Proceedings of the
2007 IEEE Intelligent Vehicles Symposium
Istanbul, Turkey, June 13-15, 2007

Lane Recognition on Country Roads

U. Franke, H. Loose, C. Knöppel
DaimlerChrysler AG

Abstract — **Most of the common lane recognition systems are designed to work on well structured roads and rely on the existence of markings. In this paper we present a lane recognition scheme for country roads. Our novel approach works even in the absence of markings. The parameter estimation is formulated as a maximum-a-posteriori estimation task fusing color, texture, and edges. The framework can easily be extended by additional features not considered here. The optimization is carried out by means of a particle filter. Efficient computation schemes allow running the system in video real-time using a standard PC. The proposed algorithm can cope with varying feature statistics. Practical tests prove the robustness on marked as well as unmarked roads.**

I. INTRODUCTION

THE reliable recognition of the road course and the relative position of the vehicle is a prerequisite for many driver assistance systems such as lane departure warning, lane departure protection, and lane keeping. Most systems developed until now follow the classical principle introduced by Dickmanns [1]. Lane markings are detected and a Kalman Filter is used to determine the parameters of a clothoidal road model. Much effort is carried out in order to make these systems operational if the visibility of the markings degrades. Pomerleau tried to circumvent this undesired dependency with the RALPH [2] system which has become part of Assist Ware's lane departure warning system. Zhang and Nagel [3] classified the road region using texture features and pixel locations.

Particle filter (PF) (see e.g. [4] for an introduction) based estimation schemes have gained interest in recent years due to their ability to exploit information such as "this pixel is more likely to belong to the road than to the background". Such information cannot be used in Kalman Filters. While Southall [5] describes a PF-based system that only considered markings, the approach investigated by Apostoloff and Zelinsky [6] fuses markings with color and edge information. Recently, Smuda et al. [7] used a PF to extend the visibility range of a standard lane recognition system by adding texture information in the long range. Meis and Schneider [8] use a PF to obtain road course information from the data of a 32 beam radar sensor. Hummel et al. use stereo, texture and color for path planning in the Grand Challenge scenario [9].

In the future, lane recognition systems will be required that are not only restricted to well marked highways and secondary roads, but work on unmarked or badly marked country roads as well. At the same time, these systems should be able to exploit markings when available. For lane departure warning, we need principles to estimate at least the lateral vehicle position using cues other than markings. If one aims at more sophisticated assistance systems, the estimation of the road course is desirable.

The most prominent features that support the estimation of the road course are intensity, color, texture, edge strength, edge direction and height-over-ground. We propose to treat the road recognition as a maximum-a-posteriori (MAP) estimation task that optimizes the parameters of a road model given an image sequence. This optimization is carried out by means of a particle filter for simplicity and the lack of faster schemes. The statistics of the pixel features intensity, color and texture for the road as well as the non-road area are estimated on-line and used to determine the most probable road course using a mathematic framework. The edge information is also used probabilistically without any threshold that suppresses information.

Section II describes the basic MAP estimation system for pixel features and a computationally efficient formulation necessary to achieve real-time capability of the PF. Section III discusses the use of the pixel features intensity, color and texture. Section IV presents a novel generalized distance transform and a proper way to exploit the edge direction in a probabilistic framework. Results are presented in section V.

Fig. 1: Country road with two possible road hypotheses. Obviously, the dark marked hypothesis should result in a higher probability than the bright one.

Correspondence address: uwe.franke@daimlerchrysler.com

II. MAP BASED ROAD RECOGNITION

Given an image I, the goal is to determine the optimal parameter vector $\vec{p}$ of an appropriate road model that maximizes the probability $p(\vec{p} \mid I)$. For example, the dark marked model in Fig. 1 should result in a higher probability than the bright one. In order to maintain maximum compatibility with standard lane-marking based systems we stick to the common clothoidal road model:

$$x(L) = \pm 0.5b - x_{offset} - \Delta\psi L + \frac{1}{2}c_0 L^2 + \frac{1}{6}c_1 L^3 . \qquad (1)$$

The meaning of the parameters is as follows. b : road width, x_{offset} : lateral position of the vehicle relative to the road center, $\Delta\psi$: vehicle's yaw angle relative to the road, c_0 and c_1 : curvature and curvature change, respectively and L : look ahead distance.

