Pacific Graphics (2014)
J. Keyser, Y. J. Kim, and P. Wonka (Editors)

Sky Browser: Search for HDR Sky Maps

Andrew Chalmers[1,2], John Lewis[1,2], Peter Hillman[2], Charlie Tait[2], and Taehyun Rhee[1,3]

[1]Victoria University of Wellington
[2]Weta Digital
[3]taehyun.rhee@ecs.vuw.ac.nz

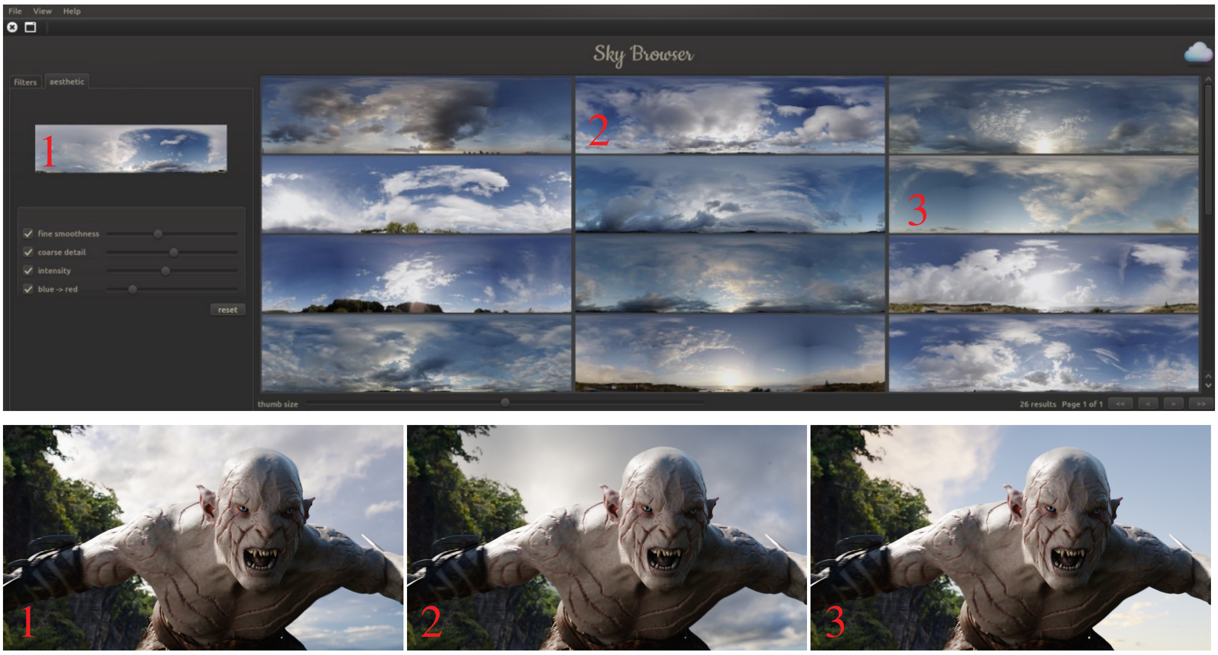

Figure 1: Sky Browser: example of the sky search, as well as the application of switching the sky backdrop using different sky maps. The names given to the sliders are for artistic purposes, where GLCM correlation, entropy of the Laplacian, intensity, and spherical harmonics red/blue maps to "fine smoothness", "coarse detail", "intensity", and "blue -> red" respectively.

Abstract

In a visual effects studio for movie production, sky maps play an important role for acting as a sky backdrop to a scene. The backdrop to a scene is often represented using a high-resolution sky map. This motivates the need for a large collection of sky maps to match various moods and lighting conditions. A comprehensive collection of images is not useful however, without a method of searching for desired images within that database. In this paper we define a feature space that supports an interactive search function for HDR sky maps, allowing users to find ideal images based on its appearance. The set of features are automatically extracted from the sky maps in an offline pre-processing step, and are queried in real time for progressive browsing. The system uses unsupervised learning techniques, discarding the need for labelling a large set of existing sky maps.

