

Defocus Morphing in Real Aperture Images

Subhasis Chaudhuri

Department of Electrical Engineering

Indian Institute of Technology, Bombay

Powai, Mumbai - 400 076, INDIA

Email: sc@ee.iitb.ac.in

Abstract

We introduce a new concept called defocus morphing in real aperture images in this paper. View morphing is an existing example of shape preserving image morphing based on the motion cue. We prove that images can also be morphed based on the depth related defocus cue. This illustrates that the morphing operation is not necessarily a geometric process alone; one can also perform a photometry based morphing wherein the shape information is implicitly buried in the image intensity field. A theoretical understanding of the defocus morphing process has been presented in this paper. Mathematically we show that, given two observations of a 3D scene for different camera parameter settings, we can obtain a virtual observation for any camera parameter setting through a simple nonlinear combination of these observations.

OCIS Codes

100.1830, 100.6890, 110.2990, 330.3790, 330.6110

I. INTRODUCTION

A great deal of research efforts have been reported in the area of view interpolation [1] and view morphing [2]. The basic idea here is to generate a new view of a scene such that the geometric structure of the scene in the rendered view is preserved, given two or more different views of the scene. This has many practical applications in virtual view generation. The key advantage of view morphing is that the scene structure does not have to be computed, albeit a knowledge of camera parameters and the feature point correspondence is required. The new view can be generated using the Beier and Neely algorithm [3], and the homography relationship for the novel view can be shown to be a convex combination of homographies of the respective views.

Since the work of Seitz and Dyer [2], a large amount of research literature is available that

performs shape preserving novel view generation. For example, Seitz and Dyer also derive the conditions under which a given set of sparse views of a scene is good enough to generate a continuous range of view points [4]. Lhuillier and Quan explores a robust method to handle two main difficult tasks in view morphing, namely pixel matching and visibility handling [5]. Kang *et al.* discusses the trade off issues related to the need for accurate 3-D modeling of a scene in terms of geometry and photometry while maintaining the photorealism of the rendered scene [6]. Agarwal *et al.* assume the surface to be composed of planar patches while generating the novel views [7]. In [8], the authors use the polygonal approach to map the texture in novel views. Avidan *et al.* suggest the use of trilinear tensor while synthesizing widely apart novel views [9], [10]. Vazques *et al.* investigate the optimal positions for camera placement to obtain a minimal set of views for image based rendering [11]. Laveau and Faugeras [12] address the problem of generating a new view from a reference view by interactively selecting four points in the scene to predict a new view. Svedberg and Carlsson also attempt to generate a novel view from a single, constrained scene using the symmetry property [13]. A general review on view morphing can be found in [14]. In all these cases, one changes the view point and hence the camera geometry (mostly the external parameters and in some cases the lens-to-image plane distance also) plays an important role.

The concept of view morphing has also been extended to many other applications, such as, dynamic view interpolation [15], [16], [17], model based interactive view generation [18], image coding based on view rendering [19]. Weiskopf *et al.* [20] propose an interesting application of image based rendering wherein the plenoptic function is appropriately modified to take care of relativistic aberration of light and search light effect to simulate a super-fast motion film. Ezzat and Poggio [21] applied the view morphing technique in morphing visemes for text-to-visual

speech synthesis. Radke and Richard emulate the concept in audio interpolation to create a virtual microphone [22].

The view morphing technique assumes that the camera is a pin-hole one and hence it has an infinite depth of field. However, all commercial cameras have a finite aperture and hence provide only a limited depth of field. The depth related defocus [23] also provides an important cue for structure recovery. This also makes the image photo-realistic and an object of delight to a connoisseur. Now we ask the question can we also perform a structure preserving *defocus morphing* of a scene, i.e., given two observations of an arbitrary 3D scene under two different camera parameter settings, resulting in a different amounts of blurring in the pictures, can we obtain a novel observation of the same scene for a virtual camera parameter setting without estimating the depth in the scene? Here the possible camera parameters include aperture, focal length of the lens and the lens-to-image plane distance. It may be noted that the researchers in computer vision do not consider the aperture and the focal length of the lens as part of internal camera parameters since they consider a pin-hole model. In real-aperture photography, all of them play very important roles. The novel view synthesis problem as discussed in [2] is a case of motion or stereo cue morphing, and, to be more precise, we may just call it a plenoptic transfer. Similarly, the virtual setting of internal camera parameters in real aperture images is an example of defocus morphing. Hence the morphing process is no longer restricted to a change in camera geometry alone, it may also involve photometric changes. In this paper we prove that it is, indeed, possible to have defocus morphing and this problem is, in some sense, equivalent to view morphing. Just like in view morphing technique we have difficulties with respect to correspondence generation and handling of occlusion, there will be equivalent problems while defocus morphing. Further, for the view morphing problem the homography relationship for a

novel view can be obtained as a combination of those of the given views. We show that the morphing with respect to the defocus cue follows a similar form of combination in terms of the observed data, although it is not a linear one.

There is one basic difference between the view morphing and the defocus morphing processes. In view morphing, the photometric properties of the scene is assumed to be unchanged when the camera position is changed. Thus, one tries to account for disparity adjustments while doing the plenoptic transfer. This is purely a geometric process. In case of defocus morphing, since the optical center of the camera does not move, the geometry remains unchanged (strictly speaking, there could be a small change in magnification while changing some of the camera parameters which also needs to be accounted for during defocus morphing). Hence, one needs to perform only photometric adjustments between the two views to obtain the virtual observation. In both cases of image morphing, the 3D structure of the scene is, however, preserved.

It is worth referring to the work of Kubota *et al.* [24], [25], [26], [27], where the authors use a set of linear filters for photometric adjustment of observations from a multi-focus system to obtain a fully focused, virtually rendered view. They also explain how visual special effects can be created based on the depth related defocus [26]. The purpose of using multi-focus observations is the removal of aliasing in the texture while rendering a new view as defocus removes the high frequency components. Our objective in this paper is quite different in the sense that we explain how these multi-focus observations themselves can be virtually obtained through defocus morphing. Further, Kubota *et al.* assume that the filters for each depth layer are known. We do not make any such assumption. Finally, their method has the advantage of being a linear one unlike that of ours, making an analysis of the performance an easier proposition. However, the method proposed in this paper provides a better insight to the concept of defocus morphing

with respect to camera parameters.

The organization of the paper is as follows. We introduce the concept of defocus morphing in Section II along with the necessary physical interpretation. We show that the morphing process can be explained through the change of any of the following internal camera parameters – aperture, focal length and the lens-to-image plane distance. We do present some experimental results on defocus morphing in Section III. The paper concludes in Section IV.

II. VIRTUAL DEFOCUS RENDERING

We consider the following problem: given two observations E_1 and E_2 of the same scene under two different settings of camera parameters C_1 and C_2 , obtain a rendered view of the same scene as a function of only E_1 and E_2 for any virtual camera parameter setting C which is a convex combination of C_1 and C_2 . The imaging model is no longer a pin hole one. We consider the real aperture imaging principles. We prove that one does not have to find the depth at each point in the scene to render the view.

