Computational Aesthetics in Graphics, Visualization, and Imaging (2009)
O. Deussen and P. Hall (Editors)

ARTcams: Attributed Rational Tensor Cameras

Chuan Li, Peter Hall and Philip Willis

Media Technology Research Centre
University of Bath

Abstract

Non-linear camera models are playing an increasingly important role in computer graphics, especially in image based rendering and non-photorealistic rendering. We introduce ARTcams as simple non-linear cameras, which are unique in combining both geometric projection and non-geometric attributes such as colour into a single model. The geometric component of an ARTcam subsumes many contemporary non-linear cameras, including General Linear Cameras, push-broom cameras, and X-slit cameras. The colour component generalises compositing operations. ARTcams, though, by combining geometry and other attributes generalise yet further. ARTcams can be thought of as lenses (or mirrors) that can reproduce a wide variety of real effects, including aerial perspective, depth of field, as well as both geometric and chromatic aberrations. They can be calibrated for both geometry and colour against real optical devices. It is possible to specify ARTcams by drawing alone. This paper explains and demonstrates the ARTcam model.

Categories and Subject Descriptors (according to ACM CCS): Computer Graphics [I.3.3]: Image processing software—

1. Introduction

We introduce a simple non-linear camera model we call the Attributed Rational Tensor Camera (ARTcam). The camera operates on points, and works equally well in both object space and image space. In this paper we use images as input, so the points are pixels. ARTcams can be thought of as lenses (or mirrors, or general optical devices). ARTcams are unique in combining geometry with non-geometric attributes, so that they can affect both the shape and colour of an object.

This paper briefly reviews the literature on camera models. It concentrates on non-linear camera models, which cause objects to look warped, because these transformations play an important role and are the foundation of ARTcams. The paper then outlines the basic mathematics underpinning the ARTcams model, including calibration. We then show examples to demonstrate its use and versatility in practice.

Renderers often use the ideal pin-hole camera as a projection model. The pin-hole camera is well understood mathematically and is conveniently simple to integrate into the rendering pipeline because it can be written as a matrix. The pin-hole camera needs no lens at all — a tiny hole in the side of a box is enough. The consequence of this is, of course, that lens aberrations cannot be reproduced. This problem has already been addressed in Computer Graphics. One of the earliest lens models, by Potmesil and Chakravarty [PC81], operated in image space. A little later, object space algorithms were published. Distributed ray-tracing [CPC84] is one of the best known and allowed lens effects such as depth of field and blur to be produced. Sophisticated models of real camera lenses appeared later [KMH95, HS97]. Nonetheless, all of these assume pin-hole projection; that is, a linear camera. It is only more recently that non-linear projection has been considered — principally in the Non-Photorealistic Rendering and Computer Vision literatures.

Non-Photorealistic Rendering (NPR) is a recent aim of Computer Graphics concerned with synthesising artwork. NPR algorithms come in two classes: (a) algorithms that use three-dimensional models as input, so act in object space; (b) algorithms that use photographs or video of the real world as input and so operate in image space. Both classes have been centrally concerned with modeling painting media (oil paint or pencil, for example) and placing marks (typically to retain important detail). Only researchers in class (a) — NPR from

DOI: 10.2312/COMPAESTH/COMPAESTH09/073-081

models — have published research into non-linear cameras [CS04, YM04, AZM00, Lev98]. Research in NPR from photographs, class (b), has tacitly accepted the projective system of the real camera, which is well approximated by a pin-hole camera, at least up to lens aberrations.

We are most interested in class (b), which we call Non-Photorealism from Photographs (NPRP). However, we are not interested in making or placing marks. Instead, our interests concentrate upon the problem of introducing novel lens dependent effects into photographs. We output images of photographic quality, which gives two benefits. (1) We can use our techniques for tasks such as inserting graphics into real world images. (2) We can use existing NPRP methods to make marks upon our photographic output. This division between projection and mark-making reflects the division between the *projective system* and *denotational system* in art, introduced by Art commentator John Willats [Wil97]. Willats shows how the style of different schools of Art depends as much on their projective model as it does in the way they make marks and the physical media they use. It is, therefore, important for NPRP to take projective systems into account — as we do in this paper.

The aim of our research is to find a simple camera model capable of emulating both real world cameras and the projection systems used by artists. Moreover, we want to do this not just for geometrical elements, which capture distortions, but for non-geometrical elements too. We want to be able both to analyse and to synthesise common optical effects such as aerial perspective and depth of field. Our camera should be capable of calibration to specific real optical devices. It should also synthesise non-geometric effects no real camera can, but which are useful to artists. Ideally, we want a camera that treats photorealism as a special case of non-photorealism. Finally, our camera should be mathematically well principled so we can analyse and understand both it and its relation to other non-linear camera models.

