HAL
open science

Finding Images with Similar Lighting Conditions in Large Photo Collections

Mauricio Díaz, Peter Sturm

To cite this version:

Mauricio Díaz, Peter Sturm. Finding Images with Similar Lighting Conditions in Large Photo Collections. CIARP 2009 - 14th Iberoamerican Congress on Pattern Recognition, Mexican Association for Computer Vision, Neural Computing and Robotics (MACVNR), Nov 2009, Guadalajara, Mexico. pp.53-60, 10.1007/978-3-642-10268-4_6 . inria-00434337

HAL Id: inria-00434337

https://inria.hal.science/inria-00434337

Submitted on 24 May 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Finding Images with Similar Lighting Conditions in Large Photo Collections

Mauricio Díaz[1]* and Peter Sturm[2]

[1] Laboratoire Jean Kuntzmann, 38400 Saint Martin d'Hères, France
[2] INRIA Grenoble Rhône-Alpes, 38330 Montbonnot, France
{Mauricio.Diaz,Peter.Sturm}@inrialpes.fr

Abstract. When we look at images taken from outdoor scenes, much of the information perceived is due to the ligthing conditions. In these scenes, the solar beams interact with the atmosphere and create a global illumination that determines the way we perceive objets in the world. Lately, exploration of the sky like the main illuminance component has began to be explored in Computer Vision. Some of these studies could be classified like color-based algorithms while some others fall in the physics-based category. However most of them assume that the photometric and geometric camera parameters are constant, or at least, that they could be determined. This work presents a simple and effective method in order to find images with similar lighting conditions. This method is based on a Gaussian mixture model of sky pixels represented by a 3D histogram in the La^*b^* color space.

1 Introduction

Nowadays, the Internet has become an interactive system that allows anyone with a compact camera to share their visual experiences. This fact has been the starting point for large online image databases such as Flickr or Picasaweb. Computer vision scientists have not taken too much time to make use of these tools, and the development of new algorithms that exploit them is an activity in progress. Among these works, an excellent example is the PhotoTourism project [16]. The main goal is to make a sparse 3D reconstruction of a popular monument from thousands of images using *structure-from-motion* type algorithms. Applications like that allow the user to explore unstructured photo collections.

The work presented in this article is also motivated by another kind of application that could exploit the richness of the photometric information available. For example in photomontage applications human intervention is often necessary to create realistic images. If one wants to insert an object from a photo into an environment determined in a second photo, both images should show up as if they had been taken under the same conditions. In that case, a human expert could retouch the images and force them to match illumination conditions and

* Supported by the Programme Alβan, the European Union Programme of High Level Scholarships for Latin America, scholarship No. E07D402742CO.

shadows [1, 3]. Our work is developed on the basis that we can exploit the above freely accessible collections and try to find a group of images that depict similar illumination conditions.

In a general context, one of the most successful approximations to extract information about the illumination is to use a spherical object in such a way that it can reflect and capture the light distribution when the photo is taken [17]. The main disadvantage of this method is that it requires access to the scene at the moment of the click. In his thesis [13], Love did some observations that let us think that the sky is a very important illumination source and, as consequence, one of the essential factors involved in the color perception of outdoor scenes. Some other researchers have proposed to find the behavior of the sky illumination using only images for example, using physically-based models [12], or other ones using advanced techniques for color correlation [11, 10]. If one does not limit analysis to photos including sky portions, there are several works in the image completion context (*hole filling* or *inpainting*) [6, 19] and in the context of color classification using raw pixels [14, 18]. Some of the researchs that explore the sky as information source show restricted constraints, for example the use of a single camera or an *a priori* knowledge about the calibration (in the photometric and geometric sense). These constraints try to reduce the number of variables involved in the color perception process (surface reflectance, light sources, sensor response). For instance, the work of Lalonde *et al.* [12] proposes an algorithm to find similar sky images given the intrinsic camera parameters and a physically-based sky model. The main idea consists in finding a sky model for each given image and to compare the features that define these models. This process can only be applied to images taken with a static camera, for example time lapse sequences. In [11], an automatic photomontage system that uses a presegmented object library is described. The method used by the authors to determine the global illumination matching consists of calculating the χ^2 distance between 3D-histograms that represent the sky. Short distances are recognized as possible matches. In our experiments, this metric presents a low performance due to the high dependance to the color workspace.