A. The defocusing Process

Let us now look at the image formation process in a real aperture camera employing a thin lens. When a point light source is in focus, all light rays that are radiated from the object point and intercepted by the lens converge at a point on the image plane. When the point is not in focus, its image on the image plane is no longer a point but a circular patch of radius σ that defines the amount of defocus associated with the depth of the point in the scene. It can be shown [23] that

$$\sigma = \kappa r v \left(\frac{1}{F} - \frac{1}{v} - \frac{1}{Z} \right) \quad (1)$$

where r is the radius of the aperture, v is the lens-to-image plane distance, F is the focal length of the lens, Z is the depth (distance from the optical center) at that point and κ is a camera constant that depends on the sampling resolution on the image plane (see Fig. 1 for illustration). If the aperture $r = 0$, it corresponds to a pin-hole image when every point in the scene is in focus signifying $\sigma = 0$. Let us use the symbol $I(x, y)$ to represent the pin-hole (focused everywhere) image of the scene. It may be noted that $I(x, y)$ is not a physically observable quantity in a real aperture camera when there is depth variation in the scene. From eqn (1), we note that $C = \{r, F, v\}$ defines the internal camera parameters each of which may be changed to effect a different amount of defocus blur for a fixed depth Z .

The depth related defocus process is linear but not space invariant. Assuming a diffraction-limited lens system, the point spread function of the camera system at a point (x, y) may be approximately modeled as a circularly symmetric 2-D Gaussian function [28], [29], [30]

$$h(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (2)$$

where the blur parameter $\sigma(x, y)$ has been defined in eqn (1). Assuming the depth to be constant (we shall relax this assumption later), the observed defocus image $E(x, y)$ is given by

$$E(x, y) = I(x, y) * h(x, y) \quad (3)$$

where the operator $*$ represents convolution. It is easier to represent the above relationship in the spectral domain. Denoting the Fourier transform of a function $f(x, y)$ by $\hat{f}(\omega_x, \omega_y)$ we obtain

$$\begin{aligned}\hat{E}(\omega_x, \omega_y) &= \hat{I}(\omega_x, \omega_y) \hat{h}(\omega_x, \omega_y) \\ &= \hat{I}(\omega_x, \omega_y) \exp\left(-\frac{\sigma^2(\omega_x^2 + \omega_y^2)}{2}\right).\end{aligned}\tag{4}$$

It may be mentioned that if one follows geometric optics, then the blur kernel $h(x, y)$ has a circular shape and the corresponding frequency response of the lens system is given by the Bessel function (see Hopkins [31] for a theoretical analysis and Baker [32] for an experimental validation). In this paper, we restrict ourselves to using the Gaussian blur model valid for a diffraction limited lens (see [28]).

B. Defocus Morphing

Let us now get back to the image morphing problem. For a given scene, one can have two defocused observations E_1 and E_2 corresponding to two different camera parameter settings C_1 and C_2 such that the resulting blur parameters are σ_1 and σ_2 , respectively. Since the defocus blur is characterized by the parameter σ^2 in eqn (4), let us select the morphing relationship

$$\sigma^2 = \alpha\sigma_1^2 + (1 - \alpha)\sigma_2^2\tag{5}$$

where $0 \leq \alpha \leq 1$ and σ_i^2 corresponds to the blur in image E_i . Why we select the convex combination in terms of σ_1^2 and σ_2^2 and not σ_1 and σ_2 will be explained in the next subsection.

From eqns (4 and 5) we obtain

$$\hat{E}(\omega_x, \omega_y) = \hat{I}(\omega_x, \omega_y) \exp\left[-\frac{1}{2}\{\alpha\sigma_1^2 + (1 - \alpha)\sigma_2^2\}(\omega_x^2 + \omega_y^2)\right]$$

$$= \{\hat{I}(\omega_x, \omega_y) \exp[-\frac{\sigma_1^2(\omega_x^2 + \omega_y^2)}{2}]\}^\alpha \{\hat{I}(\omega_x, \omega_y) \exp[-\frac{\sigma_2^2(\omega_x^2 + \omega_y^2)}{2}]\}^{1-\alpha}$$

or,

$$\hat{E}(\omega_x, \omega_y) = \hat{E}_1^\alpha(\omega_x, \omega_y) \hat{E}_2^{(1-\alpha)}(\omega_x, \omega_y). \quad (6)$$

Taking logarithms of both sides,

$$\ln \hat{E}(\omega_x, \omega_y) = \alpha \ln \hat{E}_1(\omega_x, \omega_y) + (1 - \alpha) \ln \hat{E}_2(\omega_x, \omega_y) \quad (7)$$

Compare eqn (7) with the corresponding rendering equation for the view morphing problem in [2] - they are very similar except that the convexity relation holds good in the logarithmic scale after having transformed the images into the spectral domain. Hence we observe that any two defocused observations can be combined using the relationship given in eqn (6) to render the scene for a virtual camera parameter setting such that the resulting blurring parameter σ^2 lies in the range $\sigma_1^2 \leq \sigma^2 \leq \sigma_2^2$. Further, there is no need to calculate the depth or structure in the scene.

The morphing relationship as given in eqn (6) does not consider the presence of noise in observations. Let us now analyze what happens if noise is present. Let N_i be the sensor noise in the i th noisy observation denoted as E_{ni} . In the Fourier domain, $\hat{E}_{ni} = \hat{E}_i + \hat{N}_i$. Let E_n be the result of morphing two noisy observations E_{n1} and E_{n2} . Now eqn (6) can be written as

$$\begin{aligned} \hat{E}_n(\omega_x, \omega_y) &= \hat{E}_{n1}^\alpha(\omega_x, \omega_y) \hat{E}_{n2}^{(1-\alpha)}(\omega_x, \omega_y) \\ &= (\hat{E}_1(\omega_x, \omega_y) + \hat{N}_1(\omega_x, \omega_y))^\alpha (\hat{E}_2(\omega_x, \omega_y) + \hat{N}_2(\omega_x, \omega_y))^{(1-\alpha)} \\ &= \hat{E}_1^\alpha(\omega_x, \omega_y) \left(1 + \frac{\hat{N}_1(\omega_x, \omega_y)}{\hat{E}_1(\omega_x, \omega_y)}\right)^\alpha \hat{E}_2^{1-\alpha}(\omega_x, \omega_y) \left(1 + \frac{\hat{N}_2(\omega_x, \omega_y)}{\hat{E}_2(\omega_x, \omega_y)}\right)^{(1-\alpha)} \\ &= \hat{E}_1^\alpha(\omega_x, \omega_y) (1 + \beta_1(\omega_x, \omega_y))^\alpha \hat{E}_2^{1-\alpha}(\omega_x, \omega_y) (1 + \beta_2(\omega_x, \omega_y))^{(1-\alpha)} \end{aligned}$$

where β_1 and β_2 are measures of inverse signal to noise ratio (SNR) at various spatial frequencies. Since the above equation is nonlinear it is difficult to quantify the effect of noise on the morphed image analytically. We do the analysis for two special cases, namely when the SNR is high and when it is very low. Typically at low frequencies, the SNR is quite high when $|\beta_1|, |\beta_2| \ll 1$. The above equation can be simplified (the arguments in the equations are also dropped for simplicity) to

$$\hat{E}_n = \hat{E}_1^\alpha \hat{E}_2^{1-\alpha} (1 + \alpha\beta_1)(1 + (1 - \alpha)\beta_2). \quad (8)$$

