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

Automatic Views of Natural Scenes

M. Bratkova[1] and W. B. Thompson[1] and P. Shirley[1,2]

[1]University of Utah, Salt Lake City, UT, USA
[2]NVIDIA Research, Salt Lake City, UT, USA

Abstract

Automatic generation of well-composed and purposeful images is useful in a variety of computer graphics applications. In this work, we explore a set of criteria based on utility, perception, and aesthetics applicable to natural outdoor scenes. We also propose a method that uses the criteria to produce renderings of terrain scenes automatically.

Categories and Subject Descriptors (according to ACM CCS): Computer Graphics [I.3.3]: Picture/Image Generation—[Viewing algorithms]

1. Introduction

There are infinitel many 2D images of a given 3D object or scene. How do we make a good one? Intuitively, it seems that an image from a good viewpoint is one that provides as much information about the world of interest as possible. In practice, this definitio is too general as "information" is application dependent. For example, in recognition, a good view, also referred to as a "canonical view", is one that aids people identify an object in the least amount of time. A good view could also be the one that is more "aesthetic" or subjectively visually pleasing, as opposed to a bad view (Figure 1).

We note that measures of view quality are ultimately subjective. However, we believe automatically generating good initial images of scenes, or selecting representative viewpoints can be beneficia for meaningful scene visualization. Gaining insight from an image is hard for large or complex scenes, especially when interactive exploration is not possible. Therefore, our goal is to provide a general framework that computes automatic viewpoints of natural scenes by taking into account important features, preserving perceptually relevant 3D information during the image projection, and presenting the result in a visually pleasing manner.

In this work we present a set of criteria for viewpoint quality estimation that is based on utility, perception, and aesthetics. We also describe a method that uses the criteria to automatically produce reasonably good images of outdoor terrain scenes.

2. Background

2.1. Viewpoint Selection

Viewpoint selection is a widely researched topic in a variety of fields including computer vision, cinematography, and image modeling. More recently it has also gained momentum in the field of computer graphics and visualization.

Most viewpoint selection methods in computer graphics use the heuristic that the best view is the one that provides the user with the most information (depending on the application). For virtual world exploration and camera path planning, the goal is to automatically determine camera positions which when connected produce camera path trajectories that explore the space in the most informative manner (see [CO06] for an in-depth overview). In scene understanding, it is desirable to choose individual viewpoints that are most representative and indicative of the scene [BDP00], while in image-based rendering, the intent is to fin a minimal set of viewpoints that can see the object of interest and allow its reconstruction [FCOL99]. For mesh-saliency, an attempt is made to fin views that emphasize the most important features of an object [LVJ05, PPB*05, FSG08].

Viewpoint quality is commonly computed as a sum of visible surface quantities, i.e. projected area of non-degenerate faces or angle between the look direction and the face normal [KK88, BDP00]. Viewpoint entropy is an information theory based measure that represents distribution of visibility and is applied to the unconnected geometry faces. A good

© The Eurographics Association 2009.

DOI: 10.2312/COMPAESTH/COMPAESTH09/033-040

delivered by
EUROGRAPHICS
DIGITAL LIBRARY
www.eg.org
diglib.eg.org

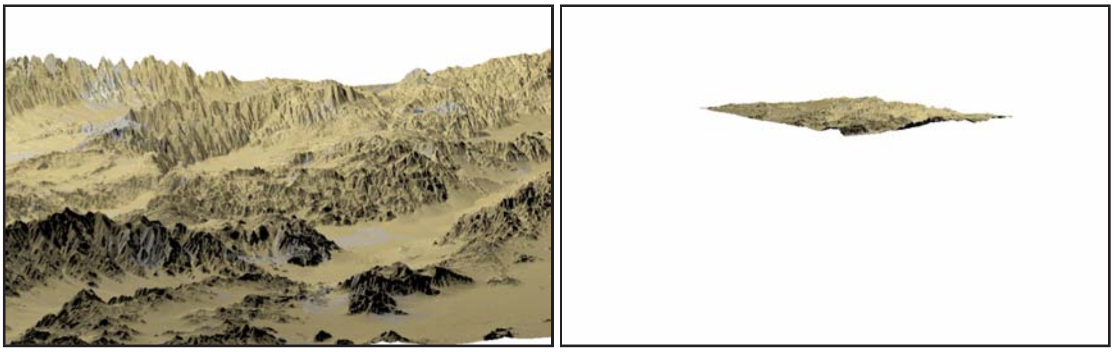

Figure 1: *Good view of an outdoor scene (left). Bad view of the same scene (right).*

