

Suppression of Range Ambiguities in Synthetic Aperture Radar Systems

Abstract— One of the principle limitations of synthetic aperture radar (SAR) systems is the restriction on resolution and coverage. In a conventional SAR simultaneous wide swath (coverage) and high cross range resolution is not possible, since the two requirements contradict each other. In practice only a semi-optimum solution is obtained, which depends on the relation between the two performance parameters. This paper investigates possible enhancements of a conventional SAR system in order to overcome this contradiction. Special attention is paid to the analytic calculations necessary to evaluate the performance of the new methods and the feasibility of the hardware realization.

I. INTRODUCTION

The idea behind Synthetic Aperture Radar (SAR) is to replace the small real antenna aperture by a large *synthetic* aperture taking advantage of the radar's motion. The principal geometry of such a system is shown in Fig. 1.

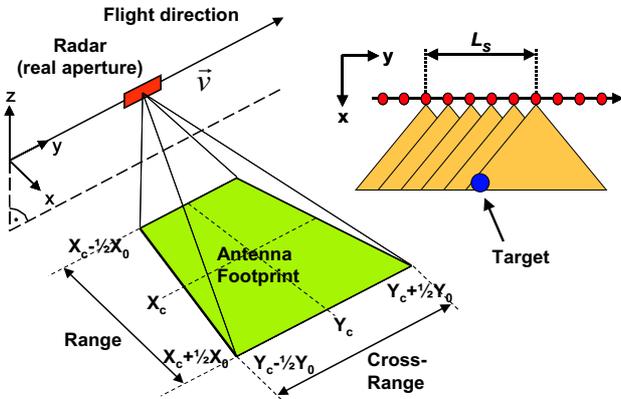


Fig. 1. Principal geometry of a SAR system

Data for a target on the ground is collected during the time it is illuminated by the antenna footprint. The distance covered by the radar during this time is known as the *synthetic aperture length* L_s . During the time a target is *seen* by the radar many pulses are transmitted and the scattered signals received by the radar. As the radar is moving, there is a Doppler shift in the received signals. Two steps are necessary in order to obtain radar images out of the collected data[1]: in the first step, *range compression* is used to obtain

the distance to each pixel in the image area. In the second step, known as *azimuth compression*, the exact position of the target in the cross-range direction is determined by evaluating the Doppler information of the received signals.

The cross range (azimuth) resolution Δy of a conventional SAR system is determined by the antenna length [1]:

$$\Delta y \approx \frac{L_a}{2} \quad (1)$$

where L_a is the along-track length of the antenna. Reducing the antenna length results in a wide beamwidth which improves the azimuth resolution Δy . The return echo signal is recorded whenever the platform moves a distance Δu . The maximum cross range spatial sample separation Δu required to process the measured signals without aliasing effects (azimuth ambiguities) is given by [2]:

$$\Delta u \leq \frac{L_a}{2} \quad (2)$$

For each azimuth sample, one chirped pulse signal is transmitted by the radar and the echo recorded by the receiver. The spatial separation of the samples is thus determined by the platform velocity and the pulse repetition frequency PRF . However, the PRF also determines the maximum swath width – the higher the PRF the smaller the swath:

$$PRF \leq \frac{c}{2X_0} \quad (3)$$

where X_0 describes the length of the target area in range direction. Neglecting this crucial condition leads to the appearance of range ambiguities in the processed radar images resulting in shadow targets that can not be distinguished from real targets.

Thus two contradicting requirements have to be met when designing a SAR-system. According to (1), a high azimuth resolution can be achieved through a smaller antenna length, this results in a higher sample spacing, and thus a smaller value for Δu as given by (2). For a given platform velocity a smaller Δu requires a higher pulse rate and this a higher PRF , which in turn results in a smaller ambiguous free swath width according to (3). It can be said that for a conventional SAR system a high azimuth resolution and wide swath can not be achieved simultaneously.

