M.I.T Media Laboratory Perceptual Computing Section Technical Report No. 259
Appears, Proc. First IEEE Int. Conf. on Image Proc., Austin TX, November 1994.

VIRTUAL BELLOWS: CONSTRUCTING HIGH QUALITY STILLS FROM VIDEO *

Steve Mann and Rosalind W. Picard
MIT Media Laboratory; 20 Ames Street; Cambridge, MA 02139
steve@media.mit.edu, picard@media.mit.edu

Abstract

Cameras with bellows give photographers flexibility for controlling perspective, but once the picture is taken, its perspective is set. We introduce 'virtual bellows' to provide control over perspective **after** a picture has been taken. Virtual bellows can be used to align images taken from different viewpoints, an important initial step in applications such as creating a high-resolution still image from video.

We show how the virtual bellows, which implements the projective group, is an exact model fit to both pan and tilt. Specifically, we identify two important classes of image sequences accomodated by the virtual bellows.

Examples of constructing high-quality stills are shown for the two cases: multiple frames taken of a flat object, and multiple frames taken from a fixed point.

1 INTRODUCTION

High-quality still images are not yet obtainable from portable electronic still-image cameras. However, multi-frame resolution enhancement can be used to provide a high-quality still image from an inexpensive video camera. The enhancement relies on the fact that typically there is some relative movement between the scene and the camera; movement is exploited in some way, using multiple frames of an image sequence to make a new image with higher resolution.

Tekalp, Ozkan, and Sezan [1] have assumed the movement between frames is translation: in their model, they assume that the image sensor has shifted by some small amount between frames, so one frame may be used to fill in some of the spaces between the samples of another. They note that image noise reduction can be applied either after this filling-in process, or concurrently, to further enhance the image. Others have assumed affine motion (six parameters) between frames [2] [3].

Consider an "ideal"[1] pinhole camera. We identify two cases where the affine model correctly describes the relationship between frames of an image sequence: **1**, *arbitrary static scene, camera at fixed location but free to rotate about its optical axis, lens free to zoom*; **2**, *single planar scene (e.g. flat terrain in aerial photography), both center of projection (COP) and image sensor free to translate (both laterally resulting in image shift, or along the optical axis resulting in a change in magnification) or rotate, provided that the plane of the image sensor remains parallel to the planar scene.* The affine model, however, does not correctly account for camera pan or tilt.

The projective model correctly describes two broader cases: **1**, *arbitrary static scene, camera at fixed location, camera free to rotate about its center of projection, (e.g. camera free to rotate about its optical axis, and to pan and tilt), lens free to zoom* ; **2**, *planar scene free to move arbitrarily, both center of projection and image sensor free to move arbitrarily.*

Tsai and Huang [4] have also explored the group structure associated with images of a 3-D rigid planar patch, as well as the associated *Lie algebra*, although they assume that explicit features have been located and that the correspondence problem has already been solved.

Previous work has been done to simply blend multiple pictures of a single scene [5], using a 2-D projective model. This work involved a search over the 8-parameter space to minimize the mean-square error (or maximize the inner product) between one frame and a 2-D projective coordinate transformation of the next frame, and did not rely on explicit feature correspondences. Szeliski and Coughlan [6] have more recently proposed a similar blending of images using an 8-parameter projective model.

In this paper we propose a means of resolution enhancement that does not require the tracking and correspondence of explicit features, yet runs fast enough to be computationally practical.

While our goal is to enhance 2-D projections (images) taken from a 3-D world, we first consider the problem of obtaining an enhanced 1-D "image", given multiple 1-D "images" each formed by projection from a 2-D world, where the "camera" consists of a center of projection and a line ("film"). The 1-D "images" are confined to a line within a planar world, which we call "Flatland" after the title of Abbott's book [7], which is a story about an alien culture living in a 2-D world. We explore the underlying *group* structure and provide some new insight, first for this simpler 2-D world; the extension to the 3-D world is then discussed, and experiments are demonstrated on video from 3-D scenes.