Non-linear cameras remove the requirement that rays of light must pass through a single focal point. In fact non-linear cameras may contain any number of real-valued focal points, including zero or infinitely many. They are the subject of a considerable volume of contemporary study and are used in many different ways in both Computer Graphics and Computer Vision. Push-broom cameras [GH97] use a line of foci, rather than a single point. These are related to strip-cameras in which a moving camera acquires very thin pictures, mosaicing them into a final image [AAC*06, RL06, RGL04, RB88]. X-slit cameras use two lines of foci. They are the equivalent of two strip cameras, and so find use in stereopsis, but strip cameras can be treated as single entity. X-slit cameras can be used for novel mosaicing too [ZFPW03]. General Linear Cameras (GLCs) [YM04] subsume both push-broom and X-slit cameras. GLCs erect a light field by specifying three points, each with a ray — the remaining rays are located and oriented by interpolating barycentric coordinates. GLCs map planar surfaces to biplanar surfaces, and so rays of light appear to bend, but they cannot account for radial distortions that appear in real cameras, such as pin-cushion and barrel distortions. Rational Function cameras (RFcams) [CF05] are the ratio of quadratic functions and are more general than GLCs. They can be calibrated to match the radial distortion of a real camera, so the effect can be either removed or added, as desired. Calibrating such cameras more generally, using tensors, is an open issue being addressed [RSL05].

We build on non-linear cameras designed to warp geometry, called the Rational Tensor camera or RTcam [HCS*07]. The RTcam subsumes the linear camera and many non-linear cameras as special cases, including GLC, bush-broom, X-slit and others. It can be calibrated to emulate real cameras and reproduces non-physical projective systems used by artists. However, like all the non-linear cameras it affects geometry only, by which we mean it appears to warp objects (so geometry is synonymous with location in this paper). No other attribute of an object is affected. In this paper we introduce a generalisation of the RTcam model. We introduce the Attributed Rational Tensor camera, or ARTcam which makes the following contributions:

- We established the dependence between general attributes, both geometric and non-geometric. Hence a much wider gamut of effects can be created in the ARTcam model.
- We provide accurate calibration, through which many of the non-linear lens effects can be captured, allowing ARTcams to insert graphics into real scenes.
- We allow artists to specify ARTcams by drawing; the analogy to lenses means the drawings are easy and intuitive.
- The camera is mathematically well founded. It subsumes not only many prior non-linear camera models but various compositing models too.
- The camera operates equally well in object space and image space.

Figure 1 shows an ARTcam acting on a model of a Rubik cube to produce a variety of different effects. In this case the ARTcams were specified by the user explicitly setting numbers — a method useful for expert users, though we provide calibration and drawing as preferred alternatives.

ARTcams operate on points in either image space (two dimensions of space plus one homogeneous dimension) or object space (with an extra spatial dimension). The model of the cube in Figure 1 is a cloud of coloured points in object space, produced from original photographs using stereo reconstruction techniques from the Computer Vision literature. Object space points could come from any other suitable source. Nearly all other examples in this paper act on image space points, that is pixels in homogeneous space. We now describe the Attributed Rational Tensor camera.

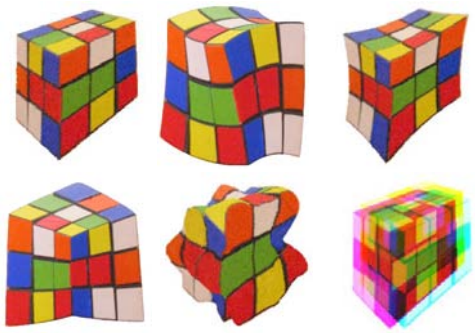

Figure 1: *ARTcams applied to a three-dimensional point cloud, constructed using stereo computer vision. Top left is an original cube. Top middle and top right show non-linear warping in 3D; twist and pin-cushion distortion respectively. Bottom left shows inverse perspective, in which objects magnify as the recede. Middle bottom shows we can use colour to control scale. We might interpret colour as heat – warm colour expands the object and cold colour shrinks the object. Bottom right shows a chromatic aberration effect, found in some lenses and occasionally seen in mis-aligned prints.*

2. ARTcams

We denote an n-dimensional real space by $\Re^n$ and the corresponding projective space by P^n: if $\mathbf{x} \in P^n$, then it has $n+1$ elements, only n of which are independent for $\mathbf{x} \equiv \lambda\mathbf{x}$ for all $\lambda \neq 0$. Typically we set $\lambda = 1/x_{n+1}$ to transform a point in proejctive space to one in real space.