view is one with high viewpoint entropy and representing high number of faces [VFSH01, SPFG05]. Viewpoint mutual information, a different information theory measure, is calculated on “information channels" between all viewpoints and all object polygons [FSG08]. It is more robust to polygonal discretization (unlike viewpoint entropy). Stoev and Strasser [SS02] note that such methods generate top-down views that lack depth for scenes whose normals point in a similar direction, e.g. terrain models, and advocate the need for a measure that accounts for scene depth. More recent work explores the connectivity of the geometry of the visible surfaces in addition to using primitive elements of the 3D model. Sokolov and Plemenos consider the total curvature of visible surfaces as an amount of information appropriate for a specifi viewpoint [SP05].

Viewpoint selection criteria have commonly been aimed at determining the best view of a standalone object, as opposed to scenes. Higher-level methods have also been proposed, where not only the visibility of faces, but also the visibility of individual objects has been deemed appropriate [SPT06, SP08].

The search for best viewpoint requires the evaluation of specifi viewpoint parameters. Most commonly, viewpoint position is chosen from a discrete set selected on the bounding sphere of the object or scene, with direction aimed at the center of the sphere. The rendered view is given a weight, and the best view is the image with the best criteria result. Some authors introduce heuristics that limit the search space. For objects, Polonsky et al. assume an up direction, three-quarter views, and normal clustering [PPB*05]. For scenes, Sokolov and Plemenos choose a set of discrete camera positions, fi ed at a pre-determined height, as well as a discrete set of view directions [SP08]. The algorithm performs a greedy search through the discretized parameter space.

2.2. Viewing

Many rendering systems use a perfect pinhole camera and place the image plane a unit distance away from the pinhole. Introducing bounds to the projection plane produces a fi ed dimension *image* of the world. Since a pinhole camera never projects along parallel lines, we always have a perspective, as opposed to parallel, projection. While pinhole cameras do not have focal length, f, for practical reasons focal length is used to indicate the distance between the pinhole and the image. In graphics, pinhole cameras usually have zero aperture, and therefore depth of fiel is rarely a parameter.

Geometry is typically specifie in world coordinates. When projecting the geometry into the camera image plane, it is useful to express it in the camera's coordinate frame. A standard rigid transformation described by a rotation matrix, $\mathbf{R}$, and a translation vector, $\mathbf{t}$, takes a general world coordinate system frame, $\mathbf{F_w} = (O_w, U_w, V_w, W_w)$, into the pinhole camera coordinate system frame, $\mathbf{F_c} = (O_c, U_c, V_c, W_c)$.

Parameters that specify the position (via translation w.r.t. the world frame), O_c, and orientation (via rotation w.r.t. the world frame), W_c, of the camera are called *extrinsic*. Constructing the camera frame assumes a standard *up* direction. The fiel of view of the pinhole determines the bounds of the image in the image plane, as it controls what portion of the scene projects to it. It is usually measured as an absolute angle in the horizontal ($hFov$) and vertical ($vFov$) directions centered around the optical center of the camera. Assuming unit focal length, the dimension of the image plane is $w \times h$, where $w = 2 * \tan\ (hFov/2)$ and $h = 2 * \tan\ (vFov/2)$. Similarly to wide angle lens in photographic cameras, the wider the fiel of view, the bigger the portion of the world that is recorded in the image.

2.3. Image Composition

Modifying the viewing parameters of a virtual camera will produce images of diverse quality. Without doubt, aesthetics is subjective. Nonetheless, artists commonly use heuristic compositional rules. And they can be especially helpful when automatically creating images in computer graphics.

In the graphics literature, Gooch et al. [GRMS01] propose a simple system based on artistic heuristics that determines

© The Eurographics Association 2009.

the format, viewpoint, and layout for an image of a 3D object. Grimm discusses composition techniques such as camera angle, screen placement, and color adjustment as well as their use in a post-rendering step [Gri01].

3. Criteria for Automatic View Selection and Image Composition of Outdoor Scenes

As there is no agreement among researchers for common criteria that determine the “best view" of a complex scene, we will discuss three criteria that we believe are important in the context of outdoor terrain scenes.

While we are interested in automatically determining good viewpoints for complex scenes in terms of their usefulness (i.e. can we see the objects that matter), we think it is also important that we aid the perceptual system in recovering depth from the image, and finally - that the image produced is aesthetically pleasing.

