CSANG: Continuous Scale Anisotropic Gaussians for Robust Linear Structure Extraction

V. B. Surya Prasath*, Sruthikesh Surineni†, Ke Gao*, Guna Seetharaman‡, Kannappan Palaniappan*

*Computational Imaging and VisAnalysis (CIVA) Lab
Department of Computer Science, University of Missouri-Columbia, MO 65211 USA
Emails: {prasaths, palaniappan}@missouri.edu, kg954@mail.missouri.edu
†Networking and Parallel Systems Lab
Department of Computer Science, University of Missouri-Columbia, MO 65211 USA.
Email: ssurineni@mail.missouri.edu
‡US Naval Research Laboratory, Washington D.C., USA

Abstract—Robust estimation of linear structures such as edges and junction features in digital images is an important problem. In this paper, we use an adaptive robust structure tensor (ARST) method where the local adaptation process uses spatially varying adaptive Gaussian kernel that is initialized using the total least-squares structure tensor solution. An iterative scheme is designed with size, orientation, and weights of the Gaussian kernel adaptively changed at each iteration step. Such an adaptation and continuous scale anisotropic Gaussian kernel change for local orientation estimation helps us obtain robust edge and junction features. We consider an efficient graphical processing unit (GPU) implementation which obtained \(30\times\) speed improvement over traditional central processing unit (CPU) based implementations. Experimental results on noisy synthetic and natural images indicate that we obtain robust edge detections and further comparison with other edge detectors shows that our approach obtains better localization and accuracy under noisy conditions.

1. Introduction

Edge detection and feature extraction are important image processing tasks. Digital image edge detection has a long history and has been the object of study for many decades now [1], [2], [3]. Many of these edge detectors are based on image derivatives and gradients and thus require usually pre-smoothing with a smoothing filter for e.g. convolution with a Gaussian kernel. One of the classical and most famous edge detector is due to Canny [3], which is still widely used in many image processing pipelines. Though Canny edge detector includes Gaussian pre-smoothing, it can still obtain spurious edge pixels under noise (see e.g. Figure 1) or double edges, and various improvements have been proposed over the years. Structure tensors are alternative to gradient computations and recently they are advocated in various image processing tasks where edges, orientations are required to be computed accurately and robustly under noisy conditions [4], [5].

Nath and Palaniappan [6] considered an adaptive robust structure tensor (ARST) approach for orientation estimation and image segmentation. The ARST is derived via robust statistics and oriented Gaussian kernel adaption in an iterative manner. By utilizing a spatially varying adaptive Gaussian kernel, improved accuracy in finding edge and corner locations were obtained. Further, robust error function such as Geman-McClure is utilized for estimating ARST in local neighborhoods which leads to better localization of low level linear structure extraction. Due to the nonlinear dependency on the orientation of the patch and iterative nature of the ARST estimation the method requires high computational cost. In this work, we consider utilizing graphical processing unit (GPU) for the main computational iterative scheme of ARST. We show that GPU ARST for fast edge detection with an average \(30\times\) speed improvement over traditional central processing unit (CPU) based implementations. We provide comparative experimental results on edge detection with other well-known edge detectors on synthetic, natural images.

Rest of the paper is organized as follows. In Section 2 we review the adaptive robust structure tensor (ARST) and
corresponding edge detection approach. Section 3 provides experimental results on various synthetic, real images and comparison with other edge detectors. Finally, Section 4 concludes the paper.

2. Adaptive Robust Structure Tensor for Edge Detection

2.1. Robust Estimator for Gradient Estimation

We follow the notations and conventions from [6] and detail the iterative method for the robust estimator based continuous scale anisotropic Gaussians adaptation. Let \( v(x) \) be the true (latent) gradient of an image patch \( \Omega(y) \), centered at pixel \( x \). The norm of the error vector between the estimated gradient \( g(y) \) at location \( y \) and \( v(x) \) is given by \( e(x,y) \) as

\[
||e(x,y)|| = ||g(y) - (g^T(v(x))v(x)||.
\]

The aim here is to estimate \( v \), hence the straightforward way is to minimize an error functional \( \rho(||e(x,y)||^2) \) integrated over the image patch \( \Omega \),

\[
\epsilon_{LS}(x) = \int_{\Omega} \rho(||e(x,y)||^2) W(x,y) dy.
\]  

(2)