In the following, we will concentrate on pixel-based features such as intensity. According to Bayes rule, for any feature image $F(x,y) = f(I(x,y))$ the probability $p(\vec{p} \mid F)$ can be rewritten as

$$p(\vec{p} | F) = \frac{p(F | \vec{p})\, p(\vec{p})}{p(F)} . \qquad (2)$$

Maximizing of (2) is equivalent to maximizing its enumerator. If we assume that single image points are independent, $p(F \mid \vec{p})$ can be expressed by:

$$p(F|\vec{p}) = \prod_{(x,y)\in lane} p_1(F(x,y)) \cdot \prod_{(x,y)\notin lane} p_0(F(x,y)) \qquad (3)$$

with p_1 representing the probability density of the considered feature on the road (foreground) and p_0 representing the probability density of the feature in the background.

Obviously, the calculation of these products is computationally expensive and – even worse – numerically unstable. Thanks to the Boltzmann statistics, probability and energy are coupled according to

$$p \propto e^{-\beta E} \qquad (4)$$

with a constant β. This relationship allows switching between both perceptions. Maximization of the product (3) is equivalent to minimization of the negative logarithm (energy):

$$-\log p(F|\vec{p}) = -\sum_{(x,y)\in lane} \log(p_1(F(x,y))) - \sum_{(x,y)\notin lane} \log(p_0(F(x,y))) . \qquad (5)$$

Within the PF framework, this formula has to be evaluated for each particle. This fact would prohibit a large number of particles, which in turn is desirable to get good estimates. However, a close look reveals that the summation can be carried out efficiently using line-oriented integral images which have to be generated only once. With

$$F_{\text{int}}(x,y) = \sum_{i=0,..,x} F(i,y) \qquad (6)$$

the summations in (5) are reduced to only one difference for the first and two differences for the second sum of eq.5, respectively.

After normalization to the number of considered pixels the computed energy is reconverted into a meaningful probability. The fusion of statistically independent cues leads to the joint probability

$$p(F_1 \ldots F_n \mid \vec{p}) = \prod_{f=1,\ldots,n} p_f(F_f \mid \vec{p}) \qquad (7)$$

where p_f denotes the probability of a road model considering feature f .

III. PIXEL FEATURES

The above estimation works with various features. In the following we will consider pixel intensity, texture, and color.

A. Intensity

In [9] Hummel assumes the intensities of the road to be Gaussian distributed and proposes an equal distribution for the non-road area, since no additional a-priori information is available. In cooperative situations (as shown in Fig. 1) this approach works well. However, in more complex situations, e.g. with shadows on the road (see Fig. 2), this assumption is ill-suited.

Fig. 2: Unmarked road with shadows that cause violation of a Gaussian gray value probability assumption. The statistics for foreground and background are taken in the gray and black marked areas, respectively.

Fig. 3 shows the gray value histograms $h(f)$ of the road area (marked gray in Fig. 2) and the non-road area (marked black in Fig. 2). Instead of using a parametric model, we use a low-pass filter and normalize these histograms and use them as conditional probability densities. In order to ensure that no factor in (3) has a value equal to zero, which would rate the current hypothesis as absolutely impossible, we add a small constant c. Thus we get

$$p(f) = \frac{1-c}{N} h(f) + c/256 \qquad (8)$$

for 8 bit images. Fig. 4 shows the "probability images" for the road assumption $p_1(I(x,y))$ and non-road assumption $p_0(I(x,y))$.

It is evident that computing the logarithm in (5) can be efficiently implemented by means of a lookup-table. The integral images are built upon the mapped log-images.

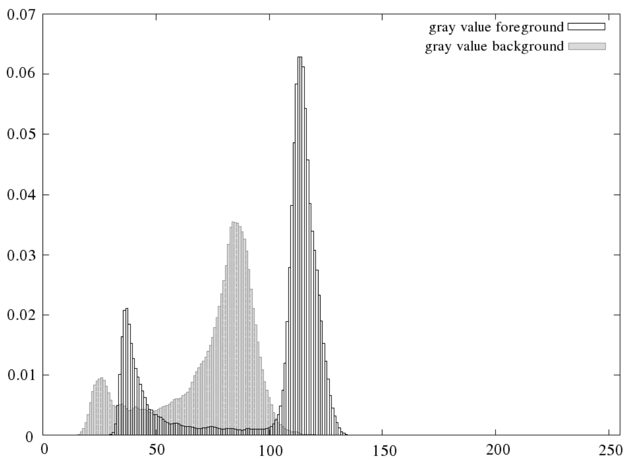

Fig. 3: Conditional intensity histograms for the road area and the non-road area shown in Fig. 2