Categories and Subject Descriptors (according to ACM CCS): I.4.8 [Image Processing and Computer Vision]: Scene Analysis—Shading; I.3.4 [Computer Graphics]: Graphics Utilities—Picture description languages

DOI: 10.2312/pgs.20141266

1. Introduction

Skies are a well established and long-standing research area in computer graphics. Significant and ongoing research has addressed topics in simulation, rendering, and identification of sky images ([NSTN93, DNKY97, TYS09, YIC*10, DIO*12] and others). High-resolution HDR sky maps (HSMs) are frequently used in visual effects studio production for live action movies, where the HSM is used to provide either a backdrop to the scene, or lighting for synthetic objects to be composited into the scene. In our case, we are focusing on the former purpose. Two of the common scenarios in which HSMs are needed are:

Scenes filmed on set: most scenes in a live action movie are recorded on a stage, rather than in a pre-existing natural environment. In these scenes, only the actors and foreground objects will remain in the final image. Distant objects are generated with computer graphics, physical models, or matte paintings, and the sky may be obtained from a HSM. VFX artists must find sky images that have various desired qualities as requested by the artistic supervisor.

Scenes filmed in nature: in this case, the video images already have distant landscapes and sky. Nevertheless, it is often necessary to replace the sky, for various purposes. For example, it may be necessary to provide continuity with a different scene that was shot at the same location. Frequently, the supervisor may request a sky with a different appearance.

Given this desire for alternate HSMs, a large database is needed for providing a wide range of images that differ in appearance. The problem arises when the number of HSMs in the database is overwhelming, it then becomes highly unlikely that the ideal image is ever seen, thus rendering the database of images redundant. Previous work in this domain typically uses non-spherical low dynamic range (LDR) images of skies and classifies them using labelled data. We focus our work on defining a set of features without labelled data, as well as including features suitable for HDR images.

We propose a system that allows the artist to intuitively navigate the space of images, assisting in finding the ideal image. This includes the ability to use a given image as a query to find another image in the database. The search function finds certain qualities of an image, for example, a clear sky with occasional fluffy clouds, or an overcast sky with large, dark clouds. The search function can also take an input HSM, and search the database for images that range in resemblance from very similar to completely different.

The criteria is difficult for a human to verbalise, and similar images are probably rated differently by different people. For example, the supervisor may request clouds that are "ominous" or "peaceful". How these words relate to images is subjective. For this reason, it is not effective to manually tag various images in the sky database with descriptive words such as "wispy, fluffy, peaceful, angry". This also means that we are not able to apply supervised learning methods to the search problem - the labels in a training set would be both subjective and hard to define.

For these reasons we formulate our problem as one of providing a feature space in which distance reflects the visual similarity of the images. The chosen features are low-level image and textural features. This side-steps the issue of defining what various descriptive words mean for different people. Instead, the artist simply navigates across images in the feature space. Our system, *Sky Browser*, is now in use at a visual effects facility.

2. Related Work

The sky is a common component of many images, and selecting the right sky is important to suggest the time of day, weather and mood. Since Klassen [Kla87] presented his work on sky visualisation, skies have been an important subject in computer graphics with many associated research publications [NSTN93, DNKY97]. Research in these areas is ongoing, and we refer to [YIC*10, DIO*12] for an entrance to this literature, and refer to [SJW*06, Deb98] on methods for capturing HDR environment maps.

Given an outdoor scene, changing the background with a better sky image is a common task in 2D image processing as well as visual effects for live action movies. The main task is to search for alternate images using an appropriate query. Generic content-based image retrieval (CBIR) methods can be used, but these systems rely on features (shape descriptors and interest points) that are not appropriate for clouds [DJLW08]. As well, many of these systems require supervised learning. Therefore, we only focus our survey on papers which are highly related to our main topic, the sky.

Proper labelling of an image or a part of it can guide searching in many applications [JGJJ*06, LHE*07, LRT*14]. However, verbalising the criteria is difficult for a human due to inconsistent meaning of subjective terms such as "peaceful" or "smooth". With this in mind, we formulate our problem as feature extraction and searching the feature space, where relative distance reflects the similarity of the images. Other methods to get around subjective labelling include crowdsourcing, as found in [LRT*14].