2 BACKGROUND: AFFINE GROUP

The affine model is often used as an approximation to perspective projection, and is somewhat adequate when the effective focal length of the lens is sufficiently large, and the camera is not panning excessively.

The affine mapping from $g \in G(\mathbb{R}^1)$, to $h \in H(\mathbb{R}^1)$ may be described by a change of coordinates: $x_2 = ax_1 + b$. The coordinate transformation so described is given by a *dilation* by amount a followed by a *translation* by an amount b. Such a mapping from function space G to function space H is

*This work sponsored in part by HP Labs and BT, PLC.
[1] Assumes infinite depth of field, and ignores diffraction.

known as an *operator*. This affine *operator* may itself be regarded as a function, namely a straight line of slope a and intercept b (Fig. 3(a)).

The set of all affine-coordinate transformations for which $a \neq 0$ forms a group, the *affine group*. This group of operators together with the set of 1-D images (operands) form a *group operation*[2]. The new set of images that result from applying all possible operators from the group, to a particular image from the original set, is called the *orbit* of that image under the group action. Given a set of images that lie in the same *orbit* of the affine group operation, we may find for each image pair, that element of the affine group, $\mathbf{p} = \{a, b\}$ which takes one image to the other image. In practice, due to noise, interpolation error and end effects, no element of the affine group will take one image to the other, so we find the parameters that make one image most closely match another: those parameters identify the desired group element.

Given corresponding sets of features in both images, we can derive the parameter, $\mathbf{p} = \{a, b\}$, that gives the mapping from x_1 to x_2 (e.g. find the slope and intercept of the line in Fig. 3(a) by linear regression).

When the change from one image to another is small, optical flow [9] is often used. A pair of 1-D images are related by a quantity, u at each point in one of the images, that indicates the *flow* to the next image (e.g. units of velocity):

$$uE_x + E_t + h.o.t. = 0 \tag{1}$$

where E_x and E_t are the spatial and temporal derivatives respectively, and $h.o.t.$ denotes higher order terms in the underlying Taylor series representation upon which *optical flow* is based.

Assuming the affine model, $u = ax + b$, in (1), summing the squared error over the whole image, differentiating, and equating the result to zero, gives a linear solution for both a and b.

3 PROJECTIVE GROUP

Suppose we take two (1-D) pictures of the same scene from a common location, where the camera is free to pan and zoom between taking the two pictures, but it cannot move to a new location. We define the common COP at the origin of our coordinate system in the plane. In Fig. 1 we have depicted a single camera that takes two pictures in succession as two cameras shown together in the same figure. Let $Z_k, k \in \{1, 2\}$ represent the distances, along each optical axis, to an arbitrary point in the scene, P, and let X_k represent the distances from P to each of the optical axes. The principal distances are denoted z_k. In the example of Fig. 1, we are *zooming in* (increased magnification) as we go from frame 1 to frame 2.

The geometry of Fig. 1 defines a mapping from x_1 to x_2, given by [10],[11]:

$$\begin{aligned} x_2 &= z_2 \tan(\arctan(x_1/z_1) - \theta), \quad \forall x_1 \neq o_1 \\ &= (ax_1 + b)/(cx_1 + 1), \quad \forall x_1 \neq o_1 \end{aligned} \tag{2}$$

where $a = z_2/z_1$, $b = -z_2 \tan(\theta)$, $c = \tan(\theta)/z_1$, and $o_1 = z_1 \tan(\pi/2 + \theta) = -1/c$, is the location of the singularity in the domain ('appearing point[3]' [11]). We should

