Auther identification information:

Peter Højholt
National Survey & Cadastre
Rentemestervej 8-10
DK-2400 København N
Denmark
Tel: (+45) 35 87 53 62
Fax:(+45) 35 87 50 51
E-mail phs@kms.dk

Peter Højholt holds a Ph.D. in Mechanical Engineering and serve as head of the Topographic Section of the Informatics Department in the National Survey & Cadastre, Denmark.

From 1:10.000 to 1:50.000 at 300 km2 pr
Peter Højholt

The danish vector map registration standard is 1: 10.000. Next summer the whole country will have been measured and the updating proces is already underway.

Once the 1:10.000 map coveres all of Denmark the next question will be which other vector maps we should have. 1: 50.000 and 1: 100.000 will propably be the next maps.
This paper discuss the possiblilities of generalization of maps from 1:10.000 to 1: 50.000 at very high automation rates.

Although Denmark has had scale 1:50.000 paper maps for a long time, there is no correspon-ding vectormap, and the proces described below is therefore the "first time" proces of making a 1:50.000 vector map extraced from and consistent with the existing 1:10.000 map. No extra collection of topographic data is a allowed, but the use of external electronic data sources are allowed (as described below). The old papermaps have been computer scanned and can therefore be used in head-up displays to assist manual operations. Pattern recognisition techniques and other electronic data retrival processes from the scanned material which could possibly assist the automation proces are explicitly excluded in this paper, but will be considered before the start of the production.

Most of the elements in the proces have already been tested seperately but now needs to be put together.

Nature, registration practice and cartography in Denmark

The task of generalizing a map from 1:10.000 to 1: 50.000 will be different depending on details in the physical area being mapped, the registration done at 1:10.000 and the specific cartography chosen in scale 1: 50.000.

The physical area to be mapped.
Land in Denmark is highly cultured and among the objects mapped most objects and boun-daries between objects are man made or regulated. Very few map objects represent things that have not been created or regulated by man. One major exception here is the coastline which in most places still has its "natural" shape. The danish landscape is a gentle landscape so even where objects remain in their natural shape boundaries as presented in scale 1:10.000 will not be extremely irregular. Even where nature has influenced man made objects as in the case of road-centerlines the influence is small.

Man made objects have been made with a level of detail that suited mans purpose and mans need for regularity. Except for buildings which can enclose details right down to the basic registration accuracy of 1 meter most object have already been "generalized" at their very creation. This is for example the case for boundaries of build up areas and forests, hedges and ditches.

Very few road centerlines have curves or irregularites below a scale of say 20-30 meters. Road sidelines on the other hand include many details below this accuracy and therefore needs some geometric generalization.

Registration at 1:10.000
The basic registration i scale 1:10.000 was designed primarily to support GIS use of the map by major organisations as municipalities and network operators.

This shows as a careful registration of land use polygons in the map. The land use polygons are divided into low-rise areas, high-rise areas, industrial areas, recreational areas, city center, forest, moose etc. Thes data are organized in a way that makes it easy to point out a specific area or objects inside or next to an area. For these land use areas a typical minimum size is defined to be 2500 m2 and a typical minimum width is defined to 10 meters and interruptions less than 10 meters in width are not registrated.

Roads are not considered areas, but are registrated with centerlines. Where there is no areal polygon registration along the road, road sidelines are registrated.

Buildings are allowed to be either inside or outside these areas, but a building is not allowed to be both inside and outside an area.

An example of the 1:10.000 map is shown on figure 1 (in a smaller scale).

The cartographic expression in 1:50.000
The cartography of the danish 1:50.000 map is not to be described here but figure 2 shows an example which also covers the area presented in figure 1.

Planning the generalization proces

It was from the start considered impossible to achieve an acceptable data quality if only automatic procedures were allowed to be used, and we are therefore planning a proces of both manual and automatic procedures. The results up till now have indicated that a higher degree of automation than anticipated will be possible. However there will still exist some processes that will need manual corrections.

It is important that manual corrections to the automatic processes are introduced before the errors made by the automatic process can corrupt subsequent automatic processes. This would indicate that manual processes should be inserted directly after each automatic porcesss that have a low rate of succes and/or which can have a major influence on the subsequent processes.

On the other hand past experience at our institute strongly suggests that shifts between manual and automatic processes are troublesome and errorprone. If needed they should be kept within the same computer program, using the same computer operating system, and made on the same computer. Processing and saving data in one computer program and reading and correcting data in another computer/computer program requires data handling and increase the load of process knowledge placed on the operator.

Required manual interceptions in the automatic process should therefore not only be kept at a minimum but also joint in only one or two intermediate steps in the process so that the amount of shifts between manual and automatic operations are minimized.