A Rational Tensor camera (RTcam [HCS*07]) is a homogeneous mapping that carries a vector $\mathbf{x} \in P^n$ to another vector $\mathbf{y} \in P^n$, n. Each element of $\mathbf{y}$ is defined as the ratio of quadratic functions over the elements in $\mathbf{x}$. This is conveniently specified using an array of matrices, $\mathbf{Q}_i$. An RTcam is therefore defined by the mapping

$$y_i = \frac{\mathbf{x}\mathbf{Q}_i\mathbf{x}^T}{\mathbf{x}\mathbf{Q}_n\mathbf{x}^T} \tag{1}$$

The matrices $\mathbf{Q}_i$ are the planes in a cube of numbers, which is a tensor. For example, given the homogeneous location of a pixel in image space, $\mathbf{x} = (x_1, x_2, x_3)$, so $n = 3$, we get

$$y_i = \frac{\Sigma_{j=1}^3 \Sigma_{k=1}^3 Q_{ijk} x_j x_k}{\Sigma_{j=1}^3 \Sigma_{k=1}^3 Q_{3jk} x_j x_k} \tag{2}$$

as the ith coordinate of the output pixel. This mapping is clearly a projection. The standard linear camera, push-broom cameras, X-slit camera, GLCs, and RFcams are all contained as special cases, as proven in [HCS*07].

2.1. Extend RTcams to General Attributes

Attributed Rational Tensor cameras (ARTcams) are similarly defined, but the inclusion of terms of different kinds (geometry, colour, etc) creates a problem — different components require division by different homogeneous terms. This problem can be solved in a simple and obvious way. First consider the mapping from n-dimensional space to n-dimensional space:

$$z_i = \mathbf{x}\mathbf{Q}_i\mathbf{x}^T \tag{3}$$

Now, just divide each component by its appropriate homogeneous element. For example, suppose we are interested in spatial location and colour attributes. The vector input to the ARTcam $\mathbf{x}$, can now be decomposed into two parts

$$\mathbf{x} = [\mathbf{p}|\mathbf{c}] \tag{4}$$

in which $\mathbf{p}$ is the location in homogeneous coordinates and $\mathbf{c}$ is colour, also in homogeneous coordinates. The output is now just the vector $\mathbf{z}$, as given by Equation 3, which shares exactly the same partition as $\mathbf{x}$. Let us write $\mathbf{z} = [\mathbf{q}|\mathbf{d}]$. Now, we divide $\mathbf{q}$ and $\mathbf{d}$ by their homogeneous elements, q_h and d_h respectively to obtain a final result

$$\mathbf{u} = \left[\frac{\mathbf{q}}{q_h} | \frac{\mathbf{d}}{d_h}\right] \tag{5}$$

We note $\mathbf{y} = \mathbf{q}/q_h$, so the RTcam model is a special case of the ARTcam model.

There is already a precedent for using colour with homogeneous terms: the projective alpha colour model described by Willis [Wil07], so both $\mathbf{p}$ and $\mathbf{c}$ belong to projective spaces. Now we will show that the ARTcam model generalises not only the RTcam model, but projective alpha colour too. We begin again by partitioning the input vector into location and colour $\mathbf{x} = [\mathbf{p}|\mathbf{c}]$ We also partition the matrices of the ARTcam into four sub-matrices

$$\mathbf{Q}_i = \begin{bmatrix} \mathbf{Q}_{ipp} & \mathbf{Q}_{ipc} \\ \mathbf{Q}_{icp} & \mathbf{Q}_{icc} \end{bmatrix} \tag{6}$$

so that we can easily see how geometry and colour relate

$$z_i = \mathbf{p}\mathbf{Q}_{ipp}\mathbf{p}^T + \mathbf{p}\mathbf{Q}_{ipc}\mathbf{c}^T + \mathbf{c}\mathbf{Q}_{icp}\mathbf{p}^T + \mathbf{c}\mathbf{Q}_{icc}\mathbf{c}^T \tag{7}$$

Clearly $\mathbf{Q}_{ipp}$ affects geometry alone. This is the RTcam component. The $\mathbf{Q}_{icc}$ affects colour alone, and so can be used for both linear and non-linear compositing operations. This component, therefore, generalises projective alpha colour. It is a novel contribution, so far as we know. The $\mathbf{Q}_{ipc}$ and $\mathbf{Q}_{icp}$ mix geometry and colour, which is a second novel component so far as we know. It is these components which create a dependency between attributes of different kinds, geometry and colour in this example.