The present work aims to find the matches between sky zones of multiple images in a large image collection using minimum information about the camera parameters. For this, we propose the union and improvement of three stages previously developed in other contexts. The first one, a preamble stage that allows the sky pixels segmentation. A second stage where the pixels are represented by a sum of Gaussians in the La^*b^* space (section 2.1). Finally a third stage that compares the estimated models (section 2.2). The final part of this article (section 3) presents some results and discusses the algorithm proposed.

2 Representation and Comparison of Sky Regions

2.1 Sky Model in the La^*b^* color space

In the literature, it is common to find physically-based models that express the sky luminance as a function of diverse variables, among them, sun position and

atmospheric turbulence [9, 15]. These models have been used in Computer Vision under some constraints, like a fixed point of view. In the present work the images are extracted from a large database and came from different cameras. This fact increases dramatically the problem complexity because we do not have any information about the geometry of the scene. Moreover, we do not know how the data captured by the sensors was modified during the acquisition process (most cameras apply post-processing algorithms to enhance the "perceptual quality" of the images). Our methodology is based on the camera's final output, a *jpeg* file, assuming that most of the time this image does not represent faithfully the real world illumination at the acquisition time. To accomplish the goal of this work, the first stage consists in extracting the sky pixels for each image. The application described in [8] allows us to make an automatic segmentation of the image sky.

To determine a model from sky pixels is a decisive step in the formulation of our problem. For that reason, the choice of the color space plays an important role. The CIE (Commission Internationale de l'Éclairage) [4] have created several standards and nowadays, the last model published (CIECAM02) has reached an excellent performance as well as a high degree of complexity [5]. This color appereance model allows a trustworthy representation of the real world. However the parameters required are, in some cases, impossible to acquire. The La^*b^* space and the xyY space used by [11] and [12] are simpler models derived from the CIE's works. Although these color spaces are used in a vast number of works, it is important to take precautions when they are applied. For example, the spaces above mentioned are always attached to a predetermined "reference white", usually unknown. In this work we use the La^*b^* color space because of its good results in distinguishing color differences, but under the assumption that all the images were taken using a natural global illumination. That means that the "reference white" used in the transformation is the same for all images (Illuminant D65). Following the work done by Lalonde *et al.* [11], we build color histograms for the sky region in each image. These histograms are probability distributions of the colors in terms of the variables L, a^* and b^*. Figure 1 shows some of those histograms in their 3D space. Color and point size represent the bin magnitude[3]. One can note that the pixels are spread mainly on the negative part of the b^* axis which corresponds to variations between yellow and blue. On the other hand, the histograms of images, where the sky is white and/or partially cloudy, are distributed throughout the space and they are not easily distinguishable using this representation. We believe that these histograms have different modes, and that they could be modeled using a mixture of Gaussians.

According to our observations, the sky in each image is modeled following the next equation:

$$M(x) = \sum_{k=1}^{K} \pi_{\mathbf{k}} \mathcal{N}\left(\mathbf{x} | \mu_{\mathbf{k}}, \mathbf{\Sigma_k}\right) \quad , \tag{1}$$

for an undetermined K (in our implementation we limit this value to 4). One well-known method to find the parameters in equation (1) is the Expectation-

[3] For figures, the reader is kindly invited so see the electronic version.