The general method of the search function investigates machine learning and texture classification to define feature extraction techniques. Haralick et al. [HSD73] describes easily computable textural features based on grey tone spatial dependencies. This is often referred to as the grey level co-occurrence matrix (GLCM). Gu et al. [GDR*89] compares techniques for measuring cloud textures. They use GLCM features to measure spatial properties, where they found that entropy based features gave good results for frequency properties. Chethan et al. [CRK09] consider textural features based on the Gabor transform to classify clouds, as well as using a support vector machine (SVM) as their method of classification. Mazzoni et al. [MHG*05] label parts of images as a clear sky, or a type of cloud. They use the Multi-angle Imaging Spectro Radiometer (MISR), an instrument used by NASA, to study clouds and aerosols. Heinle et al. [HMS10] classifies skies into seven different categories. They used the k-nearest neighbour method for classification, and the colour and tonal variation of an image as features.

Recently, Tao et al. [TYS09] developed an interactive search system for finding sky photographs using supervised learning techniques. Their method allows for offline computation as well as an interactive user interface. Most similar to our work is [ODY11], which extracts four features to characterise the images in the database. Other recent works include [MF12, LRT*14]. Our method fundamentally differs from these papers on four points: We focus on HSMs (an industry standard image category), we remove the ambiguity and manual labour of labelling data required for supervised learning, we explicitly define two textural properties for more artistic control over clouds, and we use the spherical harmonics as a novel tonal feature. These four components target a specific and important area of computer graphics and the movie industry.

3. Searching

The search is based on observable but subjective image properties. For example, such properties may include how blue the sky is, or how patchy the clouds are. There are tonal properties such as the contrast or brightness of an image, and there are textural properties, such as the bumpiness of a cloud. Further, skies tend to have strong properties relating to how blue, white or red the sky is, such as clear blue skies, large bright white clouds, or a red sunset.

Clouds come in a wide variety of forms that can differ depending on the atmosphere and temperature. We considered the possibility of using the scientific names of clouds as their class labels, and attempted to find features that could categorise them so. For our application, a major disadvantage of supervised learning is the requirement of labelled data. Human labelling of the sky images is not only expensive but conceptually difficult as well. It is difficult to identify a set of labels (such as "wispy", "romantic", etc.) that are useful and consistently interpreted. Further, in our experience the desired labelling is simply not done in some cases.

Instead, we define a search space that does not require a labelled dataset, but that can be visually traversed with no prior training. This requires a set of features that capture perceptually relevant information while ignoring information that is not important or even imperceptible. In addition, we require a minimal set of features, in order to avoid the curse of dimensionality.

Unfortunately the number of possible features is large (it is some fraction of the number of possible programs that take an image patch and output a number), and choosing a best subset is not feasible due to the combinatorics. Choosing a small set of features therefore requires intuition and experience with the problem. We discarded keypoint features such as SIFT [Low] because skies are more appropriately considered as random textures than as images of objects with common and reproducable parts. Instead we explored features such as GLCMs that have proven successful for texture modeling and classification.

After several months of experimentation on actual datasets, we chose the following four-dimensional feature space: *GLCM correlation*, the *entropy of the Laplacian* (EL), a *ratio between red and blue spherical harmonic coefficients*, and the *mean intensity*. The result is a small set of features that define a visually searchable space.

There are numerous advantages to this approach: Firstly, it does not require manually labelling data (or finding a readily available labelled dataset). Secondly, the classification method is not defined by scientific labels that would have to be learned. Instead the search space simply relies on visual perception of the images. Finally, images often fall between scientific labels (such as a single sky having two cloud types), so removing scientific labelled data gives more artistic freedom for defining a continuous feature space for skies. For example, the ability to move from "very patchy" clouds to "somewhat patchy", to "not patchy at all", while maintaining other key features, for example, "a very blue sky" in conjunction with the varying levels of patchy clouds.

3.1. Features

The followings are the set of features we use to define our search space. The HDR images are in a linear colour space. In a pre-processing step, we scale the images to a standard resolution of 360x160. For the spherical harmonics feature, we reduced the input image to 512x256.