This paper investigates an approach to overcome the above mentioned contradicting requirements in SAR systems. The basic idea of the approach is described in section II. In section III a figure-of-merit for performance evaluation is introduced. Section IV develops the analytic expressions needed for the evaluation. Two different methods to further improve the performance are investigated in sections V and VI. Finally, a SAR simulator and image generator is used in section VII to validate the result for the different methods.

II. BASIC APPROACH TO PERFORMANCE ENHANCEMENT

The basic idea for increasing the *PRF* without diminishing the swath is the ability to *mark* the pulses. The system must be able to transmit signals with different marks and to identify the scattered signals respectively. The azimuth sample separation Δu is given by the – small – time interval between two successive samples of different type, while the swath is confined by the – large – time interval between similar pulses. In the case of two marks, the *PRF* can be doubled for the same swath or alternatively the swath doubled while maintaining the *PRF*.

Fig. 2 explains this principle. The marked signals are denoted by two different colours, each colour representing a special mark. The receiver consists of two parallel paths each containing a matched filter for one marked signal. The respective unwanted signal is suppressed. Since the *PRF* within each receive branch is halved with respect to the incoming signals, the swath can be doubled. In reality there exists several possibil-

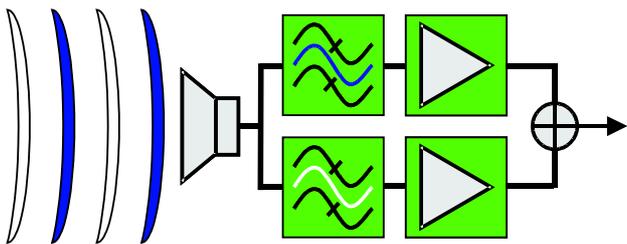


Fig. 2. Using marked pulses to differentiate the received signals

ities to mark pulses. The approach given in this paper is to make use of different chirp signals (up- and down-chirps). As chirp signals are commonly used in SAR systems, only minor modifications in the hardware of the transmitter path are necessary in order to transmit alternate up- and down-chirps. In the receiving part an additional branch with a matched filter has

to be added. Since the received signals are commonly processed using digital signal processing the modification on the receive side reduces to a simply software update.

III. FIGURE-OF-MERIT FOR SYSTEM PERFORMANCE

A crucial point is how good the different received signals can be differentiated from each other. This decides whether the approach is useful in practice or not. The matched filters in the receive branches compresses the wanted signals while suppressing or smearing the unwanted signals. A matched filter represents a correlation between the transmitted and the received signals.

For a wanted received signal, the output of the correlation is given by CCF^{xx} and for the correlation with an unwanted signal by CCF^{xy} later on. In order to validate this system, the levels of these functions have to be calculated. The figure-of-merit is given by the ratio between the two correlation functions

$$\Delta_{CCF} = 10 \log \left(\frac{\max\{|CCF^{xx}|\}}{\max\{|CCF^{xy}|\}} \right) \quad (4)$$

where a high value for Δ_{CCF} means a good performance.

IV. RECEIVED SIGNAL REPRESENTATION

The signals commonly used in SAR systems are linearly frequency modulated (*chirp signals*) given by

$$s_1(t) = \exp(j\omega_1 t + \alpha_1 t^2) \cdot \text{rect}[0, T_p] \quad (5)$$

The rectangular window $\text{rect}[0, T_p]$ limits the signal to the transmitting interval $0 \leq t \leq T_p$, where T_p denotes the pulse duration. The properties of this signal are determined by the parameters ω_1 (start frequency) and α_1 (chirp rate).

