<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper presented at *IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), 2014*.

Citation for the original published paper:

Kristensen, J., Nilsson, J. (2014)

Fast triangular binning kernel approximations for weighted gradient histogram creation.

In:

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-136275>

Fast triangular binning kernel approximations for weighted gradient histogram creation

Johan Kristensen and John-Olof Nilsson

Department of Signal Processing, ACCESS Linnaeus Centre, KTH Royal Institute of Technology, Stockholm, Sweden

Abstract—The implementation of weighted gradient histograms are studied. Such histograms are commonly used in computer vision methods, and their creation can make up a significant portion of the computational cost. Further, due to potentially severe aliasing, non-uniform binning kernels are desirable. We show that previously presented fast methods for uniform binning kernels can be extended to non-uniform binning, and that the triangular kernel can be well approximated for common weighting strategies. The approximation is implemented with sums and products of projections of the gradient samples on specially chosen vectors. Consequently, only a few standard arithmetic operations are required, and therefore, the suggested implementation has a significantly lower computational cost when compared with an implementation in which the gradient argument and magnitude are explicitly evaluated. Finally, the frequency components of the different kernels are studied to quantify the fundamental gain achieved by using triangular kernels instead of uniform kernels.

I. INTRODUCTION

Weighted gradient histograms are used in many computer vision methods to describe local image regions, see e.g. [1,2,3]. The creation of such histograms can make up a significant portion of the computational cost for these methods. Consequently, fast methods for creating these histograms are of interest. The computational cost primarily comes from evaluating the argument/phase and the magnitude of the gradient samples, requiring an arctangent and a square-root function evaluation, respectively. The argument is needed for the binning and the magnitude for the weighting. Recently, however, it has been shown that the binning can be achieved without explicitly computing the argument (argument-free binning) [4,5]. Further, we have previously argued for using a quadratic magnitude weight in some applications [6,7], eliminating the need for the square-root of the magnitude calculation as well.

Unfortunately, the argument-free binning as described in [4,5], gives a uniform (nearest neighbour) binning kernel. This kernel is acceptable when a large number of bins or samples are used. However, when comparing image regions, low dimensional descriptors (few bins) are desirable and a low number of samples may contribute to each descriptor. Using a uniform kernel in such cases, can introduce severe quantization errors. The effect is illustrated in Fig. 1, where the upper plot shows the implicit gradient distribution prior to binning (sampling). As is evident, a small change in how the bins are positioned relative to the samples can make a large difference in the acquired histogram. Therefore, exploiting a more complex binning kernel instead, like the triangular kernel

Fig. 1: The figures show the implicit gradient distribution prior to histogram binning (sampling). The upper part shows the result of using a nearest neighbour binning kernel (uniform), and the lower figure shows a linear interpolation kernel between the two nearest bins (triangle). Clearly, sampling of the smoother lower distribution will result in less quantization errors compared to that of the jagged upper distribution.

shown in the lower plot of Fig. 1, could be highly beneficial. Further, the quadratic magnitude weight previously suggested can amplify signal noise, and therefore, it is not desirable under all circumstances. Consequently, in this article we describe how the argument-free binning from [4] can be extended to an approximate triangular kernel. For the proportional and quadratic magnitude weighting, different approximative solutions are suggested. Both solutions only exploit standard arithmetic operations and no high-level function evaluations are required. Therefore, a significant computational saving could be achieved, and the solutions are suitable for low-level system implementations. Due to the wide-spread use of gradient histograms in computer vision and image processing, the solutions are believed to have a wide applicability. Finally, the frequency responses of the different kernels are studied to quantify the gains of using a triangular kernel instead of a uniform kernel.