3. ARTcam examples

ARTcams can be used to create a picture with any number of focal points from zero to infinity. One way to think about this is to begin with a collection of photographs of a scene, each taken with a real camera. We will allow ourselves to believe the real-cameras have a single point as their focus. Now we can consider the collection of cameras as a single

Figure 2: *A stereo pair is input to an ARTcam which creates a picture from a pair of orthogonal views. The output is a photographic quality image (not shown) that can be processed to some NPR style, here crayon is shown.*

Figure 3: *ARTcams creating optical and atmospheric effects. Top left: original photograph. Top right: with aerial perspective. Bottom left: with depth-of-field. Bottom right: with depth-of-field and aerial perspective.*

Figure 4: *Colour aberration as an example of interacting between geometry and colour. On the left is the original input image; in the middle is the colour aberration image whose red, green and blue layers have been shifted by different amount using RTcams; on the right is another colour aberration image made by ARTcams, who shift pixels based on their colour. The difference between RTcams and ARTcams can be seen from the geometric distortion which only exists in the right image.*

light-collecting device; this device has many focal points. If these focal points all lie on a quadric surface, then the device can be modeled by an ARTcam. If the quadric surface is at infinity, then the device has no real focal points, resulting in orthogonal projection.

The example in 2 shows that the projective system of real cameras can be switched to orthogonal projection. We input a stereo pair and output an "unfolded view" of the walls. The roof was stretched to fit into the unfolded walls. The output was of photographic quality, but this was then further processed to produce the crayon effect shown here. The crayoning was chosen to emphasise the child-like projective system. Apart from marking up the walls, this process was automatic — and it could not easily be done by simply warping a single photograph. This example is more fully explained elsewhere [HCS*07], along others examples and some theory that underpins ARTcam projective geometry.

The above example uses only the geometric part of the ARTcam. But other ARTcam terms link geometry to colour, or indeed other attributes we choose. To assist the readers to understand what kind of effect can be produced by ARTcams, we will first give a example in which there is a direct dependence between location and colour.

Aerial perspective is the visual effect whereby the colour of distant objects appears more blue and less saturated. This is an atmospheric effect that requires a complicated light scattering calculation to model accurately, which would be an object-space operation: a non-geometric attribute depends on a geometry attribute. The $\mathbf{Q}_{ipc}$ and $\mathbf{Q}_{icp}$ terms capture this, and because $\mathbf{Q}_{ipc} = \mathbf{Q}_{icp}^T$ we need only discuss one of them.

For simplicity, we consider contrast only. This will fall with the logarithm of distance. We fitted an ARTcam to approximate a logarithmic function (so a line of equally spaced input points map to a line of logarithmically spaced points). Here the user sets scene depth under the assumption that it is built from planar surfaces — the wall and floor in Figure 3. The arbitraryness of this process is of no consequence here, but could be removed using stereopsis. Figure 3 shows the result of modifying contract with distance.

Depth of field is a convolution effect, readily emulated with a Gaussian filter. We simply make the width and standard deviation of the filter proportional to depth. Under ARTcam control the changes can be non-linear. The result is seen in Figure 3. As mentioned above, depth of field has been emulated in model-based computer graphics using distributed ray-tracing [CPC84] — an object space operation — and also in image space [PC81]. Depth of field over photographs, has been emulated before [BTHC03] but using specialised algorithms. The ARTcam approach is not only simpler, but can be integrated with other effects such as aerial perspective, as seen in Figure 3. The advantage of a unified camera model is that we need only change the way the camera is specified to produce any combination of effect. We do not need to use many different models.

ARTcams can simulate colour aberration effect, as shown in Figure 4. A simple miss-printing picture can be made using three RTcams, which successively shift the red, green and blue layers of the input image by different amount. A more interesting result can be made using ARTcams that control geometric transformation based on colour. In this case, we use the $\mathbf{Q}_{ipc}$ components of the tensor. For example, the ARTcams applied on the red layer is defined as:

$$\mathbf{Q}_{1pc} = \begin{bmatrix} s_{red} & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{8}$$

$$\mathbf{Q}_{2pc} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ s_{red} & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{9}$$

From equation 7, 8 and 9 we have $z_1 = \mathbf{p}\mathbf{Q}_{1pc}\mathbf{c}^T = \mathbf{p}_1 * (s_{red} * \mathbf{c}_1 + 1)$ and $z_2 = \mathbf{p}\mathbf{Q}_{2pc}\mathbf{c}^T = \mathbf{p}_2 * (s_{red} * \mathbf{c}_1 + 1)$. Here s_{red} is a scaler that controls the interaction between geometry and colour – positive s_{red} give us a lens expands red objects. This indicates the change of a pixel's location is depended on its colour. And we can set different values to s_{red}, s_{green} and s_{blue} to get strong miss-printing effects.

