
Statistical Analysis and Transfer of Pictorial Styles

Soonmin Bae
Frédéric Durand

SMBAE@MIT.EDU
FREDO@MIT.EDU

MIT Computer Science and Artificial Intelligence Laboratory, 32 Vassar Street, Cambridge MA, 02139 USA

Abstract

Humans read content from images. However, images have not only content, but also a notion of style. Some pictorial styles differ noticeably and can be distinguished without careful inspection of the picture. This paper analyzes the style that can be noticeable at a pre-attentive level, using image statistics.

Since the way images are decoded in the human visual system is related to the structure of visual data in the images, we examine the statistical properties of images. It has been found that visual data in natural images is highly structured and statistical tools have been frequently used to parameterize the structure. We extend the range of statistical analysis to parameterize and transfer simple notions of pictorial style and ambiance in images.

Furthermore, it has been shown that images are processed in the human visual system using oriented linear filters. Therefore, this paper will exploit steerable pyramids, which have multiscale and oriented subbands resembling simple cell receptive fields. We show that the average of the absolute steerable coefficients, as a function of the scale, characterizes simple notions of composition and style.

1. Introduction

Digital cameras have brought a revolution by facilitating the way we take and share pictures. Moreover, digital processing will improve the quality of pictures and the flexibility with which we can interact with them, whereas traditional photographic skill has relied on complex lighting, tedious processing in the dark room, and the eyes of talented artists to create compelling pictures. In the same way as traditional techniques have allowed skilled photographers to achieve mood, tonal balance, contrast, and the emphasis of important parts of the picture, this project provides novice users with stylish enhancement for black-and-white or color photography and video, developing a tool which captures pictorial styles from master photographers. This tool will help users to create, manipulate and enhance pictures to achieve compelling pictorial styles.

This paper attempts to link the statistical properties of images with their aesthetic characteristics, which are hard to quantify. Because of the difficulty in measuring aesthetic characteristics, other studies have characterized pictorial styles in terms of history, artists, artists' nationalities and so forth. Instead, this paper explores pictorial styles using explicit quantitative measurements to parameterize notions of pictorial style prevailing over the image. However, we do not claim that quantitative criteria can capture all aspects of style; they will certainly not capture genius or the notion of art. Nevertheless, they will afford a better understanding of certain dimensions of style.

The intuition behind this project is that some pictorial styles differ noticeably and can be distinguished even at a pre-attentive level, that is, without carefully scrutinizing the picture. For example, the difference between a Pointillist painting and a Renaissance picture is obvious because the former contains much more detailed texture than the latter. Therefore, we want to capture such stylistic features that are easily distinguishable at a pre-attentive level.

In this paper, image-processing and statistical tools, which have been used to study the statistical structures in natural images, are applied to characterize pictorial styles and to demonstrate that statistical estimators on color and multiscale content correlate with pictorial styles. We further validate the relevance of these statistical properties by transferring the statistical properties from one picture to another, which confirms that the statistical properties correspond to major visual aspects of these styles.

1.1 Related Work

1.1.1 STATISTICS OF PICTORIAL STYLE

Previous studies on statistical characterization of pictorial styles have focused on classification, either between the styles of individual painters or between photographs and paintings. Most approaches to painting style classification have worked on images labeled by individual artist, and have used supervised machine learning techniques based on various image features such as wavelet coefficients, color histograms, or oriented filters.

Style classification is also relevant to content-based image retrieval that uses similar features. The results of Leykin, Cutzu and Hammoud (Leykin & Cutzu, 2003; Cutzu & Hammoud, 2003; Leykin & Cutzu, 2002; Farid & Lyu, 2003) are close to our work since their quantitative model of style corresponds to explicit visual qualities. They made their own assumptions on what made paintings different from photographs and used these assumptions to build the features used for classification. However, while their approaches rely on parametric models, they do not provide synthesis or manipulation capabilities.

1.1.2 STYLE TRANSFER

Recently, Shannon's n-grams (Shannon, 1948) have inspired approaches that transfer textural aspects of pictorial styles by imitating local textures from a source image to a destination image or between strokes. Although these approaches can reproduce the detailed texture better than our approach, they do not provide similarity metrics between styles, and being non-parametric, they do not afford natural interpolation between styles.

1.2 Overview

This paper makes contributions in two areas: the statistical analysis of pictorial style and style transfer between images. Both contributions rely on a new set of statistical characterizations of the distinctive visual quality of pictures. We focus on statistical features having a clear visual impact and those that can be naturally transferred.

In particular, steerable pyramids (Freeman & Adelson, 1991; Simoncelli et al., 1992; Simoncelli & Freeman, 1995) were used to compute image statistics, since they provide oriented multiscale decomposition and have good reconstruction properties; they are free of aliases and use a smooth reconstruction basis. Furthermore, this paper demonstrates its validity by transferring captured statistical properties between paintings and photographs.