II. IMAGE REGION DESCRIPTION

To create a descriptor for an image region, gradient samples from the region are commonly binned with respect to their argument, and weighted based on their magnitude and potentially a spatial weight. This will create a weighted gradient

argument histogram. Formally, a histogram h is created from a set of gradient samples $\nabla s_k \in \mathbb{R}^2$ indexed by k , and the i th gradient argument histogram bin h_i (i is in the range from 0 to $N - 1$ where N is the number of bins) is determined by

$$h_i = \sum_k w_m(\|\nabla s_k\|) w_b(\phi_i - \arg(\nabla s_k)) \rho_k \quad (1)$$

where $w_m(\cdot)$ is a weight depending on the gradient magnitude $\|\nabla s_k\|$, $w_b(\cdot)$ is a 2π -periodic binning kernel evaluated with respect to the difference between the gradient argument/phase $\arg(\nabla s_k)$ and the i th bin center $\phi_i = 2\pi \cdot i/N$, and ρ_k is a spatial weight effectively determining the described region. The magnitude weight is often just the magnitude, the binning kernel is often the uniform kernel or the triangular kernel, and the spatial weight is typically monotonously decreasing from some central point. See [8] for definitions and a detailed treatment of binning kernels; and see [7] for details about different weighting strategies. Evaluating (1) is equivalent to sampling the implicit distributions illustrated in Fig. 1. Typically, the weights are normalized such that $\sum_i h_i = 1$.

Writing the histogram in terms of the value of single bins, as shown in (1), gives a nice expression, but does not reflect how the histogram is created. For typical binning kernels, the contribution of a gradient sample is only non-zero for one or two bins. Consequently, instead of working with a bin at a time, one suitable work with a gradient sample at a time, finds the bins it will contribute to, and evaluates and adds the resulting weights. Since the spatial weight is independent of the value of the gradient sample, in essence this means evaluating $w_m(\|\nabla s_k\|) w_b(\phi_i - \arg(\nabla s_k))$ for one or two identified i , for which the weight is non-zero. If the argument $\arg(\nabla s_k)$ is evaluated explicitly, or in the case of uniform binning kernels $w_b(\cdot)$, finding the bin(s) i and evaluating the weight $w_m(\|\nabla s_k\|)$ becomes decoupled. In the latter case, this is due to the fast argument quantization methods described in [4], which enables us to find i without calculating $\arg(\nabla s_k)$. However, as previously argued, evaluating the argument explicitly is costly, and using a uniform kernel may introduce significant quantization errors. Consequently, we seek methods for evaluating $w_m(\|\nabla s_k\|) w_b(\phi_i - \arg(\nabla s_k))$ when $w_b(\cdot)$ is non-uniform without taking the path over $\arg(\nabla s_k)$. This cannot be done in general, but for specific kernels $w_b(\cdot)$ and magnitude weights $w_m(\|\nabla s_k\|)$ it is possible. In the following two sections, two such examples are given.

III. TRIANGULAR KERNEL AND PROPORTIONAL WEIGHTING

The problem in circumventing evaluating $\arg(\nabla s_k)$ is the difference $\phi_i - \arg(\nabla s_k)$. For a general kernel $w_b(\phi_i - \arg(\nabla s_k))$, we have to evaluate $\phi_i - \arg(\nabla s_k)$. However, we can construct kernel components dependent on $\phi_i - \arg(\nabla s_k)$, which can be evaluated without finding the argument. From these we can construct suitable binning kernels.

For notational compactness, let $r_k = \|\nabla s_k\|$ and $\theta_k = \arg(\nabla s_k)$. We can now write our k th gradient sample as

$$\nabla s_k = r_k (\cos(\theta_k), \sin(\theta_k)).$$

Fig. 2: Illustration of the triangular kernel approximation with a magnitude weighting (solid lines) and the corresponding triangular kernels (dotted lines) for different numbers of bins. For 8 or more bins, the agreement with the triangular kernel is good.

