EUSAR 2021 . ISBN 978-3-8007-5457-1 / ISSN 2197-4403 . ©VDE VERLAG GMBH . Berlin . Offenbach

A scheme for stabilizing the image generation for VideoSAR

Fabian Hartmann, Aron Sommer, Ulrike Pestel-Schiller, and Jörn Ostermann
Institut für Informationsverarbeitung, Leibniz Universität Hannover, Appelstr. 9A, 30167 Hannover, Germany,
hartmann@tnt.uni-hannover.de, sommer@tnt.uni-hannover.de, pestel@tnt.uni-hannover.de, office@tnt.uni-hannover.de

Abstract

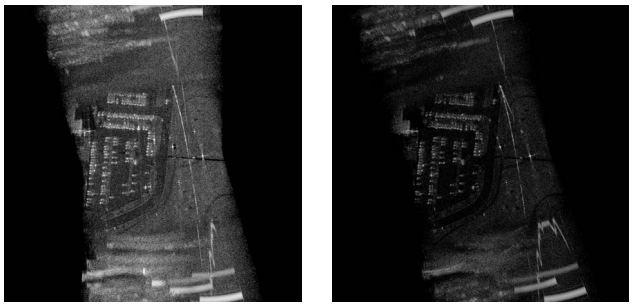
Using methods for SAR images in order to create VideoSAR raises the question of their temporal consistency with the scene. Some unadjusted normalizations and croppings even introduce artifacts into the videos. In this paper we present a scheme for stabilizing the generation of SAR videos to prevent artificial flickering of large parts of the frames of the video. The scheme is based on statistics of the video and we show that our method leads to a stable video using real data.

1 Introduction

Many years ago, sequences of SAR images were used for object detection and velocity estimation by Kirscht [1]. The name video synthetic aperture radar (VideoSAR) used in this paper for continuously concatinating a sequence of SAR images into a video is due to Wells et al. [2]. In this technique, the images are computed from partly overlapping windows of the data. SAR videos allow for better detection of moving targets. They require the computation of several SAR images, thus needing a multiple of computations. However this can be mitigated by reusing already computed amplitude frames when using Backprojection algorithms.
The transition from single SAR images to SAR videos requires the use of temporarily consistent methods. These methods have to specifically account for the different temporal behaviour of the pixel of the images concatenated to a video or be robust to this effect. The different temporal behaviour can be caused by changes occuring only in a part of the scene or by different changes in parts of the scene.
In the Literature, this problem was seldom addressed. Commonly, the SAR images are displayed in decibel. Doerry [3] detailed this and other methods for scaling and displaying SAR images. Pestel-Schiller and Ostermann [4] described a scheme for cropping the amplitudes of SAR images. But neither addressed problems introduced by creating SAR videos.
In this paper, we propose a scheme for automatically producing temporarily consistent frames in SAR videos. This is achieved by using statistics of the amplitude image for each frame separately to adapt the cropping, normalization and quantization of the amplitudes. Since this approach only uses the computed amplitude image, it can be used on a sequence of SAR images generated by Omega-K algorithms or by Backprojection algorithms.
The paper is structured as follows. In section 2 we discuss the problem adressed in this paper and in section 3 we describe the proposed processing scheme. In section 4 we describe the experiments conducted to test the proposed system and in section 5 we analyse the scheme regarding the statistics and the conversion method used.

2 Problem description

In this section, we address the need for cropping and normalisation and exemplify occuring problems.
Amplitude SAR images have a large dynamic range. Thus for displaying them, their dynamic range has to be limited. This is done by cropping their range of amplitudes.
Furthermor, like in ordinary images, the contrast of an image patch in SAR images depends on the distribution of amplitudes of all pixels in the frame. Thus the images have to be normalized.
Since SAR videos are a concatenation of SAR images, these frames have to be cropped too. But care has to be taken, not to introduce artifacts into the video by using cropping and normalization inconsistent with the actual changes in the scene. Most obvious in the video, this can lead to the frames flickering in brightness.



(a) Frame 1974 of Example 1 **(b)** Frame 2024 of Example 1

Figure 1 Example of video flickering: The normalization and the cropping change with time.

For comparison, Figure 1 shows the flickering between frames of a scene. The amplitude frames from Example 1 are each normalized to their respective maximum, their range cropped to $[-30\,\mathrm{dB}, -10\,\mathrm{dB}]$ and quantized to 8-bit with equal quantization steps.