The instantaneous frequency of the chirp is the derivative of the phase with respect to time:

$$\begin{aligned} \omega_{inst}(t) &= \frac{d}{dt} (\omega_1 t + \alpha_1 t^2) \\ &= \omega_1 + 2\alpha_1 t \end{aligned} \quad (6)$$

The baseband bandwidth in the radar signal is

$$B_0 = 2\alpha_1 T_p \quad (7)$$

A. Cross-Correlation of Chirp Signals

The output of the correlator is given by [3]

$$CCF(\tau) = \lim_{T_p \rightarrow \infty} \frac{1}{T_p} \int_{-T_p/2}^{T_p/2} s_1^*(t) \cdot s_2(t + \tau) dt \quad (8)$$

For two chirp signals

$$s_1(t) = \exp(j2\pi f_1 t + j2\pi\alpha_1 t^2) \cdot \text{rect}[0, T_p] \quad (9)$$

$$s_2(t) = \exp(j2\pi f_2 t + j2\pi\alpha_2 t^2) \cdot \text{rect}[0, T_p] \quad (10)$$

it is given by

$$CCF^{xy}(\tau) = \frac{U_1 U_2}{2T_p \sqrt{\alpha_2 - \alpha_1}} \cdot \exp\left(j\pi \left(2f_2\tau + 2\alpha_2\tau^2 - \frac{(f_1 - 2\alpha_2\tau - f_2)^2}{2(\alpha_2 - \alpha_1)}\right)\right) \cdot \left[(C(\nu_2) - C(\nu_1)) + j(S(\nu_2) - S(\nu_1)) \right] \quad (11)$$

with $C(\nu)$ and $S(\nu)$ representing the real Fresnel-integrals (a simple approximation of these integrals is given in [4]). The parameter ν is a function of the chirp rates α , start frequencies f , and the integration limits. Since the signals $s_1(t)$ and $s_2(t)$ differ from each other, the result of the above equation (11) represents the cross-correlation CCF^{xy} .

B. Auto-Correlation of Chirp Signals

The Auto-Correlation can not be calculated using (11), since the denominator vanishes for $s_1(t) = s_2(t)$. Solving (8) for $s_2(t) = s_1(t)$ the correlation function becomes

$$CCF^{xx}(\tau) = \frac{U_1^2}{4\pi\alpha_1 T_p \tau} \exp(j2\pi(f_1\tau + \alpha_1\tau^2)) \cdot \left(\exp(j4\pi\alpha_1\tau t_u) - \exp(j4\pi\alpha_1\tau t_o) \right) \quad (12)$$

t_o and t_u represent the integration limits. These limits differ for $\tau < 0$ and $\tau > 0$ and are given by:

$$\tau < 0 : \begin{cases} t_u = -\tau \\ t_o = T_p \end{cases}$$

$$\tau > 0 : \begin{cases} t_u = 0 \\ t_o = T_p - \tau \end{cases}$$

For $\tau = 0$ (12) results in an undefined expression $\frac{0}{0}$. Using l'Hospital's rule this problem can be circumvented:

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)} \quad (13)$$

Therefore the derivatives of denominator and numerator are needed. Inserting $\tau = 0$ and the integration limits results in

$$CCF^{xx}(\tau = 0) = U_1^2 \cdot j \frac{\overbrace{-j4\pi\alpha_1 T_p \exp(j4\pi\alpha_1 0 T_p)}^{=1}}{4\pi\alpha_1} + \frac{\overbrace{j4\pi\alpha_1 T_p 0 \exp(j4\pi\alpha_1 0 T_p)}^{=0}}{4\pi\alpha_1} \quad (14)$$

This can be reduced to the following simple equation:

$$CCF^{xx}(\tau = 0) = U_1^2 \cdot T_p \quad (15)$$

C. Simulation Results

By means of (11), (12), and (15) the levels of CCF^{xy} , CCF^{xx} and the ratio Δ_{CCF} can now be calculated. Since the analytic solution is available, the dependencies on various parameters can be evaluated easily.

An example of the shape of the two correlation functions is shown in Fig. 3. The shape of these functions is typical of CCF^{xx} and CCF^{xy} of signals covering the same frequency band: the CCF^{xx} -function shows an explicit peak that indicates a point target, the CCF^{xy} -function shows a nearly constant level over the entire correlation length $2T_p$. The ratio of the levels always refers to the maxima of both functions as indicated in this figure.