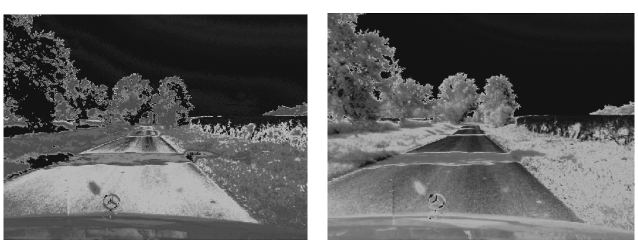

Fig. 4: Conditional probability images based on intensity. Left: hypothesis "road" ($p_1(I(x,y))$), right: hypothesis "non-road" ($p_0(I(x,y))$). The brightness is proportional to the probability.

B. Color

Whereas roads are usually gray, the non-road area is often covered with grass and plants. For this reason, color can help to determine the road parameters. It is obvious, that the RGB color space is not well suited. We convert the RGB-color to the HSI color space and use the hue as a second feature. Fig. 5 shows a colored road image and the corresponding hue channel. This feature is used in the same way as the intensity. Unfortunately, tests with today's automotive qualified CMOS color cameras gave less satisfying results than one might expect. One reason is that these sensors deliver weaker color saturation than consumer CCD cameras. Secondly, we did the tests in late fall when all colors were faded out.

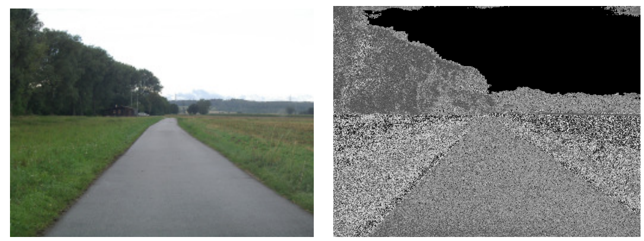

Fig. 5: Left: original image, right: intensity encoded hue

C. Texture

Roads are generally less structured than the background. Therefore, texture is a good cue for road recognition. Among the various ways to measure texture we select the well known structure tensor given by

$$G = \frac{1}{N} \sum_{W} \begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix}. \qquad (9)$$

This tensor was already used for road segmentation in [3]. In unstructured areas, both eigenvalues of the structure tensor are small. At edges, one eigenvalue is large whereas the other is small. In textured areas, both eigenvalues have values significantly different to zero. Following the principle used for intensity and color, we compute the histograms for the smaller eigenvalue in the road area and the non-road area and use them as estimates for the conditional probability density functions. This allows the usage of the same mathematics and the fast implementation as described in section II.

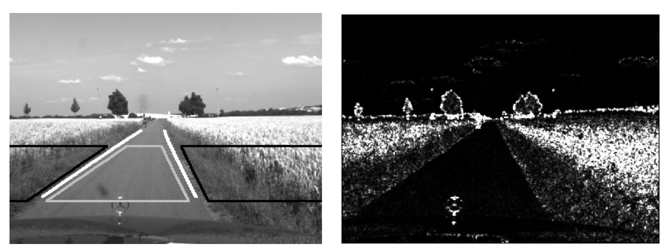

Fig. 6: Texture is an important information cue for road recognition. The right image shows the intensity encoded smaller eigenvalue of the structure tensor.

IV. Edge Features

Even if markings are missing, at most times the borders of paved roads generate visible edges. It is a must to exploit this information. If there is no doubt which edge describes the road boundary, a standard Kalman Filter approach offers a robust and fast estimation and is therefore the favoured solution. However, on country roads the road edges may be weak and ambiguous. In this case, special care is necessary and the fusion of edge information with other cues is beneficial.

Edge information is only meaningful in a probabilistic framework if probabilities or energies can be derived for each road parameter vector. Obviously, the probability should be higher (energy lower, respectively)

1. the closer the expected road boundary is to the edge,
2. the higher the gradient of that edge is, and
3. the less the direction of that edge deviates from the expected orientation.