3.1. Utility Criteria

When determining if an image presents a useful view of a scene, we feel it is important that objects and features of interest are clearly visible and recognizable. For outdoor terrain scenes, such features could include any of the following types: *topographic*, *landmarks*, or *man-made*.

Topographic features include *ridges*, *valleys*, *peaks*, etc. Such features are part of the geology of the terrain. They are meaningful, as their visual recognition provides us with vital information about the essence of the geometry of the space.

Landmarks are prominent or well-known markers that help us determine spatial directions and relations. A well-known mountain (i.e. the Grand Teton), or a famous peak (i.e. Denali), are semantically useful. In fact, many people navigate through complex terrain based on landmark features alone.

Man-made features include roads, buildings, etc. They are not naturally part of the geology of the terrain, but are also important in navigation and self-localization.

3.2. Perceptual Criteria

If the only objective is to make sure that all objects of interest are in view, an obvious solution is an orthographic aerial view from above. Such a view is indeed useful in cartography, and has led to the development of the topographic map. However, the view from above does not provide any cues that can be used to reconstruct the three dimensionality of the real terrain. Since the goal of visualization is to provide a meaningful and perhaps, enhanced, rendition of the geometry, in this work our goal is to determine a view from which the image projection of the 3D geometry allows the viewer to maintain the sense of depth present in the real scene.

We propose the following three metrics: *occlusion*, *depth variability*, and *horizon visibility*, as they directly affect our depth perception of the geometry of a scene from an image

When foreground surfaces obstruct the view of background surfaces, our perceptual system can interpret the *occlusion* cues as an indication of real depth. We need to be careful not to take this to the extreme (e.g. having near features completely replace more distant features in the image), as our goal here is the visualization of the entire scene, not just a subset of it.

A complementary metric to occlusion is *depth variability*. The more occluded surfaces are, the smaller the range of scene depth for the objects visible in the image. Also, the higher the depth variability between the visible near and far surfaces in the scene, the more pictorial depth cues will be maintained later when 3D shading and texturing are applied, making depth perception easier.

Finally, ability to see the *horizon* further helps establish depth, as objects near the horizon are perceived to be farther away than objects away from it.

3.3. Aesthetic Criteria

Composition is the visual structural arrangement of elements in an image. Successful compositions are achieved by selecting a meaningful viewpoint and then *framing* the objects in the scene such that they form an image with a balanced visual *layout* [GRMS01].

Framing determines the image boundaries of the depicted scene and is usually specified by *format* and *aspect ratio*. The framing *format* deals with the proportions of the final image. The appropriateness of the format depends on the visual elements depicted [Arn82, LUS*05]. An object that has more horizontal extent will be better visualized in an image that is of a landscape format, as it fits better in the image frame. Similarly, objects that have more vertical extent, will look better in a portrait format. The framing *aspect ratio*, the ratio between the width and height of an image, is commonly determined from existing standards. For example, 35 mm film constraints the aspect ratio to $(3:2)$, while HDTV uses $(16:9)$. In art, however, the aspect ratio is not controlled by a standard. Instead, the artist uses it to increase the emotional and compositional impact of the work. Elongated aspect ratios are perceived to be more dynamic, while aspect ratios close to one - more static [Her71]. Certain aspect ratios, however, seem to be preferred more than others. The golden section, ϕ, is a commonly employed, aesthetically pleasing aspect ratio, that has been used in art and architecture since the Ancient Greeks [Ela01], and perhaps earlier. It is expressed algebraically as follows:

$$\phi = \frac{a}{b} = \frac{a+b}{a} = \frac{1+\sqrt{5}}{2} \approx 1.618$$

Layout ensures the balanced distribution of important elements in the image plane. There are no real rules that can be

© The Eurographics Association 2009.

used here, only heuristics. A variety of approaches have been recommended by artists [Gou04, Kra05, LUS*05, Sch06]. In its simplest form, creating a balanced image requires that all elements balance their visual weight, which is commonly determined from their *luminance*, *color*, and *placement in the frame*. We note that brighter objects have a stronger perceived visual weight than darker objects. Similarly, certain hues are perceived to be brighter than others, despite the same luminance. Among photographers, equal division of the visual elements for *placement in the frame* is considered bad, and thirds and fifth are considered better than halves and fourths [Cli73]. These are the so called *rule of thirds* or *rule of fifth* [Kra05, LUS*05, Sch06], which work in the following way: when the aspect ratio of the fina image has been determined, the visual elements are placed onto an imaginary 3x3 (for rule of thirds) and 5x5 (for rule of fifths grid in the fina image. These rules are only a guide and not a silver bullet. In fact, it is suggested that variety and unity [Cli73] are essential to a good composition.

