USING THE GENERALIZED RADON TRANSFORM FOR DETECTION OF CURVES IN NOISY IMAGES

Peter A. Toft

Electronics Institute, Technical University of Denmark, DK-2800 Lyngby.
Email ptoft@ei.dtu.dk. Phone +45 45934207

ABSTRACT

In this paper the discrete generalized Radon transform will be investigated as a tool for detection of curves in noisy digital images. The discrete generalized Radon transform maps an image into a parameter domain, where curves following a specific parameterized curve form will correspond to a peak in the parameter domain. A major advantage of the generalized Radon transform is that the curves are allowed to intersect. This enables a thresholding algorithm in the parameter domain for simultaneous detection of curve parameters. A threshold level based on the noise level in the image is derived. A numerical example is presented to illustrate the theory.

1. INTRODUCTION

In recent years the Hough transform [1] and the related Radon transform [2] have received much attention. These two transforms are able to transform two dimensional images with lines into a domain of possible line parameters, where each line in the image will give a peak positioned at the corresponding line parameters. This have lead to many line detection applications within image processing, computer vision, and seismics.

A natural expansion of the Radon transform is the (discrete) generalized Radon transform (GRT) [3, 4, 5]. Analogous to the linear Radon transform, the GRT transforms curves in the image into a discrete multi dimensional parameter domain producing peaks positioned at the corresponding curve parameters. In this way the GRT converts a difficult global detection problem into a more easily solved local peak detection problem. A major advantage of the GRT is that curves are allowed to intersect. Another major advantage that will be demonstrated in this paper, is that the GRT is very robust to noise.

In this paper a probabilistic approach is used to show that the GRT can be used for curve detection if the noise in the image is below a certain level compared to the signal values on the curves. If noise is added to an image containing curves, the problem is that peaks in the parameter domain may or may not correspond to actual curve parameters. A threshold level, based on the noise level, is derived and applied for separation of noise and curve information in the parameter domain. A numerical example is provided to illustrate the presented theory.

2. THE GENERALIZED RADON TRANSFORM

The Generalized Radon transform, GRT, of a digital image can be defined in many ways. One way is

$$\breve{g}(\boldsymbol{j}) = \sum_{l=0}^{L-1} g(\phi_m(l,\boldsymbol{j}), \phi_n(l,\boldsymbol{j})) \qquad (1)$$

where $\breve{g}$ denote the GRT of the image $g(m,n)$ and $\boldsymbol{j}$ is a multi dimensional vector containing the curve parameters. The two curve functions $\phi_m(l,\boldsymbol{j})$ and $\phi_n(l,\boldsymbol{j})$ define the curve type and are (in principle) arbitrary[1]. A popular choice is the linear curve functions; e.g., normal parameters $\boldsymbol{j} = (\rho, \theta)$. Another frequent choice is the (τ, p)-parameters (known as slant stacking in seismics), where $\phi_m(l,\tau,p) = l$ and $\phi_n(l,\tau,p) = p\ l + \tau$.

Even though the GRT can be applied to any given image, the main feature of the GRT is that an image, which contains a discrete curve matching the curve functions at one parameter vector $\boldsymbol{j}^*$ implies that the parameter domain $\breve{g}(\boldsymbol{j})$, will show a peak at that specific parameter vector $\boldsymbol{j} = \boldsymbol{j}^*$. The linearity of the GRT implies that each curve in the image will be transformed into a peak in the parameter domain. In this paper a curve in an image is defined by large image values of the same sign on the curve and otherwise (approximately) zero.

Initially only two values of the GRT will be considered. The first, $\breve{g}(\boldsymbol{j}^*)$, corresponds to a curve in the

[1] An interpolation scheme is assumed implicitly; e.g., by rounding the functions $\phi_m(l,\boldsymbol{j})$ and $\phi_n(l,\boldsymbol{j})$ to the nearest sample point.

image. Another, $\breve{g}(\boldsymbol{j}^-)$, corresponds to a parameter vector, that does not match a curve in the image. It is assumed that $\breve{g}(\boldsymbol{j}^*)$ is the sum of a mean signal value μ over all L samples, and $\breve{g}(\boldsymbol{j}^-)$ covers (approximately) no samples of the curve(s) in the image. Both values of $\breve{g}$ are contaminated with noise due to noise in the image. Assume the noise in the image is nearly uncorrelated with zero mean (e.g., by subtracting a DC-value from the image) and variance σ^2.