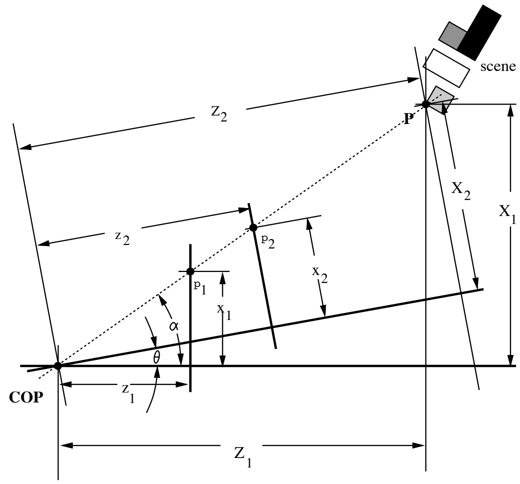

Figure 1: Camera at a fixed location. An arbitrary scene is photographed twice, each time with a different camera orientation, and a different principal distance (zoom setting). In both cases the camera is located at the same place (COP). The dotted line denotes a ray of light traveling from an arbitrary point, **P**, in the scene, to the **COP**. Heavy lines denote both camera optical axes in each of the two orientations as well as the image sensor in each of its two pan and zoom positions. The two image sensors are in front of the camera to simplify mathematical derivations.

emphasize here that if we set $c = 0$ we arrive at the affine group, and that c, the degree of perspective, has been given the interpretation of a chirp-rate [10] – uniformly spaced points in x_1 are mapped to "chirped" points in x_2.

Let $\mathbf{p} \in \mathbf{P}$ denote a particular mapping from x_1 to x_2, governed by the three parameters $\mathbf{p}' = [z_1, z_2, \theta]$, or equivalently by a, b and c from (2).

Proposition 1 *The set of all possible operators,* $\mathbf{P}_1$*, given by the coordinate transformations (2),* $\forall a \neq bc$*, acting on a set of 1-D images, forms a group-operation.*

Proof: A pair of images produced by a particular camera rotation and change in principal distance (depicted in Fig. 1) is an operator that takes any function g on image line 1, to a function, h on image line 2:

$$\begin{aligned} h(x_2) &= g(x_1) = g((-x_2 + b)/(cx_2 - a)), \quad \forall x_2 \neq o_2 \\ &= g \circ x_1 = g \circ \mathbf{p}^{-1} \circ x_2 \end{aligned} \tag{3}$$

where $\mathbf{p} \circ x = (ax + b)/(cx + 1)$ and $o_2 = a/c$. As long as $a \neq bc$, each operator, $\mathbf{p}$, has an inverse, namely that given by composing the inverse coordinate transformation:

$$x_1 = (b - x_2)/(cx_2 - a), \quad \forall x_2 \neq o_2 \tag{4}$$

with the function $h()$ to obtain $g = h \circ \mathbf{p}$. The identity operation is given by $g = g \circ e$, where e is given by $a = 1$, $b = 0$, and $c = 0$.

In complex analysis, (see for example, Ahlfors [12]) the form $(az + b)/(cz + d)$ is known as a linear fractional transformation. Although our mapping is from $\mathbb{R}$ to $\mathbb{R}$ (as opposed to theirs from $\mathbb{C}$ to $\mathbb{C}$), we can still borrow the concepts of complex analysis. In particular, a simple group-representation is provided using the 2×2 matrices, $\mathbf{p} = [a, b; c, 1] \in \mathbb{R}^2 \times \mathbb{R}^2$. Closure[4] and associativity are obtained by using the usual laws of matrix multiplication followed with dividing the resulting vector's first element by its second element. □

[2] also known as a *group action* or *G-set* [8].

[3] Single quotes denote terms coined by the authors here or elsewhere in the literature.

[4] Also know as *law of composition* [8]

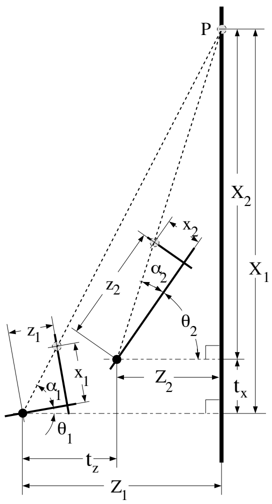

Figure 2: Two pictures of a flat (straight) object. The point P is imaged twice, each time with a different camera orientation, a different principal distance (zoom setting), and different camera location (resolved into components parallel and perpendicular to the object).