At a first glance the analytic solution for the correlation functions seems to depend on various parameters, e.g. the start frequencies f_1 and f_2 , chirp rates

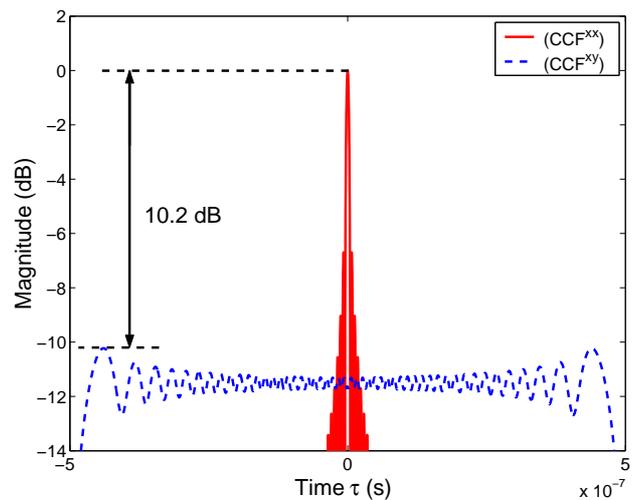


Fig. 3. Example of level difference of CCF^{xx} and CCF^{xy} for $T_p = 0.5 \mu\text{s}$ and $B_0 = 200 \text{ MHz}$

α_1 and α_2 and the pulse duration T_p . Some parameters however are interdependent and as a result the levels of the correlation functions only depend on the pulse duration T_p and bandwidth B_0 (time-bandwidth product). Increasing one (or both) of these parameters results in a better suppression of the unwanted signals. The dependencies are shown in Fig. 4 where the level difference Δ_{CCF} is plotted over the bandwidth B_0 for three different pulse durations T_p .

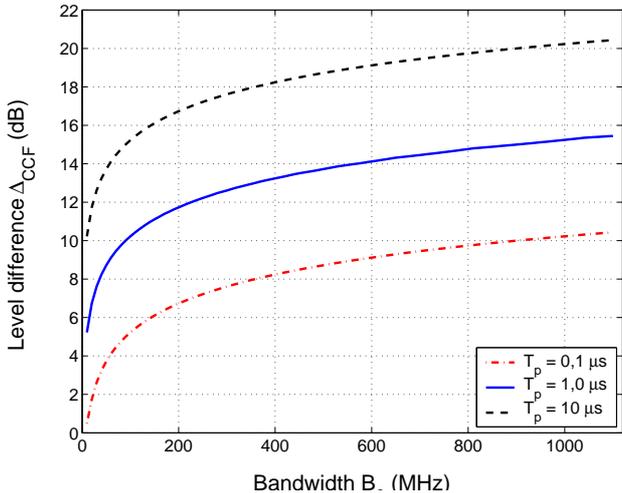


Fig. 4. Level difference Δ_{CCF} as a function of the bandwidth B_0

Based on this results large bandwidths ($B_0 \geq 1000$ MHz) and long pulse durations $T_p \geq 10 \mu\text{s}$ result in level differences higher than 20 dB. The Terra SAR-X system [5] uses a bandwidth of 150 MHz and a pulse duration of $25 \mu\text{s}$. With these system parameters a level difference of about $\Delta_{CCF} = 18$ dB can be reached.

In practice, level differences of about 20 dB are not sufficient to process high quality radar images. In order to distinguish between the wanted and unwanted signal parts and being able to detect targets with small radar cross section it is necessary to increase the level differences considerably.

V. SPLIT BANDWIDTH METHOD

The CCF^{xy} can be further reduced by using different frequency bands. The total available bandwidth of the radar B_0 is split into two sub-bands with the respective bandwidth $B_0/2$. Now the different chirp forms are assigned to the different sub-bands, e.g. the up-chirp to sub-band 1 and the down-chirp to sub-band 2. Consequently the correlation of two signals of different sub-bands will be denoted as CCF^{u1d2} and the correlation between two signals in the same

sub-band as CCF^{u1u1} (u denoting up-chirp, d down-chirp). For comparison the pulse durations are kept constant, since the sub-band bandwidth is only half as large, the chirp rates of the split chirps are halved.

