Computers & Graphics 38 (2014) 150–157

Contents lists available at ScienceDirect

Computers & Graphics

journal homepage: www.elsevier.com/locate/cag

CAD/Graphics 2013

Feature-preserving filtering with L_0 gradient minimization

Xuan Cheng [a], Ming Zeng [b], Xinguo Liu [a,*]

[a] State Key Lab of CAD&CG, Zhejiang University, China
[b] Software School of Xiamen University, China

ARTICLE INFO

Article history:
Received 5 August 2013
Received in revised form
4 October 2013
Accepted 19 October 2013
Available online 31 October 2013

Keywords:
Feature-preserving
L_0 norm
Gradient
Fused coordinate descent

ABSTRACT

Feature-preserving filtering is a fundamental tool in computer vision and graphics, which can smooth input signal while preserving its sharp features. Recently, a piecewise smooth model called L_0 gradient minimization, has been proposed for feature-preserving filtering. Through optimizing an energy function involving gradient sparsity prior, L_0 gradient minimization model has strong ability to keep sharp features. Meanwhile, due to the non-convex property of L_0 term, it is a challenge to solve the L_0 gradient minimization problem. The main contribution of this paper is a novel and efficient approximation algorithm for it. The energy function is optimized in a fused coordinate descent framework, where only one variable is optimized at a time, and the neighboring variables are fused together once their values are equal. We apply the L_0 gradient minimization in two applications: (i) edge-preserving image smoothing (ii) feature-preserving surface smoothing, and demonstrate its good performance.

1. Introduction

Many problems in computer vision and graphics require estimating some spatially varying quantity from noisy raw data. A main property of such quantity is piecewise smoothing: they should vary smoothly almost everywhere except at the sharp features that should be preserved. For this purpose, feature-preserving filtering was proposed, and became a fundamental tool in many applications.

Meanwhile, feature-preserving filtering is inherently challenging, because it is difficult to distinguish features from noise. There exist some different feature-preserving filtering methods, which follow the same energy minimization framework with a data term and a smooth term. The data term measures the disagreement between the filtered signal and the original signal, while the smooth term measures the extent to which the filtered signal is not piecewise smooth. The design of the data term is usually straight forward. For instance, the squared L_2 distance between the filtered signal and the original signal is often used. The choice of the smooth term is a critical issue.

A key observation is that, if a signal is piecewise smooth, most of the gradients tend to be small or even zero, and large gradients only appear at the sharp features. This gradient sparsity prior brings us a great help to design the smooth term. A representative work is total variation [1], where the smooth term is the L_1 norm of gradient.

Recently, Xu et al. [2] use L_0 term instead of L_1 term to directly measure the gradient sparsity in the context of image smoothing, and achieve some promising results. Compared with L_1 norm, L_0 norm can obtain more sparse solution. Some mathematical analysis [3] have shown that, under certain conditions, there is a formal equivalence between L_1 norm and L_0 norm. But this equivalence does not hold in our case. Gradient with L_1 norm or L_0 norm will lead to completely different solution, as validated by our experimental results.

Although L_0 gradient minimization is very suitable for feature-preserving filtering, it is indeed a non-convex problem. Xu et al. [2] give an approximation of this problem with a variational method. Through introducing a set of auxiliary variables related to gradients, the original non-convex problem is decomposed to a sequence of computationally tractable L_0–L_2 problems. The sparsity of auxiliary variables obtained by L_0 minimization is expected to be transferred to the gradients through L_2 minimization. However, since L_2 norm tends to severely penalize outliers and propagate the residual uniformly, the sparsity may be corrupted in transferring. Thus the gradients obtained by their method are not sparse enough.

In this paper, we propose a new approximation algorithm for L_0 gradient minimization problem. Our algorithm is based on a fused coordinate descent framework. It can obtain a solution with good gradient sparsity and sufficiently close to the original input. We apply the L_0 gradient minimization in edge-preserving image smoothing and feature-preserving surface smoothing. We compare our method with some existing feature-preserving filtering methods, which show that our method produces better results.

The rest of paper is organized as follows. Section 2 reviews some related work, Section 3 introduces our algorithm for L_0 gradient minimization, and Section 4 presents some applications in image smoothing and surface smoothing. Finally, we conclude our work in Section 5.

* Corresponding author. Tel.: +86 571 8820 6681x526; fax: +86 571 8820 6680
E-mail address: xgliu@cad.zju.edu.cn (X.Liu).

0097-8493/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.cag.2013.10.025

2. Related work

Edge-preserving image smoothing: A good edge-preserving image smoothing method should not blur the edges that are vital for neural interpretation to make the best sense of the scene, while smooth the regions between such edges. Bilateral filtering [4] is a traditional method dealing with this problem, which extends the concept of Gaussian smoothing by re-weighting the filter coefficients with their corresponding relative pixel intensities. And several variants of the bilateral filtering have been proposed [5–7]. Anisotropic diffusion [8] achieves edge-preserving smoothing by involving an edge-stopping function to make smoothing take place only in the interior of regions without crossing edges. Farbman et al. [9] formulate this problem in a weighted least square framework, which is more flexible compared with local filtering such as bilateral filtering. Rudin et al. [1] propose total variation, which utilizes the gradient sparsity enforced by an L_1 penalty term to do edge-preserving smoothing. Xu et al. [2] use an L_0 penalty term to directly measure gradient sparsity, and their method has a stronger ability to preserve edges. There also exist some methods depending on local features [10,11]. Our method is a global estimation process.