Assuming the noise N_1 and N_2 to be independent and zero mean, it is easy to show that

$$\mathcal{E}(\hat{E}_n | \hat{I}) = \hat{E}_1^\alpha \hat{E}_2^{1-\alpha} = \hat{E}, \quad (9)$$

where \mathcal{E} is the expectation operator. Thus we observe that one does not have much problem dealing with observation noise at lower frequencies. However at higher frequencies, the SNR is small and hence $|\beta_1|$ and $|\beta_2|$ may not be very small compared to unity. Thus

$$\mathcal{E}(\hat{E}_n | \hat{I}) \neq \hat{E}_1^\alpha \hat{E}_2^{1-\alpha}. \quad (10)$$

Hence the estimate of the high frequency components in the morphed image will be biased, resulting in a degradation in the quality of the synthesized view. However, if $\alpha \approx 0$ or 1 , the linear approximation as shown in eqn (8) is still quite valid, signifying that as one moves away from the given two observations E_{n1} and E_{n2} while synthesizing the virtual view, the effect of noise in the high frequency components may start becoming noticeable. However, since $\alpha \in [0, 1]$, the effect of noise is very minimal and can be neglected for all practical purposes.

A few more comments are in order now. In eqn (6) the term $|\hat{E}_1|^\alpha(\omega_x, \omega_y)$ refers to the image enhancement technique known as root filtering [33]. It boosts up the high frequency components, thereby it reduces the amount of blurring. In case the observations E_1 and E_2 have a good

amount of sensor noise, the rendered image $E(x, y)$ may suffer from noise enhancement problem. However, we are not exactly performing root filtering as the phase information is preserved in root filtering. But in the expression $\hat{E}_1^\alpha(\omega_x, \omega_y)$, the phase is altered based on the value of α . It is interesting to note that the structure of the root filter (one should not confuse this with the square root filtering technique used in subspace updating or Kalman filtering applications) is not well understood in time/spatial domain and a very little information is available in the literature.

The derivation of the rendering relationship given in eqn (6) is obtained by assuming that the scene has a constant depth. Now we relax the above assumption. When there is depth variation in the scene, eqn (3) is no longer valid as the blurring process becomes shift-variant. In a typical depth from defocus (DFD) problem, this issue is handled by forming a small $M \times M$ window about a point over which the depth can be assumed to be constant. The depth computation is done locally at the center pixel through the eqn (3), and a dense depth map is obtained by sliding the window by a single pixel. The same procedure can be repeated while morphing the images E_1 and E_2 . The morphing is performed at every point using the relationship defined in eqn (6) by considering a small $M \times M$ neighborhood over which the Fourier transforms are computed.

A comparison with the view morphing process results in an understanding that the knowledge of homography is equivalent to knowing the thin lens camera parameters. The need to establish the feature point correspondences in view morphing now translates into an equivalent problem of having to handle a shift varying system. Similarly, the problems of *holes* and *folds* do persist as the defocus process is known to suffer from partial self occlusion [34]. Wider the aperture, more pronounced is the effect of occlusion.

Now that we have derived the blur morphing relationship, we discuss the physical interpretation for morphing with respect to each of the camera parameters C and/or the scene structure.

C. Aperture Morphing

The defocus blur is a function of three different camera parameters, namely the aperture (r), lens-to-image plane distance (v) and the focal length (F) of the lens used. It is also a function of the depth (Z) in the scene. We now look at the physical interpretation of the morphing relationship $\sigma^2 = \alpha\sigma_1^2 + (1 - \alpha)\sigma_2^2$. Assuming all other camera parameters remaining constant, using eqn (1) that defines the blur radius σ in terms of the camera parameters, we get

$$r^2 = \alpha r_1^2 + (1 - \alpha)r_2^2. \quad (11)$$

Multiplying both sides by $\pi/2$ and noting that this refers to the aperture area A for the lens, the morphing relationship can be written as

$$A = \alpha A_1 + (1 - \alpha)A_2. \quad (12)$$

Thus if the observations for two given aperture areas A_1 and A_2 are given, an image for any aperture $A_1 \leq A \leq A_2$ can be synthesized. This also explains why we started our morphing relationship in terms of σ^2 and not σ . The amount of light collected at the image plane is a linear function of the aperture area (A) and not the aperture radius. The basic morphing process with respect to blurring due to aperture change is illustrated in Fig. 1. From the figure it is clear that any virtual aperture in the range (r_1, r_2) can be synthesized given the observations E_1 and E_2 corresponding to the apertures r_1 and r_2 , respectively.

If one changes any of the camera parameters, apart from the change in blur radius, there are other effects in the captured image such as the magnification. Due to change in aperture area, there is no change in magnification as the equivalent pin hole geometry (*i.e.*, the point of convergence of the incident rays) is preserved. However, the image plane receives a different

amount of light energy that is usually compensated in a camera by appropriately changing the shutter time. Since this option is not available while morphing these images, we recommend a re-normalizing operation of gray levels of the morphed image based on the input images.

D. Lens-to-Image Plane Distance Morphing

For the given blur parameters σ_1^2 and σ_2^2 , we now want to find out what would be the equivalent morphing relationship in terms of corresponding lens-to-image plane distances v_1 and v_2 , instead of a variation in the aperture. Since, in the expression

$$\sigma_i = \kappa r [v_i (\frac{1}{F} - \frac{1}{Z}) - 1] \quad (13)$$

the term $(1/F - 1/Z)$ refers to the actual lens-to-image plane distance for which the image will be sharply focused, we denote this term as $1/v_o$. The morphing relationship $\sigma^2 = \alpha\sigma_1^2 + (1 - \alpha)\sigma_2^2$, now becomes

$$\alpha v_1^2 + (1 - \alpha)v_2^2 - v^2 = 2v_o(\alpha v_1 + (1 - \alpha)v_2 - v). \quad (14)$$

On simplification, this takes a nice form

$$(v - v_o)^2 = \alpha(v_1 - v_o)^2 + (1 - \alpha)(v_2 - v_o)^2. \quad (15)$$

The above relationship resembles very closely to the morphing relationship used in this study. The only departure is that the lens-to-image plane distances should be measured relative to the actual distance v_o for which the scene point is perfectly in focus. This is now illustrated in Fig. 2. A virtual image plane can be synthesized between the lines A_1B_1 and A_2B_2 given the corresponding observations E_1 and E_2 . From the figure we observe that the above relationship

can also be very easily derived using the properties of similar triangles. The same thing can be said about the eqn (11), the one used for aperture morphing, while obtaining it from the basic blur morphing relationship given in eqn (7). Hence the physical significance is the same as what was discussed in the previous subsection.

It is to be noted that since the aperture remains constant, there is no need for camera gain adjustment while morphing. However, there is a change in magnification (see Fig. 2). Thus the images E_1 and E_2 should be magnification adjusted while morphing them. A linear isotropic scaling achieves the purpose.