The advantage of using this method is shown in Fig. 5. In the upper part the course of the instantaneous frequency f is plotted over time t . The assignment of the chirp forms to the respective sub-bands becomes clear. The lower part of Fig. 5 shows a direct comparison of the CCF between the original pulses (sharing the available bandwidth B_0) and the halved pulses (covering only half the bandwidth). The level of CCF^{u1d2} is significantly lower than the original level. Besides the lower level over the whole length of the CCF the shape of CCF^{u1d2} shows a more favourable run of the curve: there is only one explicit maximum and the other values decrease continuously to the left and the right to a very low level at the beginning and the end of the correlation length.

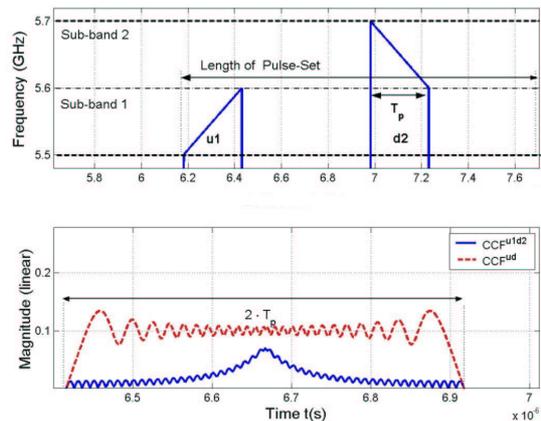


Fig. 5. Frequency setup for split bandwidth. Comparison of CCF^{u1d2} and CCF^{ud} for $B_0 = 200$ MHz and $T_p = 0.25 \mu\text{s}$

Simply evaluating and processing these halved pulses leads to a better suppression of the unwanted signals, but at the same time the resolution in range dimension Δx is also halved, since the resolution Δx solely depends on the bandwidth B

$$\Delta x = \frac{c}{2B} \quad (16)$$

VI. OFFSET BANDWIDTH METHOD

A further possibility of increasing the level difference between wanted and unwanted signals is to allow a frequency gap between the two sub-bands. The influence of this “guard”-band is shown in Fig. 6, where

the effect of the gap is shown as a function of the gap width. The entire bandwidth B_0 remains constant for increasing gap widths. However, the active bandwidth decreases: with a gap of 20% of the total bandwidth B_0 only 40% of B_0 remains for each chirp signal.

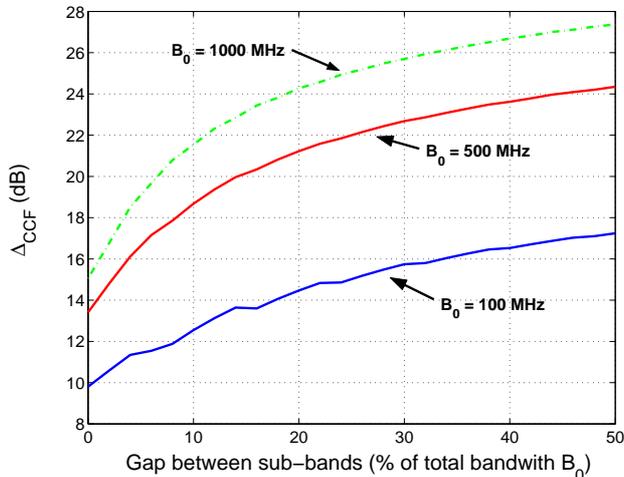


Fig. 6. Influence of a gap on the level difference between CCF^{xx} and CCF^{xy} ($T_p = 0, 25 \mu s$)

For example, a system with 500 MHz bandwidth the level difference can be improved from original 14 dB (no gap, the sub-bands are adjacent to each other) to 24 dB by using a gap width of 45%. Consequently the gap covers 225 MHz, the remaining 275 MHz is split between the different pulse forms. This means a decreasing resolution in x -direction as the gap width increases. Despite this fact the use of frequency gaps can be an attractive alternative for systems with large bandwidth.