3 Description of Scheme

In this section, we detail our proposed scheme for producing time consistent SAR videos. The scheme automatically adjusts the normalization and quantization for each frame to prevent artifacts from these steps when concatenating them to a video. These are caused by changes in the data, which the normalization and quantization did not adjust for.

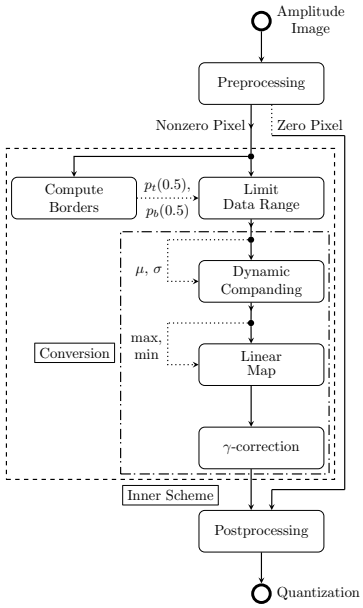


Figure 2 The proposed scheme, using the conversion from [4].

Figure 2 shows the block diagram of our proposed scheme. Amplitude images, generated by either Omaga-K algorithms or Backprojection algorithms, are separately fed into the scheme. In the Preprossesing step in Figure 2, potential pixels with zero amplitude are split apart from the other pixel. They are recombined with the processed Nonzero Pixel in the Postprocessing step. Since the functions used in the Inner Scheme of Figure 2 depend on the total number of pixel, adding blank pixel to the image changes them. Thus splitting apart Zero Pixel is necessary, since in both examples of section 4, parts of the szene are not in the antenna footprint. Therefore they have zero amplitude.

The Nonzero Pixel are fed into the inner scheme and follow the left path to the Compute Borders step, where the top and bottom 0.5-percentiles $p_t(0.5)$ and $p_b(0.5)$ are computed, where for $x \in [0,100]$

$$p_t(x) := \sup\left\{ y \in \mathbb{R} \,\middle|\, \frac{\#\{i|\text{Amplitude}(i) < y\}}{\#\{i\}} \geq x\% \right\},$$
$$p_b(x) := \inf\left\{ y \in \mathbb{R} \,\middle|\, \frac{\#\{i|\text{Amplitude}(i) > y\}}{\#\{i\}} \geq x\% \right\}$$

with $\{i\}$ the set of pixel in a frame.

These values control the Limit Data Range step on the right path of the Inner Scheme of Figure 2, where the amplitudes of the Nonzero Pixel are cropped to the range between these two values.

After this step, the pixel are processed by the Conversion scheme propoesd by Pestel-Schiller and Ostermann [4]. In the Dynamic Companding step, the amplitudes are cropped below the value $\mu + a \cdot \sigma$, where the parameter a is set to 16. In the Linear Map step of Figure 2, the amplitudes are linearly mapped from the range between their minimum and maximum to the range from zero to the square of the number of quantization levels, i.e. the quantization levels of the 8-bit greyscale output frames used for Figure 1. This is done, to align the lowest data with zero for the Postprossesing step and the highest with the highest quantization level after the next step. In that step, the amplitudes are γ-corrected. The effects of our scheme on in the examples of section 4 are discussed in subsection 5.2.

4 Experiments

In this section, the scheme is demonstrated on the data from the Gotcha Public Release Dataset [5] and the Air Force Research Laboratory (AFRL) Gotcha airborne SAR-based GMTI data set [6].

Table 1 Parameters of used X-band SAR data

Parameter	**Example 1 [5]**	**Example 2 [6]**
Carrier frequency	9.6 GHz	9.6 GHz
Chirp bandwidth	640 MHz	640 MHz
Sampling frequency	624 MHz	640 MHz
Squint angle	0°	0.13°
Depression angle	45°	45°
Number of pulses	42208	8192
Image size in pixel	1024x1024	2048x1024
Percentage of Zero Pixel	[54%, 62%]	[15%, 17%]

Table 1 shows the parameter of the data of the examples. The frames are each computed from 2000 pulses and they each start 16 pulses after the other, thus overlapping 99.2% in data from one frame to the next.

4.1 Example 1

We use the HH channel of the whole loop of pass 1 without autofocussing. Figure 1 provides examples of the output of the scheme from this example.