GLCM Correlation: The GLCM is a commonly used technique in texture classification [HSD73]. The method involves finding a co-occurrence histogram of an image, and running various formulas across the histogram. The histogram counts how often two intensity values in a greyscale image co-occur with some spatial relation (dx,dy), for example, the number of times that a pixel with value 5 is to the right of a pixel with value 20.

The GLCM histogram is a lot of information – potentially much more than that image itself, depending on how many spatial relations are considered. For this reason, various summary statistics are often used [Alb08]. After experimenting with several GLCM summary statistics, we selected correlation since it is minimially correlated with a second textural feature (described in the next section). For a particular spatial relation, the correlation is

$$\sum_{i,j=0}^{255} P_{i,j}\left(\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}}\right)$$

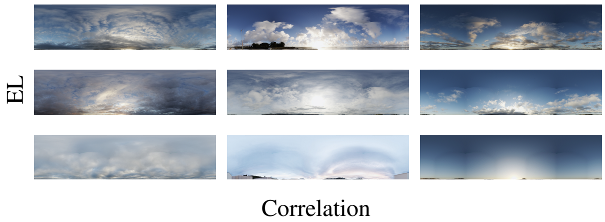

Figure 2: Textural feature space examples.

where i, j are indices of the GLCM corresponding to pixel values on an 8-bit scale, μ is the mean, σ is the variance, and $P_{i,j}$ is the GLCM histogram normalised to serve as a probability. The correlation characterises an approximate smoothness or correlatedness in a particular spatial direction. We used the eight nearest-neighbour directions as the spatial relations. The computation is accelerated restricting the computation to a number of windows evenly placed across the sky area in the HSM. 27 windows are selected by our experimental tests on images with a resolution of 2048x1024. The windows are uniformly placed across the image (9x3).

Entropy of the Laplacian (EL): While the GLCM correlation captures a type of roughness or smoothness, it does not say anything about the distribution of changes. To account for this, we introduce the EL. The Laplacian ∇^2 is an approximate scalar curvature measure. We form the histogram of the Laplacian values at all pixels. The entropy of the normalised histogram distinguishes whether the curvature is concentrated in a few values (low entropy) or takes on many possible values. The feature is computed as

$$-\sum_{i=0}^{N} P_i(\ln P_i)$$

where P_i are values of the normalised histogram of Laplacian values. The latter are computed by a standard finite difference stencil.

The GLCM correlation measure in combination with the EL can differentiate between clouds with difficult to describe textural qualities. Figure 2 shows an example of images distributed in the feature space. The images in row 3 are all coherent images in a sense, as described by their low EL measure. Yet the GLCM correlation measure separates the images; the image in row 3, column 3 is very smooth, and the image in row 3, column 1 has a lumpy texture.

Spherical Harmonics Ratio of Red and Blue: The two features defined above capture tonal and textural properties of an image. An artist also looks for images with certain colour properties. Red and blue values are salient among skies, for example, it is often the case we find vibrant blue skies or red tainted clouds.

As a starting point, we consider the ratio of the amount of red and blue in the sky image. As discussed in [HMS10], the ratio of red and blue defines how much cloud is in the sky, so this ratio has the additional effect of defining cloud cover. Artists can increase the amount of red to find more clouds in the image, as well as increasing it further to find red skies or clouds. We found that green is correlated with the red coefficient in sky images, thus it did not add any useful information in the search function. Furthermore, the ratio is independent of intensity, a desirable property. We can see the distribution in Figure 4.

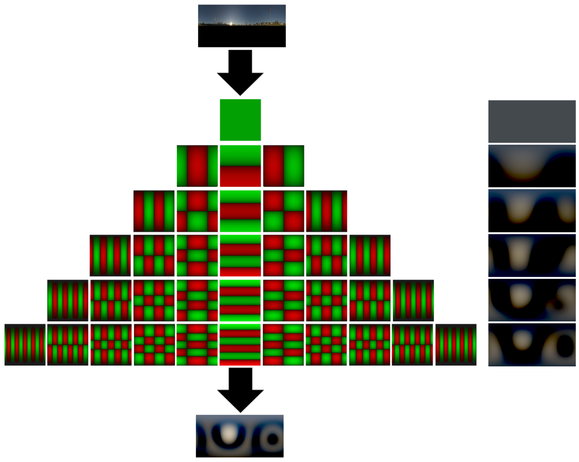

Figure 3: The spherical harmonics expansion. Above the expansion is the input image, and below is the approximation of the input. To the right of each band is the corresponding approximation of the input image.