4. Calibrating ARTcams

Although parameters of ARTcams can be directly set by experienced user, calibration makes them easier to use. As a simple example, consider the example in Figure 5. A user has drawn a grid of nine points, which is used to calibrate an ARTcam that warps the underlying photograph.

Of course, many commercial systems allow such warping; but ARTcams are not just for warping photographs — ARTcams model non-linear cameras and so can be fitted to real optical devices, as later examples show. We are not the first to calibrate to non-linear projection. Claus and Fitzgibbon [CF05] correct for barrel distortion in a stero-pair using a camera model exactly equivalent to the geometric ARTcam.

To calibrate we must match source and target points, $\mathbf{x}$ and $\mathbf{y}$ respectively. It is easiest to think these represent location, but colour and other information can be used equally well. Section 2 partitioned an ARTcam into independent parts (each a projective subspace), seen in Equation 7. We calibrate each part independently, so here $\mathbf{Q}$ refers to an any ARTcam part. Recall the basic definition

$$y_i = \frac{\mathbf{x}\mathbf{Q}_i\mathbf{x}^T}{\mathbf{x}\mathbf{Q}_n\mathbf{x}^T}$$

Figure 5: *Example applications of solving rational tensor for geometry distortion: Left: The Original picture with source and target points marked in different colour; Right: The warped picture rendered in Non-photorealistic style.*

This leads directly to the set of homogeneous linear equations

$$\mathbf{x}\mathbf{Q}_i\mathbf{x}^T - y_i(\mathbf{x}\mathbf{Q}_n\mathbf{x}^T) = 0$$

which are a basis for calibration. The idea is to factor out the unknown $\mathbf{Q}_i$; then match a sufficient number of source points with target points to solve the linear system. We continue by factoring out the unknowns.

Using Einstein's tensor notation, in which repeated indices denote summation, $\mathbf{x}\mathbf{Q}_i\mathbf{x}^T = \mathbf{Q}_{ijk}x_jx_k$. The term x_jx_k is a matrix, which is given by the outer product $\mathbf{x}^T\mathbf{x}$, so we set $\chi = \mathbf{x}^T\mathbf{x}$ We see that $\mathbf{Q}_{ijk}\chi_{jk}$ is just an inner product — corresponding terms are multiplied together and summed to get a final result. Hence we can re-write the homogeneous linear system as

$$\mathbf{Q}_{ijk}\chi_{jk} - y_i\mathbf{Q}_{njk}\chi_{jk} = 0$$

The inner product $\mathbf{Q}_{ijk}\chi_{jk}$ can be written in vector form $\mathbf{Q}_i\chi$; set a single index $l = j + (k-1)n$, for example. This allows us to factor out the unknowns, giving

$$[\mathbf{Q}_i\mathbf{Q}_n]\begin{bmatrix} \chi \\ -y_i\chi \end{bmatrix} = 0$$

as the homogeneous linear system to be solved. Before "vectorising", though, we notice that because $\chi_{jk} = \chi_{kj}$ the matrix $\mathbf{Q}_i$ can always be replaced by a symmetric matrix, with no loss of generality. We therefore set

$$\mathbf{R}_i = \frac{1}{2}(\mathbf{Q}_i + \mathbf{Q}_i^T) \tag{10}$$

However, this complicates "vectorisation" because we need only estimate the upper triangular part of $\mathbf{R}_i$. To vectorise we therefore just use the upper triangular section of both $\mathbf{R}_i$ and χ_{jk}, with a scale factor of 2 on all off-diagonal elements of χ_{jk}. This gives a vector of unknowns which we will call $\mathbf{r}_i$, and a vector made up of elements in a source point χ. Along with the elements in a target point we now have

$$[\mathbf{r}_i\ \mathbf{r}_n]\begin{bmatrix} \chi \\ -y_i\chi \end{bmatrix} = 0$$

which is a smaller system than that just written. In fact, it reduces the number of unknowns from n^3 to $n^2(n+1)/2$. This governs the minimum number of matches between source and target points; since each match contributes $n-1$ constraints we will need at least

$$m = \frac{n^2(n+1)}{2(n-1)}$$

matches.