2. Multiscale Content of Images

2.1 Multiscale Analysis of Pictorial Styles

This project used steerable pyramids for their orientation content and good reconstruction capabilities. While Fourier analysis would be the first choice to study frequency content, it suffers from poor spatial localization that would result in reconstruction artifacts. For each image, a steerable pyramid with four orientations is generated and the average of the absolute coefficient of the pyramid is calculated for each scale with four subbands. The average coefficients are plotted in Fig. 1, 2 and 3 and the coefficients are normalized by $2s$, where s is the scale to obtain equal energy at all scales.

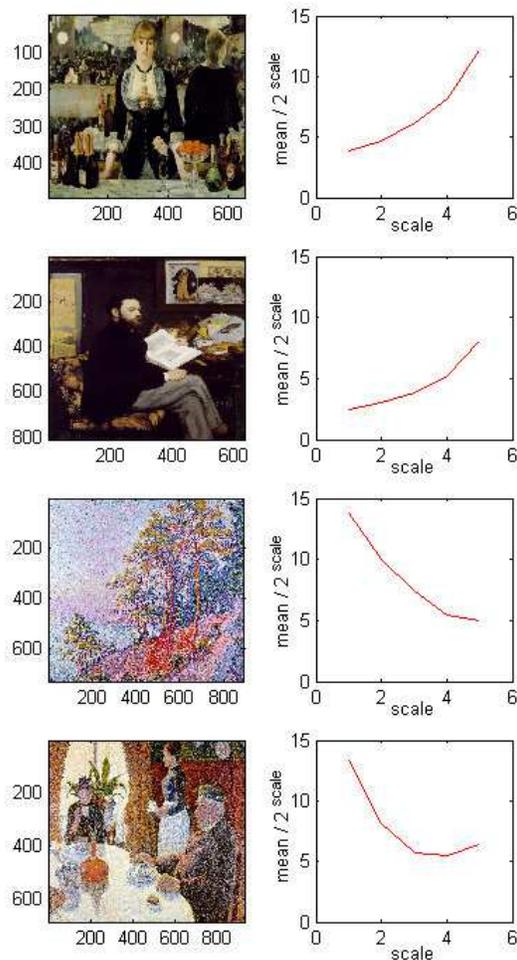


Figure 1. The mean of the absolute values of the detail coefficients from the steerable pyramid. On the left is small scale and the right is large scale. The first two plots are for Manet's *A Bar at the Folies-Bergeres* and *Portrait d'Emile Zola*, and the last two plots are for Signac's *Above Saint-Tropez* and *The Dining Room*.

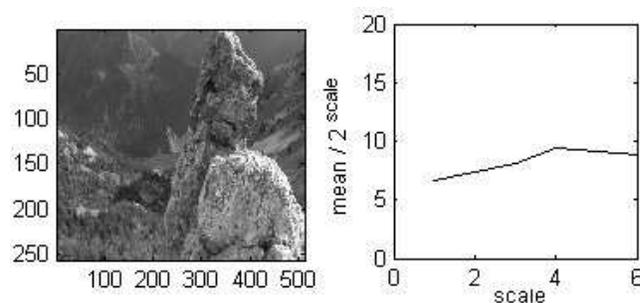


Figure 2. A plot for the means of the absolute values of the detail coefficients at scales from 1 to 5 in the steerable transform of the natural image.

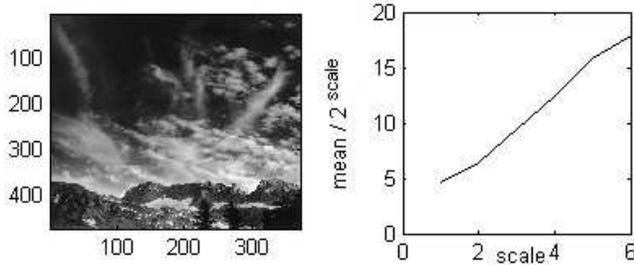


Figure 3. A plot of the means of the absolute values of the detail coefficients at scales from 1 to 5 in the steerable transform of the artistic photograph.

Manet's paintings have low values at the fine scale and high values at the coarse scale, whereas Signac's paintings have high values at the fine scale. This points out that Signac used many small points and that Manet emphasized the large-scale contrast between large regions of the paintings.

Comparing Fig. 3 with Fig. 2, we can observe that the photograph by Ansel Adams has less power in the high frequency and more in the low frequency, whereas the natural image in Fig. 2 has a flat curve. This points out that Ansel Adams emphasizes the large-scale contrast between large regions of the photograph.