E. Focal Length Morphing

In the depth from defocus problem, one does have the option of changing the focal length (F) of the lens to obtain the depth related defocus. Let us now find out the equivalent morphing relationship in terms of the focal lengths of the two lenses (F_1 and F_2), for which the observations E_1 and E_2 have been obtained.

For a given depth (Z) and the lens-to-image plane distance (v), let F_0 be the focal length of a lens that would focus the object sharply on the image plane. The morphing relationship $\sigma^2 = \alpha\sigma_1^2 + (1 - \alpha)\sigma_2^2$ can now be written as

$$\left(\frac{1}{F} - \frac{1}{F_0}\right)^2 = \alpha\left(\frac{1}{F_1} - \frac{1}{F_0}\right)^2 + (1 - \alpha)\left(\frac{1}{F_2} - \frac{1}{F_0}\right)^2$$

or,

$$\left(\frac{F_0}{F} - 1\right)^2 = \alpha\left(\frac{F_0}{F_1} - 1\right)^2 + (1 - \alpha)\left(\frac{F_0}{F_2} - 1\right)^2. \quad (16)$$

Hence the equivalent focal length morphing depends on the squares of the harmonic difference of focal length setting of the observations with respect to the true focal length needed to have the

image captured in sharp focus. This case is illustrated graphically in Fig. 3. One can construct a virtual lens having a focal length anywhere between F_2 and F_1 .

The physical significance of the above relationship is that the point of convergence of all rays emanating from a scene point changes due to the change in the focal length of the lens. The amount of defocus depends on the divergence of the beam from the focused imaged point. Likewise in the previous case, there is a change in magnification which needs to be accounted for while morphing the observations.

F. Depth Morphing

In the last three subsections we explained the defocus rendering process with respect to the virtual settings of the camera parameters. Now we look at a different hypothetical situation which would also let us morph the observations in a precisely same manner. We assume that the camera parameters are unchanged but the object can be moved toward or away from the camera (see Fig. 4 for illustration). In order to keep the geometry unchanged, i.e., the image of the point in the scene should remain unchanged, the point in the scene should be moved along the principal ray (the line $ABCO$ joining the scene point and the optical center) as illustrated in Fig. 4. Here the morphing operation allows us to virtually move a point in the scene anywhere between the line segment AB , where the points A and B denote the location of the same scene point in two observations E_1 and E_2 , respectively. The movement of a point along the line $ABCO$ demands that the image be magnified as the depth increases so that the camera continues to observe the same part of the scene through the adjustment of its field of view. By substituting eqn (1) in the blur morphing relationship defined in eqn (5), we obtain

$$\left(\frac{1}{F} - \frac{1}{v} - \frac{1}{Z}\right)^2 = \alpha\left(\frac{1}{F} - \frac{1}{v} - \frac{1}{Z_1}\right)^2 + (1 - \alpha)\left(\frac{1}{F} - \frac{1}{v} - \frac{1}{Z_2}\right)^2.$$

Using the basic lens equation, where Z_0 denotes the scene depth for which the object would have been in perfect focus on the image plane, we obtain

$$\left(1 - \frac{Z_0}{Z}\right)^2 = \alpha\left(1 - \frac{Z_0}{Z_1}\right)^2 + (1 - \alpha)\left(1 - \frac{Z_0}{Z_2}\right)^2. \quad (17)$$

The above expression is quite similar to what was obtained in the previous section on focal length morphing, i.e., the depth morphing follows a harmonic difference relationship. Thus, given the pictures of a scene at two different depths, the scene can be rendered with geometric correctness for an intermediate depth provided the pictures are adjusted for scaling changes.

Thus we summarize that the basic defocus blur morphing relationship

$$\hat{E}(\omega_x, \omega_y) = \hat{E}_1^\alpha(\omega_x, \omega_y) \hat{E}_2^{(1-\alpha)}(\omega_x, \omega_y)$$

can be explained in terms of morphing through changes in any of the camera parameters. All these equivalent morphing relationships given in eqns (11, 15, 16, and 17) produce the same effect of morphing on the observed data.

III. EXPERIMENTAL RESULTS

Although the primary motivation in this paper is to develop a mathematical theory on virtual camera parameter adjustment, we do provide some relevant experimental results for illustration. In this section we show the results of shape preserving defocus morphing given two arbitrary observations. For brevity we present the results for only a few values of the morphing parameter α . Experimentations have been performed on both real and simulated data sets.

Figs. 5(a, b) show the observations of a scene where there is a large amount of depth variation, under two different settings of the camera parameters. The first one is better focused than the second observation. Due to a large change in the depth field, different parts of the scene have different amounts of defocus. Clearly, an assumption of a constant depth in the scene is not valid and hence the morphing should not be done globally based on the expression given in eqn (6). The corresponding result for $\alpha = 0.5$ is given in Fig 5(c) and one notices a fudge pattern almost everywhere in the image. This is due to the fact that eqn (6) is not valid for a scene having depth variations. However, the locally morphed image that does not assume a constant depth in the scene with a neighborhood size of 8×8 , as shown in Fig 5(d), does not suffer from this problem. The penalty for obtaining the improved result is in terms of increased computation. The order of computation goes up from $O(N^2 \log N)$ to $O(N^2 M^2 \log M)$ where $N \times N$ is the size of the image and $M \times M$ is the window size over which the depth is assumed to be constant. As expected, the amount of blurring in the morphed image is somewhere in between those in the two given observations. One can render the same scene for any other camera parameter setting by suitably adjusting the parameter α using one of the equations (11, 15, 16, and 17).

The experiments are repeated for another set of observations given in Fig. 6(a, b). Here there is a gradual variation of depth in the scene. The results of global and local morphing for $\alpha = 0.7$ are given in Fig. 6(c and d), respectively. Notice that the globally morphed image that assumes a constant depth in the scene is not as poor as it was in the previous example due to a small variation in the scene depth. However, the results of the local morphing method is definitely of better quality even in this experiment. Hence we perform a structure preserving blur morphing without actually estimating the depth in the scene. Both of the above experiments were carried out by changing the camera aperture to effect the depth related defocus in the observations.

Hence there was no need for scale (magnification) adjustment. The intensity variation due to aperture changes was removed using gray scale normalization in the images.

The purpose of the next experiment is to evaluate the performance of the proposed morphing operation by matching the synthesized view to an actual observation. Figures 7(a,b) show two views of a scene for two different settings of the lens-to-image plane distance. The first one is very poorly focused, and the second one is also partly defocused. We also capture another image shown in Fig. 7(c) for a camera parameter setting that is somewhere in between the values used for the first two observations. Can the image shown in Fig. 7(c) be obtained from those given in Fig. 7(a and b)? Since we are using an uncalibrated camera, we do not know what value of α will generate this particular view. Hence we synthesize all possible views for $\alpha \in [0, 1]$ in steps of size 0.1 and calculate the corresponding PSNR with respect to the desired view. The corresponding PSNR plot is shown in Fig. 8(f) and we find that $\alpha = 0.2$ provides the best match. There is a clear peak in the plot that suggests that the synthesized view matches well with the intended view. The corresponding synthesized view is shown in Fig. 7(d). A comparison of these two images demonstrates that the synthesized view is quite close to the true observation even in presence of existing sensor noise.