A classical curve detection algorithm is to determine the parameter vectors from the positions of peaks in the parameter space

$$\boldsymbol{j}^* = \arg\{|\breve{g}(\boldsymbol{j})| \geq L\,\mu^*\} \tag{2}$$

The reason for choosing the significance level in this way is that Eq. 1 consists of a summation over L samples, and μ^* is a lower positive bound on the mean signal level on the curve; e.g, found by estimation. The purpose of the following is to estimate whether curves having the signal level μ can be detected using Eq. 2, if the image is contaminated with the described noise.

Due to the linearity, $\breve{g}$ consists of a curve part and a noise part. If $L \gg 1$ the sum of the noise terms $\breve{g}_{\text{noise}}$ will approximately be Gaussian distributed with zero mean and variance $L\sigma^2$ due to the Central Limit Theorem. This implies that the two considered values of the GRT are distributed as $\breve{g}(\boldsymbol{j}^*) = \mu L + \breve{g}^*_{\text{noise}} \in \mathcal{N}(\mu L, L\sigma^2)$ and $\breve{g}(\boldsymbol{j}^-) = \breve{g}^-_{\text{noise}} \in \mathcal{N}(0, L\sigma^2)$. Since Eq. 2 selects the large values in the parameter domain, an important issue is the probability of detecting the correct parameter vector of the two considered

$$\mathrm{P}_{\text{det 2}} \equiv \mathrm{P}\{|\breve{g}(\boldsymbol{j}^*)| > |\breve{g}(\boldsymbol{j}^-)|\} \tag{3}$$

Assuming that $\breve{g}(\boldsymbol{j}^*)$ and $\breve{g}(\boldsymbol{j}^-)$ are independent, inserting the joint probability distribution function, i.e, the product of the two individual Gaussian probability distribution functions, and using the trick of rotating the coordinate system 45 degrees, the integrals separate into one dimensional integrals which easily can be expressed by the error-function, $\operatorname{erf}(\cdot)$

$$\mathrm{P}_{\text{det 2}} = \frac{1}{2}\left[1 + \left(\operatorname{erf}\left(\frac{\lambda}{2}\right)\right)^2\right] \quad \text{and} \quad \lambda = \frac{\mu\sqrt{L}}{\sigma} \tag{4}$$

Note that in this case $\mathrm{P}_{\text{det 2}}$, shown on Fig. 1, only depends on one parameter λ. Note that $|\lambda| > 4$ gives an almost certain detection. This is the case if $\sigma \to 0$ or $\mu \to \infty$.

When using the GRT to detect curves then the discrete parameter domain will not only have two, but $\mathcal{J}$ different parameter vectors, where $\mathcal{J}$ is the number of samples in the parameter domain. It is assumed that all the noise sources in the parameter domain are independent and in the following, the detection of a single curve is analyzed. Selecting the position of the highest peak in the parameter domain, the probability of the selected parameter vector being correct, can be approximated by

$$\mathrm{P}_{\text{det all}} \cong \prod_{i=2}^{\mathcal{J}} \mathrm{P}_{\text{det 2}} \cong \mathrm{P}_{\text{det 2}}^{\mathcal{J}-1} =$$
$$\left(\frac{1}{2}\left[1 + \left(\operatorname{erf}\left(\frac{\lambda}{2}\right)\right)^2\right]\right)^{\mathcal{J}-1} \cong 1 - \frac{2\mathcal{J}}{\lambda\sqrt{\pi}}\,\mathrm{e}^{-\lambda^2/4} \tag{5}$$

The last simple approximation is valid if the detection probability is close to 1 as seen from Fig. 2. Several characteristics can be noted: The Figure shows a narrow transition from low to high detection probability as a function of λ, and $\mathcal{J}$ does not change the shape of $\mathrm{P}_{\text{det all}}$ significantly. If demanding a high $\mathrm{P}_{\text{det all}}$ then Eq. 5 and Fig. 2 demonstrate that $\mathcal{J}$ should be held low; i.e., by reducing the number of samples in the parameter domain to a minimum. It should be noted that this will involve a compromise on the range of parameter vectors.