Each match is written into an $n-1$ column of a design matrix $\mathbf{D}$. For example, consider a problem matching two-dimensional points. We have $n = 3$ as the dimension of homogeneous space. For each match we obtain

$$[\mathbf{r}_1\ \mathbf{r}_2\ \mathbf{r}_1]\begin{bmatrix} \chi & \mathbf{0} \\ \mathbf{0} & \chi \\ -y_1\chi & -y_2\chi \end{bmatrix} = \mathbf{0}$$

as our homogeneous linear system. Each new match can easily be included by adding two new columns to the left-hand matrix, which is the design matrix $\mathbf{D}$. For two-dimensional points, as in pixel locations, the homogeneous space has dimension $n = 3$, so we need at least $m = 9$ matches. Once all matches are determined the resulting homogeneous linear equations can be written

$$[\mathbf{r}_1 \ldots \mathbf{r}_n]\mathbf{D} = 0$$

The ARTcam parameters, in the vector $[\mathbf{r}_1 \ldots \mathbf{r}_n]$, are in the null space of $\mathbf{D}$, so a least-squares solution can be found via singular value decomposition. The resulting ARTcam is unique up to scale, but since ARTcams are homogeneous this scale ambiguity has no effect on the mapping.

We are now in a position to apply the ARTcam in practice.

4.1. Colour ARTcams

Colour calibration for ARTcams is analogous to geometric calibration. Coloured ARTcams is a generalised form of the current parametric colour constancy models ([Fin96] [IW05] [Jos04]) and we provide automatically selection over the complexity of ARTcams which improves the robusty. For details please refer to [LH07].

The need for general non-linearity is suggested in Figure 6. It shows a silver coffee pot and a green Christmas bauble, captured under identical conditions. The reflections of these two reflectors are both specular. We aim accurately to transform the coffee pot's silver surface to the bauble's green surface, including the reflection. To achieve this, sufficient colour samples and a non-linear transformation are needed. This is because the highlights remain approximately a constant colour regardless of the colour of the reflector, but coloured regions away from highlights change to a considerable degree. As showed in Figure 6, we can see both the highlight and general colour transformation are accurately captured.

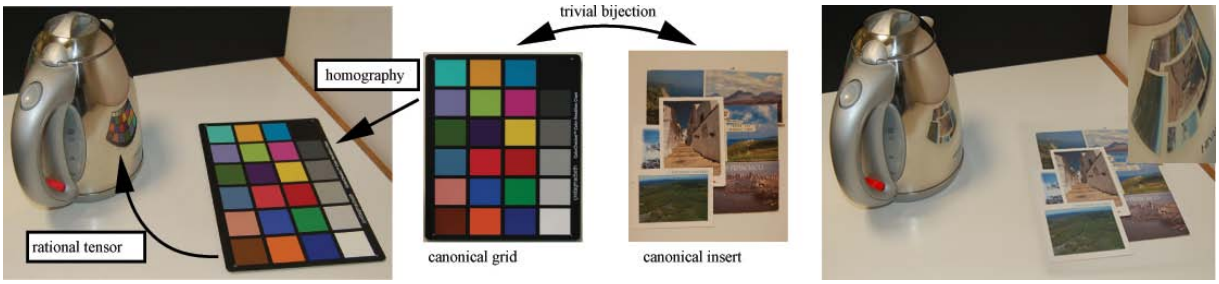

Figure 7: *Inserting an object into a scene and matting a section of the environment. (a) A calibration grid lying on a table is reflected in a coffee pot; the mapping from grid to reflection is calibrated by an RTcam. (b) A canonical view of the grid is obtained via a homography. This is in a bijective relation with the required graphical insert. (c) The mapping from the canonical graphical insert to both its synthetic image and its synthetic reflection is now fully defined, so that a novel scene can be created. Both the geometry and colour are realistically synthesised.*

Figure 6: *Non-linearity in specular reflection. Using colour ARTcams ([LH07]), we can relight the coffee pot's silver surface to the bauble's green surface*

5. Applications: environment matting

Inserting graphical objects into real world imagery is a contemporary aim that arises in many applications. There are many problems. The one of interest here occurs when the scene contains curved reflecting/refracting objects: the insert must somehow be reflected/refracted too. ARTcams can solve this problem, as Figures 7 and 8 show. These are examples of environment matting. Like Zongker *et al.* [ZWCS99] and Chuang *et al.* [CZH*00], but unlike Wexler *et al.* [WFZ02], we use a calibration grid to set up our mapping. Uniquely though, we aim to calibrate a generalised camera model that has uses beyond modeling reflection and refraction.