As seen in Figure 3, the statistics show four distinct spikes centered roughly around frame 600, frame 1250, frame 1900 and the last would be centered just right outside the Figure. Those correspond to the viewing angle being aligned to the orientation of the cars on the parking lot by a multiple of 90° and the cars brightening in the video. This causes other parts of the image to appear relatively dark.

4.2 Example 2

The part used consists of the first 8192 pulses with phase corrections.

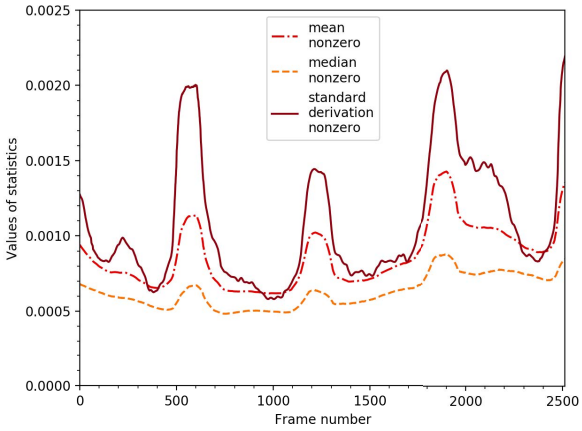


Figure 3 Comparison of the mean (red, '−·'), the median (orange, '−−') and standard deviation (crimson, '−') of the amplitudes of the Nonzero Pixel in Example 1 when entering the inner scheme.

5 Analysis

5.1 Statistics

In Figure 3 we compare the mean, the median and standard deviation of the Nonzero Pixel of each frame for Example 1. In this subsection, we analyse the Zero-normalized cross-correlation (ZNCC) of the graphs of several statistics of the frames, to determine the best choice of variables for cropping the range of amplitudes of the frames for VideoSAR. The ZNCC of two statistics x_1 and x_2 is defined as

$$\mathrm{ZNCC}(x_1,x_2) := \frac{1}{N}\sum_{i=1}^{N}\frac{x_1(i)-\mu_{x_1}}{\sigma_{x_1}}\frac{x_2(i)-\mu_{x_2}}{\sigma_{x_2}}$$

where μ_{x_j} is the mean of the framewise statistic $x_j(i)$ over all frames i, σ_{x_j} the corresponding standard deviation and N the number of frames.

Figure 5 lists the ZNCC of the maximum, the minimum, the mean, the standard deviation ('std'), the top and bottom 0.5-percentile, the variance ('var'), the median and the limit used by Pestel-Schiller and Ostermann [4] ('EuMW18'). Since the ZNCC is symmetric in its arguments, the upper half shows the values for Example 2 and the bottom half for Example 1. Figure 5 provides the statistics with and without the Zero Pixel and the later are labeled '*nonzero'. As we stated in Example 2, the resulting frames from that dataset contain several pixel with zero amplitude. Therefore some statistics with the Zero Pixel are omitted, since they are zero for each frame. Furthermore, the statistics are provided for the Nonzero Pixels with amplitude cropped to the range between the top and bottom 0.5-percentile. They are labeled '*limited'. In general, in Figure 5 we omit statistics which are equal to other statistics. For example, the maximum doesn't change by excluding the Zero Pixel and thus 'max nonzero' is equal to 'max' and therefore omited. Several ZNCC in Figure 5 are colored, for grouping them by color for the following analysis.

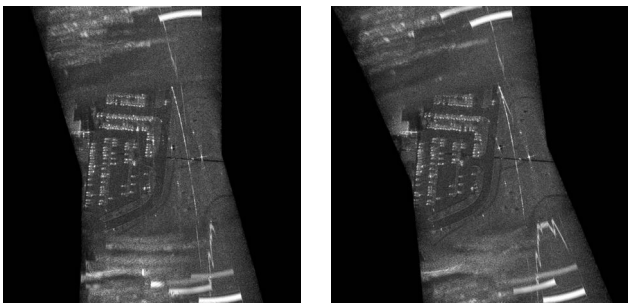
In Figure 5, the bottom and top 0.5-percentile are provided in comparison to the minimum and maximum, to compare different statistics for cropping the range of the amplitudes. We further examin the mean, standard deviation, variance and median as statistics to monitor the overall evolution of the frames. This is done, because the flickering in Figure 1 occurs over most of the image.