The purpose of the next experiment is twofold. The reader may remember that we have assumed a Gaussian shaped blur kernel to derive the morphing relationship. However, a few researchers prefer to use the circular blur kernel as obtained from geometric optics [23]. In this experiment we study the validity of the Gaussian assumption when the blurring is, indeed, circular. Further, the defocus morphing relationship given in eqn 6 implicitly assumes that we are working on the same side of the blur cone as explained in Fig. 2, ie, all rays are either converging or diverging. Thus, the amount of blurring changes monotonically between the observations

for $0 \leq \alpha \leq 1$, and one cannot perform deblurring while synthesizing a new view. We also demonstrate this property through this experiment. We synthetically generate two views of a scene with varying depth using the povray rendering toolkit. For the first view shown in Fig. 8(a), we adjust the focusing distance to be somewhere in front of the nearest (left most) pawn. Hence the left most pawn is partly blurred but the distant pawn is very poorly in focus. For the second view in Fig. 8(b), the focusing distance is adjusted to be somewhere in between the left pawn and the middle pawn. The left and the middle pawns are relatively in better focus compared to the distant pawn. Quite naturally, the observations correspond to same side of the blur cone (diverging rays) for the middle and the right pawns but both sides of the cone (convergent in a, and divergent in b) for the left pawn. The aperture in this experiment has been kept fairly large to reduce the depth of field, making the defocus blur quite prominent. Since the povray toolkit uses a standard ray tracing method to render the scene, it follows the geometric optics and hence the defocus blur is circular in shape. Note that we do not add any noise to the rendered images other than what is inherent during 8-bit quantization of the intensity field. We now simulate the third view of the same scene shown in Fig. 8(c) with the focal distance so adjusted that the parallel rays converge at a point just in front of the left most pawn signifying that the left most pawn is quite close to the second observation but the distant pawn is closer to the first observation. We now try to synthesize this view from the given two observations (a and b). We try for various values of α in the range $[0, 1]$ in steps of 0.1 and compare the quality of the synthesized view against the given observation Fig. 8(c) by plotting the PSNR. From the PSNR plot shown in Fig. 8(f), we observe that there are two peaks – one at $\alpha = 0.2$ and the other at $\alpha = 0.8$. The corresponding virtual views are shown in Figs. 8(d, and e), respectively. The reason we get two peaks instead of a single peak is that the synthesized views are not very

close to the intended view Fig 8(c). In the first case, the left pawn matches well with that in the intended view and in the second case, the distant pawn matches well. Hence there are two peaks. Since we violate the assumption of monotonicity of the blur in two views, there does not exist a single value of α , constant over the entire image, that will generate the given view. On the other hand we find that the assumption of a Gaussian shaped blur kernel in place of a circular blur does not create much of a problem as the reconstruction of the individual pawns is quite good.

IV. CONCLUSIONS

In this paper we have introduced the basic concept of shape preserving defocus morphing. There is no need to estimate the structure in the scene. The morphing relationships in both cases of view point change or internal camera parameter change have been shown to be equivalent; in view morphing problem the homographies are to be morphed while in defocus morphing the image intensity fields are to be morphed appropriately. We have also presented the physical interpretation of the morphing process in terms of various camera parameters. The primary emphasis in this paper has been in proving the existence of an equivalent class of morphing relationships, be it geometric or photometric. Although the defocus morphing does not generate the same kind of dramatic effect as the view morphing does, it is nonetheless very important in understanding how to restore and analyze real-aperture images. Currently we are looking for practical applications of the morphing process in real aperture imaging. We are also investigating how all these different types of cues, namely motion and defocus, can be combined to have a simultaneous multiple cue morphing.

ACKNOWLEDGMENTS

Financial assistance in the form of a research grant under the *Swarnajayanti Fellowship* and the programming help from Vinay P.N. are gratefully acknowledged.

REFERENCES

- [1] S. E. Chen and L. Williams, "View Interpolation for Image Synthesis", *Proc. SIGGRAPH 93*, pp. 279-288, Anaheim, CA, 1993.
- [2] S. M. Seitz and C. R. Dyer, "View Morphing", *Proc. SIGGRAPH 96*, pp. 21-30, New Orleans, LA, 1996.
- [3] T. Beier and S. Neely, "Feature-based Image Metamorphosis", *Proc. SIGGRAPH 92*, pp 35-42, Chicago, IL, 1992.
- [4] S. Seitz and C. R. Dyer, "View Morphing: Uniquely Predicting Scene Appearance from Basis Images", *Proc. Image Understanding Workshop*, pp. 881 - 887, New Orleans, LA, 1997.
- [5] M. Lhuillier and L. Quan, "Image Interpolation by Joint View Triangulation", *Proc. of the Conference on Computer Vision and Pattern Recognition*, Fort Collins, Colorado, USA, Vol. 2, pp. 139-145, June, 1999.
- [6] S. B. Kang, R. Szeliski and P. Anandan, "The Geometry-Image Representation Tradeoff for Rendering", *Proc. International Conference on Image Processing*, Vancouver, September, 2000.
- [7] D. Rathi, G. Agarwal, P. Kalra and S. Banerjee, "A System for Image Based Rendering of Walk-throughs", *Computer Graphics International (CGI2002)*, Bradford, UK, July, 2002.
- [8] P. Debevec, G. Borshukov, Y. Yu, "Efficient View-Dependent Image-Based Rendering with Projective Texture-Mapping", *9th Eurographics Rendering Workshop*, Vienna, Austria, June, 1998.
- [9] S. Avidan, T. Evgeniou, A. Shashua and T. Poggio, "Image-Based View Synthesis", *Artificial Intelligence (AI) Lab, AIM-1603, CBCL paper-145*, MIT, Cambridge, Mass. 1997.
- [10] S. Avidan and A. Shashua, "Novel View Synthesis in Tensor Space", *Proc. Computer Vision and Pattern Recognition*, Puerto Rico, pp. 1034-1040, June 1997.
- [11] P. Vazques, M. Feixas, M. Sbert and W. Heidrich, "Image Based Modeling Using Viewpoint Entropy", *Computer Graphics International CGI 2002*, Bradford, UK, July, 2002.
- [12] S. Laveau and O. D. Faugeras, "3D Scene Representation as a Collection of Images and Fundamental Matrices", *Technical Report No. 2205*, Institut National de Recherche en Informatique et en Automatique (INRIA), Sophia Antipolis, France, 1994.
- [13] D. Svedberg and S. Carlsson, "Calibration, Pose and Novel Views from Single Images of Constrained Scenes", *Proc. of the 11th Scandinavian Conference on Image Analysis (SCIA'99)*, Kangerlussuaq, Greenland, pp. 111-117, June, 1999.