To reflect a graphics insert in a curved reflector, we must determine a mapping. Typically this mapping is from homogeneous image space to homogeneous image space; that is from one part of an image to another (from where the object is, to its reflection). In a single operation, the mapping must be able to translate, rotate, scale, add perspective distortion, and warp straight lines into curves. ARTcams are capable of such an operation. As explained in Section 4, given a sufficient number of matched points, we can determine the required mapping as an ARTcam. In the example, Figure 7 we want to reflect objects on a tray into a coffee pot. We used a test object — a grid of coloured squares — to assist the process of matching points. We matched by hand via a simple user interface. After calibration we have an ARTcam $\mathbf{Q}$ which comprises a set of matrices $\mathbf{Q}_i$. This ARTcam is plugged into Equation 3 which maps source points (in the tray) to target points (in the coffee pot).

However, it is more convenient to map a graphics insert from a standard view; one that does not have the perspective distortion seen in the photograph of Figure 7. This requires a second mapping, from the grid in a standard view to the perspective view. This mapping is also calibrated by matching points. We now have two mappings, one from the standard view, in which the test grid looks rectangular, and a second from the perspective view of this grid into the coffee pot (see Figure 7). Applying one after the other is sufficient to enact the complete mapping.

We can do better than apply two ARTcams in succession. The mapping from the standard view to the tray is a linear homography $\mathbf{H}$, which is a single matrix. This homography maps a point $\mathbf{x}$ to $\mathbf{xH}$. Consequently we can easily compute a single ARTcam to map directly from the standard view into the coffee pot. We just make the substitution $\mathbf{x} \leftarrow \mathbf{xH}$ in Equation 3 to get a new set of matrices $\mathbf{Q}_i \leftarrow \mathbf{H}^T \mathbf{Q}_i \mathbf{H}$, which is the new ARTcam we want. Figure 7 shows we can replace the test grid with some other object that is conveniently defined in a standard view. The same figure magnifies the reflection for the reader's benefit.

Refracting graphics inserts is also possible, as shown in Figure 8. Again we use a calibration grid and proceed in much the same way as for reflection. Refraction, though, tends to be poorly modeled by a single rational quadratic function, which is what an ARTcam is. We can measure the error by refracting the grid with our ARTcam and comparing

(a) (b) (c) (d)

Figure 8: *Result of parametric mapping for refraction. (a): we divide the environment (calibration grids) into small pieces. Each of them will be mapped to the refraction by a unique RTcam. (b): the real refraction scene.(c): The calibration grid warped by the calibrated RTcams. (d): A new environment was mapped to the refraction.*

it to the real image. Of course, we could redefine an ARTcam to use a higher order tensor, but the number of terms in the tensor rises exponentially with its order; moreover we will encounter over-fitting problems. Instead, we choose to use different ARTcams in different parts of the image. The image region that undergoes refraction is divided into patches. A distinct ARTcam is calibrated to each, so that the maximum error in any patch is pleasingly small. In fact we have a piece-wise camera. Calibration ensures continuity between patches.

Figure 8 shows the result of refracting a graphics insert through a glass of water. This required six patches to refract though air and another sixteen patches to refract through water. We have left in place the reflection of the test grid on the table, which shows that the graphics insert is a true insert. We could of course remove this using the reflection methods already described.

6. Conclusion

The ARTcam model is a powerful new tool for projection. It extends projection from linear to non-linear, and unifies geometric and non-geometric quantities. Its versatility has been demonstrated through application to environment mapping, colour mapping, aerial perspective and depth of field. This paper has space for only a few examples. ARTcams can be specified directly, through drawing, or by calibration to real optical devices.

The ARTcam model subsumes many alternative camera models. Elsewhere we prove prove that the model subsumes X-slit, push-broom, RFcams, GLCs and the traditional pinhole camera [HCS*07]. That earlier works further demonstrates the versatility of ARTcams by non-linear warping, creating novel "round the corner" views, and non-linear mosaicing, all using stereo pairs of images as input. The addition of non-geometric attributes, in this paper, means ARTcams subsume colour transformations, such as colour constancy models ([Fin96] [IW05] [Jos04]), and provides even more robust performance.

As currently defined, ARTcams are rational quadratic projectors. This places a fundamental limit of the class of operations they can perform. One way to overcome this is to raise the order of the tensor, so that ARTcams become the ratio of polynomials of greater order. However, we have found it easier to patch together several ARTcams — the equivalent of approximating a surface in piecewise fashion. The most important lesson, though, is that by unifying geometry and non-geometry in a single model provides the computer graphics user with a very powerful tool, and one which is open to significant future development.