2.2 Multiscale Transfer of Pictorial Styles

The importance of these curves of multiscale average coefficients can be assessed by modifying images by matching the curves. Using a casual black-and-white photograph of a landscape, its multiscale coefficients were changed to match the average curves of a photograph by Signac or Manet, and of a photograph by Ansel Adams (Fig. 4). Note in particular how the image that matches the multiscale characteristics of Ansel Adams obtained more interesting contrasts between large regions of the image.

3. Conclusions

Our technique exploits multiscale decompositions of images to parameterize and transfer simple notions of pictorial style and ambiance in images which can be noticeable at a pre-attentive level. To achieve this goal, a multiscale decomposition is performed using steerable pyramids. Not only were the pyramids selected related to behaviors of the human visual system, but also for its good reconstruction capabilities, since our approach is both to transfer and to analyze stylistic aspects.

The results show that the distributions of energy over the image, which are represented by the average coefficient curves of steerable pyramids, are related to pictorial styles. In addition, by matching the average curves, transfer of the

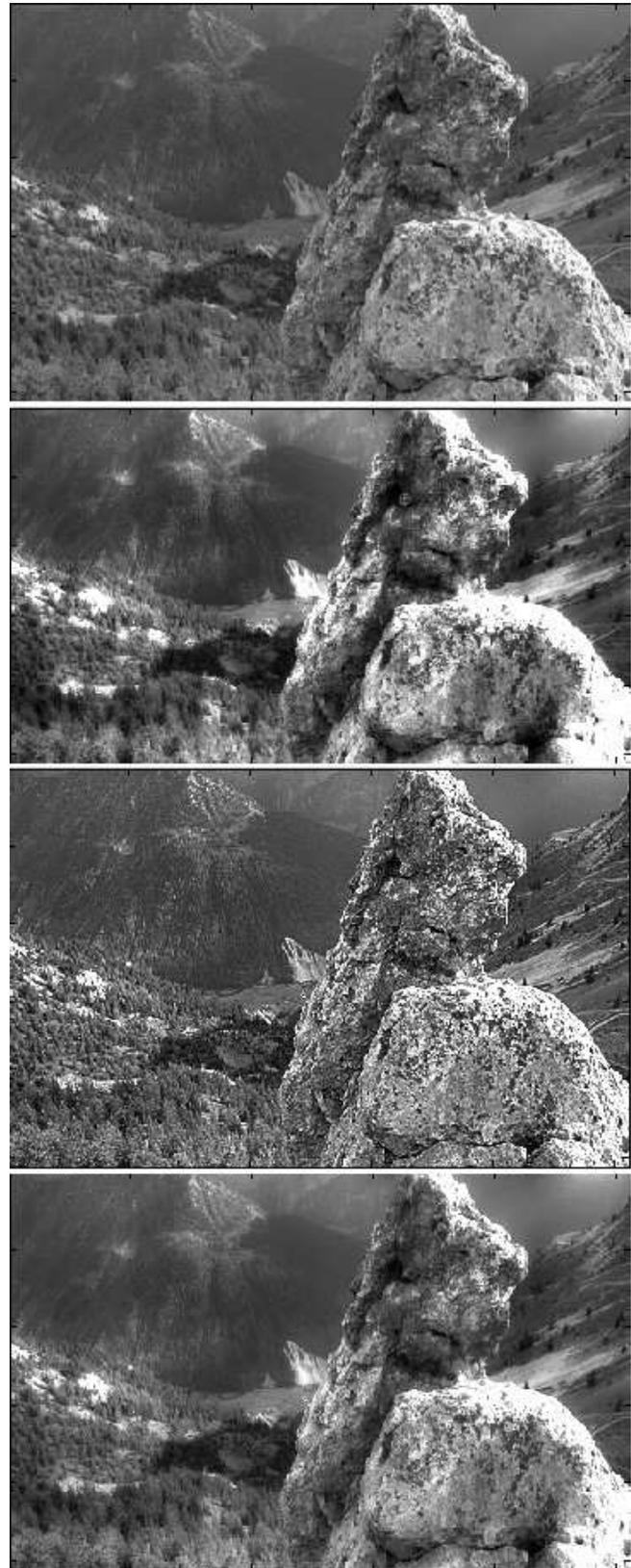


Figure 4. The top is a natural image, the second was modified with the Manet histogram, the third with the Signac histogram, and the bottom with the Adams Lab histogram.