- [14] Z. Zhang, "Image-Based Geometrically-Correct Photorealistic Scene/Object Modeling (IBPhM): A Review", *Proc. of the Asian Conference on Computer Vision (ACCV 98)*, Hong Kong, January 8-11, 1998.
- [15] S. Vedula, S. Baker and T. Kanade, "Spatio-Temporal View Interpolation", *Proc. of the 13th ACM Eurographics Workshop on Rendering*, pp. 1-11, SaarBruecken, Germany, 2002.
- [16] S. Baba, H. Saito, S. Vedula, K. M. Cheng and T. Kanade, "Appearance-Based Virtual View Generation for Fly Through in a Real Dynamic Scene", *Proc. of the Joint Eurographics and IEEE TVCG Symposium on Visualization*, pp. 179-188, Interlaken, Switzerland, 2000.
- [17] R. A. Manning and C. R. Dyer, "Interpolating View and Scene Motion by Dynamic View Morphing", *Proc. of the Conference on Computer Vision and Pattern Recognition*, Fort Collins, Colorado, USA, Vol. 1, pp. 388-394, June, 1999.
- [18] D. Cohen-Or, "Model-Based View-Extrapolation for Interactive VR Web-systems", *Proc. Computer Graphics International (CGI97)*, Belgium, pp. 104-112, June, 1997.
- [19] M. Magnor, P. Ramanathan and B. Girod, "Multi-View Coding for Image-based Rendering using 3-D Scene Geometry", *IEEE Transactions on Circuits and Systems for Video Technology*, vol-13, no-11, pp 1092-1106, 2003.
- [20] D. Weiskopf, D. Kobras and H. Ruder, "An Image-Based Approach to special relativistic Rendering", *IEEE Visualization*, pp. 303-310, October, 2000.
- [21] T. Ezzat and T. Poggio, "Visual Speech Synthesis by Morphing Visemes", *International Journal of Computer Vision*, Vol. 38, pp. 45-57, 2000.
- [22] R. J. Radke and S. Rickard, "Audio Interpolation", *Proc. of the International Conference on Virtual, Synthetic and Entertainment Audio*, Espoo, Finland, June, 2002.
- [23] S. Chaudhuri and A. N. Rajagopalan, *Depth from Defocus: A Real Aperture Imaging approach*, Springer Verlag, NY, 1999.
- [24] A. Kubota, K. Takahashi, K. Aizawa and T. Chen, "All-focused Light Field Rendering", *Proc. Eurographics Symp. on Rendering*, Norrköping, Sweden, 2004.
- [25] A. Kubota and K. Aizawa, "Inverse filters for reconstruction of arbitrarily focused images from two differently focused images", *Proc. IEEE Intl. Conf. Image Processing*, pp 101-104, Vancouver, 2000.
- [26] K. Aizawa, K. Kodama and A. Kubota, "Producing object based special effects by fusing multiple differently focused images", *IEEE Trans. CSVT*, vol-10(2), pp 323-330, 2000.
- [27] A. Kubota, K. Aizawa and T. Chen, "Virtual view synthesis through linear processing without geometry", *Proc. IEEE Intl. Conf. Image Processing*, Singapore, 2004.
- [28] M. Born and E. Wolf, *Principles of Optics*, Pergamon, London, 1965.
- [29] A. P. Pentland, "A New Sense for Depth of Field", *IEEE Transactions on Pattern Analysis Machine Intelligence*, Vol. 9, No. 4, pp. 523-531, July, 1987.

- [30] M. Subbarao, "Parallel Depth Recovery by Changing Camera Parameters", *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, Florida, USA, pp. 149-155, 1988.
- [31] H.H. Hopkins, "The Frequency Response of a Defocused Optical System", *Proc. Royal Soc. London A*, vol-231, pp 91-103, 1955.
- [32] L.R. Baker, "An Interferometer for Measuring the Spatial Frequency Response of a Lens System", *Proc. Phys. Soc. B*, vol-68, pp 871-880, 1955.
- [33] A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice Hall, Englewood Cliffs, NJ, 1988.
- [34] S. Bhasin and S. Chaudhuri, "Depth from Defocus in Presence of Partial Self Occlusion", *Proc. IEEE International Conference on Computer Vision*, Vancouver, Canada, pp. 488-493, July, 2001.

List of Figures

1. Illustration of aperture morphing. The blur is increased by increasing the aperture. Aperture morphing is possible within the range (r_1, r_2) .
2. Illustration of lens-to-image plane distance morphing. A virtual image plane between A_1B_1 and A_2B_2 can be constructed.
3. Illustration of focal length morphing. A virtual lens with the focal length in the range (F_1, F_2) can be constructed.
4. Illustration of depth morphing. For a fixed camera parameter set up and for the principal ray $ABCO$, a point in the scene can be virtually moved anywhere between the line segment AB .
5. (a,b) Views of an audience during a seminar presentation under two different focusing adjustments. Results of blur morphing with $\alpha = 0.5$ when performed (c) globally and (d) locally.
6. (a,b): Views of a scene with slowly varying depth under different camera parameter settings, (c,d): Results of global and local morphing for $\alpha = 0.7$.
7. (a,b) Two views of a scene under different focus settings. (c) Same scene captured with a focus setting lying in between the above two observations. (d) Result of synthetically generating the view in (c) using images (a and b) through blur morphing.
8. (a,b) Two synthetically rendered views of a scene under different aperture settings. (c) Another rendered view for an aperture lying in between the above two settings. (d,e) Results of synthetically generating the view in (c) using images (a and b) through blur morphing. (f) The plot of PSNR against the morphing parameter α for both the current and the previous experiment.

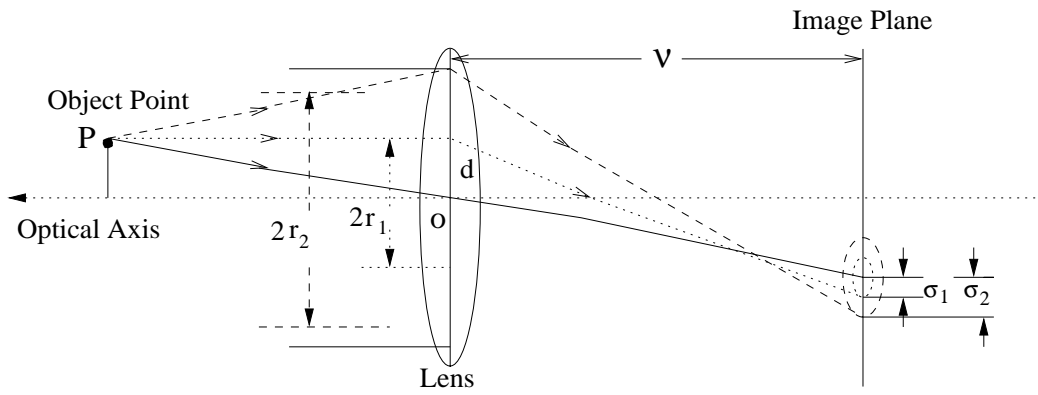


Fig. 1. Illustration of aperture morphing. The blur is increased by increasing the aperture. Aperture morphing is possible within the range (r_1, r_2) .