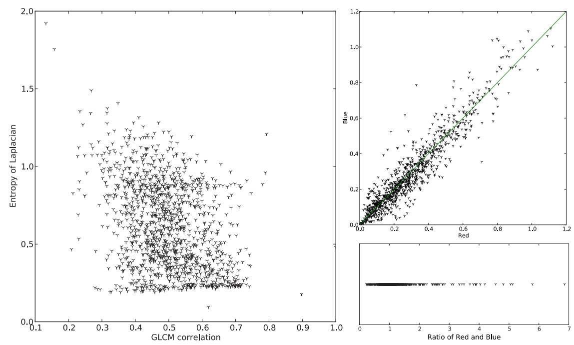

Figure 4: Feature space distribution, where each point represents an image. Left: the textural feature space. Right-top: the red and blue spherical harmonic. Right-bottom: the colour of the sky in one dimension by taking the ratio of the red and blue value.

The redness of an image is somewhat subjective however, as redness can be interpreted for the entire image, or the sun itself. To handle this choice, we consider the problem in the spherical harmonic frequency domain (Figure 3). The first band of spherical harmonics coefficients reflects the image mean, so the red/blue (R/B) ratio can be computed from these. There is no reason to restrict ourselves to just the first band however. The contribution of the sun appears in higher bands, since the sun (when visible) is a nearly a delta function. This is due to the relative values in the HDR image: the sun may have a value much higher than 1, whereas clouds are below one. As our final feature space, we use the R/B ratio of the spectral energy (sum-square of coefficients) [KFR03] in the first three bands. A weighted mixture of the ratios of the first three bands gives the artist control over whether they are seeking the colour of the overall sky, or the colour of bright sky adjacent to the sun.

Mean Intensity: We use the mean intensity of the image as a fourth feature. These four features (correlation, entropy, red/blue ratio, mean intensity) define the space that is navigated by artists.

3.2. Feature Correlation

Pearson's Correlation Coefficient shows that our features all have low correlation with one another (Table 1). GLCM Correlation and the EL act as our textural measures, which allows the user to define the variation of cloud types. We observe from Figure 2, 4 and Table 1 that there is weak-negative relationship with the two features, as the lower end of the GLCM correlation measure defines a lack of smoothness, which can be interpreted similarly as a high EL value. Given that the correlation is weak, the combination of these features is useful, as shown in Figure 2.

Table 1: Feature Correlation

	GLCM	Entropy	Intensity	SH
GLCM	1.00	-	-	-
Entropy	-0.36	1.00	-	-
Intensity	-0.24	0.17	1.00	-
SH	-0.05	0.11	-0.22	1.00
	≥0.0	≥0.1	≥0.3	≥0.5

4. Results

Figure 5 shows the exploration results of *Sky Browser* on a database of 1300 environment maps. It demonstrates a user searching through the feature space using the sliders. The user can begin the search by using an image as a query which has properties that are similar to what they're looking for, as shown in Figure 5 (1st column), where the user has used an image to define the starting slider values of the features. Following from this, the user can adjust parameters to move towards their ideal images. For example, in Figure 5, the 2nd and 3rd columns show intensity and colour changes respectively, while maintaining the textural properties. The 4th column moves towards patchy images, where as the 5th column maintains edges (EL feature) but increases the smoothness (correlation feature), returning large and apparent clouds. The 6th column maintains the smoothness, and removes edges, thus producing clearer skies. Figure 5 shows just one example of finding clouds. There are many other possibilities, for example moving towards blue in column 5's state can bring in smooth blue skies and smaller distinct clouds, instead of large smooth clouds.