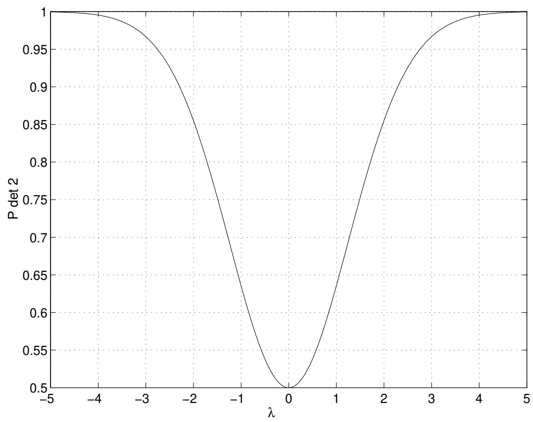

Figure 1: The probability of detecting the right curve parameters of two possible as a function of λ.

Eq. 5 can be used to set requirements on, e.g., the absolute mean signal level μ^* of the curve(s) to be detected. Demanding a detection probability $\mathrm{P}_{\text{det all}}$ greater than P^* implies that $\mu^* = \sigma/(\lambda^*\sqrt{L})$, where λ^* can be found by from Eq. 5 with a given detection probability, summation length L, number of samples in the parameter domain $\mathcal{J}$, and the standard deviation σ (e.g., by found by estimation). Any $|\mu|$ less than the threshold level, μ^*, can be considered as noise. In this way it is possible to give a statistically based estimate on the thresholding level in Equation 2.

Even though the above theory is developed by analyzing one curve in the image, the theory can be used

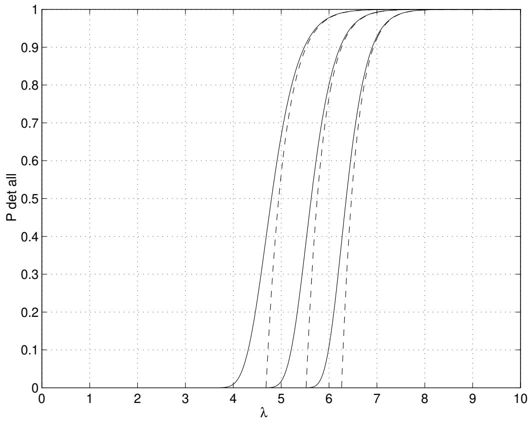

Figure 2: Solid lines show the probability of detecting the right curve parameter as a function of λ, and dashed lines show the simple approximation. From left to right $\mathcal{J}$=1000, $\mathcal{J}$=10000 and rightmost $\mathcal{J}$=100000.

if the image contains few curves. Instead of having one peak in the parameter domain representing one curve in the image, each of the Z curves in the image, where $Z \ll \mathcal{J}$, will give a peak in the parameter domain, even if the curves cross each other. With Z curves each of corresponding Z $\breve{g}$-values must be larger than the rest in the parameter domain. If only a few curves are present, the rest of the parameter domain is dominated by noise, and the probability of detection for each of the Z curves can be found from Eq. 5.

The theory used to derive Eq. 5 is somewhat pessimistic in estimation of the influence of the noise. This is partly due to the assumption that all the GRT-values include summing noise over L samples. Normally some of the GRT-values will require summing up over a curve partially outside the image, where the image must be assumed equal to zero. Furthermore some correlation must be expected in the parameter domain, especially if the number of dimensions in the parameter domain is higher than two. Depending on the sampling parameters, this implies that an effective $\mathcal{J}$ (less than the number of samples in the parameter domain) must be used in Eq. 5.