4. Automatic Viewpoint Selection and Image Composition

Our goal is to automatically determine viewpoints that form meaningful and well-composed images of terrain scenes. We do not want to modify the geometry of the underlying scene in order to improve the composition, nor do we want to modify the lighting and the surface materials and textures.

Though there are different ways one may approach this problem, our strategy is to perform a global optimization through the multidimensional search space of camera parameters. Let **C** be the space of all possible camera configu rations, and **F**, the objective function we are trying to maximize. Then, we can defin the problem mathematically as trying to fin a camera configuration $\mathbf{c} \in \mathbf{C}$, that maximizes an objective function, **F**, as follows:

$$maximize \;\; \mathbf{F}(f_1(\mathbf{c}),\; f_2(\mathbf{c}),\; \ldots,\; f_n(\mathbf{c})), \;\; s.t.\; \mathbf{c} \in \mathbf{C}$$

where the function f_i measures the quality of an image for a specifi criteria. In its simplest case, **F** is a linear combinations of scalar weighted functions:

$$\mathbf{F}\,(f_1(\mathbf{x}),\; f_2(\mathbf{x}),\; \ldots,\; f_n(\mathbf{x})) \;=\; \sum_{i=1}^{n} w_i f_i(x)$$

4.1. Finding a solution

Optimization algorithms often require the availability of derivatives, as this determines the minimization direction in the local neighborhood. In our setup we cannot express the objective function analytically, and therefore we can not directly fin its derivative. It is possible to express the function, as well as to approximate its derivative, numerically. However, such an approximation would be a major bottleneck for the evaluation function, as each derivative estimation will require 2 evaluations per parameter (assuming central differences).

Stochastic search techniques are useful when the structure of the search space is not well understood, or is not smooth (as is our case). Simulated Annealing [KGV83] is a stochastic search technique that addresses these complications. It does not require the presence of derivatives and is applicable to multi-dimensional, global optimization problems. It is resilient to local minima, and converges to a plausible global solution in a reasonable time.

4.2. Setup

Our camera setup uses a perfect pinhole camera positioned a unit distance away from the image plane. The image is formed symmetrically around the optical axis. Assuming a default up orientation, we specify the extrinsic parameters of the camera via 6 independent degrees of freedom (DOF): 3 for the direction vector (using cartesian coordinates), and 3 for the position. Our images are rendered using the golden ratio. Vertical field-of-vi w is used to control the image formation. Therefore, each camera configuratio is controlled by 7 independent parameters.

Given a set of fi ed weights for the different criteria discussed in Section 3, we start with an initial set of camera parameters determined from the bounding box of the scene. The camera is positioned at the maximum extent of the bounding box and is oriented towards the center of the bounding box. The heightfiel scene is rendered with a ray tracer that uses a dynamic BVH implementation [WBS07]. We wrap this in a Simulated Annealing optimizer that employs the ray tracer to render an image for each set of camera parameters and compute its corresponding objective function. Whether a rasterizer or a ray tracer is used, however, is a matter of preference, and should not affect results.

4.3. The Objective Function

In our discussion of meaningful criteria for outdoor scenes in Section 3 we advocated that a “best image” objective function should satisfy, to some reasonable degree, utility, perceptual, and artistic considerations. Specificall , the image produced should maximize the normalized projected area of the desired features. The view should maximize scene depth, prefer views that provide greater depth variability, reward a certain level of element occlusion, and make sure to include the horizon in the fina rendering. Finally, the image should evaluate the position and balance of the functional elements, and should choose an image plane that can be framed according to compositional constraints.

In our evaluation we attempt to use a representative subset of these criteria. In addition, all criteria metrics are normalized to the $[0,1]$ range, so that their behavior can be meaningfully controlled by the desired weight factors.

© The Eurographics Association 2009.

As a matter of notation, an image produced by a camera configuratio $\mathbf{c_i}$ is indicated by $\mathbf{I_i}$. The quality of such an image is measured by $\mathbf{F}\left(U(\mathbf{I_i})\right), P(\mathbf{I_i}), A(\mathbf{I_i})$), where $U(\mathbf{I_i})$ represents *utility*, $P(\mathbf{I_i})$ *perceptual*, and $A(\mathbf{I_i})$ *aesthetic* criteria.

UTILITY CRITERIA

A simple binary test that determines whether a feature can be seen in an image is not useful if the landmark projects to a single pixel. Instead, a quantitative measure should be used to ensure that the object of interest is "visible enough".