The key tool in this article is the scalar product (projection). By constructing vectors

$$\mathbf{p}_i = (\cos(\phi_i), \sin(\phi_i)),$$

where the indices i are assumed periodic with period N , and taking the scalar product with the gradient, we get the components

$$\nabla s_k \cdot \mathbf{p}_i = r_k \cos(\phi_i - \theta_k).$$

Let j and $j+1$ denote the two bins nearest in argument to θ_k , such that $\phi_j \leq \theta_k < \phi_{j+1}$. This pair of indices can easily be found with previously presented projection and bisection methods [4]. Now we want to find weights according to some $w_m(r_k) w_b(\phi_i - \theta_k)$ for j and $j+1$ and some sensible kernel. Assume that the magnitude weight is some power of the magnitude, $w_m(r_k) = r_k^p$. Then the triangular binning kernel, giving the individual weights $w_{k,i}^{(p)}$ for the nearest two bins are

$$\begin{aligned} w_{k,j}^{(p)} &= r_k^p (1 - (\phi_{j+1} - \phi_j)^{-1} |\theta_k - \phi_j|) \\ &= r_k^p (1 - \frac{N}{2\pi} (\theta_k - \phi_j)) \\ w_{k,j+1}^{(p)} &= r_k^p (1 - \frac{N}{2\pi} (\phi_{j+1} - \theta_k)), \end{aligned}$$

where the arguments $\theta_k - \phi_i$ are assumed to be subtracted in a 2π periodic fashion. We propose that this two bin weight distribution kernel is approximated according to

$$\begin{aligned} \hat{w}_{k,j}^{(1)} &= \nabla s_k \cdot \mathbf{p}_j - (\mathbf{p}_j \cdot \mathbf{p}_{j+1}) (\nabla s_k \cdot \mathbf{p}_{j+1}) \\ &= r_k (\cos(\theta_k - \phi_j) - \cos(\phi_j - \phi_{j+1}) \cos(\theta_k - \phi_{j+1})) \end{aligned} \quad (2)$$

$$\begin{aligned} \hat{w}_{k,j+1}^{(1)} &= \nabla s_k \cdot \mathbf{p}_{j+1} - (\mathbf{p}_{j+1} \cdot \mathbf{p}_j) (\nabla s_k \cdot \mathbf{p}_j) \\ &= r_k (\cos(\theta_k - \phi_{j+1}) - \cos(\phi_{j+1} - \phi_j) \cos(\theta_k - \phi_j)). \end{aligned} \quad (3)$$

The resulting kernels for different numbers of bins are illustrated in Fig. 2, together with the corresponding ideal triangular kernels.

The rational behind the approximation in (2) and (3) may not be obvious. However, it can be understood in the way that

Fig. 3: Weighting functions (2) and (3) and (4) and (5) for 8 bins plotted over a range extending outside their applicable regions. Outside the region marked with the dashed lines, the bins j and $j+1$ will not be the closest pair of bins. The binning methods in [4] are seen to select approximately linear portions of the functions.

the logical operations in the binning methods presented in [4] (finding bin j and $j+1$) select a small portion of a cosine-function. Then, a weighted sum of the cosine functions of different phases is used to select the approximately linear part $\cos(\frac{\pi}{2} - \Phi)$ for small Φ . To illustrate this fact, Fig. 3 shows a plot of (2) and (3) for a larger range of arguments. The delimiting vertical bars are seen to mark a region containing the almost triangular portion of the weights.

The binning as described by (2) and (3) can be implemented with only two scalar products, two multiplications by a constant ($\mathbf{p}_{j+1} \cdot \mathbf{p}_j$), and two subtractions. This is indeed a low computational cost compared to explicitly calculating the gradient argument $\arg(\nabla s_k)$ and magnitude $\|\nabla s_k\|$ followed by a binning function $w_b(\cdot)$. The argument is suitably implemented with some polynomial approximation which may require a large number of arithmetic and logical operations. See for example [5,9]. The magnitude will essentially require a scalar product of the gradient by itself, followed by a square-root. Finally, a triangular binning function will require multiplication by a constant, an absolute value, and an addition.