3. AN EXAMPLE OF LINE DETECTION IN A VERY NOISY IMAGE

A noise free image containing eight lines with limited slope is created. The image, shown in Fig. 3, has 101×101 samples. The curve sampling functions are chosen to $\phi_m = l - 50$ and $\phi_n = p\ (l - 50) + \tau$ and L is set to 101. The offset is made in order to lower the sampling requirements in the parameter domain. The sampling distances in the parameter domain is set to $\Delta\tau = 1$ and $\Delta p = 0.01$. The line parameters are listed in Tab. 1.

No	p	τ	μ	No	p	τ	μ
1	0.10	60	1.5	5	0.00	30	1.0
2	0.25	50	-1.5	6	0.40	45	-1.0
3	-0.10	80	-1.0	7	-0.36	52	-1.0
4	-0.25	30	1.0	8	0.10	80	0.5

Table 1: Line parameters. p is the slope, τ is the offset, and μ is the curve amplitude.

To illustrate the potential of the GRT a very noisy image is generated, by adding Gaussian noise to the noise free image with zero mean and standard deviation $\sigma = 1$. It can be seen from Figure 4, that the lines are hard to identify. Choosing $\mathrm{P}_{\mathrm{det\ all}} = 0.95$, Eq. 5 gives $\lambda^* = 6.53$; i.e., only lines with $|\mu| > 0.65$ should be detectable. This implies that all but line number eight should be detectable. The absolute value of the parameter domain obtained by the use of the GRT to the noisy image is shown in Fig. 5.

Since the noisy image contains few lines with absolute curve amplitude $|\mu|$ being of the same order of magnitude as σ and has approximately zero mean, σ was estimated from the image using the ordinary central variance estimator, which gave $\hat{\sigma} = 1.04$. Setting $\mathrm{P}_{\mathrm{det\ all}}$ to 0.95, Eq. 5 results in $L\ \mu^* = 65.7$. This is used for thresholding of the parameter domain as shown in Figure 6. Seven of the eight line parameters are found despite the poor signal to noise ratio in the image. Note that some of the lines will be represented by a few neighbor parameter vectors. This error be can corrected by clustering neighbor parameter vectors. The error is due to the sampling of the parameter domain and the finite image size.

The theory predicted that only seven lines could be detected. The eighth line can be detected if the curve length can be increased or the noise variance can be reduced. If the theory is used with $\mathrm{P}_{\mathrm{det\ all}}$ very low, $L\mu^*$ get lower and noise peaks will appear in the parameter domain along with parameters of the eighth line. In Fig. 7 the threshold level has been reduced to, e.g., $0.7L\mu^* = 46.0$. As it can be seen, noise will now give parameter vectors which do not represent a curve. As seen from Fig. 8 a further reduction of the threshold level to, e.g., $0.5L\mu^* = 32.9$ gives a parameter domain, where all eight lines are present. Due to the noise level many false parameter vectors can also be observed.

4. CONCLUSION

A statistically based noise analysis of the generalized Radon transform has been presented, which was used to derive a threshold level in order to separate curve information and noise in the parameter domain. A numerical example was provided to illustrate the theory.

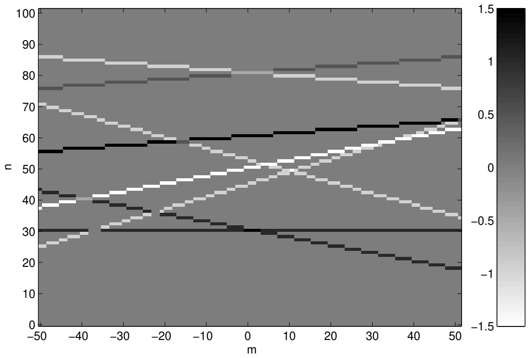

Figure 3: Noise free image with eight lines.

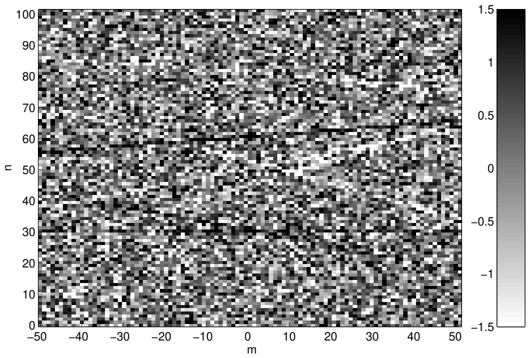

Figure 4: The same image contaminated with additive Gaussian noise.