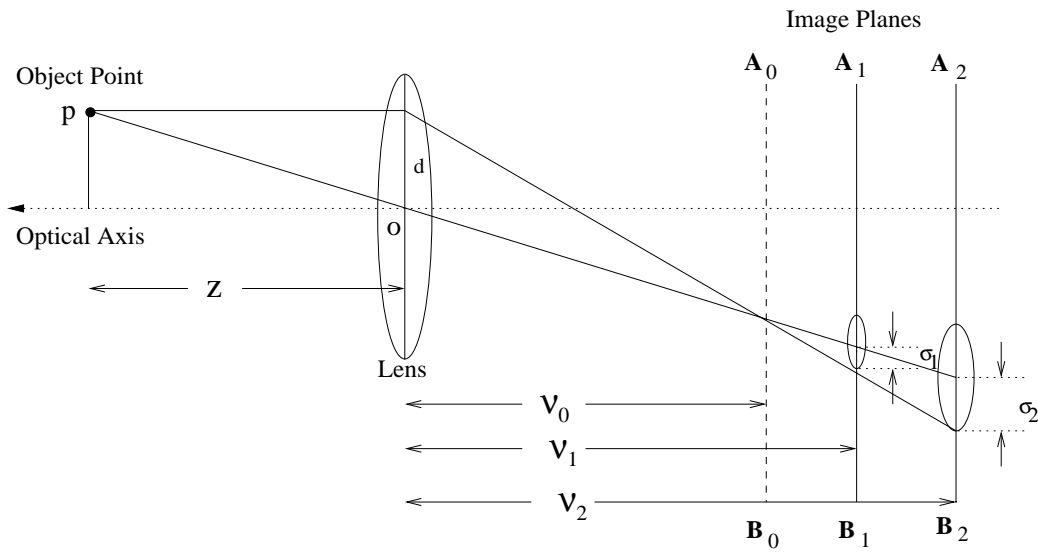


Fig. 2. Illustration of lens-to-image plane distance morphing. A virtual image plane between A_1B_1 and A_2B_2 can be constructed.

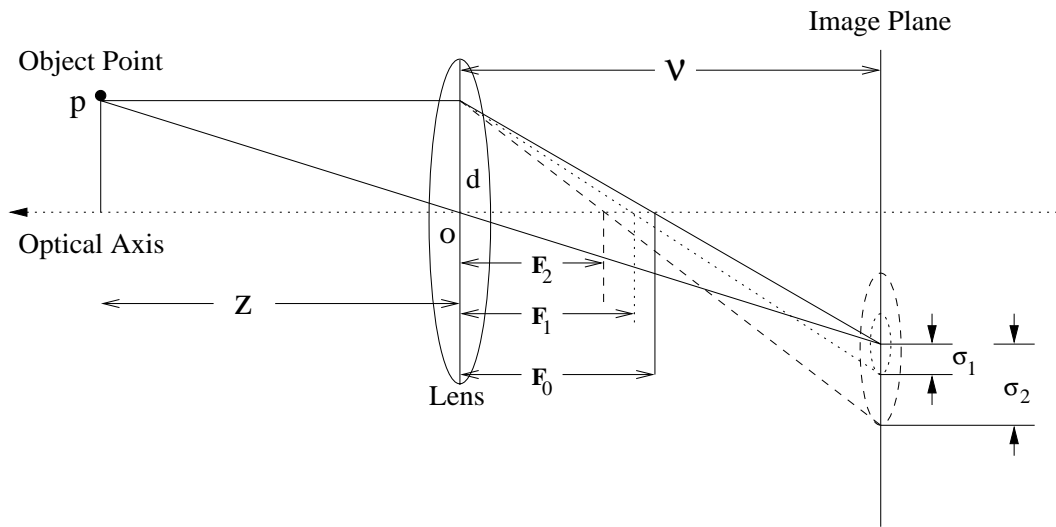


Fig. 3. Illustration of focal length morphing. A virtual lens with the focal length in the range (F_1, F_2) can be constructed.

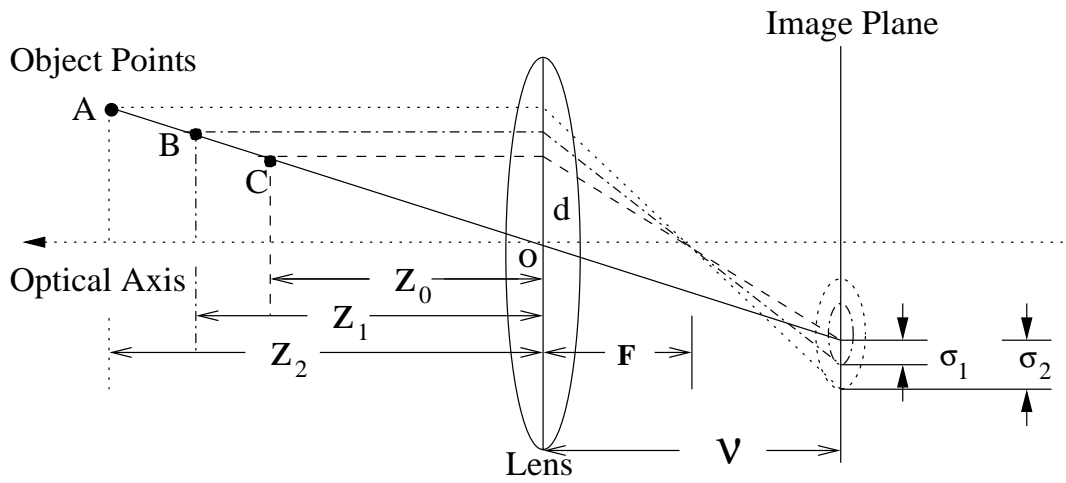


Fig. 4. Illustration of depth morphing. For a fixed camera parameter set up and for the principal ray ABCO, a point in the scene can be virtually moved anywhere between the line segment AB.