The aimed formulation has to meet these three requirements. If an adequate energy value $e(x, y)$ can be given for each pixel, the total energy of a certain road hypothesis is computed as

$$E_e(\vec{p}) = \sum_{(x,y)\in boundary} e(x, y)\,. \qquad (10)$$

A common way for binary correlation is the distance transform (DT, e.g. [10]). Starting with a binary edge image $g(x, y)$, an image $DT(x, y)$ is generated encoding the distance to the closest edge for every pixel. This distance can be interpreted as energy, thus fulfilling requirement 1. In the one-dimensional case, the DT is defined by

$$DT(x) = \min_{u\in row,\, g(u)\neq 0} \|x-u\| \qquad (11)$$

with a certain norm, usually L_1 or L_2. Fig. 8 shows the DT image (left) for the road situation displayed in Fig. 1. The disadvantage of this correlation scheme is that it is based on a binary edge image which includes a critical threshold. Consequently, DT is not robust in the sense that small signal changes can result in large changes of the output. Moreover, sometimes the road edges are so weak that any standard threshold would be too high.

In [11] a generalization of the DT is proposed by adding a data dependent term to the optimization:

$$DT(x) = \min_{u\in row} (\|x-u\| + g(u))\,. \qquad (12)$$

An algorithm that computes this transform in $O(n)$ time is presented in the reference.

For the considered problem, we slightly modified this equation according to

$$DT(x) = \min_{u\in row} G_u(x-u)\,. \qquad (13)$$

This allows the usage of the energy function G_u shown in Fig. 7 that has a maximum energy level E_0 and a minimum $E_{\min}$ depending on the edge gradient at position zero. The width of the function is kept constant in order to get a fixed effective range of any edge, independent of its amplitude. This distance transform is computed using a simple two-pass algorithm. Since the slope of road boundaries is usually far away from horizontal, it is sufficient in the considered application to compute the DT image one-dimensionally line by line. Using (3) the obtained total energy can be reconverted to an applicable probability.

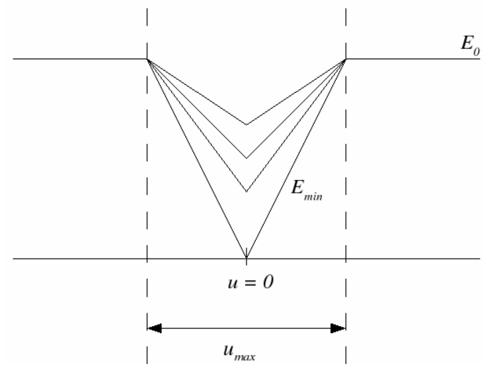

Fig. 7: Energy function G_u used for the generalized distance transform. The energy minimum depends on the gradient magnitude.

This form fulfills the requirements one and two. In order to meet the requirement three, we store the phase of the edge that causes the minimum energy at each image point. This can easily be done while computing the generalized distance transform. The probability p_φ of an image point to belong to the boundary of a hypothetic road depends on the angular difference $\Delta\varphi$ between the expected and assigned angle. An appropriate formulation is

$$p_\varphi(x, y) = \cos^{2n}(\Delta\varphi(x, y))\,. \qquad (14)$$

Good results have been obtained for n=5, while n=1 leads to a weak rating. In order to avoid costly computations, the angular difference is quantized and the probability is stored in a lookup-table.

Fig. 8: Gray value encoded distance transforms. Left: standard DT, black encodes distance zero; white encodes distance greater 20 pixel. Right: generalized DT gained by (13) with $u_{\max} = 40$.

V. Results

In this section we show results of a PF based road recognition system using the described features. During initialization, the particles are equally distributed within the six-dimensional parameter space. Tests with a reduced parameter space (e.g. fixed tilt angle, no clothoid parameter) were less satisfying in practice. Yaw rate and speed are taken from inertial sensors. The pixel feature statistics are taken from small windows in front of the car and at the image borders.

During convergence of the PF the sizes of the statistic windows are increased up to the final size shown in the result images. The final decision for the best parameter vector is drawn from the 40% best weighted particles. A prior for a certain road width can be added.
The PF usually converges within 20-40 frames. The distribution of the particles is analyzed in order to verify convergence. In order to achieve real-time performance, the number of particles is restricted to 500-1000.