Figure 5 shows the low correlation of the maximum with the other variables (colored in red). Furthermore, the median is less correlated to the other statistical variables (colored in orange). Within the limits of accuracy, the limit by Pestel-Schiller and Ostermann [4] is perfectly correlated to the standard derivation. Lastly the mean, the top 0.5-percentile and the standard deviation, respectively the variance, have the highest cross correlation (colored in dark yellow). The top 0.5-percentile was chosen as the limit to crop the data because of its high correlation and the comparison with the limit by Pestel-Schiller and Ostermann [4] in subsection 5.2.

Table 2 Dynamic ranges of statistics for videos, normalized to the mean of the maximum of the resptective amplitude frames

Statistic	Example 1 [5]	Example 2 [6]
Maximum	$[4.48,27.27]10^{-1}$	$[8.42,12.45]10^{-1}$
Top 0.5-percentile of nonzero pixel	$[1.59,6.03]10^{-2}$	$[4.83,6.04]10^{-2}$
Bottom 0.5-percentile of nonzero pixel	$[1.23,2.32]10^{-4}$	$[1.26,1.67]10^{-4}$
Mean of cropped nonzero pixel	$[2.98,6.81]10^{-3}$	$[4.65,4.92]10^{-3}$
Standard deviation of cropped nonzero pixel	$[2.33,8.13]10^{-3}$	$[5.79,6.79]10^{-3}$
Limit of dynamic companding	$[4.01,13.68]10^{-2}$	$[9.72,11.36]10^{-2}$
Dynamic Range of nonzero pixel	$[19.98,25.23]$dB	$[24.64,26.51]$dB

Table 2 shows the minimum and maximum of the statistics used in our scheme (Figure 2) for Example 1 and Example 2. They are normalized to the mean of the mean maximum of the amplitude images of the respective example. In addition, the limit of the Dynamic Companding step is provided. Since it exceeds the top 0.5-percentile, the Dynamic Companding step doesn't crop the frames. Thus the Conversion step of Figure 2 is effectively just the Linear Map