Projected area is a direct, quantitative measurement of how visible a feature is. Since some objects are quite large, their relative projected area is a more useful measure of whether or not "enough of the object" is visible.

Ridges and valleys are the interesting features of the terrain and accentuate peaks and canyons, while landmarks mark areas of semantic importance to the user. Our utility criteria is therefore evaluated by the normalized projected area of ridges (R_n), valleys (V_n), and user-specifie landmark features (F_n), with weights w_r, w_v, and w_f respectively, where:

$$U\left(\mathbf{I_i}\right) = w_r * R_n\left(\mathbf{I_i}\right) + w_v * V_n\left(\mathbf{I_i}\right) + w_f * F_n\left(I_i\right)$$

PERCEPTUAL CRITERIA

Our images need to maintain a sense of depth. However, it is also desirable that they show much of the extent of the terrain we are trying to visualize.

The depth variability metric addresses both requirements, as it rewards scene depth, but also encourages seeing the scene in its entirety. Another useful depth-related metric is occlusion. Since it is based on the properties of the topology, not simply the distance to parts of the terrain, it nicely complements scene variability

We evaluate our perceptual criteria by using the normalized depth variability (D_n) and occlusion (O) with weights w_d and w_o:

$$P\left(\mathbf{I_i}\right) = w_d * D_n\left(\mathbf{I_i}\right) + w_o * \sqrt{O\left(\mathbf{I_i}\right)}$$

Though O has a range of $[0,1]$, its values vary non-linearly with respect to variety of views and topologies. For example, terrain occlusion of 20% does not produce twice as good a result as having only 10% of the terrain occluded. We try to make the effect of each metric of the objective function as linear as possible. Therefore, we scale this metric non-linearly to correct for its contribution.

AESTHETIC CRITERIA

Since our scenes are made of large terrains that typically have a wide horizontal extent, it is natural to select a landscape format for our renderings. We choose not to optimize for an aspect ratio, as artists commonly make that choice before producing a fina image. Instead we fi the aspect ratio to the golden section for all rendered images.

In the interest of reducing the complexity of our analysis, we currently ignore the effects of illumination and color on the layout. Our images are produced by only controlling for frame placement, specificall - we use the rule of fifth to place the horizon. The frame placement of the horizon is rewarded by fillin the upper fift of the image with pixels marked as sky ($\mathbf{S_n}$), and fillin the rest of the image with land ($\mathbf{L_n}$) with weights w_s and w_l:

$$A\left(\mathbf{I_i}\right) = w_s * S_n\left(\mathbf{I_i}\right) + w_l * L_n\left(\mathbf{I_i}\right)^2$$

We feel people are especially sensitive to large gaps of open space, where land is instead expected. To penalize such behavior, we non-linearly scale L_n.

WEIGHTS

Clearly, not all of the individual criteria are of equal importance or interest. Their weighting should be left as fl xible input that can be set by the user of the optimization system.

However, we believe that all three types of criteria should have similar importance. Since we mix the objective function weights in a linear fashion, we expect each of the three groups to contribute similarly. Our tests support this observation. Refer to Table 1 for a summary of weight ranges that worked fairly well for our scenes.

Metric	Weight Range %	Weight % Used For Results
Ridges (w_r)	8 - 12	8
Valleys (w_v)	7 - 12	7
Features (w_f)	10 - 15	10
Depth Variability (w_d)	10 - 25	12
Occlusion (w_o)	12 - 25	28
Sky (w_s)	15 - 20	17.5
Land (w_l)	15 - 20	17.5

Table 1: *Criteria weight ranges that proved useful in practice, as well as weights used for all the results in this paper.*

5. Results

Our system is implemented in C++ on a 3GHz Mac Pro System with 8GB of memory. We use three different heightfiel data sets with 30 m resolution data - Yellowstone NP, Rocky Mountain NP, and Death Valley NP. All our images are rendered at resolution 900×557, and the optimizations run for 210 iterations.

We used the same weights for all optimizations (see Table 1) . The only user input consisted of texture maps that marked features of interest for each of the three scenes. The values of the individual metrics are evaluated automatically. For each dataset, we ran 10 independent evaluations, with the same initial camera parameters. For rendering statistics, please refer to Table 2. Our results show the image with the highest energy at the fina iteration.

© The Eurographics Association 2009.