A. Well Marked Roads

As mentioned in the introduction, compatibility with the common marking based lane recognition is a necessity. If a prior on the expected lane width is added, practical results on marked winding roads are promising. Fig. 9 shows such a road and the overlaid result of the PF. If any lane width up to 10 m is of equal probability, the PF will usually generate two possible solutions: the "correct" one and an alternative with double width. The sequence has a length of 20 sec and the road parameters are precisely estimated throughout the whole sequence.

Fig. 9: Recognition result for a marked road using the features intensity, edge gradient and edge direction.

B. Badly Marked Roads

The algorithm has been designed to allow road recognition on roads with low marking quality. Fig. 10 gives an example. The road is tracked throughout the sequence although a car is approaching on the opposite lane. Nevertheless, it would be beneficial if the occluded image region could be excluded from the evaluation by separate detection schemes.

C. Unmarked Roads

Fig. 11a and Fig. 11b give two examples for unmarked paved roads. In the first image the left border is clearly visible in the edge image, whereas the right border is mainly defined by texture differences. Fortunately, the probability estimation scheme tends to concentrate on features that give clear hints. It should be pointed out that the intensity of the road is far from a Gaussian distribution. The shown road is robustly tracked throughout the whole sequence.

Fig. 10: 4 frames of a winter country road with the overlaid estimated road course. Used features are intensity, edge gradient, and edge direction.

The second image exhibits low contrast and only the combination of all cues gives satisfying results. Fig. 12 shows plots for lateral offset, yaw angle and curvature. The road starts with a slight curve to the left, followed by a right curve. In the beginning, the driver turned the steering wheel quickly in order to prove the correct estimation of the yaw angle. The results are comparable to standard Kalman Filter results obtained for marked roads.

Fig. 11a and 11b: Road recognition result on unmarked roads. Used features are intensity, edge gradient, edge direction and texture.

D. CPU Time

The parameter estimation consists of three main steps:

1. preparation,
2. scoring of the hypotheses,
3. analysis of the scored particles and prediction.

While the computational load of step 2 grows linearly with the number of particles, step one has to be carried out only once. Therefore, our strategy was to move the workload to the first step and to make step two as fast as possible.
According to the features that should be used, step one includes the computation of

- gray value and hue statistics and integral images,
- structure tensor including eigenvalue calculation, histogram calculation and integral images,
- edge magnitude and phase plus distance transform.

Using Intel's IPP-support [12], these low-level operations are carried out very efficiently on a Pentium IV. As described above, the scoring according to (5), (10) and (14) is done efficiently by using lookup-tables.

It is evident that the road boundaries on unmarked roads are not as clearly defined as on marked roads. Therefore it is sufficient to work on quarter VGA images. The measured CPU times (Pentium IV 3.2 GHz) for preparation and scoring are given in Tab. I. Note that the scoring times depend linearly on the number of particles used.

TABLE I: COMPUTATION TIME

	Preparation	1000 particles
Gray value	0.5 ms	7.5 ms
Edge features	8.6 ms	9.7 ms
Texture	5.7 ms	7.5 ms