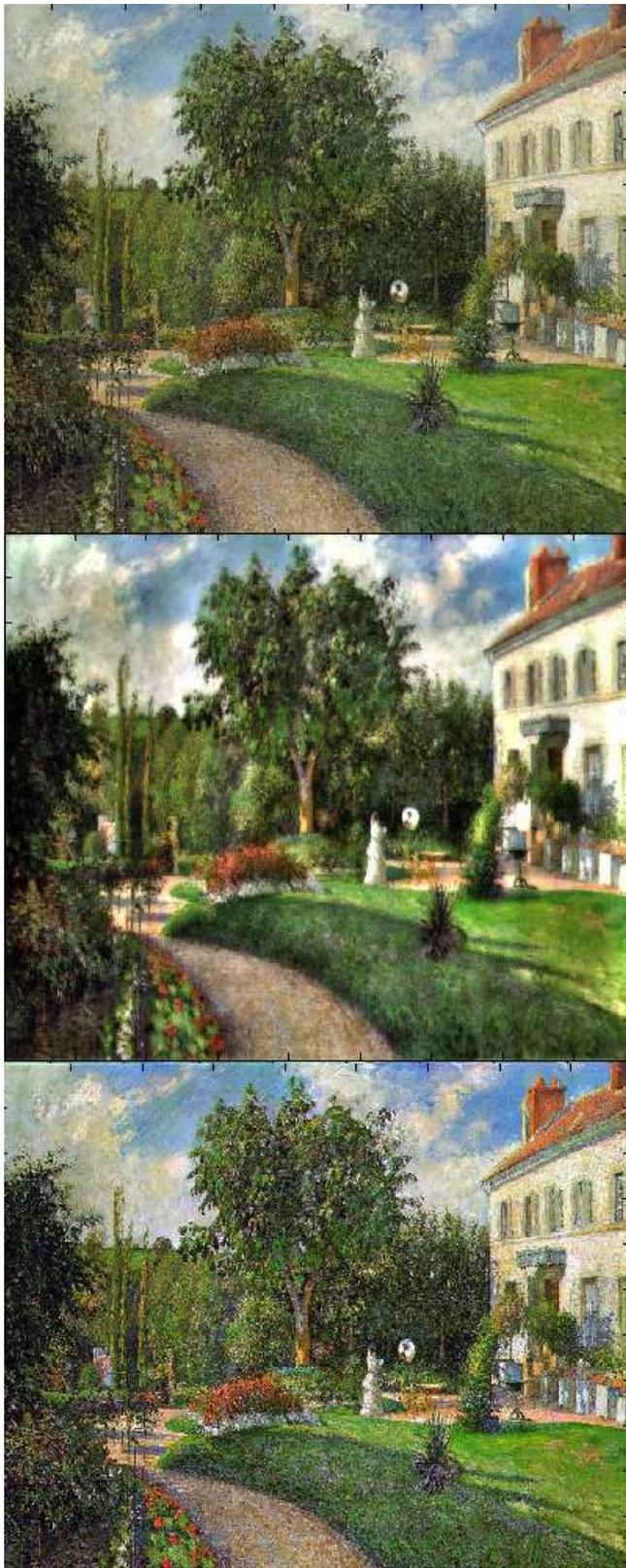


Figure 5. Pissarro's *Garden of Les Mathurins at Pontoise* original, with the Manet histogram and with the Signac histogram.

pictorial styles can be achieved.

The practical contribution of this research is that the transfer tool can be used to provide novice users with stylish enhancement for black-and-white or color photography and video, since the tool can capture pictorial styles from master photographers and can help users to create, manipulate and enhance pictures to achieve compelling pictorial styles.

References

- Cutzu, F., & Hammoud, R. (2003). Estimating the photo-realism of images: Distinguishing paintings from photographs. *Proc. of CVPR*.
- Farid, H., & Lyu, S. (2003). Higher-order wavelet statistics and their application to digital forensics. *IEEE Workshop on Statistical Analysis in Computer Vision (in conjunction with CVPR)*.
- Freeman, W. T., & Adelson, E. H. (1991). The design and use of steerable filters. *IEEE Trans. Patt. Anal. and Machine Intell.*, 13.
- Huang, J., & Mumford, D. (1999). Statistics of natural images and models. *Proc. Conf. CVPR* (pp. 541–7).
- Leykin, A., & Cutzu, F. (2002). Distinguishing paintings from photographs. *Journal of Vision*, 2.
- Leykin, A., & Cutzu, F. (2003). Differences of edge properties in photographs and paintings. *IEEE International Conference on Image Processing*.
- Shannon, C. (1948). The mathematical theory of communication. *Bell System Technical Journal*.
- Simoncelli, E. P., & Freeman, W. T. (1995). The steerable pyramid : A flexible architecture for multi-scale derivative computation. *Second Int'l Conf. on Image Proc.* (pp. 444–447).
- Simoncelli, E. P., Freeman, W. T., Adelson, E. H., & Heeger, D. J. (1992). Shifttable multi-scale transforms. 38, 587–607.