Sky Browser searches HSMs in feature space without labelled images. Therefore, it is difficult to evaluate the result quantitatively. Instead, we conducted a subjective test with 11 professional visual effects studio artists. They were asked to use *Sky Browser* to find suitable images, both by using an existing image as a query and by interactively browsing. Following this test, they were asked following question: *"Given the input, did the Sky Browser return similar results?"*, where their input would involve searching with an image as well setting the features manually. They answered with a score between 1 and 5, where 1 is *very dissimilar*, 2 is *dissimilar*, 3 is *uncertain*, 4 is *similar* and 5 is *very similar*. The mean score of the qualitative evaluation by these professional artists is 4.0.

5. Conclusion

The focus of this paper is on defining a minimal set of features capable of unsupervised classification. The set of features describe the appearance properties to make up the *Sky Browser* application. The features are minimally correlated and thus define a search space that is intuitively navigated by the artist. To navigate the search space, the artist has control over the features as parameters. These parameters are useful to describe images based on tonal and textural properties. We can search the space using a nearest-neighbour approach, and the current system runs at interactive rates with a database of 1300 HSMs. The search function is scalable using parallel processing [ML14]. A weakness of the present system is that larger slider movement is needed to navigate in areas that are sparsely populated with data. Future work may investigate user interface issues such as this. Our feature space for *Sky Browser* is defined based on the evaluation by professional visual effects artists, and additional features can be applied for particular purposes.

References

[Alb08] ALBREGTSEN F.: Statistical texture measures computed from gray level coocurrence matrices, 2008. accessed 14 June 2014. 3

[CRK09] CHETHAN H., RAGHAVENDRA R., KUMAR G.: Texture based approach for cloud classification using SVM. Proc. of ARTCom 2009, 2009. 2

[Deb98] DEBEVEC P.: Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with

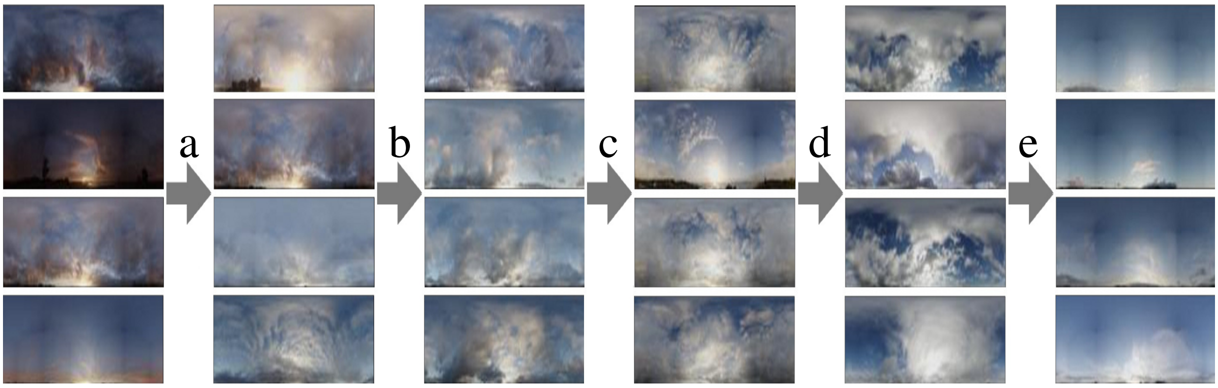

Figure 5: The search results: navigating the feature space. The transition sequence is as follows: (a) increase intensity, (b) slight move away from red, (c) decrease smoothness and increase edges, (d) increase smoothness, (e) large decrease of edges.

global illumination and high dynamic range photography. In *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques* (New York, NY, USA, 1998), SIGGRAPH '98, ACM, pp. 189–198. 2

[DIO*12] DOBASHI Y., IWASAKI W., ONO A., YAMAMOTO T., YUE Y., NISHITA T.: An inverse problem approach for automatically adjusting the parameters for rendering clouds using photographs. *ACM Trans. Graph. 31*, 6 (Nov. 2012), 145:1–145:10. 2

[DJLW08] DATTA R., JOSHI D., LI J., WANG J. Z.: Image retrieval: Ideas, influences, and trends of the new age. *ACM Comput. Surv. 40*, 2 (May 2008), 5:1–5:60. 2