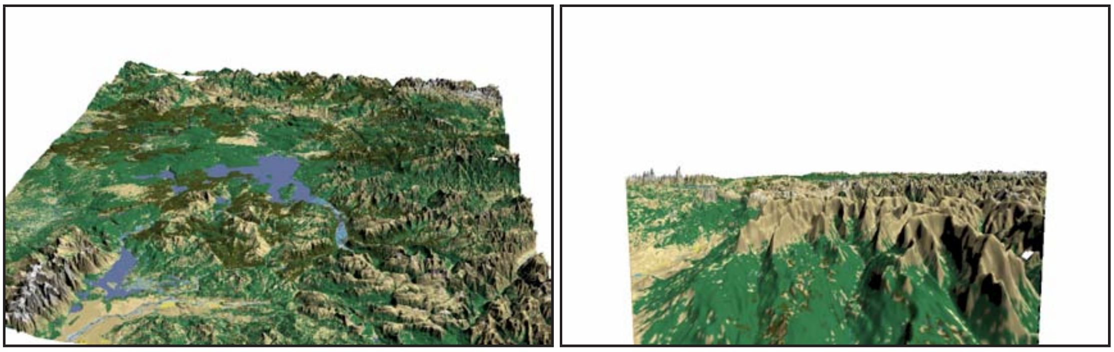

Figure 2: *Rendering of Yellowstone NP produced after optimization, with energy 0.71 (left), and with initial camera parameters (right).*

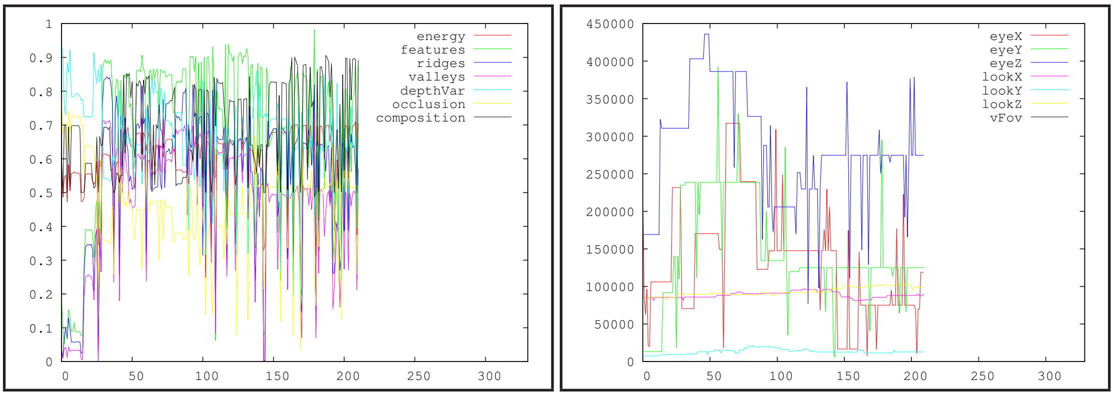

Figure 3: *Value of the objective function and its individual components during the optimization of Yellowstone NP (left), as well as corresponding values of the camera parameters (right). The horizontal axis represents number of iterations.*

Terrain Dataset	Dataset Dimensions	Total Time
Yellowstone (Fig. 2)	1889 x 1880	20.59 min
Rocky Mountain (Fig. 4)	927 x 767	12.62 min
Death Valley (Fig. 5)	2596 x 2537	21.27 min

Table 2: *Rendering statistics for our solution. Total time includes processing of the dataset, building of the bvh, and rendering as well as evaluation of each of the 210 images.*

Figures 2, 4, and 5 are renderings produced by our system. The image to the left is rendered with the optimized viewpoint solution; the one to the right - with the initial camera setup. We feel our results demonstrate promising results and consistently validate our framework.

Figure 3 displays the value of the objective function, as well as its individual components during the optimization (left), and the values of the camera parameters (right) for Yellowstone National Park. The values of the objective function reveal that the initial camera setup (iteration 1) has an acceptable composition score and great depth variability. However, it fails to visualize features of interest, and only minimally displays the ridges and valleys present in the terrain. This is easy to see in the initial setup rendering (Figure 2, right).

After the optimization, our requested features of interest, the Yellowstone Lake (center) and the Grand Tetons (bottom left) are clearly visible (Figure 2, left). The image maintains a good sense of depth - we can see the near as well as the distant parts of the terrain. Considering the finit dimensions of the terrain, the land is framed such that if fill most of the bottom of the image, while leaving a nice area at the top for the sky.

In Figure 4, we can clearly see Grand Lake (top left), and the city of Estes Park (center). The prominent mountain ridge is easy to spot, and the position of the camera is well oriented so that we can see along the fork of the Big Thompson River (center right). The image has a great sense of depth, allows us to see the extent of the terrain, and is framed exactly to our expectations. Similarly, in Figure 5,

© The Eurographics Association 2009.