Proposition 1 says that an element of the $(ax+b)/(cx+1)$ group can be used to align any two frames of the (1-D) image sequence provided that the COP remains fixed.

Proposition 2 *The set of operators that take nonsingular projections of a straight object to one another form a group,* $\mathbf{P_2}$.

A "straight" object is one which lies on a straight line in Flatland.

Proof: Consider a geometric argument. The mapping from the first (1-D) frame of an image sequence, $g(x_1)$ to the next frame, $h(x_2)$ is parameterized by the following: camera translation perpendicular to the object, t_z; camera translation parallel to the object, t_x; pan of frame 1, θ_1; pan of frame 2, θ_2; zoom of frame 1, z_1; and zoom of frame 2, z_2. (See Fig. 2.) We want to obtain the mapping from x_1 to x_2. Let's begin with the mapping from X_2 to x_2:

$$x_2 = z_2 \tan(\arctan(X_2/Z_2) - \theta_2) = \frac{a_2 X_2 + b_2}{c_2 X_2 + 1} \quad (5)$$

which can be represented by the matrix $\mathbf{p}_2 = [a_2, b_2; c_2, 1]$, so that $x_2 = \mathbf{p}_2 \circ X_2$. Now $X_2 = X_1 - t_x$ and it is clear that this coordinate transformation is inside the group, for there exists the choice of $a = 1$, $b = -t_x$, and $c = 0$ that describe it: $X_2 = \mathbf{p}_t \circ X_1$, where $\mathbf{p}_t = [1, -t_x; 0, 1]$. Finally, $x_1 = z_1 \tan(\arctan(X_1/Z_1) - \theta) = \mathbf{p}_1 \circ X_1$. Let $\mathbf{p}_1 = [a_1, b_1; c_1, 1]$. Then $\mathbf{p} = \mathbf{p}_2 \circ \mathbf{p}_t \circ \mathbf{p}_1^{-1}$ is in the group by the law of composition. Hence, the operators that take one frame into another, $x_2 = \mathbf{p} \circ x_1$, form a group. □

Proposition 2 says that an element of the $(ax+b)/(cx+1)$ group can be used to align any two images of linear objects in flatland, regardless of camera movement.

Proposition 3 *The two groups* $\mathbf{P_1}$ *and* $\mathbf{P_2}$ *are isomorphic; a group-representation for both is given by the* 2×2 *square matrix* $[a, b; c, 1]$.

Isomorphism follows because $\mathbf{P}_1$ and $\mathbf{P}_2$ have the same group representation. The $(ax+b)/(cx+1)$ operators in the above propositions form the *projective group* $\mathbf{P}$ in Flatland.

Previously we emphasized the fact that the affine operator that takes a function space G to a function space H may

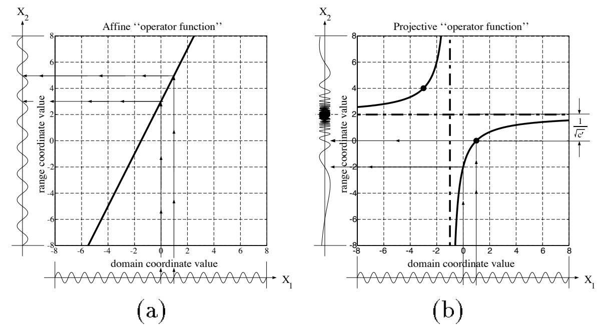

Figure 3: Comparison of 1-D affine and projective coordinate transformations, in terms of their 'operator functions', acting on a sinusoidal image. (a) Orthographic projection is equivalent to affine coordinate transformation, $y = ax + b$. In this example, $a = 2$ and $b = 3$. (b) Perspective projection for a particular fixed value of $\mathbf{p}' = \{1, 2, 45°\}$. Note that the plot is a rectangular hyperbola like $x_2 = 1/(c' x_1)$ but with asymptotes at the shifted origin $(-1, 2)$. Here $g(x_1) = \sin(2\pi x_1)$. The arrows indicate how a chosen cycle of this sine wave is mapped to the corresponding cycle of the 'P-chirp', $h(x_2)$.