Feature-preserving surface smoothing: To preserve the sharp features of surface, anisotropic methods are often used. Bajaj and Xu [12] extend the anisotropic diffusion [8] in image processing to 3D surface. Hildebrandt and Polthier [13] propose a prescribed mean curvature flow to preserve surface features. Bilateral filtering [4] is also an anisotropic smoothing method, and it has been extended to surface smoothing by Fleishman et al. [14] and Jones et al. [15]. Their methods average the vertex position in a neighborhood using a weight function of both spatial difference and vertex difference. There is a body of work that decouple normal and vertex, i.e. firstly filter surface normals and then update vertex positions from the filtered normal field. Compared with vertex updating, normal filtering has a greater impact on final result. Yagou et al. propose to use mean and median filters [16] for facet normals, and later use alpha-trimming filters [17]. Sun et al. [18] filter the facet normals within local neighborhood, weighted by the normal difference. Zheng et al. [19] improve on this method by applying bilateral facet normal filtering, considering both normal and spatial difference. Most recently, He and Schaefer [20] introduce an area-based edge operator, and adopt it in L_0 minimization using Xu et al.'s solver [2]. Our mesh smoothing method belong to normal-vertex decoupling scheme, and L_0 gradient minimization is applied in the facet normal filtering step.

3. L_0 gradient minimization

Let g be the input signal and f be its filtered result. The gradients of f are denoted by ∇f. The energy function of L_0 gradient minimization is defined as follows:

$$\min_f |f-g|^2+\lambda|\nabla f|_0 \tag{1}$$

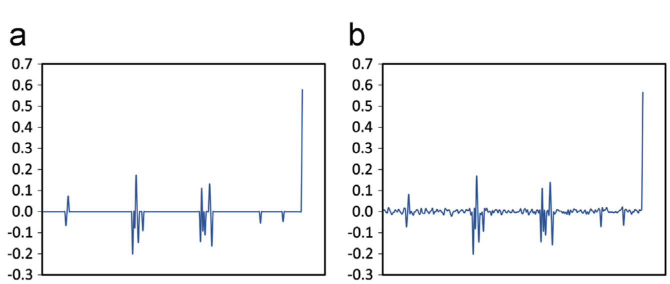

Fig. 1. The final value of variables δ and ∇f obtained by L_0–L_2 iteration algorithm.

where the data term is the squared L_2 distance between f and g, the smooth term is the L_0 norm of ∇f, and λ is a non-negative parameter controlling the weight of the smooth term. The L_0 norm of a vector is the number of non-zero value, which directly measures the sparsity. However, this term is difficult to optimize due to its non-convex and non-derivative nature. Traditional optimization techniques such as gradient descent are no longer applicable.

In Section 3.1, we briefly review L_0–L_2 iteration algorithm proposed by Xu et al. [2] for Eq. (1). In Section 3.2, we introduce our algorithm. Finally, we show the experiments and comparisons with some existing feature-preserving methods in Section 3.3.

3.1. L_0–L_2 iteration algorithm

Through introducing a set of auxiliary variables δ, the original minimization problem (Eq. (1)) then becomes

$$\min_{f,\delta} |f-g|^2+\beta|\nabla f-\delta|^2+\lambda|\delta|_0 \tag{2}$$

where β is a parameter to control the similarity between the auxiliary variables δ and their corresponding gradients ∇f.

Eq. (2) can be solved with an alternating optimization. First, δ is optimized with f fixed

$$\min_\delta |\nabla f-\delta|^2+\frac{\lambda}{\beta}|\delta|_0 \tag{3}$$

This equation can be spatially decomposed to a set of single variable function minimization. And each element δ_i has the following closed form:

$$\delta_i=\begin{cases}0 & \text{if } \nabla f_i<\sqrt{\lambda/\beta}\\ \nabla f_i & \text{otherwise}\end{cases} \tag{4}$$

Then, f is optimized with δ fixed

$$\min_f |f-g|^2+\beta|\nabla f-\delta|^2 \tag{5}$$

The equation is quadratic and a global minimum can be easily found by gradient descent.

L_0 minimization (Eq. (3)) and L_2 minimization (Eq. (5)) alternate until convergence is reached. After L_0 minimization, δ has a very high degree of sparsity. Next, L_2 minimization attempts to force ∇f to match δ. Indeed, the L_2 term tends to severely penalize outliers and propagate the residual in the energy function uniformly. This property of the L_2 term will lead to failure of sparsity transfer from δ to ∇f in Eq. (5).

To substantiate our claim, we input a one-dimensional noisy signal ranging from 0 to 1, and set λ to 0.001, β to 5. After iterations of L_0–L_2 minimization until convergence, the auxiliary variables δ are shown in Fig. 1(a) and the gradients of the filtered signal ∇f are shown in Fig. 1(b). It is clear that ∇f do not share the good sparsity with δ. In Xu et al.'s experiments [2], β is set as a large fixed value 1e5 to enhance the sparsity, which means most energy of the function is spent on matching ∇f with δ. In this paper, we provide another idea for solving Eq. (1) and we will give details in the next subsection.

3.2. Our algorithm

Our approximation algorithm here is basically built on coordinate descent [21]. Coordinate descent minimizes the energy function by solving a sequence of scalar minimization subproblems cyclically. Each subproblem performs line search along one coordinate direction with the others fixed. Coordinate descent is efficient in the situation where subproblems can be solved quickly.

However, according to the proposition of Bertsekas [22], every minimum of successive coordinate minimization of a continuously