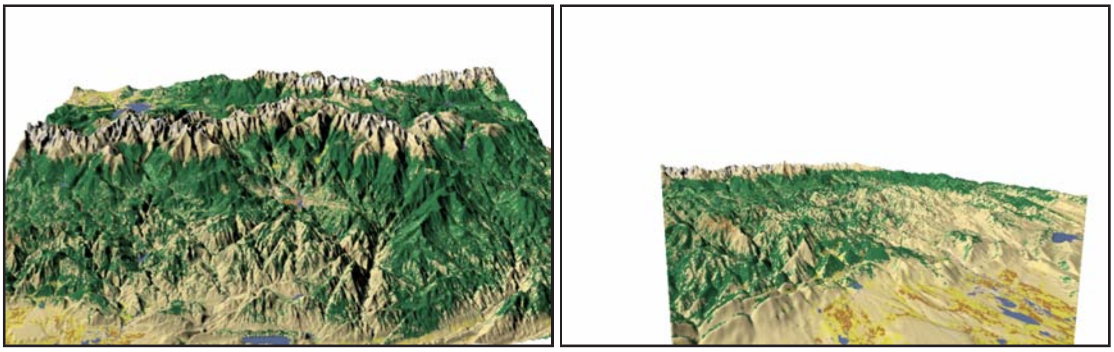

Figure 4: *Rendering of Rocky Mountain NP produced after optimization, with energy 0.63 (left), and with initial camera parameters (right).*

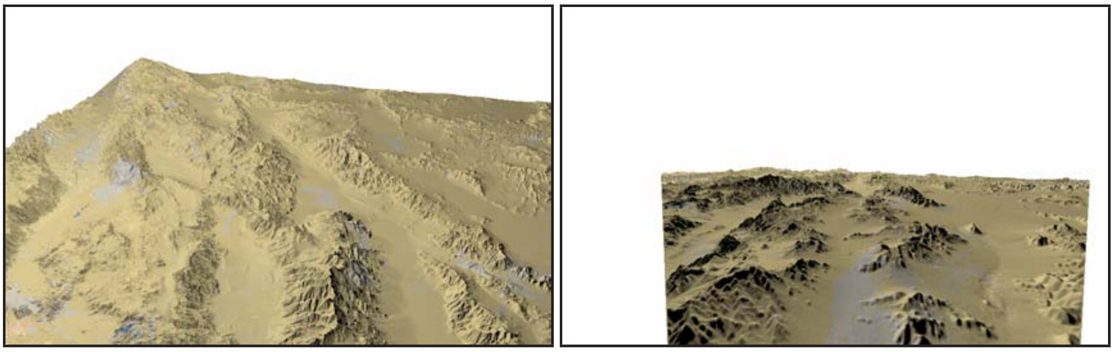

Figure 5: *Rendering of Death Valley NP produced after optimization, with energy 0.66 (left), and with initial camera parameters (right).*

the rendered image features prominent ridges and valleys, a reasonable sense of depth, and meets our framing criteria.

6. Limitations

As with any stochastic global method, we note that we cannot guarantee the image produced by a specifi optimization will be good. However, we can evaluate its goodness and repeat the optimization, until a user-specifie minimum energy is reached.

Global optimization is never fast, especially when we have terrain datasets as big as ours. Occlusion calculations on large terrains with undulating topology are also quite costly.

While speed has not been a great concern for us, optimizing and parallelizing our raytracer, as well as utilizing the power of the GPU will most certainly improve the speed of the evaluation drastically. In particular, we prefer to consider the number of frames necessary to fin a good solution rather than the particular time necessary. Faster rendering systems will therefore produce good views more quickly, but will still require the same number of iterations. In addition, there are also a number of ways to speed up the rendering. We can perform the evaluation on smaller images, or we can apply mesh simplificatio on the terrain datasets.

7. Conclusion

In this work we examine a set of criteria useful for automatic view selection and image composition of outdoor scenes. Following that, we discuss how one may create an objective function formed by linear combination of scalar weighted functions representing the criteria discussed and use it for solving the multidimensional camera parameter optimization. We present results for three different large terrain datasets.

We believe the set of metrics we have advocated for is a good starting point, as it produces fairly successful images with no manual intervention. Additional metrics are likely to improve the quality of the results, most notably, addressing the effects of color and lighting variations on image layout. However, that is beyond the scope of this paper. Future work should investigate a faster method for rendering the geometry. It will also be useful to perform a user study that vali-

© The Eurographics Association 2009.

dates our proposed criteria and approach, as well as allow us to fine-tun criteria weights based on perceived importance.