Unfortunately, a minor problem with (2) and (3) is that the total weight for a gradient sample varies with the argument. The total contribution from a sample k is

$$\begin{aligned} \frac{1}{r_k} (\hat{w}_{k,j}^{(1)} + \hat{w}_{k,j+1}^{(1)}) &= \cos(\theta_k - \phi_j) - \cos(\phi_j - \phi_{j+1}) \cos(\theta_k - \phi_{j+1}) \\ &\quad + \cos(\theta_k - \phi_{j+1}) - \cos(\phi_{j+1} - \phi_j) \cos(\theta_k - \phi_j) \\ &= (1 - \cos(\phi_{j+1} - \phi_j)) (\cos(\theta_k - \phi_j) + \cos(\theta_k - \phi_{j+1})) \end{aligned}$$

which obviously is dependent on θ_k . In the interval $[\phi_j, \phi_{j+1}]$, $\frac{1}{r_k} (\hat{w}_{k,j}^{(1)} + \hat{w}_{k,j+1}^{(1)})$ there will be minima at ϕ_j and ϕ_{j+1} , and a

Fig. 4: Illustration of the triangular kernel approximation with a magnitude weighting (solid lines) and the corresponding triangular kernels (dotted lines) for different numbers of bins. For 8 or more bins, the agreement with the triangular kernel is good.

maximum at $\theta_k = \frac{\phi_j + \phi_{j+1}}{2}$. The ratio of these extrema is

$$\begin{aligned} \frac{\sum_i \hat{w}_{k,i}^{(1)} |_{\theta_k = \frac{\phi_j + \phi_{j+1}}{2}}}{\sum_i \hat{w}_{k,i}^{(1)} |_{\theta_k = \phi_j}} &= \frac{2 \cos(\frac{\phi_j - \phi_{j+1}}{2})}{1 + \cos(\phi_j - \phi_{j+1})} = \frac{2 \cos(\frac{\phi_j - \phi_{j+1}}{2})}{2 \cos(\frac{\phi_j - \phi_{j+1}}{2})^2} \\ &= \cos^{-1} \left(\frac{\phi_j - \phi_{j+1}}{2} \right) = \cos^{-1} (\pi/N). \end{aligned}$$

For 4 bins, this gives a potential difference of 41%; for 8 bins this gives a potential difference of 8%; and for 16 bins this gives a potential difference of 2%. Consequently, using the approximation is only sensible for 8 or more bins.

IV. TRIANGULAR KERNEL AND QUADRATIC WEIGHTING

For the proportional magnitude weighting, we got a suitable triangular binning kernel approximation by (2) and (3). These can easily be modified to get a quadratic weight. Simply multiplying these weights by the projection again gives the weights

$$\begin{aligned} \hat{w}_{k,j}^{(2)} &= (\nabla s_k \cdot \mathbf{p}_j) \hat{w}_{k,j}^{(1)} \\ &= r_k^2 \cos(\phi_j - \theta_k) \\ &\quad \cdot (\cos(\theta_k - \phi_j) - \cos(\phi_j - \phi_{j+1}) \cos(\theta_k - \phi_{j+1})) \end{aligned} \quad (4)$$

$$\begin{aligned} \hat{w}_{k,j+1}^{(2)} &= (\nabla s_k \cdot \mathbf{p}_{j+1}) \hat{w}_{k,j+1}^{(1)} \\ &= r_k^2 \cos(\phi_{j+1} - \theta_k) \\ &\quad \cdot (\cos(\theta_k - \phi_{j+1}) - \cos(\phi_{j+1} - \phi_j) \cos(\theta_k - \phi_j)) \end{aligned} \quad (5)$$

which yield an approximation of the triangular kernel with quadratic magnitude weighting. The resulting kernels for different numbers of bins are shown in Fig. 4. The functions outside their applicable range are seen in Fig. 3. Similar to the case for the proportional magnitude weighting, the initial binning is seen to select a range where the functions are approximately linear, and together make up a triangular kernel.

In contrast to the case of proportional magnitude weighting, the sum of these weights is independent of the argument. By

trigonometric relations,

$$\begin{aligned} \frac{1}{r_k^2} (\hat{w}_{k,j}^{(2)} + \hat{w}_{k,j+1}^{(2)}) &= \cos(\theta_k - \phi_j)^2 + \cos(\theta_k - \phi_{j+1})^2 + \\ &\quad - 2 \cos(\phi_j - \phi_{j+1}) \cos(\theta_k - \phi_{j+1}) \cos(\theta_k - \phi_j) \\ &= \frac{1}{2} (1 - \cos(2(\phi_j - \phi_{j+1}))) = \frac{1}{2} (1 - \cos(4\pi/N)) \end{aligned}$$

which obviously only depends on the number of bins N .