Here, \( W(x,y) \) is a spatially invariant weighting function, for example the Gaussian kernel,

\[
W(x,y) = (2\pi\sigma^2)^{-\frac{1}{2}} e^{-\frac{1}{2\sigma^2}(y-x)^2}.
\]

This weighting function emphasizes the gradient at the central pixel within a small neighborhood, and we utilize an adaptive size and orientation based weighting function estimation, see Section 2.2. One of the well-known and basic formulation is based on the method of least-squares which corresponds to setting \( \rho(s) = s^2 \), thus Eqn. (2) becomes,

\[
\epsilon_{LS}(x,y) = \int_{\Omega} (g^Tg) W(x,y) dy - \int_{\Omega} (v^T(g^Tg)v) W(x,y) dy,
\]

subject to the condition \(^1 \) that \( ||v|| = 1 \). We utilize the Lagrange multipliers to write the unconstrained formulation as,

\[
\mathcal{E}_{LS}(x,y) = v^T (\int_{\Omega} (g^Tg) W(x,y) dy) v + \lambda(1-v^Tv).
\]

(5)

Differentiating \( \mathcal{E}_{LS}(x,y) \) to find the extremum leads to the standard eigenvalue problem for solving the best estimate of \( v \), given by \( \hat{v} \).

\[
\mathbf{J}(x,W) v = \lambda v,
\]

where

\[
\mathbf{J}(x,W) = \int_{\Omega} (g^Tg) W(x,y) dy.
\]

We note that this is the least-squares structure tensor at position \( x \) using the weighting kernel \( W \). The maximum eigenvector solution of Eqn. (6) gives the least-squares estimate for the gradient at the pixel \( x \) using the surrounding gradient information\(^2 \).

In [6], robust error measures are advocated instead of the least squares due to their noise tolerant property. Robust error functions impose smaller penalties on outliers. One of the important example for robust error functions is the Geman-McClure function \([7]\) and defined as,

\[
\rho(||e(x,y)||, m) = \frac{||e(x,y)||^2}{m^2 + ||e(x,y)||^2} = 1 - \frac{m^2}{m^2 + ||e(x,y)||^2},
\]

(7)

where, \( m \) is a parameter that determines the amount of penalty imposed on large errors. If we utilize the Geman-McClure robust error function in Eqn. (2) we get,

\[
\epsilon_{GM}(x,y) = \int_{\Omega} W(x,y) dy - \int_{\Omega} \frac{m^2}{(g^Tg - v^T(g^Tg)v + m^2)} W(x,y) dy.
\]

(8)

Using Lagrange multipliers we can write the unconstrained version as follows,

\[
\mathcal{E}_{GM}(x,y) = \int_{\Omega} \frac{m^2}{(g^Tg - v^T(g^Tg)v + m^2)} W(x,y) dy + \lambda(1-v^Tv).
\]

(9)

Differentiating \( \mathcal{E}_{GM}(x,y) \), with respect to \( v \), and setting it to zero gives

\[
\mathbf{J}(x,v,W) v = \lambda v,
\]

(10)

where

\[
\mathbf{J}(x,v,W) = \int_{\Omega} \frac{m^2}{(g^Tg - v^T(g^Tg)v + m^2)^2} (g^Tg) W(x,y) dy.
\]

(11)

is the Geman-McClure robust structure tensor. The following iterative equation,

\[
\mathbf{J}(x,v_i,W) v_{i+1} = \lambda v_i,
\]

(12)

is a fixed-point functional iteration scheme for numerically solving \( \lambda, v \) in Eqn. (10) that usually converges to a local minimum \([8]\). We use total least-squares solution to initialize the iterative process in Eqn. (12).

2.2. Continuous Scale Anisotropic Gaussians

Following [6] we use a soft Gaussian convolution function as the weighting function \( W(x,y) \) in our iterative formulation (12) with adaptive size and orientation within \( \Omega \).