itself be viewed as a function. Let us now construct a similar plot for a member of the group of operators, $\mathbf{p} \in \mathbf{P}$, in particular, the operator $\mathbf{p} = [2, -2; 1, 1]$ which corresponds to $\mathbf{p}' = \{1, 2, 45°\} \in \mathbf{P}_1$. We have also depicted the result of mapping $g(x_1) = \sin(2\pi x_1)$ to $h(x_2)$. When G is the space of Fourier analysis functions (harmonic oscillations), then H is a family of functions known as P-chirps [10], adapted to a particular *vanishing point*, o_2 and 'normalized chirp-rate', $c' = c^2/(bc - a)$ [11]. Fig. 3(b), is a *rectangular hyperbola* (e.g. $x_2 = \frac{1}{c' x_1}$) with an origin that has been shifted from $(0, 0)$ to (o_1, o_2).

3.1 Feature matching

Given at least three correspondences between point pairs in the two images, we can use a simple "feature matching" procedure to find the element, $\mathbf{p} = \{a, b, c\} \in \mathbf{P}$ that maps the second image into the first. Let $x(k), k = 1, 2, 3, ...$ be the points in one image, and let $u(k)$ be the corresponding points in the other image. Then:

$$u(k) = \frac{a x(k) + b}{c x(k) + 1} \quad (6)$$

can be rearranged into k linear equations in the 3 unknowns, a, b, and c:

$$\begin{bmatrix} x(k) & 1 & -u(k)x(k) \end{bmatrix} \begin{bmatrix} a & b & c \end{bmatrix}^T = \begin{bmatrix} u(k) \end{bmatrix} \quad (7)$$

and solved using least squares if there are at least three correspondence points.

3.2 A new perspective on optical flow

Applying the projective group model, $u = (ax+b)/(cx+1)$, to the optical flow equation (1), we obtain a set of equations that is difficult to solve. We may, however, expand $u = (ax+b)/(cx+1)$ in its own Taylor series about the identity (the point $a = 1, b = 0, c = 0$), which turns out to give the same result as the univariate Taylor series about $x = 0$:

$$u = b + (a - bc)x + (bc - a)cx^2 + (a - bc)c^2x^3 + ... \quad (8)$$

Taking only the first 3 terms, we have a representation for u that contains enough degrees of freedom to account for the 3 parameters being estimated. Letting $\epsilon = \sum(-h.o.t.)^2 = \sum((b + (a - bc)x + (bc - a)cx^2)E_x + E_t)^2$, and differentiating

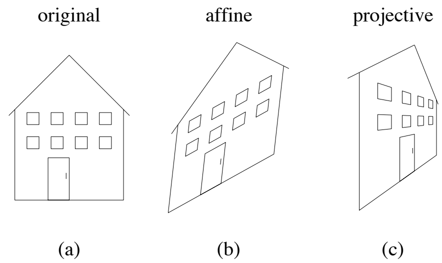

Figure 4: 2-D coordinate transformations. (a) Original, (b) Affine coordinate transformation leaves parallel lines parallel and periodic structures periodic: an equally-spaced row of windows remains equally-spaced. (c) The 'virtual bellwos' (projective coordinate transformation) does not in general preserve either parallelism or periodicity. It "chirps" periodic structures. In this example, the spatial frequency increases from left to right.

with respect to each of the 3 parameters, setting the derivatives equal to zero, and verifying with the second derivatives, gives the linear system of equations:

$$\begin{bmatrix} \sum x^4 E_x & \sum x^3 E_x & \sum x^2 E_x \\ \sum x^3 E_x & \sum x^2 E_x & \sum x E_x \\ \sum x^2 E_x & \sum x E_x & \sum E_x \end{bmatrix} \begin{bmatrix} (bc-a)c \\ a-bc \\ b \end{bmatrix} = - \begin{bmatrix} \sum x^2 E_t \\ \sum x E_t \\ \sum E_t \end{bmatrix} \quad (9)$$

3.3 Projective group model for 2-D

The incorporation of perspective into registration and resolution enhancement applies readily to 2-D images and video. The usual affine model is augmented with an additional parameter, $\mathbf{c} = [c_x, c_y]$, which, again has the interpretation of mapping a uniform pattern to a "chirp" [10] but now with two components.