[DNKY97] DOBASHI Y., NISHITA T., KANEDA K., YAMASHITA H.: A fast display method of sky color using basis functions. In *The Journal of Visualization and Computer Graphics* (1997), pp. 115–127. 2

[GDR*89] GU Z., DUNCAN C., RENSHAW E., MUGGLESTONE M., COWAN C., GRANT P.: Comparison of techniques for measuring cloud texture in remotely sense satellite meteorological image data. IEE Proceedings F (Radar and Signal Processing), Volume 136, Issue 5, 1989. 2

[HMS10] HEINLE A., MACKE A., SRIVASTAV S.: Automatic cloud classification of whole sky images. Atmospheric Measurement Techniques, Volume 3, Issue 3, 2010. 3, 4

[HSD73] HARALICK R., SHANMUGAM K., DINSTEIN I.: Textural features for image classification. *IEEE Trans. Systems, Man and Cybernetics SMC-3*, 6 (1973), 610–621. 2, 3

[JGJJ*06] JOHNSON M. B., G. J. S., J. A., O. KWATRA V., CIPOLLA R.: Semantic photo synthesis. *Computer Graphics Forum 25*, 2 (2006), 407–412. 2

[KFR03] KAZHDAN M., FUNKHOUSER T., RUSINKIEWICZ S.: Rotation invariant spherical harmonic representation of 3d shape descriptors. Proceedings of the 2003 Eurographics/ACM SIGGRAPH Symposium on Geometry Processing, 2003. 5

[Kla87] KLASSEN R. V.: Modeling the effect of the atmosphere on light. *ACM Trans. Graph. 6*, 3 (July 1987), 215–237. 2

[LHE*07] LALONDE J.-F., HOIEM D., EFROS A. A., ROTHER C., WINN J., CRIMINISI A.: Photo clip art. *ACM Trans. Graph. 26*, 3 (July 2007). 2

[Low] LOWE D. G.: Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision 2004 60*, 2, 91–110. 3

[LRT*14] LAFFONT P.-Y., REN Z., TAO X., QIAN C., HAYS J.: Transient attributes for high-level understanding and editing of outdoor scenes. *ACM Transactions on Graphics (proceedings of SIGGRAPH 2014) 33*, 4 (2014). 2, 3

[MF12] MITANI T., FUJISHIRO I.: Cosmicai: Generating sky backgrounds through content-based search and flexible composition. In *ACM SIGGRAPH 2012 Posters* (2012), SIGGRAPH '12, ACM, pp. 52:1–52:1. 3

[MHG*05] MAZZONI D., HORVATH A., GARAY M., TANG B., DAVIES R.: A MISR cloud-type classifier using reduced support vector machines. Eighth Workshop on Mining Scientific, 2005. 2

[ML14] MUJA M., LOWE D. G.: Scalable nearest neighbor algorithms for high dimensional data. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 36* (2014). 5

[NSTN93] NISHITA T., SIRAI T., TADAMURA K., NAKAMAE E.: Display of the earth taking into account atmospheric scattering. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques* (1993), SIGGRAPH '93, ACM, pp. 175–182. 2

[ODY11] ONO A., DOBASHI Y., YAMAMOTO T.: A system for editing sky images using an image database. In *SIGGRAPH Asia 2011 Sketches* (2011), SA '11, ACM, pp. 38:1–38:2. 3

[SJW*06] STUMPFEL J., JONES A., WENGER A., TCHOU C., HAWKINS T., DEBEVEC P.: Direct hdr capture of the sun and sky. In *ACM SIGGRAPH 2006 Courses* (New York, NY, USA, 2006), SIGGRAPH '06, ACM. 2

[TYS09] TAO L., YUAN L., SUN J.: Skyfinder: Attribute-based sky image search. In *ACM SIGGRAPH 2009 Papers* (2009), SIGGRAPH '09, ACM, pp. 68:1–68:5. 2, 3

[YIC*10] YUE Y., IWASAKI K., CHEN B.-Y., DOBASHI Y., NISHITA T.: Unbiased, adaptive stochastic sampling for rendering inhomogeneous participating media. *ACM Trans. Graph. 29*, 6 (Dec. 2010), 177:1–177:8. 2