2. Note that, \( v^T(x) \) could be determined by using the minimum eigenvector, however for an ideal edge, the smaller eigenvalue will be zero. Thus, it is numerically more reliable to estimate the maximum eigenvector.
Sparsely varying adaptation of the kernel (local neighborhood shape and coefficients) is beneficial for improving the estimation of oriented image structures. The neighborhood $Ω$ is initialized as a circular region and subsequently adapted to be an oriented elliptical region. Figure 2 shows the adaptive algorithm at an ideal edge. In this figure, the dashed-line elliptical region is oriented along the gradient while the solid-line elliptical region (that is scaled and rotated by $90^\circ$) is oriented along the edge. The spatially varying kernel $W_0(x,y)$ that is used with Eqn. (11) is defined as,

$$W_0(x,y) = K e^{-\frac{1}{2}(y-x)^T U_{1,-1} \lambda_1^2 U_{1,-1} (y-x)}$$

$$\Lambda_i = \begin{bmatrix} \sqrt{2}R_i^1 & 0 \\ 0 & \sqrt{2}R_i^2 \end{bmatrix},$$

where $K$ is a scaling factor associated with the Gaussian function. We initialize the kernel $W_0(x,y)$ as an isotropic Gaussian with $R_0^1 = R_0^2 = R_0$. A fairly large number is chosen (typically $R_0 = 3$), in order to reduce the influence of noise when evaluating the structure tensor. The columns of $U_1$ are the eigenvectors ($e_1^1$, $e_2^1$), with the columns of $U_{-1}$ initialized as the co-ordinate axes. Let $\lambda_1$ and $\lambda_2$ (with $\lambda_1 > \lambda_2$) be the eigenvalues of the structure tensor $J(x,v_{i-1},W_i)$ at the $i^{\text{th}}$ iteration. Scaled versions of these eigenvalues are used to update the semi-major and semi-minor axes for the $(i+1)^{\text{th}}$ iteration as

$$R_{i+1}^1 = \frac{\lambda_1}{\lambda_1 + \lambda_2} R_i^1 \quad \text{and} \quad R_{i+1}^2 = \frac{\lambda_2}{\lambda_1 + \lambda_2} R_i^2.$$  

(14)

The eigenvectors obtained from the current iteration, along with $(R_i^1, R_{i+1}^2)$ are used to update the kernel as follows

$$W_{i+1}(x,y) = K e^{-\frac{1}{2}(y-x)^T U_{i+1} \lambda_{i+1}^2 U_1 (y-x)},$$

$$\Lambda_{i+1} = \begin{bmatrix} \sqrt{2}R_{i+1}^1 & 0 \\ 0 & \sqrt{2}R_{i+1}^2 \end{bmatrix}. \quad \quad (15)$$

This kernel is used to compute a new structure tensor $J(x,v_{i},W_{i+1})$ as per Eqn. (11). To account for the spatially varying Gaussian kernel, Eq. 12 is modified to the following form,

$$J(x,v_{i},W_{i+1})v_{i+1} = \lambda v_{i+1}. \quad \quad (16)$$

We experimentally determined that two or three iterations were sufficient to achieve good results.

2.3. Edge Detection with CSANG

Algorithm 1 ARST - edge detection

Input: $u$, image of size $m \times n$; $t_{hi}$, $t_{lo}$, hysteresis thresholds ($t_{hi} > t_{lo}$).

Returns a binary image of size $m \times n$.

1: procedure INITIALIZATIONS
2: Compute the iterative solution $v$, See Eqn. (16)
3: Create maps:
4: $E_{mag}(x,y) : m \times n \to \mathbb{R}$
5: $E_{nms}(x,y) : m \times n \to \mathbb{R}$
6: $E_{bin}(x,y) : m \times n \to \{0,1\}$
7: for all image coordinates $(x,y) \in m \times n$ do
8: $E_{mag}(x,y) \leftarrow |v|$\n9: $E_{nms}(x,y) \leftarrow 0$
10: $E_{bin}(x,y) \leftarrow 0$
11: end for
12: procedure NONMAXIMUM SUPPRESSION
13: for $x \leftarrow 1, \ldots, m - 2$ do
14: for $x \leftarrow 1, \ldots, n - 2$ do
15: $e_{x}^+, e_{x}^-$ eigenvectors of structure tensor
16: $s \leftarrow \text{GetOrientationSector}(e_{x}^+, e_{x}^-)$
17: if $s = \text{LocalMax}(E_{mag}, x, y, s, t_{lo})$ then
18: $E_{nms}(x,y) \leftarrow E_{mag}(x,y)$
19: end if
20: end for
21: end for
22: procedure HYSTERESIS THRESHOLDING
23: for $x \leftarrow 1, \ldots, m - 2$ do
24: for $x \leftarrow 1, \ldots, n - 2$ do
25: if $E_{nms}(x,y) \geq t_{hi} \wedge E_{bin}(x,y) = 0$ then
26: $E_{bin}(x,y) \leftarrow E_{nms}(x,y)$
27: end if
28: end if
29: end for
30: end for
31: return $E_{bin}$

The solution of the iterative procedure (16) $v$ is used to obtain an edge map by applying non-maximal suppression and hysteresis thresholding steps, see Algorithm 1. These are standard steps in obtaining binary edge maps with single pixel width and originally used by Canny edge detector [3]. We next details the graphical processing unit (GPU) based implementation details of the core iterative procedure.