The fixed COP case, $\mathbf{P}_1$, extends to a 4-parameter group based on $\mathbf{c}$, rotation about the optical axis by ϕ, and zoom by z.

The planar object case, $\mathbf{P}_2$, exists inside an 8-parameter operator group based on the coordinate transformations:

$$\mathbf{u} = \begin{bmatrix} u \\ v \end{bmatrix} = \frac{\mathbf{A}[x,y]^T + \mathbf{b}}{\mathbf{c}^T[x,y]^T + 1} = \frac{\mathbf{A}\mathbf{x} + \mathbf{b}}{\mathbf{c}^T\mathbf{x} + 1} \quad (10)$$

where the *group representation* may be obtained, again using 2×2 matrices, $\mathbf{p} = [\mathbf{A}, \mathbf{b}; \mathbf{c}, 1]$, but where $\mathbf{A} \in \mathbb{R}^{2\times 2}$, and $\mathbf{b}, \mathbf{c} \in \mathbb{R}^{2\times 1}$. An example of the group-action of (10) is illustrated in Fig. 4 together with an exemplar affine group-action.

A useful interpretation of the 8 parameters of this group are as follows: 6 of the parameters correspond to the affine group, and the remaining two have the interpretation of chirp rate [10].

Analogous to the 1-D case, the element $\mathbf{p}$, that maps one image to another, may be found from explicit feature correspondences by solving a system of 8 (or more) linear equations in 8 (or more) unknowns. Four (or more) point correspondences are required in 2-D.

We may also use (10) to derive perspective optical flow for 2-D images. In particular, the multivariate Taylor series expansion of (10) takes on the form:

$$\begin{aligned} u &= p_u + p_{ux}x + p_{uy}y + p_{uxy}xy + p_{uxx}x^2 + p_{uyy}y^2 + \ldots \\ v &= p_v + p_{vx}x + p_{vy}y + p_{vxy}xy + p_{vxx}x^2 + p_{vyy}y^2 + \ldots \end{aligned} \quad (11)$$

There are different ways [11] of relating the parameters in (11) to the three parameters (eight scalar parameters) $\mathbf{A}, \mathbf{b}$, and $\mathbf{c}$ using a hierarchical and recursive (iterative) approach.

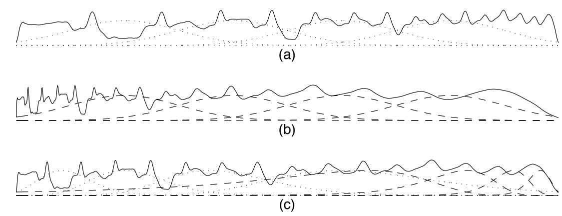

Figure 7: Two pictures of the same object, together with point spread functions for four samples (pixels). (a) Frame 1 of the image sequence. (b) Frame 2 of the sequence. (c) Frames 1 and 2 registered. Note that the periodic pixel lattice from frame 1 (dotted) is "downchirped" (decreases in spatial frequency from left to right), and the pixel lattice from frame 2 (dashed) is "upchirped".

This process amounts to automatically aligning (*registering*) the various frames without using explicit features. We should emphasize that while the Taylor series is an approximation, the recursion [11] is done using the exact projective model (approximate feedforward, exact feedback) leading to a system where very large image motions can be estimated accurately.

4 HIGH RESOLUTION STILLS FROM VIDEO

We now apply the 'virtual bellows' model to resolution enhancement for each of the two cases depicted in Figs. 1 and 2. In the first example, all the images were taken from a videotape where the video camera was swung around the center of a computer room. The camera translation was small compared to the distance to the nearest objects in the scene, so it fits the model depicted in Fig. 1. A panoramic image was constructed on the unit sphere, and a Mercator projection of the enhanced image data is presented in Fig. 5.