Any process that requires manual input should be kept non-iterative. Manual corrections usually require that much data are visually scanned by the operator, and iteration processes therefore requires time consuming repetition of the scans and is in general exhausting for the operator.

The process sequence
All processes below are given in process oreder and are automatic processes except where a manual process is explicitly indicated.

Identification of farms
In the 1: 10.000 map the registration of "low-rise" build-up areas are made both in residential areas in villages and cities and around farms in the countryside. The cartography in the 1: 50.000 map is however different around farms and in residential areas in cities/villages. Around a farm the "low-rise" polygons are dropped and instead the farm buildings are shown highly generalized and possibly enlarged. In the urban areas the "low-rise" polygons kekpt and only larger buildings are shown inside these polygons.

For this reason it is necessary to identify "low-rise" build up areas around farms as opposed to "low-rise" areas elsewhere.

The identification is made from a very simple counting of buildings. The buildings are counted in the minimum rectangle surrounding the "low-rise" polygon. If there is less than 3 buildings or if there is less than 7 and at least one building above 300 m2 a "low-rise" polygon is considered to belong to a farm.

In general this criteria was found to have a high rate of succes, but with a lower rate of succes in cases where farms are so close or villages are so small that it is difficult to see (also for me) which representtion the cartographer would choose.

As the identification of farms have significant consequences on the elimination of buildings and roads, it is considered necessary to correct the result of the identification process before these other processes are started.

Manual classification of roads and check of farm identification
In the 1:10.000 map the road are classified as "highway", "above 6 meter", "3-6 meter", and "other roads". The 1:50.000 map legend describe the same road classes, but an overlay of the two maps shows significant differences in the actual choise of road category. It is considered unlikely that the 1:50.000 cartographers will accept the road classification in the 1:10.000 map. On the other hand it will be extremely difficult for an automatic process to mimic the 1:50.000 classification based on the 1:10.000 classification, so for the moment this reclassification is left to a manual process.

Seen in just a slightly longer perspective all road segments in our maps will be coupled to the databases of the road administrative authorities in Denmark. The road admnistrative authorities have a large amount of data about every road in Denmark (and also their own ideas about which way people should drive). In the future it seems likely that this database can be the basis of a 100 % automatic road clasification in every map scale.

Identification of road sidelines
As mentioned above road sidelines area not explicitly present where roads run next to an areal feature. If the areal feature is eliminated (for example around farms) the road sideline will be missing.

To avoid this all polygon lines that runs along a road is identified and copied to a "road sideline" object. Based on past experiences [3] from similar data a hit rate of 99.8 % or better is expected in this process.

Selection of roads
Almost all roads present in the 1:10.000 map will also be present in the 1:50.000 map. The roads to be eliminated are blind roads below appr. 30 meters and roads below appr. 100 m leading only to farm courtyards (notice here the dependence on the farm identification process).

Although in principel simple the elimination can be tricky an the expected rate of succes counted as a percent of the number of roads to be eliminated is expected to be only around 80 %.

Selection of other linear features
The selcetion of other linear features are very simple because it can be made considering only the feature code. Either the feature is in the 1:50.000 map or it is not. Almost no thinning in linear features are necessary.

An exception is the feature "path" which is eliminated inside all build-up areas and kept elsewhere. Short pieces of "path" between build up areas or between road centerlines and build up areas left over in this process should also be eliminated.

Selection of buildings with special use
In general small buildings inside build-up areas are eliminated (see below), but buildings of special significance (churches, museums, inns etc) should be kept. The special buildings are identified through a coupling to our place-names/place-function database. It is expected that this identification will be 100 % correct and need no thinning.

General selection of buildings
Many buildings are two small to be represented in the 1:50.000 scale map. The area-criteria below which a building should be eliminated depends on whether the building is inside "low-rise","high-rise", "industri" etc. Buildings inside "city center" polygons are all eliminated regardless of size. Some farm buildings below the area criteria should be kept so that they can be included in the generalisation of the farm. This is not necessary elsewhere in the map.

As shown in figure 3 the building selection/elimination is rather succesfull even at this initial stage.

Agglomeration of areal features
Most areal features are either kept or eleminated considering only their size (see below). However in a few cases object from classes "nature" can be small, but lying next to another small object belonging to the same feature class. In this case the two objects should be combined and kept. This is especially likely to happen in the feature "lake" because "lake" are

registrated below the general 2500 m2 minimum area right down to a 100 m2 size.

A special search procedure should easily be able to identify such cases. Whether it is worth it to establish an automatic procedure to do the actual agglomeration has not been decided yet. Because of the extremely low number of cases and the easy identification process (which will keep them out of the subsequent elimination process) it might not be worth the trouble.

Agglomeration of areal features might also be relevant as part of a reclassification scheme. Very small "build-up" areas that lie next to another "build-up" area belong to a different sub-class might be reclassified and added to the neighbour. If necessary this process is expected to be fully automatic.