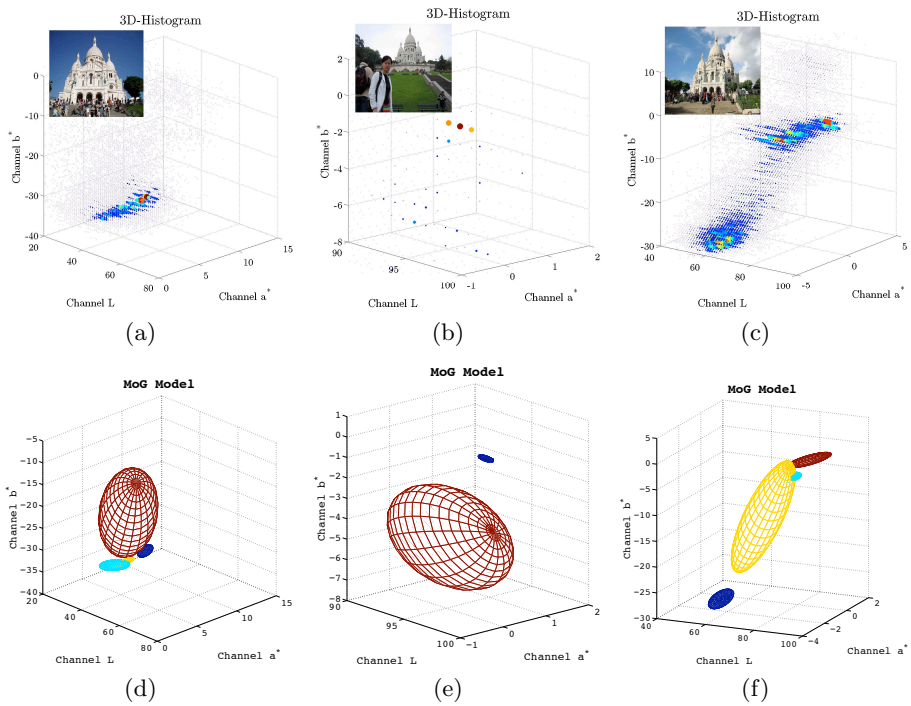

Fig. 1. Sky 3D-histograms and ellipsoids for the corresponding MoG models. (a) Image with sunny sky, (b) completely cloudy sky, (c) partially cloudy sky. (d) Model with four Gaussians. (e) Model with two Gaussians. (f) Model with four Gaussians.

Maximization algorithm (EM) [2]. The output of this method corresponds to the variables π_k, $\mu_\mathbf{k}$, $\mathbf{\Sigma_k}$ that describe the model in such a way that the probability of the solution is maximized. Our model is formulated in terms of a joint probability distribution of the sky pixels X, the latent variables Z and the goal is to find the set of values θ that maximize the likelihood function (θ represents the variables $\pi_\mathbf{k}$, $\mu_\mathbf{k}$ and $\mathbf{\Sigma_k}$). The values La^*b^* of the sky pixels $\mathbf{x_n}$ are organized in the matrix X in which the nth row is given by $\mathbf{x_n^\mathsf{T}}$. In order to briefly summarize this algorithm, an iterative 2 step process runs until it reaches a convergence parameter. The $\mathbf{E}$ step calculates the joint probability distribution $P(\mathsf{Z}|\mathsf{X}, \theta^{\text{last}})$:

$$P(\mathsf{Z}|\mathsf{X}, \theta^{\text{last}}) = \frac{\pi_\mathbf{k}\mathcal{N}\left(\mathbf{x_n}|\mu_\mathbf{k}, \mathbf{\Sigma_k}\right)}{\sum_{j=1}^{K} \pi_\mathbf{j}\mathcal{N}\left(\mathbf{x_n}|\mu_\mathbf{j}, \mathbf{\Sigma_j}\right)} \quad ,$$

and the $\mathbf{M}$ step updates the values of $\pi_\mathbf{k}$, $\mu_\mathbf{k}$, $\mathbf{\Sigma_k}$. The model dimension is a crucial factor. In our case, the number of Gaussians used in the mixture is determined based on the Akaike information criterion (AIC) [2]. Different values of K in the equation (1) are evaluated and, the model that maximizes the likelihood is chosen. In figure 1, the ellipsoids that form the Gaussian Mixture model are shown for a constant variance.