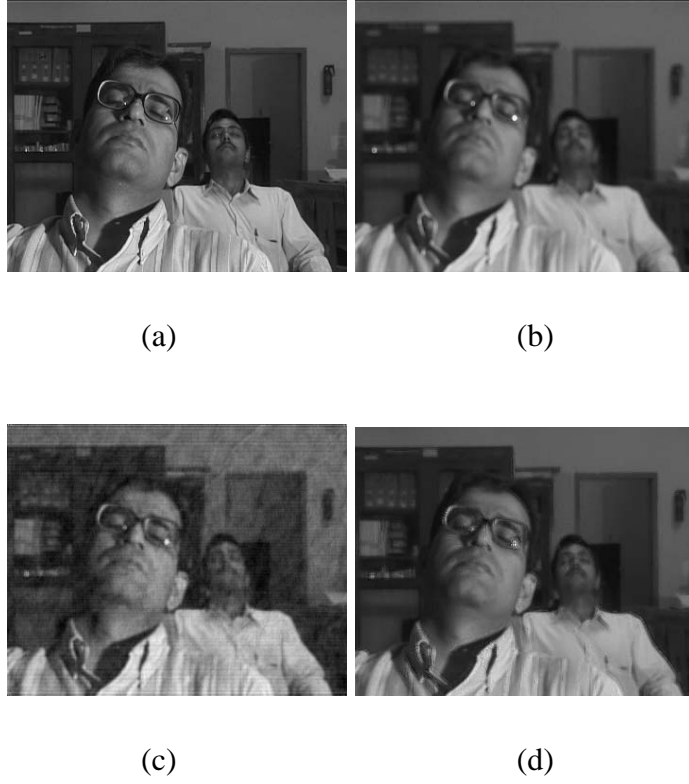
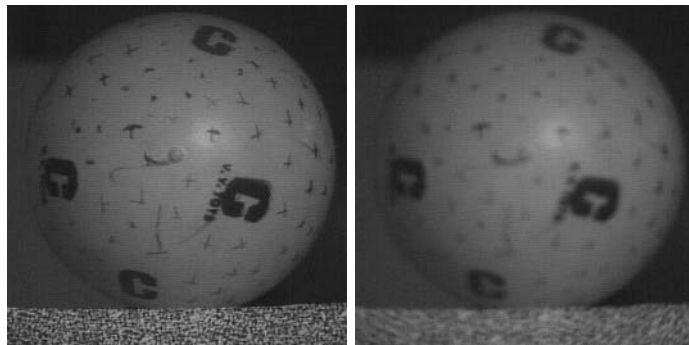
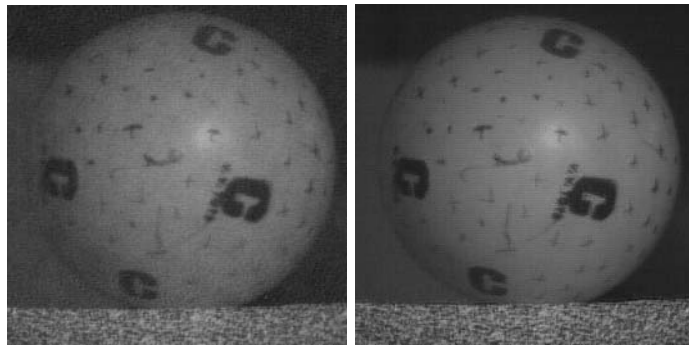


Fig. 5. (a,b) Views of an audience during a seminar presentation under two different focusing adjustments. Results of blur morphing with $\alpha = 0.5$ when performed (c) globally and (d) locally.



(a)

(b)



(c)

(d)

Fig. 6. (a,b): Views of a scene with slowly varying depth under different camera parameter settings, (c,d): Results of global and local morphing for $\alpha = 0.7$.

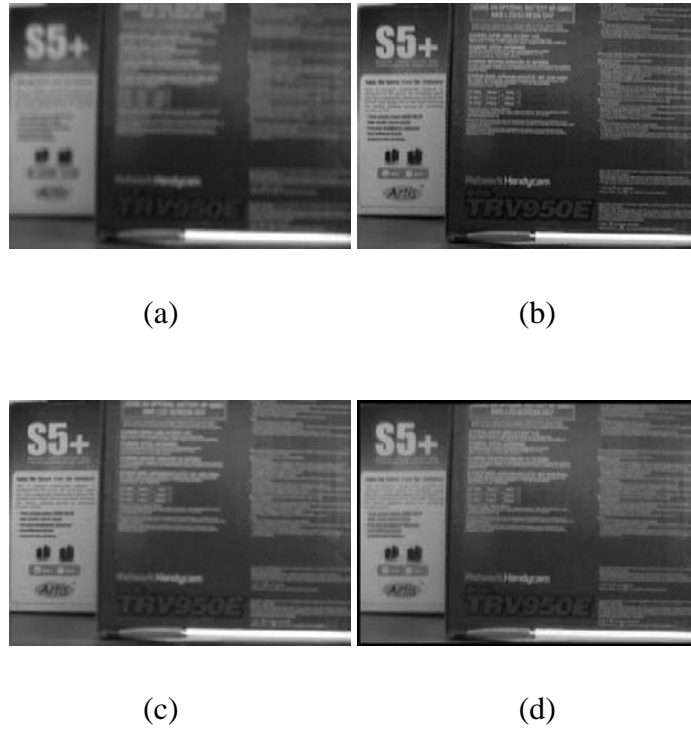


Fig. 7. (a,b) Two views of a scene under different focus settings. (c) Same scene captured with a focus setting lying in between the above two observations. (d) Result of synthetically generating the view in (c) using images (a and b) through blur morphing.

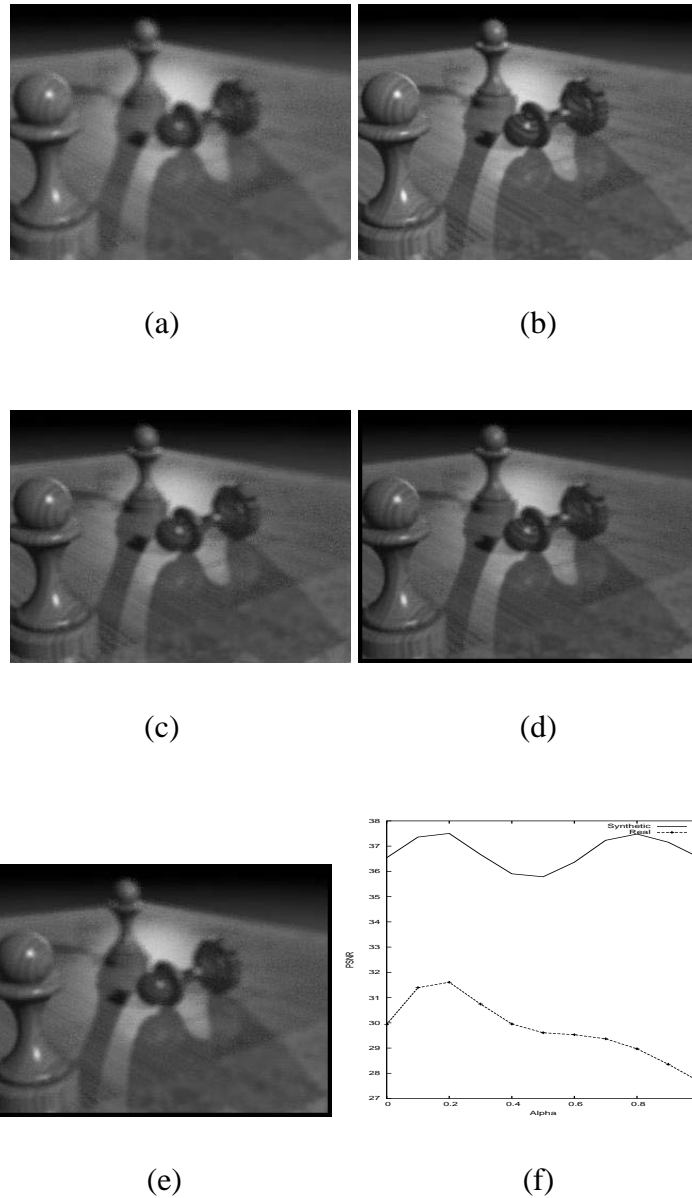


Fig. 8. (a,b) Two synthetically rendered views of a scene under different aperture settings. (c) Another rendered view for an aperture lying in between the above two settings. (d,e) Results of synthetically generating the view in (c) using images (a and b) through blur morphing. (f) The plot of PSNR against the morphing parameter α for both the current and the previous experiment.