2.4. GPU Implementation Details

2.4.1. Parallelization approach.
• **Structure Tensor**: This is one time executed kernel which loads input image into shared memory. First, we compute the gradients on shared memory data and from the computed gradients structure tensor matrix is estimated.

• **Robust Estimator**: In this kernel, we compute the GM robust estimation for smoothness and texture terms from gradients and structure tensor. All the computation equations are shown in the pseudo code. Shared memory is used for computing 2nd derivatives similar to normal gradients. Since GM robust estimation is same for both smoothness and texture we implemented it as a device function. GM robust function there is a floating point division which is very slow to make it faster we approximate division by multiplication with reciprocal.

• **Input for solver**: Inputs to Thomas solver are prepared using a couple of kernels. Inputs to these kernels are from Robust estimator and they are transposed for row-wise gradients. In the original source code, there is a data dependency on the previous value. To avoid this data dependency and make the algorithm parallel we load the input data into shared memory and re-compute the previous value so that data dependency is eliminated. The overhead for re-computation of previous value is quite minimal.

• **CuSparse batched gtsv**: Thomas solver is used from CuSparse gtsv library function. If we consider a matrix, this solver solves a column each instance. So, to make the column elements coalesced we do a matrix transpose on inputs. For multiple columns to be solved at the same time we use batched gtsv solver but this requires a considerable amount of global memory (batchCount \( \leq (4?m+2048) \) ? sizeof(_typeg_)). In some cases we ran out of memory, to solve this issue we run the solver multiple iterations with fewer batches(columns) considering the free memory available at the launch of the kernel. Gradients are updated by taking the average of row wise and column wise gradient estimation.

• **Orientation Estimation**: Finally, once all the iterations for computing the gradients are done we compute the orientation of the gradients. Orientation is estimated using atan2f, an inbuilt math function.

• **Matrix Transpose**: This is the most instanced kernel in the whole algorithm. Implementing this optimally is critical for performance. This kernel is optimized for square block size. First, we load the input matrix into shared memory. A different mapping is used to read data from shared memory and store data into output coalesced.

Figure 3(a) shows the GPU execution pipeline discussed above.

2.4.2. **Thread scheduling and shared memory utilization.**

Details about thread scheduling, shared memory utilization are described below, see Figure 3(b).

• **Thread Scheduling**: Input image could be of any resolution and thread blocks need to be traversed over the whole image. Though we have a 2D grid of blocks we convert this 2D grid into a 1D value in terms of the image width. From this image width, we compute the current iteration \( X \) and \( Y \). Shown below is the image considering grid dimensions of \( < 3, 2 > \). Thread blocks first walk the image width and once we hit the boundary of the image then we move to next row(vertical) and so on.

• **Shared memory utilization**: Shared memory is used in most of the kernels. Shared memory in matrix transpose is straight forward. Shared memory for computing gradients is shown below. For computing a gradient in \( X \) we need previous and next value in input, gradient in \( Y \) we need row above and row below value. To optimize all the overlaps we load the data into shared memory and for all the boundary elements we need an extra load. To reduce the warp divergence the threads with threadIdx.y==0, threadIdx.x==0 load two extra elements. For example, if we have a block of \( < 8, 8 > \) we load \( 10 \times 10 \) data into shared memory. Since we use __sync_threads() for barrier synchronization we make sure all the threads in a block are alive even though they might be out of the image.

3. Experimental Results

3.1. Setup and timings

We utilized synthetic images and natural images from the Berkeley segmentation data set (BSDS500) which con-
TABLE 1. Runtime in seconds for MATLAB, C, CUDA and CUDA-pre opt, for different image resolution and iterations.