2.2 Comparison between histograms of the sky

Once the sky model for each image is estimated, we proceed to compare different models. In the present case, it is necessary to measure the difference between two or more probability distributions. In our context, the well-known Kullback-Leibler divergence (KL) should be a good option, although it does not possess the property of symmetry. Given two probability distributions $p(x)$ and $q(x)$ the KL divergence is defined by:

$$\mathsf{KL}(p||q) = -\int p(x) ln \left\{ \frac{q(x)}{p(x)} \right\} dx \ . \tag{2}$$

It is widely proven that when the distributions are Gaussian, equation (2) can be expressed in closed-form. However, in the case of a Gaussian mixture, it is difficult to find an analytically tractable expression or even more, a computer algorithm to solve this problem efficiently. According to the work of Hershey and Olsen [7], the only method for estimating $\mathsf{KL}(p||q)$ with arbitrary accuracy when $p(x)$ and $q(x)$ are mixtures of Gaussians, is the Monte Carlo simulation. Nevertheless, other approximations may be valid, depending on the context. For example, a commonly used approximation is the simplification of the Gaussian mixtures $p(x)$ and $q(x)$ by simple Gaussians $\tilde{p}(x)$ and $\tilde{q}(x)$. In this case, the mean and covariance estimated are:

$$\begin{aligned} \mu_{\tilde{p}} &= \sum\nolimits_a \pi_a \mu_a \\ \Sigma_{\tilde{p}} &= \sum\nolimits_a \pi_a \left(\Sigma_a + (\mu_a - \mu_{\tilde{p}})(\mu_a - \mu_{\tilde{p}})^{\mathsf{T}} \right) \ . \end{aligned} \tag{3}$$

The KL divergence (KL_{sim}) is calculated using the estimated mean ($\mu_{\tilde{p}}$) and variance ($\Sigma_{\tilde{p}}$). Hershey and Olsen use variational methods to find a better approximation of the KL divergence. One of their contributions is a measure that satisfies the symmetry property but not the property of positivity. In this case, the approximated divergence $\mathsf{KL}_{\text{app}}(p||q)$ is given by:

$$\mathsf{KL}_{\text{app}}(p||q) = \sum_a \pi_a \log \frac{\sum_{a'} \pi_{a'} e^{-\mathsf{KL}(p||p')}}{\sum_b \omega_b e^{-\mathsf{KL}(p||q)}} \ . \tag{4}$$

This value could be seen as a measure of dissimilarity between two distributions. Hershey and Olsen's contribution allows us to compare two sky models keeping the variables for each Gaussian that composes the mixture.

3 Results

This section shows some results of the experiments developed using the models estimated and the comparison measure described in the last section. To test our method, we re-create a database with 4250 images of Sacre Cœur's Cathedral (Paris) downloaded from the Flickr website[4]. Images that do not correspond to

[4] http://www.flickr.com/

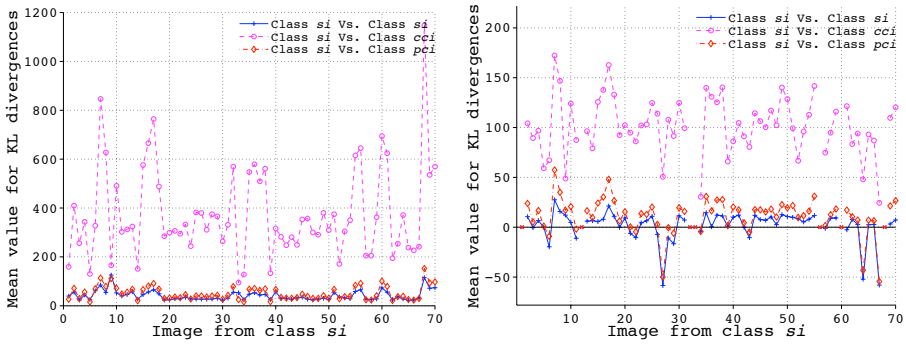

Fig. 2. Mean values for the intra-class and inter-class divergence, (a) using the $\mathsf{KL}_{\text{simple}}$ divergence and (b) using the KL_{app} divergence.