References

[AAC*06] AGARWALA A., AGRAWALA M., COHEN M., SALESIN D., SZELISKI R.: Photographing long scenes with multi-viewpoint panoramas. *SIGGRAPH* (2006), 853–861.

[AZM00] AGRAWALA M., ZORIN D., MUNZNER T.: Artistic multiprojection rendering. In *Eurographics Workshop on Rendering Techniques* (2000), pp. 125–136.

[BTHC03] BARSKY B., TOBIAS M., HORN D., CHU D.: Investigating occlusion and discretization problems in image space blurring techniques. In *Vision, Video and Graphics* (2003), pp. 97–102.

[CF05] CLAUS D., FITZGIBBON A.: A rational function lens distortion model for general cameras. In *Computer Vision and Pattern Recognition* (2005), vol. 1, pp. 213–219.

[CPC84] COOK R. L., PORTER T., CARPENTER L.: Distributed ray tracing. In *SIGGRAPH* (1984), pp. 137–145.

[CS04] COLEMAN P., SINGH K.: Ryan: rendering your animation non-linearly projected. In *Non-Photorealistic Rendering and Animation* (2004), pp. 129–138.

[CZH*00] CHUANG Y., ZONGKER D. E., HINDORFF J., CURLESS B., SALESIN D. H., SZELISKI R.: Environment matting extensions: Towards higher accuracy and real-time capture. In *SIGGRAPH* (2000), pp. 121–130.

[Fin96] FINALYSON G.: Color in perspective. In *IEEE Trans. Pattern Analysis and Machine Intelligence* (1996), vol. 18, pp. 1034–1038.

[GH97] GUPTA R., HARTLEY R.: Linear pushbroom cameras. *IEEE Transactions on Pattern Analysis and Machine Intelligence 19* (1997), 963–975.

[HCS*07] HALL P. M., COLLOMOSSE J. P., SONG Y., SHEN P., LI C.: Rtcams: A new perspective on nonphotorealistic rendering from photographs. *IEEE Transactions on Visualization and Computer Graphics 13*, 5 (2007), 966–979.

[HS97] HEIDRICH W., SEIDEL H.: An image-based model for realistic lens systems in interactive computer graphics. In *In Graphics Interface* (1997), pp. 68–75.

[IW05] ILIE A., WELCH G.: Ensuring color consistency across multiple cameras. In *ICCV* (2005), pp. 1268–1275.

[Jos04] JOSHI N.: *Color calibration for arrays of inexpensive image sensors*. Master's thesis, Stanford University Department of Computer Science, 2004.

[KMH95] KOLB C., MITCHELL D., HANRAHAN P.: A realistic camera model for computer graphics. In *SIGGRAPH* (1995), pp. 317–324.

[Lev98] LEVENE J.: *A Framework for Non-Realistic Projections*. Master's thesis, MIT, M.Eng thesis, 1998.

[LH07] LI C., HALL P.: Colour constancy based on model selection. In *BMVC* (2007), pp. 122–131.

[PC81] POTMESIL M., CHAKRAVARTY I.: A lens and aperture camera model for synthetic image generation. *SIGGRAPH* (1981), 297–305.

[RB88] RADEMACHER P., BISHOP G.: Multiple center of projection images. *SIGGRAPH* (1988), 199–206.

[RGL04] ROMAN A., GARG G., LEVOY M.: Interactive design of multi-perspective images for visualizing urban landscapes. In *IEEE Visualization* (2004), pp. 537–544.

[RL06] ROMAN A., LENSCH H.: Automatic multiperspective images. In *Eurographics Symposium on Rendering* (2006).

[RSL05] RAMALINGAM S., STURM P., LODHA S.: Towards complete generic camera calibration. In *CVPR* (2005), pp. 1093–1098.

[WFZ02] WEXLER Y., FITZGIBBON A. W., ZISSERMAN A.: Image-based environment matting. In *SIGGRAPH conference abstracts and applications* (2002), pp. 198–198.

[Wil97] WILLATS J.: *Art and Representation*. Princeton University Press, 1997.

[Wil07] WILLIS P.: Generalised compositing. In *GRAPHITE* (2007), pp. 129–134.

[YM04] YU J., MCMILLAN L.: A framework for multiperspective rendering. In *Eurographics symposium on Rendering* (2004).

[ZFPW03] ZOMET A., FELDMAN D., PELEG S., WEINSHALL D.: Mosaicing new views: The crossed-slits projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence 25*, 6 (2003), 741–754.

[ZWCS99] ZONGKER D. E., WERNER D. M., CURLESS B., SALESIN D. H.: Environment matting and compositing. In *SIGGRAPH* (1999), pp. 205–214.