VII. SIMULATED SAR IMAGES

In this section, three different kinds of radar processing will be shown. The target scenario, including point targets and extended targets with different radar cross sections, is the same for all processed images (see Fig. 7).

A. Conventional SAR Algorithm with doubled PRF

Doubling the PRF using the original SAR algorithm results range ambiguities (Fig. 8). All targets are detected twice, since the doubled PRF is not compensated by a reduced size of the target area (see equation (3)). It is not possible to distinguish between real targets and ambiguities, sometimes targets and ambiguities cover each other in the radar image.

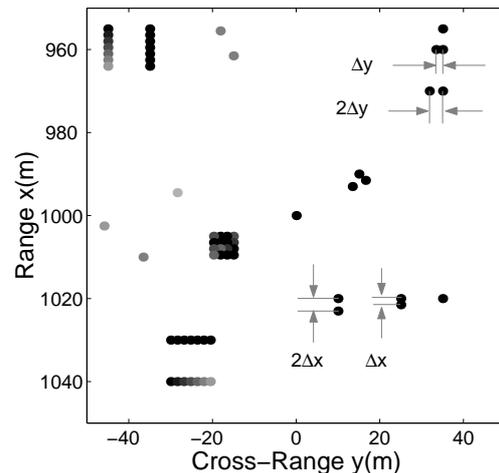


Fig. 7. Target scenario of radar images 8, 9, 10 and 11

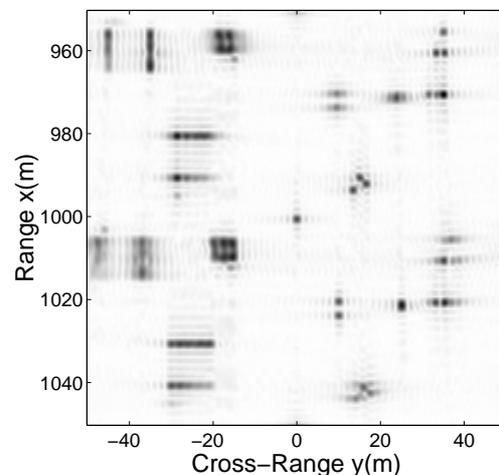


Fig. 8. Processing with the original SAR algorithm (Up-chirp only, $PRF = 3 \text{ MHz}$)

B. Up-/Down-Chirps, Same Frequency Band

The first step was the extension of the original system by two different pulses (up- and down-chirp), covering the same bandwidth B_0 . The processed images can be seen in Fig. 9. The targets with high radar cross sections can be detected without problems. The PRF was doubled as in Fig. 8, but instead of ambiguities correlation-noise appears, caused by unwanted signal parts that were not suppressed completely. This correlation-noise results in a high noise floor and limits the dynamic range which prevents the detection of weak targets.

C. Up-/Down-Chirps, Split Bandwidth

Introducing bisection of the total bandwidth B_0 into two sub-bands with assigned up- and down-chirps solves this problem: the level differences between

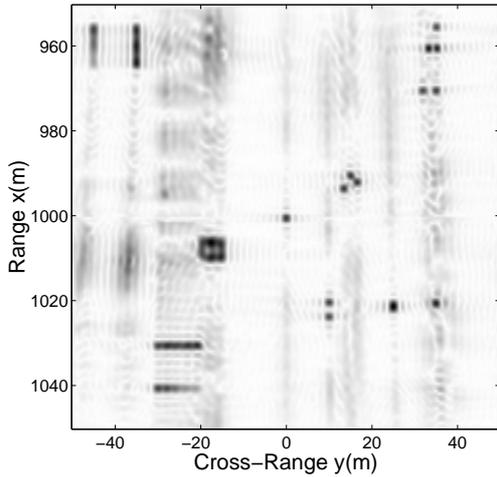


Fig. 9. Processing with up- and down-chirps covering the entire bandwidth ($PRF = 3$ MHz)