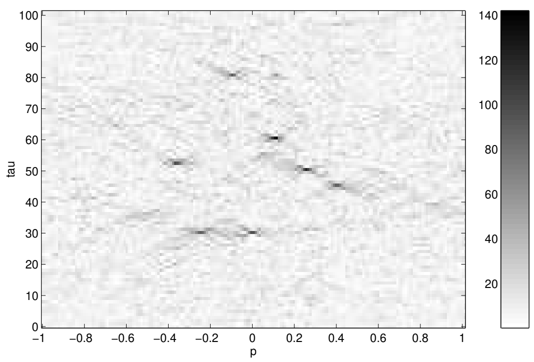

Figure 5: The absolute GRT of the noisy image. Note the peaks corresponding to the curves.

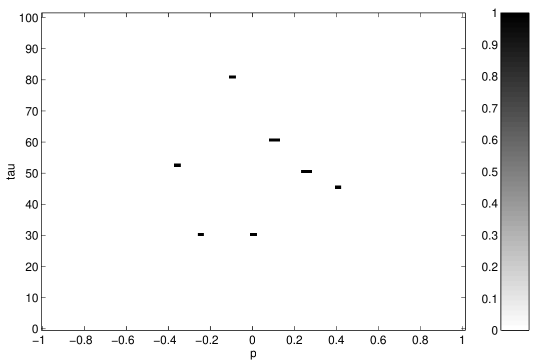

Figure 6: Threshold of the absolute GRT using the estimated threshold level.

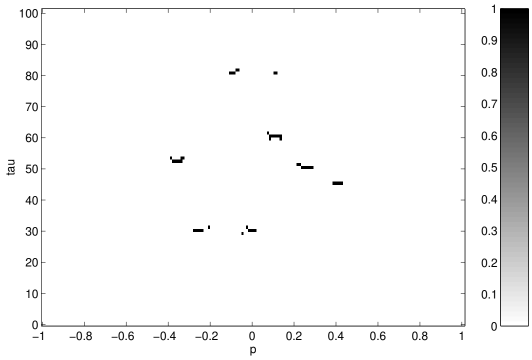

Figure 7: Threshold of the absolute GRT using 0.7 times the estimated threshold level.

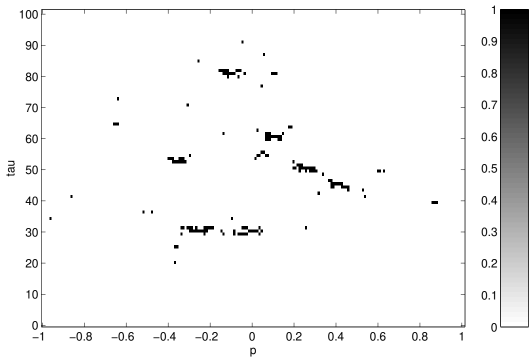

Figure 8: Threshold of the absolute GRT using 0.5 times the estimated threshold level.

5. REFERENCES

[1] J. Illingworth. A Survey of the Hough Transform. *Computer Vision, Graphics, and Image Processing*, 44, 1988.

[2] S. R. Deans. *The Radon Transform and Some of Its Applications.* Krieger Publishing Company, Malabar, Florida, 1993. Previously John Wiley & Sons Inc., 1983.

[3] G. Beylkin. Discrete Radon Transform. *IEEE ASSP*, 35(2):162–172, 1987.

[4] K. V. Hansen and P. A. Toft. Fast Curve Estimation Using Pre-Conditioned Generalized Radon Transform. Submitted for publication.

[5] P. A. Toft and K. V. Hansen. Fast Radon Transform for Detection of Seismic Reflections. In *Signal Processing VII - Theories and Applications*, volume I, pages 229–232. EURASIP EUSIPCO94, 1994.