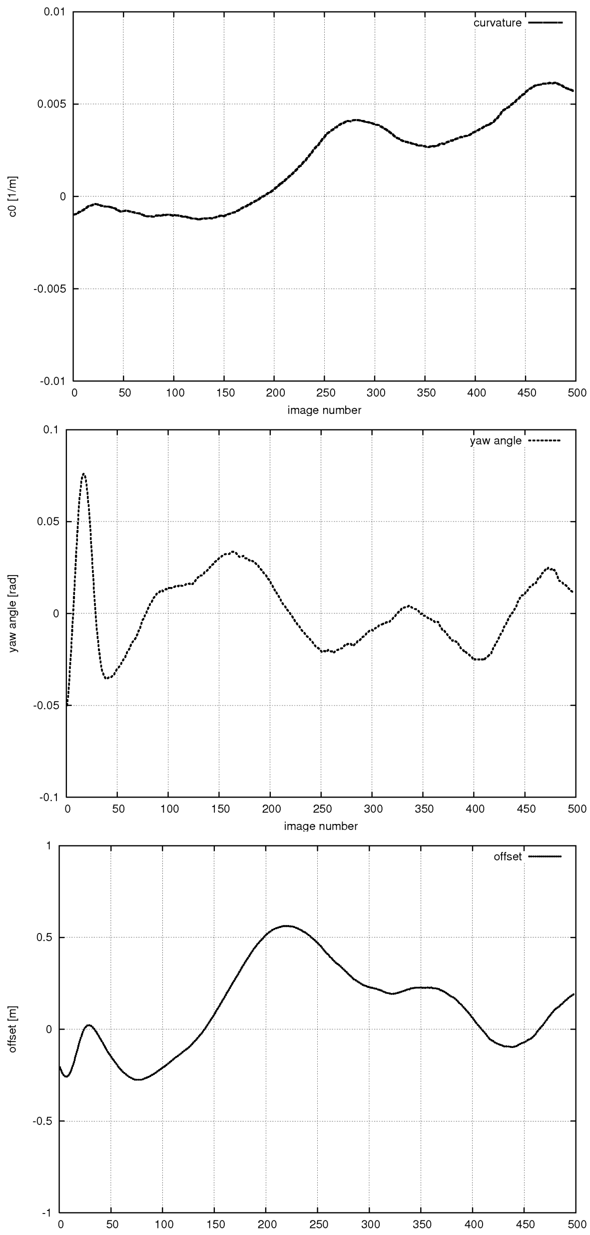

Fig. 12: Estimated offset, yaw angle and curvature for the road shown in Fig. 11b.

VI. CONCLUSIONS

Road recognition on paved country roads has been successfully formulated as a MAP estimation task. Gray value statistics in combination with edge gradient and edge direction lead to robust road recognition. Texture and color can additionally constrain the solution; markings if available are taken into account. A generalized distance transform is used that – in contrast to the classical transform – is free of any threshold.

The computational load of the used PF approach is significantly reduced by means of integral images. On a 3.2 GHz Pentium IV up to 1000 particles (road hypotheses) are evaluated in video real-time, if the estimation is based on gray value statistics and edge information (gradient and direction).

The paper describes a framework that can be extended straightforward. The use of map information e.g. local curvature is obvious. Future work will include the exploitation of stereo information.

REFERENCES

[1] E. D. Dickmanns, A. Zapp, *A Curvature-Based Scheme for Improving Road Vehicle Guidance by Computer Vision*. SPIE, vol. 727, 1986, pp. 161-168.

[2] Dean Pomerleau, “RALPH: Rapidly Adapting Lateral Position Handler”. *IEEE Symposium on Intelligent Vehicles*, 1995, pp. 506–511.

[3] J. Zhang, H.-H. Nagel, “Texture-Based Segmentation of Road Images”, *IEEE Symposium on Intelligent Vehicles*, 1994, pp. 260-265.

[4] J. MacCormick, M. Isard. “Partitioned Sampling, Articulated Objects, and Interface-Quality Hand Tracking”. In ECCV (2), 2000, pp. 3–19.

[5] B. Southall, C.J. Taylor. “Stochastic Road Shape Estimation”. In ICCV01, 2001, pp. 205-212.

[6] N. E. Apostoloff, A. Zelinsky, “Robust vision based lane tracking using multiple cues and particle filtering”. *IEEE Intelligent Vehicles Symposium,* 2003, pp. 558-563.

[7] P. Smuda, R. Schweiger, H. Neumann, W. Ritter, “Multiple Cue Data Fusion with Particle Filters for Road Course Detection in Vision Systems” *IEEE Intelligent Vehicles Symposium*, 2006, pp. 400-405.

[8] U. Meis, R. Schneider, “Radar image acquisition and interpretation for automotive applications”, *IEEE Intelligent Vehicles Symposium,* 2003, pp. 328-332.

[9] B. Hummel, S. Kammel, T. Dang, C. Duchow, C. Stiller. “Vision-based path-planning in unstructured environments”. *IEEE Intelligent Vehicle Symposium*, 2006, pp. 176-181.

[10] G. Borgefors, “Hierarchical chamfer matching: A parametric edge matching algorithm”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(6): pp.849-865, November 1988.

[11] P. F. Felzenszwalb, D. P. Huttenlocher, “Distance Transforms of Sampled Functions”. Technical Report Cornell Computing and Information Science TR2004-1963. Available: www.people.cs.uchicago.edu/~pff/papers/dt.pdf

[12] www.intel.com/cd/software/products/asmo-na/eng/302910.htm