<table>
<thead>
<tr>
<th>Resolution</th>
<th>MATLAB</th>
<th>C</th>
<th>CUDA</th>
<th>CUDA pre opt</th>
</tr>
</thead>
<tbody>
<tr>
<td>512x512-1</td>
<td>0.13</td>
<td>4.8</td>
<td>1.19</td>
<td>1.39</td>
</tr>
<tr>
<td>512x512-50</td>
<td>0.21</td>
<td>7.2</td>
<td>1.28</td>
<td>1.75</td>
</tr>
<tr>
<td>1920x1080-1</td>
<td>46</td>
<td>46</td>
<td>3.4</td>
<td>2.04</td>
</tr>
<tr>
<td>1920x1080-50</td>
<td>16.7</td>
<td>20</td>
<td>2.24</td>
<td>2.3</td>
</tr>
<tr>
<td>6600x4400-1</td>
<td>646</td>
<td>780</td>
<td>23</td>
<td>-</td>
</tr>
<tr>
<td>6600x4400-50</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Algorithm 2 Nonmaximum suppression and hysteresis thresholding used in Algorithm 1

```c
procedure GETORIENTATIONSECTOR(v_x, v_y)
Returns the orientation sector \( s_\theta \) for the 2D vector \((v_x, v_y)\)T.

\[
\begin{pmatrix}
v'_x \\
v'_y
\end{pmatrix}
= \begin{pmatrix}
\cos(\pi/8) & -\sin(\pi/8) \\
\sin(\pi/8) & \cos(\pi/8)
\end{pmatrix} \cdot \begin{pmatrix}
v_x \\
v_y
\end{pmatrix}
\]
if \( v'_y < 0 \) then
\[\begin{align*}
v'_x & \leftarrow -v'_x, \quad v'_y \\
s_\theta & \leftarrow \begin{cases}
0 & \text{if } (v'_x \geq 0) \land (v'_y \geq v'_y) \\
1 & \text{if } (v'_x \geq 0) \land (v'_y < v'_y) \\
2 & \text{if } (v'_x < 0) \land (v'_y < v'_y) \\
3 & \text{if } (v'_x < 0) \land (v'_y \geq v'_y)
\end{cases}
\end{align*}\]
return \( s_\theta \).
```

3.2. Comparison with other edge detectors

We compare our fast CSANG with traditional edge detectors due to Prewitt [1], Sobel [2], and Canny [3].

3.2.1. Synthetic images. We use a synthetic image of size 518 × 334 for testing the accuracy of different edge detectors under different noise levels [6]. We compare the edge detection results on noise-free, and noisy (obtained by adding Gaussian noise of standard deviation \( \sigma_n = 30 \)) images with other edge detectors [1], [2], [3] with default parameters. As can be seen from the Figure 4 bottom row, under noisy conditions all the traditional edge detectors pick up false positive edges while our CSANG obtained a robust result.

3.2.2. Natural images. Figure 5 shows some example comparison of edge detection results on the BSDS500 natural images with Canny edge detector and human drawn ground truth (GT). The GT was aggregated from five different humans (Figure 5(b)), and Canny edge detection results were obtained with default parameters of MATLAB (Figure 5(c)). As can be seen by comparing the edge detection results obtained with our CSANG (Figure 5(c)) we obtained better results. In Canny edge detection results we can observe lot of spurious edges while our approach obtained less spurious edges, however edge connectivity is lost in some places.

4. Conclusions

In this paper, we utilized a fast adaptive robust structure tensor for edge detection in digital images. By utilizing continuous scale anisotropic gaussians via Geman-McClure robust error estimator function we obtained efficient edge detection for noisy images. We provide an efficient GPU implementation using CUDA, and our implementation can obtain 30× speed improvement over traditional CPU implementations. Compared to different edge detectors from the
Figure 4. Comparison of different edge detectors on noise-free and noisy images. (a) Input images, (b) Prewitt [1], (c) Sobel [2], (d) Canny [3], (e) ARST. (Top row) Noise-free, (bottom row) noisy image obtained with adding Gaussian noise of standard deviation $\sigma_n = 30$.

Figure 5. Comparison on Berkeley segmentation data set (BSDS500). (a) Input images, (b) ground truth (GT), and edge detection results with (d) Canny edge detector [3], (e) our CSANG.
literature, our ARST based edge detector provides robust and accurate results. Experimental results on a variety of noisy images indicate we obtain better edge detection results and compared to other standard edge detectors. Testing other robust error estimator functions and applying the fast ARST to large scale image segmentation defines our future work.

References