8. Acknowledgment

This work was supported by NSF grant 03-12479. We thank Solomon Boulos and the CAe anonymous reviewers for their constructive comments.

References

[Arn82] ARNHEIM R.: *The Power of the Center: A Study of Composition in the Visual Arts*. University of California Press, 1982.

[BDP00] BARRAL P., DORME G., PLEMENOS D.: Scene understanding techniques using a virtual camera. *Computer Graphics Forum 19* (2000).

[Cli73] CLIFTON J.: *The Eye of the Artist*. North Light Publishers, 1973.

[CO06] CHRISTIE M., OLIVIER P.: Camera control in computer graphics. In *Eurographics* (2006), vol. 25.

[Ela01] ELAM K.: *Geometry of Design: Studies in Proportion and Composition*. Princeton Architectural Press, 2001.

[FCOL99] FLEISHMAN S., COHEN-OR D., LISCHINSKI D.: Automatic camera placement for image-based modeling. In *Pacifi Graphics* (1999), vol. 19.

[FSG08] FEIXAS M., SBERT M., GONZÁLEZ F.: A unifie information-theoretic framework for viewpoint selection and mesh saliency. In *ACM Transactions on Applied Perception* (2008).

[Gou04] GOULD T.: *An Introduction to Painting Landscapes: Techniques, Light, Color, Composition, Style*, reprinted edition; original edition 1994 by quintet publishing ed. Chartwell Books, Book Sales Inc, 2004.

[Gri01] GRIMM C.: Post-rendering composition for 3d scenes. In *Eurographics* (2001).

[GRMS01] GOOCH B., REINHARD E., MOULDING C., SHIRLEY P.: Artistic composition for image creation. *EGSR* (2001).

[Her71] HERRING J.: *The Painter's Composition Handbook*. Poor-Henry Publishing Company, 1971.

[KGV83] KIRKPATRICK S., GELATT C. D., VECCHI M. P.: Optimization by simulated annealing. *Science 220*, 4598 (1983).

[KK88] KAMADA T., KAWAI S.: A simple method for computing general position in displaying three-dimensional objects. *Computer Vision, Graphics, and Image Processing 41*, 16 (January 1988), 43–56.

[Kra05] KRAGES B.: *Photography: The Art of Composition*. Allworth Press, 2005.

[LUS*05] LONDON B., UPTON J., STONE J., KOBRE K., BRILL B.: *Photography*. Pearson Education, Inc., 2005.

[LVJ05] LEE C. H., VARSHNEY A., JACOBS D. W.: Mesh saliency. In *SIGGRAPH* (2005).

[PPB*05] POLONSKY O., PATANE G., BIASOTTI S., GOTSMAN G., SPAGNUOLO M.: What's in an image? *Visual Computer 21* (2005), 840–847.

[Sch06] SCHULZKE M.: *A Painter's Guide to Design and Composition*. North Light Books, 2006.

[SP05] SOKOLOV D., PLEMENOS D.: Viewpoint quality and scene understanding. In *VAST* (2005), Mudge M., Ryan N., Scopigno R., (Eds.).

[SP08] SOKOLOV D., PLEMENOS D.: Virtual world explorations by using topological and semantic knowledge. *The Visual Computer 24*, 3 (2008), 173–185.

[SPFG05] SBERT M., PLEMENOS D., FEIXAS M., GONZÁLEZ F.: Viewpoint quality: Measures and applications. In *Computational Aesthetics in Graphics, Visualization and Imaging* (2005), Neumann L., Sbert M., Gooch B., Purgathofer W., (Eds.).

[SPT06] SOKOLOV D., PLEMENOS D., TAMINE K.: Viewpoint quality and global scene exploration strategies. In *GRAPP* (2006).

[SS02] STOEV S. L., STRASSER W.: A case study on automatic camera placement and motion for visualizing historical data. In *IEEE Visualization 2002, Boston, MA* (2002).

[VFSH01] VAZQUEZ P.-P., FEIXAS M., SBERT M., HEIDRICH W.: Viewpoint selection using viewpoint entropy. In *VMV, Stuttgart, Germany* (2001).

[WBS07] WALD I., BOULOS S., SHIRLEY P.: Ray tracing deformable scenes using dynamic bounding volume hierarchies. *TOG 26*, 1 (2007).

© The Eurographics Association 2009.