Table 1. Average of the inter and intra-class divergences for all images in the class *si*.

	Average $\mathsf{KL}_{\text{simple}}$	Average KL_{app}
si class	45.60	4.27
pci class	351.63	97.87
cci class	50.95	13.73

the desired scenes or are taken during the night or using artificial lighting had been removed.

Our goal is to compare two or more outdoor images using the above methods applied on sky regions. To test our approach, we classified 500 images from the database by hand according to: sunny (*si*), partially cloudy (*pci*) and completely cloudy (*cci*). It is important to emphasize the subjective nature of this "ground truth". The main idea consists in comparing the distances between one selected image and the other ones that belong to the same class (intra-class distance) and those belonging to other classes (inter-class distance). Figure 2 shows the intra-class and inter-class mean distances computed for 70 images from the class *si* using the KL_{sim} divergence and the KL_{app} divergence. For these two measures, we find that the class *cci* differs clearly from the other two by medium values of greater amplitude. However, among the class *si* and the class *pci* the difference is not sufficient when using the measure KL_{sim}. As we may expect, the approximated KL_{app} divergence shows a better performance and discrimination is easier. Table 1 presents the average of the inter and intra-class divergences for all images in the class *si* according to the two measures described. The difference between images belonging to classes *si* and *cci* is clearer when we use the KL_{app} divergence. Once again, it can be explained by a better approximation of the KL divergence.

Another experiment was developed in order to find the success rate of the algorithm when running on 500 images using the KL_{app} divergence. The objective

Table 2. Success rate on 500 images.

Images	1st Image	2nd Image	3rd Image	4th Image
All	78.7%	75.4%	72.0%	71.2%
Class *cci*	77.3%	74.6%	70.8%	68.7%
Class *si*	64.0%	68.8%	64.8%	59.1%
Class *pci*	93.2%	82.0%	79.9%	84.8%

this time is to find the closest images and whether or not they belong to the same class that the query image, according to the ground truth (see table 2). Figure 3 shows the four most similar images found for three query images.

4 Conclusions

In this paper, we proposed a 3-steps pipeline (segmentation, modelisation and comparison) for grouping outdoor images with similar lighting conditions, based on techniques previously developed. The model choosen for the sky pixles has been a mixture of Gaussian, based on the observations of histograms with multiple modes. This distribution allows us to accurately model the pixels, especially for images where the sky is not of uniform color. On the other hand, the computation of the KL divergence has proven to be an adequate tool to group similar skies, despite the approximations made in the operation. These approximations must be carefully chosen in order to keep the benefits of using a multimodal distribution as model. For a finer grouping, other factors may be considered such as the position of the sun, clouds, shadows or information about the camera parameters that could be extracted from the metadata. Also, to obtain credible photomontages, information about the geometry of the scene could be valuable. The inclusion of these variables might produce natural and realistic compositions.

References

1. A. Agarwala, M. Dontcheva, M. Agrawala, S. Drucker, A. Colburn, B. Curless, D. Salesin, and M. Cohen, *Interactive digital photomontage*, SIGGRAPH 2004 (New York, NY, USA), ACM, pp. 294–302.
2. C. M. Bishop, *Pattern recognition and machine learning (information science and statistics)*, Springer, August 2006.
3. P. J. Burt and R. J. Kolczynski, *Enhanced image capture through fusion*, ICCV 1993., 173–182.
4. Committee C.T., *Spatial distribution of daylight - luminance distributions of various reference skies.*, Tech. Report CIE-110-1994, Commission Internationale de l'Éclairage (CIE), 1994.
5. ———, *Colour appearance model for colour management applications*, Tech. Report CIE-TC8-01, Commission Internationale de l'Éclairage (CIE), 2002.
6. J. Hays and A. Efros, *Scene completion using millions of photographs*, SIGGRAPH 2007 **26**, no. 3, 4.