(a) Frame 1974 of Example 1 **(b)** Frame 2024 of Example 1

Figure 4 Example of images produced by our scheme, analog to those in Figure 1

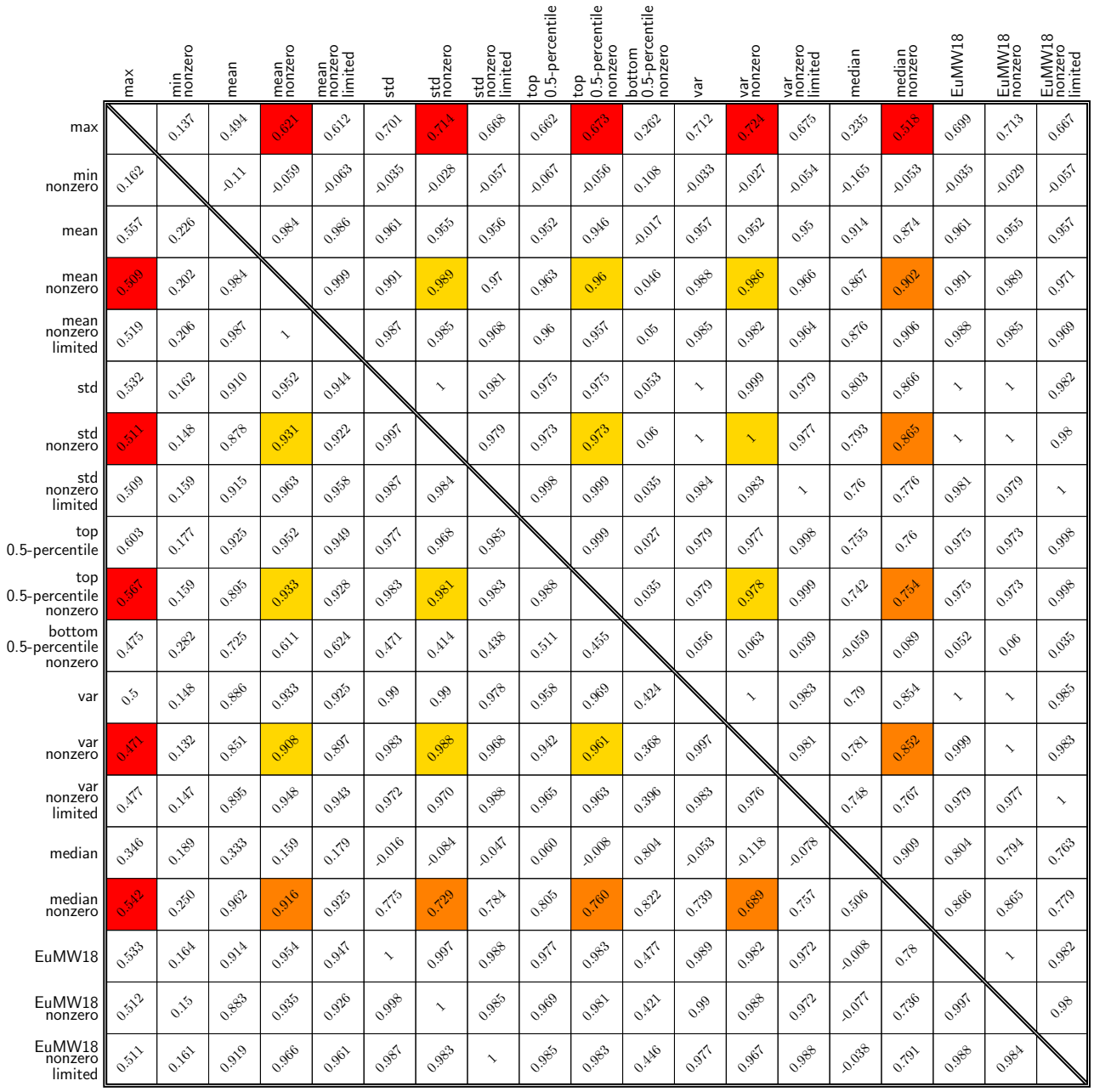


Figure 5 Zero-Normalized Cross-Correlation (ZNCC) of several statistical variables in Example 1 and Example 2. The ZNCC of the maximum with the other variables is colored in red, the ZNCC of the median with the other statistical variables is colored in orange and the ZNCC of the mean, the top 0.5-percentile and the standard deviation, respectively the variance are colored in dark yellow for the analysis. The upper half shows the values for Example 2 and the bottom half for Example 1.

and the γ-correction. We further provide the minimum and maximum of the dynamic range between the limits of the cropping in decibel.

5.2 Display

In this section we analyse the visual result of our proposed scheme. We refer to Doerry [3] for examples of the result of different conversion in place the one used in Figure 2, especially for the visual difference.

For comparison, Figure 4 shows the result of our scheme on the same frames as Figure 1. Subjectively, the video processed according to our scheme is of higher quality then the original. Furthermore, the cropping to a smaller range of amplitudes in our scheme in comparison to Pestel-Schiller and Ostermann [4] leads to a higher dynamic range for the frames, thus subjectively improving the quality.

6 Conclusion

In this paper, a scheme for generating amplitude images in VideoSAR was introduced, which is based on analysis of several statistics of amplitude frames of SAR videos. The proposed scheme accounts for different temporal behaviour of different parts of the scene. This is achieved by

copping the amplitudes of the pixel of the images framewise by there top 0.5-percentile, which is correlated to their mean and standard deviation according to our results. Subjectively, the quality was improved.

7 Literature

[1] M. Kirscht. Detection and velocity estimation of moving objects in a sequence of single-look SAR images. In *International Geoscience and Remote Sensing Symposium*, 1996.

[2] L. Wells, K. Sorensen, A. Doerry, and B. Remund. Developments in SAR and IFSAR systems and technologies at sandia national laboratories. In *Proceedings of the IEEE Aerospace Conference*, 2003.

[3] A. Doerry. Sar image scaling, dynamic range, radiometric calibration, and display. Technical report, Sandia National Laboratories, 2019.

[4] U. Pestel-Schiller and J. Ostermann. Subjective evaluation of compressed SAR images: Sometimes, compression improves usability. In *EUMW 2018*, 2018.

[5] C. Casteel, L. Gorham, M. Minardi, S. Scarborough, K. Naidu, and U. Majumder. A challenge problem for 2-D /3-D imaging of targets from a volumetric data set in an urban environment. In *Proc. SPIE—Algorithms Synthetic Aperture Radar Imagery XIV*, 2007.

[6] S. Scarborough, C. Casteel, L. Gorham, M. Minardi, U. Majumder, M. Judge, E. Zelnio, M. Bryant, H. Nichols, and D. Page. A challenge problem for sar-based gmti in urban environments. In *Proc. of SPIE - The International Society for Optical Engineering*, 2009.