V. FREQUENCY PROPERTIES

The binning kernels ultimately have the role of smoothing out the data to avoid excessive aliasing when the signal is sampled (binned). Therefore, it is of interest to study the frequency properties of the different kernels. This will illustrate the actual gain of using the suggested kernel in contrast to a uniform kernel.

Since the signal (descriptor) components are translated binning kernels, the aliasing will apply to the kernels themselves. The fourier coefficients of the triangular kernel approximations, together with the fourier coefficients of the corresponding uniform kernels are shown in Fig. 5. The Nyquist frequency is on half of the main lobe. The remaining part of the frequency response will be folded back over the main lobe when sampled.

If we assume that gradient distributions are compared in a quadratic sense, then we can use Parseval's theorem to do the comparison in the fourier coefficients of the difference c_k instead. We can analyze one coefficient at a time. One way to quantify this is to look at the relative portion of the energy which is not folded back, i.e. the utility part of the signal relative to the aliased energy:

$$\frac{c_k^2}{\sum_{k=1}^{\infty} (c_{i-kN}^2 + c_{i+kN}^2)}.$$

This gives a signal-to-noise-ratio-like figure of merit for the different kernels. This measure is plotted for the first coefficients in the fourier series for different numbers of bins in Fig. 6. Obviously, the triangular kernel and its approximations express a significantly lower aliasing effect.

VI. CONCLUSION

The creation of weighted gradient histograms potentially make up a significant portion of the computational cost of many computer vision methods. Therefore, fast methods are of interest. Unfortunately, due to potential severe aliasing, non-uniform binning is suitably used. At first this may seem to eliminate the possibility to use existing fast implementations of the binning; however, in this article we have extended earlier presented fast implementation methods and shown that the triangular kernel can be well approximated together with common magnitude weighting strategies. This gives a significantly lower computational cost compared to an implementation where the argument and the magnitude of the gradient are explicitly calculated. This in turn makes the beneficial spectral properties of the triangular weights available for many computational-resources-constrained computer vision application.

Fig. 5: Triangular kernel approximations and uniform kernels in the frequency domain (fourier coefficients). The sampling frequency is half the main lobe of the kernels. Consequently, the remaining tail will be folded back over the main lobe.

Fig. 6: Signal utility ratio for the first four fourier coefficients for the effective kernels, resulting from the projection binning and the ideal uniform and triangular kernels.

REFERENCES

- [1] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, pp. 91–110, 2004.
- [2] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *International Conference on Computer Vision & Pattern Recognition*, (INRIA Rhone-Alps, France), 20–25 June 2005.
- [3] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. PAMI*, vol. 27, pp. 1615–1630, oct. 2005.
- [4] J.-O. Nilsson and P. Händel, "Fast argument quantization," *IEEE Signal Processing Magazine*, vol. 30, pp. 169–172, nov. 2013.
- [5] J.-O. Nilsson, "Efficient implementation of data binning in sectors," Tech. Rep. TRITA-EE 2012:041, School of Electrical Engineering, KTH Royal Institute of Technology, 2012.
- [6] J.-O. Nilsson, J. Kristensen, and P. Händel, "A unifying approach to feature point orientation assignment," *Submitted manuscript*, 2013.
- [7] J.-O. Nilsson and P. Händel, "Gradient sample argument weighting for robust image region description," in *Electronics, Computing and Communication Technologies (CONECT)*, 2013 IEEE International Conference on, (Bangalore, India), 17–19 Jan. 2013.
- [8] P. Hall and M. Wand, "On the accuracy of binned kernel density estimators," *Journal of Multivariate Analysis*, vol. 56, pp. 165–184, 1996.
- [9] J. Crenshaw, *Math Toolkit for Real-Time Programming*. CMP, 2000.