The second case, that of a nearly planar object and arbitrary camera movement is shown in Fig. 6. Four frames of an image sequence of a circuit board appear together with the registered frames and the enhanced image. Since the circuit board is almost flat, it is a close fit to the model depicted in Fig. 2, and is well-described by the virtual bellows.

In both cases the method we used was similar to that of Irani and Peleg [2], except that the proposed virtual bellows (projective group) model was used, instead of the affine group model.

4.1 The "chirping" point spread function

Estimating and working with the point spread function (PSF) is an important aspect of resolution enhancement. A single point in the continuous scene affects a neighborhood in the image sensor, rather than just a single sensor point. This blurring effect is characterized by the PSF, and is often assumed to be independent of location on the image sensor.

In Fig. 7(a) and (b) we show a typical situation in 1-D where we have a continuous image that is sampled by a four-pixel camera, with a sampling much coarser than the details of the image, as is generally the case. The pixel sampling lattice would remain periodic (equally spaced samples) under affine coordinate transformations, but the virtual bellows induce a "chirping" that is evident in Fig. 7(c) where both images have been 'dechirped' (registered) to a frame of reference between the two views. As the images are 'dechirped', the PSFs associated with each image become 'rechirped' onto the no-longer periodic lattice.

Irani and Peleg [2] have proposed the use of a small dot to characterize the PSF. We found, however, in attempting

Figure 5: Mercator projection of resolution enhanced image data from multiple video frames all taken from approximately the same point in space, namely the center of a typical computer room. The resulting image is more than 4000 pixels across, and the detail is quite well-resolved, even when enlarged to a width of two or three meters.

