

早稲田大学大学院情報生産システム研究科

博士論文概要

論文題目

**Study on Constant-time
Gaussian/Bilateral Filters
and Fast Color Descriptor**

申請者

Kenjiro SUGIMOTO

情報生産システム工学専攻
イメージメディア研究

2015年4月

Digital image technologies have become common among many people and utilized for various purposes during the past two decades. As common use, an enormous volume of image/video data are mundanely shared via social networking services in the world, which is promoted by downsizing and commoditization of digital cameras. Rapid advance of digital devices has made image/video data higher-quality information resources such as higher resolution and frame-rate year by year. As professional use, higher-dimensional data scanning became essential equipment in various fields, e.g., the Magnetic Resonance Imaging (MRI) in the medical world. Computational photography, which is a novel imaging technology that integrates image processing and optical systems, is also expected to spread throughout professional fields in the future. In parallel, this recent trend leads to a growing demand for more sophisticatedly understanding the underlying information contained in a visual scene caught on an image or a wealth of images, e.g., three-dimensional reconstruction from a tremendous amount of images, object identification from images captured under tough conditions, and efficient denoising to vast quantities of high-dimensional image data, and so forth. Thus, digital image processing exerts a great important role more than ever before.

This dissertation mainly focuses on image filtering, which is one of the fundamental techniques in the fields of image processing, computer vision, and computer graphics. Even though image filtering is widely known as a key tool for denoising, it actually plays various roles in many modern applications over the last decade including object recognition, stereo vision, high-dynamic range imaging, volume data rendering, and visual saliency. Under the recent trend above mentioned, the modern applications face a major difficulty of computational efficiency. This is because higher-resolution or higher-dimensional images require much more arithmetic operations, which normally shows some orders more than linear to data size. Moreover, most state-of-the-art algorithms in image processing perform image filtering many times to understand a visual scene more sophisticatedly. In short, despite seeming old-fashioned, image filtering still plays a dominant role for the current research in computational efficiency. One may think that utilizing parallel processors such as Graphical Processing Unit (GPU) or cloud computing is a silver bullet to the performance problem but it is inapplicable to all situations, e.g. mobile devices and consumer electronics. Hence, algorithmic improvements for the problem should be considered more than hardware development.

A useful algorithmic solution is constant-time filters, which have computational complexity independent of filter window size, i.e., they run in linear or-

der to the number of pixels. This algorithmic property has a more notable effect on higher-resolution or higher-dimensional images because they require larger filter window. Many constant-time algorithms based on various concepts have been proposed in the past. We considered that most of them basically share a key framework called kernel decomposition. For example, a linear filter can be approximated by a bunch of filters that can be convolved in constant-time by some recursive ways via spatial kernel decomposition into simple-structured kernels, e.g., box kernel and one-sided attenuating kernels. Similarly, a non-linear filter can be approximated by a bunch of linear filters via range kernel decomposition, which is also replaced to constant-time filters as mentioned above. A natural question derived from this consideration is how kernel decomposition can achieve the optimal performance. Based on this aspect, we found that state-of-the-art algorithms still have room for improvement in terms of performance efficiency from a theoretical viewpoint. We tackle this question using the Karuhunen-Loeve transform (KLT), i.e., the Principal Component Analysis (PCA), and the Discrete Cosine/Sine Transforms (DCT/DST), which provide the optimal or a near-optimal performance tradeoff between computational complexity and approximate accuracy.

Chapter 1 provides a general introduction of this dissertation. We first summarize the remaining problems that our research fields have and then discuss the aforementioned remaining problems. The main contributions and organization of this dissertation are also described.

Chapter 2 presents an efficient constant-time algorithm for Gaussian and Gaussian derivative filters that provide high approximate accuracy in low computational complexity regardless of its filter window size. The proposed algorithm consists of two key techniques: second-order shift properties of the DCT/DST Type-5 (DCT/DST-5) and dual-domain error minimization for finding the optimal parameters. The former enables us to perform filtering in fewer numbers of arithmetic operations as compared with state-of-the-art algorithms. The latter enables us to find the optimal filter size that provides the most accurate filter kernel approximation. Experiments show that the proposed algorithm outperforms state-of-the-art ones in computational complexity, approximate accuracy, and accuracy stability. Specifically, the results show roughly 2.5 times faster than state-of-the-art constant-time Gaussian and Gaussian derivative filters without a loss of accuracy.

Chapter 3 proposes an efficient constant-time bilateral filter that produces a near-optimal performance tradeoff between approximate accuracy and computational complexity without any complicated parameter adjustment, called a

compressive bilateral filter (CBLF). Although many existing constant-time bilateral filters have been proposed step by step to pursue a more efficient performance tradeoff, they have less focused on the optimal tradeoff for their own frameworks. It is important to discuss this question because it can reveal whether or not a constant-time algorithm still has plenty room for improvements of performance tradeoff. This chapter tackles the question from a viewpoint of compressibility and highlights the fact that state-of-the-art algorithms have not yet touched the optimal tradeoff. The CBLF achieves a near-optimal performance tradeoff by two key ideas: approximate Gaussian range kernel through Fourier analysis and period length optimization. Moreover, it utilizes the constant-time Gaussian filter proposed in Chapter 2. Experiments demonstrate that the CBLF significantly outperforms state-of-the-art algorithms in terms of approximate accuracy, computational complexity, and usability. Specifically, it shows at most 8 times faster than a state-of-the-art constant-time bilateral filter without sacrifice in accuracy.

Chapter 4 discusses a color-based method for medicine package recognition, called a linear manifold color descriptor (LMCD). In general, color distribution, which is a set of color pixels, is compactly described as histogram or clusters. However, these approaches are oversensitive to noise and light condition. We avoid this problem by describing the color distribution of a color package image as a linear manifold in the color space, which is derived via the KLT. We then recognize an anonymous package by matching linear manifolds. Our analysis reveals that the LMCD is effective to color distribution consisting of few color clusters, which medicine packages generally satisfy this condition. Mainly due to low dimensionality of color spaces, it can provide more compact description and faster computation than description styles based on histogram and clusters. This chapter also proposes distance-based dissimilarities for linear manifold matching. Experiments on medicine package recognition validates that the LMCD outperforms competitors including MPEG-7 color descriptors in terms of description size, computational cost and recognition rate. Specifically, the rate is approximately 1.0-6.5% higher than the histogram-based method of MPEG-7.

Chapter 5 concludes this study and suggests future work. Our proposed algorithms show significantly higher efficiency than the state-of-the-art algorithms in many applications. Moreover, their fundamental concepts can be extended to other more complicated algorithms including the trilateral filter and scale-space analysis. By combining them with hardware acceleration, they will play essential roles for modern image processing applications in the future.