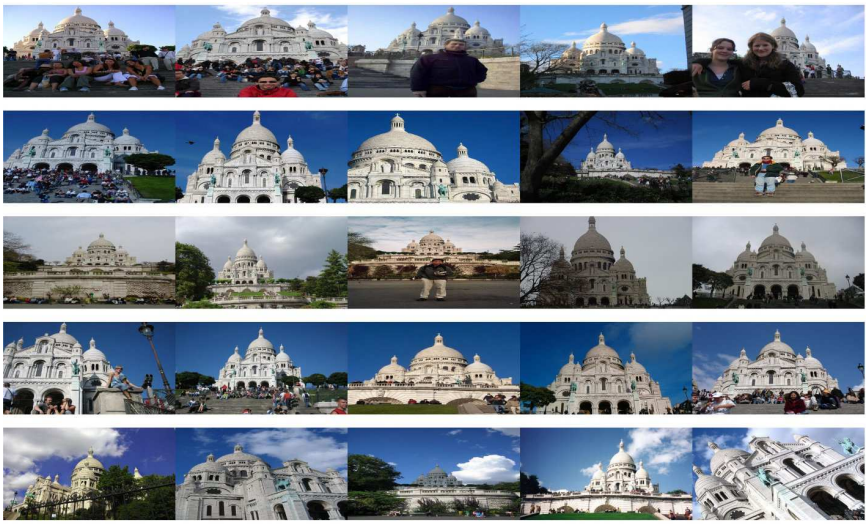

Fig. 3. Subset of typical results. In the first column we observe the query image. The following columns from left to right, show the images closest to the query image.

7. J. R. Hershey and P. A. Olsen, *Approximating the kullback leibler divergence between gaussian mixture models*, ICASSP 2007 **4**, IV–317–IV–320.
8. D. Hoiem, A. Efros, and M. Hebert, *Geometric context from a single image*, ICCV 2005, pp. 654–661.
9. N. Igawa, Y. Koga, T. Matsuzawa, and H. Nakamura, *Models of sky radiance distribution and sky luminance distribution*, Solar Energy **77** (2004), no. 2, 137 – 157.
10. N. Jacobs, N. Roman, and R. Pless, *Consistent temporal variations in many outdoor scenes*, CVPR 2007, 1–6.
11. J. Lalonde, D. Hoiem, A. Efros, C. Rother, J. Winn, and A. Criminisi, *Photo clip art*, SIGGRAPH 2007 **26**.
12. J. Lalonde, S. G. Narasimhan, and A. Efros, *What does the sky tell us about the camera?*, ECCV 2008.
13. Robert Charles Love, *Surface reflection model estimation from naturally illuminated image sequences*, Tech. report, The University of Leeds, Ph.D. thesis, 1997.
14. R. Manduchi, *Learning outdoor color classification*, IEEE Transactions on PAMI **28** (2006), no. 11, 1713–1723.
15. R. Perez, R. Seals, and J. Michalsky, *All-weather model for sky luminance distribution - preliminary configuration and validation*, Solar Energy **50(3)** (1993), 235–245.
16. N. Snavely, S. M. Seitz, and R. Szeliski, *Modeling the world from internet photo collections*, Int. J. Comput. Vision **80** (2008), no. 2, 189–210.
17. J. Stumpfel, A. Jones, An. Wenger, C. Tchou, T. Hawkins, and P. Debevec, *Direct hdr capture of the sun and sky*, AFRIGRAPH 2004, pp. 145–149.
18. K. Sunkavalli, F. Romeiro, W. Matusik, Y. Zickler, and H. Pfister, *What do color changes reveal about an outdoor scene?*, CVPR 2008, 1–8.
19. M. Wilczkowiak, G. J. Brostow, B. Tordoff, and R. Cipolla, *Hole filling through photomontage*, BMVC 2005, July, pp. 492–501.