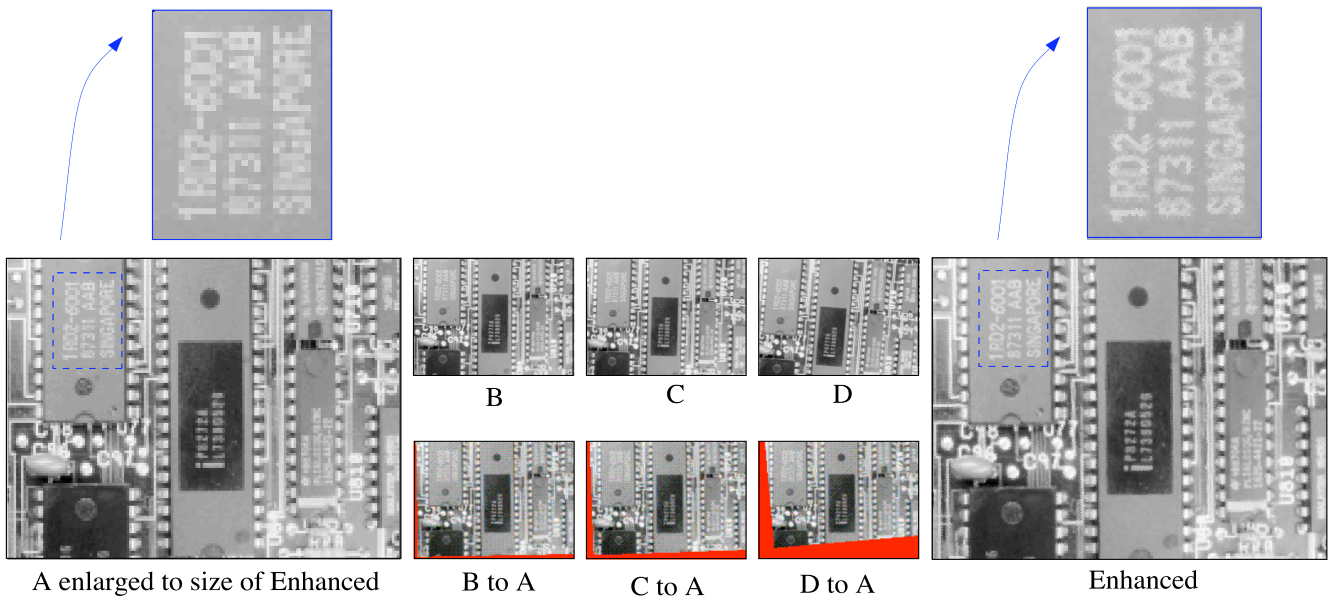

Figure 6: Example of resolution enhancement applied to a planar object: low-resolution 4-frame image sequence (left and center) is registered (center). Parameters determined from the registration process are used to map the original unregistered images onto a quarter-pixel grid (right). For comparison purposes, the first frame of the image sequence is enlarged to the same size as the enhanced image.

to measure the PSF, that the result depended greatly on the position of the dot relative to the pixel boundary. Furthermore, this method only gives discrete samples of the PSF. Therefore, we used a different approach. We attached a circular piece of white paper (filter paper used in chemistry labs is ideal because it is very neatly cut and is almost perfectly non-specular) to a piece of black cardboard. A picture of this was taken under uniform lighting. Thresholding the image, we determined the boundary of the transition from white to black. This boundary provided a step edge in all possible directions, though we used a Fourier-based approach to integrate all this information into a single continuous 2-D estimate of the PSF.

5 SUMMARY

Two cases of camera motion have been treated with the 'virtual bellows', both of which allow for rotation, zoom, pan, tilt, and translation of the film plane. In the first case, the solution is for a planar object and arbitrary COP; in the second case the solution is for an arbitrary object and fixed COP. Theory was discussed for 1-D and 2-D, and successful image alignment, overlap-mosaicing, and enhancement demonstrated on still image and video.

Acknowledgement

Thanks to Alex Drukarev, Jeanne Wiseman, and Shawn Becker for many interesting discussions.

References

[1] A.M. Tekalp, M.K. Ozkan, and M.I. Sezan. High-resolution image reconstruction from lower-resolution image sequences and space-varying image restoration. In *Proc. of the Int. Conf. on Acoust., Speech and Sig. Proc.*, pages III–169, San Francisco, CA, Mar. 23-26, 1992. IEEE.

[2] M. Irani and S. Peleg. Improving Resolution by Image Registration. *Graphical Models and Img. Proc.*, May 1991.

[3] L. Teodosio and W. Bender. Salient video stills: Content and context preserved. *Proc. ACM Multimedia Conf.*, 1993.

[4] R. Y. Tsai and T. S. Huang. Estimating Three-Dimensional Motion Parameters of a Rigid Planar Patch . *Trans. Accoust., Speech, and Sig. Proc.*, 1981.

[5] S. Mann. Compositing multiple pictures of the same scene: Generalized large-displacement 8-parameter motion. In *Proceedings of the 46th Annual IS&T Conference*, Cambridge, Massachusetts, May 9-14 1993. The Society of Imaging Science and Technology.

[6] R. Szeliski and J. Coughlan. Hierarchical spline-based image registration. *CVPR*, 1994.

[7] Edwin A. Abbott. *Flatland*. Signet Classic, June 1984.

[8] M. Artin. *Algebra*. Prentice Hall, 1991.

[9] B. Horn and B. Schunk. Determining Optical Flow. *Artificial Intelligence*, 1981.

[10] S. Mann. Wavelets and chirplets: Time–frequency perspectives, with applications. In Petriu Archibald, editor, *Advances in Machine Vision, Strategies and Applications*. World Scientific, 1992.

[11] S. Mann and R.W. Picard. The virtual bellows: A new perspective on the rigid planar patch. TR 260, M.I.T. Media Lab Perceptual Computing Section, Cambridge, Ma, Jan 1994.

[12] L.V. Ahlfors. *Complex Analysis*. International Series in Pure and Applied Mathematics. McGraw Hill, Inc., 3rd edition, 1979.