wanted and unwanted signal parts become obviously larger and the correlation-noise floor is on a lower level (see Fig. 10). All targets can be detected. In this picture, only the upchirp was processed, resulting in a deterioration of the resolution Δx in range direction, which is only dependent on the used bandwidth. Processing both up- and downchirps and combining this data into a radar image leads to Fig. 11. Here the original resolution in range direction is restored, resulting in a very clear radar image, where all targets – even those with a weak radar cross section – can be detected without problems.

Even if the chosen parameters of the above example are quite far away from reality (pulse duration was chosen quite short in order to reduce the computation time of the simulations), there are almost no

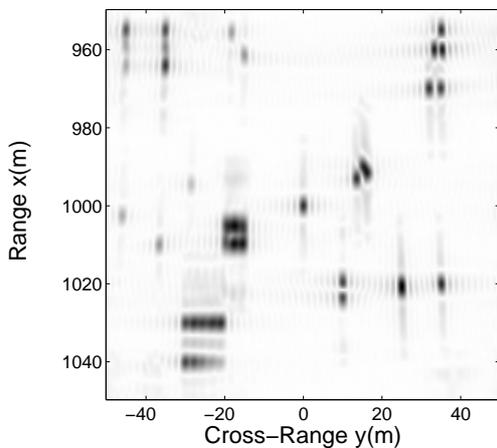


Fig. 10. Processing with halved up- and down-chirps, no pulse-combining ($PRF = 3$ MHz)

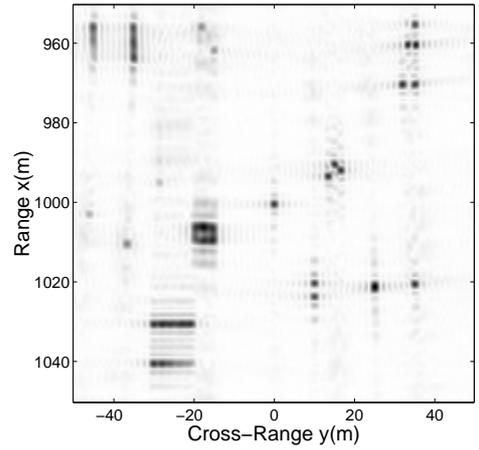


Fig. 11. Processing with halved up- and down-chirps and pulse-combining ($PRF = 3$ MHz)

differences between the original SAR processing and the enhanced system (Fig. 11). Doing the simulation of these processing with real bandwidths and pulse durations will result in radar images that cannot be distinguished by a first glance.

VIII. CONCLUSION

An improvement of conventional SAR systems by introducing differently marked pulses can lead to either better resolution in the cross range dimension or to an enlarged target area in range direction. By using up- and down-chirps as marks even existing systems can be updated quite easily. The theoretical background for validation of such improved systems is provided in this paper. Introducing a guard band between two sub-bands provides further improvements, even if the trade-off between better suppression of unwanted signals and the desired resolution Δx has to be taken into account.

REFERENCES

- [1] S. Hovanesian, *Introduction to Synthetic Array and Imaging Radars*. Artech House, 3 ed., 1985.
- [2] M. Soumekh, *Synthetic Aperture Radar Signal Processing*. John Wiley, 1999.
- [3] F. G. Stremler, *Introduction to Communication Systems*. Addison-Wesley, 2 ed., 1982.
- [4] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*. New York: Dover Publ, 1972.
- [5] R. Zahn and E. Velten, “Enabling technologies for the TerraSAR mission,” in *Proceedings Int. Geoscience and Remote Sensing Symposium IGARSS’2000, Honolulu, Hawaii*, vol. 3, pp. 1174 – 1176, 2000.