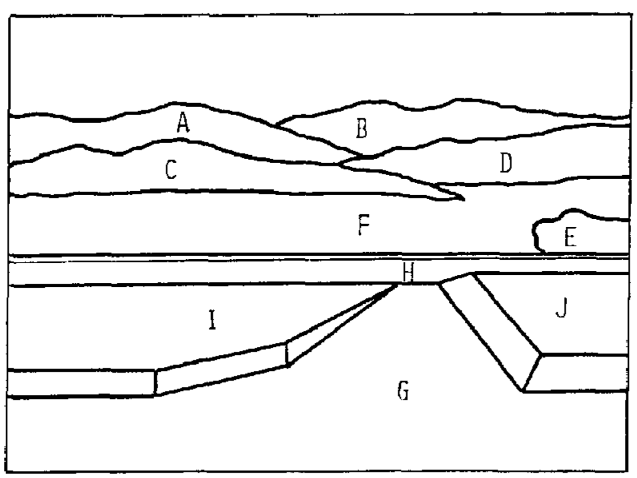


Fig. 2 Partition of the image for fog effect processing.

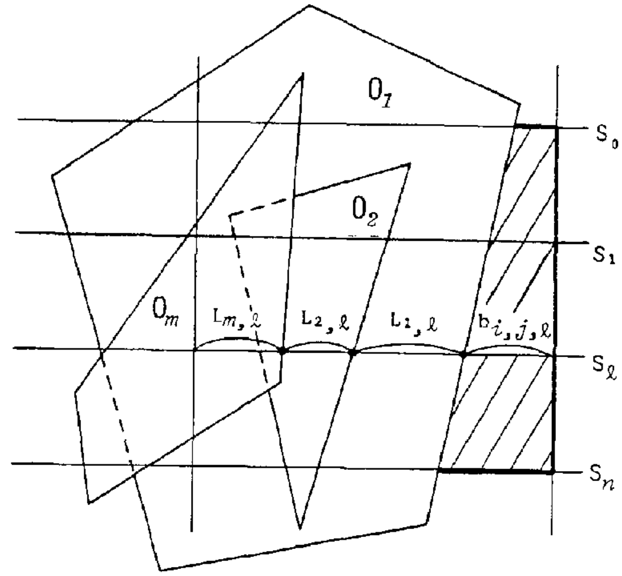


Fig. 3 Anti-aliasing operation in one pixel.

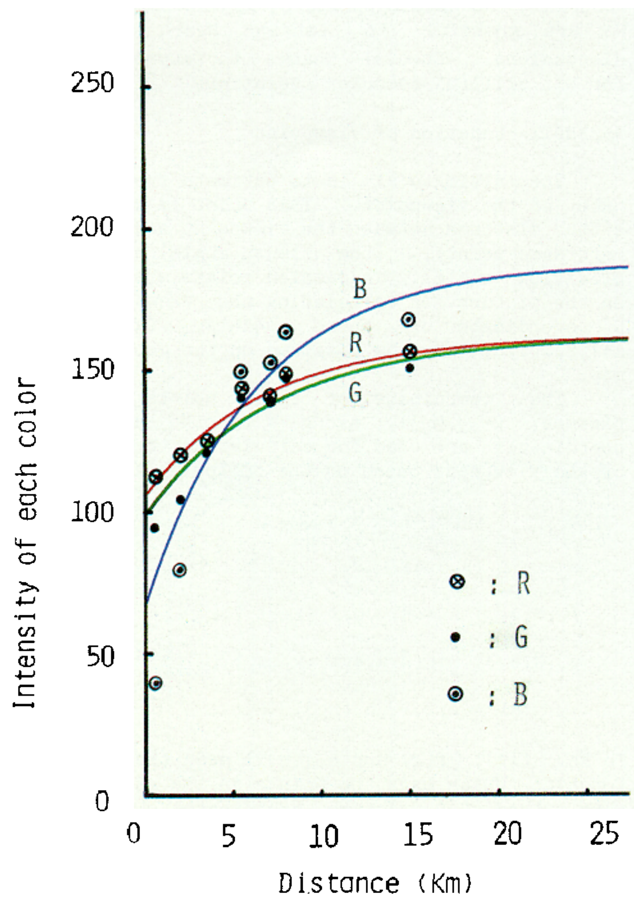
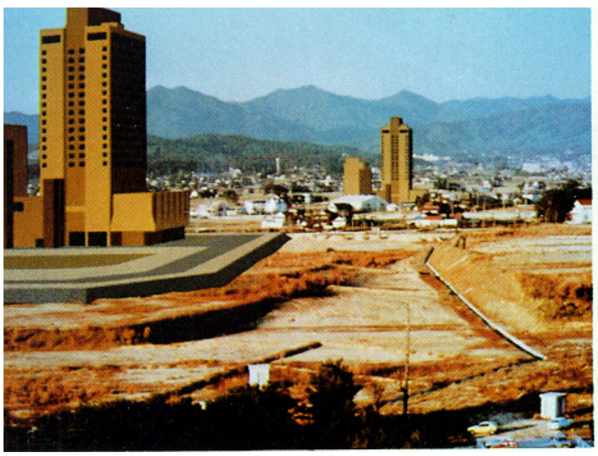


Fig. 4 Intensity variation with respect to the distance. This figure was given from the first background scene.

(a)

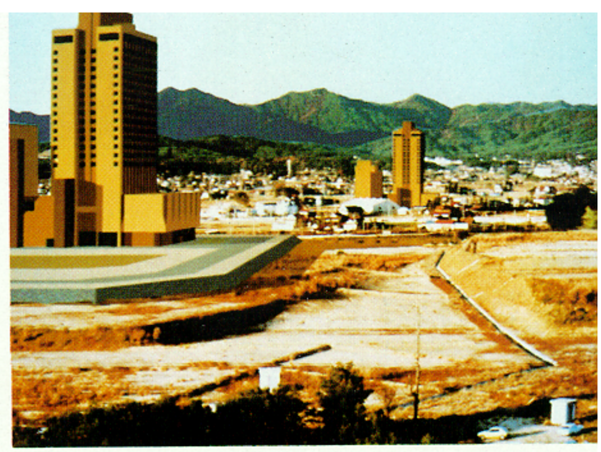
(b)



(c)

(d)

(e)



(f)

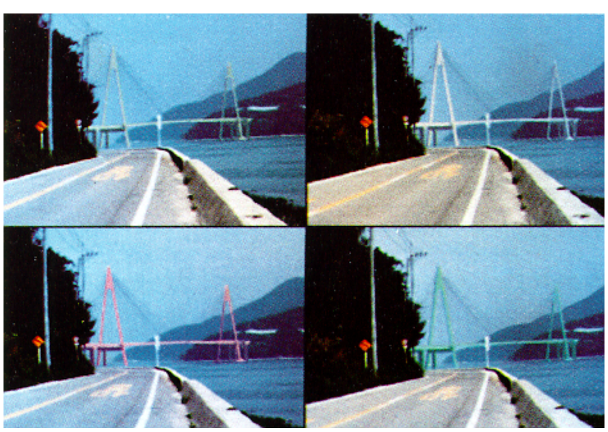
Fig. 5 (continued)

(g)

(h)

(i)

(j)



(k)

Fig. 5 Examples. Explanations are given in the text.