Selection of area features
After the agglomeration of area feature the final selection is based entirely on size criterias. An expection is the "silo/tank" feature that are eliminted in the countriside but kept inside technical areas.

Generalization of buildings
Up till now objects have been selected or deselected, but has not been geometrically generali-zed. The remaining processes are all geometric generalisations.

Buildings can enclose details right down to the basic registration accuracy of 1 meter and as the number of buildings still present in scale 1:50.000 is very significant there is a need for an automatic geralization tool for buildings.

The building generalization although different in detail is done much along the tradition of the Hannover-university, which as is well known has a long tradition of producing very capable generalizations of buildings.

At the institute we have a tool that was used to generalize buildings from scale 1:4.000 to scale 1:10.000. By only adjusting the parameters in this tool acceptable generalizations were made, see figure 4. The remaining number of errors is considered too small to justify added funcionality in the tool.

The generalization tool generalizes the buildings, and at the same time keeps the topological relations between the building and its surroundings. If the new building corners extent over other boundaries these boundaries will be pushed aside. If the old corner was connected to a hedge so will the new corner be. If the old building shared edges with other object the generalization routine will consider whether the generalized building should include shared edged with the same objects.

This generelization tool does not include agglomeration of nearby buildings. For buildings left inside "build-up"-areas this by and large seems not to be necessary.

If the tool runs into topological problems which it cannot solve, it will leave the building untouched, but mark the building as "not-generalized". Under no circumstances will the tool allow generalized buildings or building details to violate the basic topological constraints in the original map.

In the countryside generalization of farms might benefit from an agglomeration of nearby buildings, but the danish tradition dictates a generalization of farms that is cross between a generalization and a symbolization and which originated in the tools used for hand-drawing the generalized buildings.

It will be extremly difficult to mimic the traditional procedure in the computer, and as the number of farms are high a negotiation with the cartographers concerning a new cartography for farms will be necessary. The suggested cartography is the generalisation using the tool mentioned above (perhaps adding agglomeration in some form) followed by an enlargement of the generalized buildings. As seen from my point of view this should give an acceptable cartographic expression, and perhaps even give an easier building identification for the end users of the map. The enlargement of the buildings will be done as part of the finite element procedure used to generalize the road sidelines (see below). The enlargment of buildings in a finite element procedure has been tested and reported in [2].

Other geometric generalisations
The need for generalization of other features than buildings and road sidelines are extremely small. If such automatic procedures are developed they should be placed here in the process, but at the moment such procedures are not planned, and the remaining generalizations are left entirely to the manual process at the end.

Displacement of road sidelines and buildings
As the final automatic process road sidelines should be displaced to a uniform distance (depending on road classification) from the road centerline.

The displacement process should keep certain topological features:
Area or line features that share edges with the road sidelines should be displaced simulta-niously, so that the common edges are conserved as common edges..

Line features (like hedges and ditches) that share end points with a road sideline should be shortened or prolonged so that the common points are kept.

Buildings and other small objects which geometry is so close to the road centerline that the new road sideline will overlap the building or object after generalization should be moved and kept out of the road area.

In [1] and [2] a finite element method which can make these displacement and at the same time keep the topological relations to other objects is described. The results indicates that the method can be succesful. However the results in the paper also show that the road sidelines after the process is not perfect.

Most of the imperfections in the generalized road sidelines originate from an insignificant number of triangular points in the road sidelines and a lack of a specialized procedure to take road junctions into consideration. Both these problems will be mended in future versions.

A few of the mistakes stems from buildings having comming edges with the road sidelines. In general it is not acceptable that buildings are rotated just to keep this topological restraint to the road sideline, and if a building has more that one edge in common with the road sideline the road sideline criteria cannot be fulfilled without deforming the building (which is unacceptable). In the papers [1] and [2] this phenomenon is seen in a couple of buildings in the test area. To solve these problems additional null-triangles between the building and the road sideline must be introduced to give the necessary degrees of freedom.

Manual check and improvement of the total result including generalisation and selection
The final result should be manually checked and corrected. Special computer tools to assist this process will be developed, but not included in this paper.

Calculation of work load

Let's face it:
Calculation of work load and judgment of map quality are extremely political processes.

Even my inclusion of rough estimates in this working paper can give me nothing but trouble, but here they are anyway.

Out of a 7½ hour working day ½ hour is used for lunch and 1 hour will be spent on handling data and starting and waiting for automatic processes. This leaves 6 hours to do manual corrections in the map. The 1:10.000 scale map production works in block of 10x10 km. If the goal is to do 300 km2 in one day (roughly corresponding to a traditional map sheet in two days) this means one block of 10x10 km in two hours.

The amount of work that an operator has to do to make one correction vary both in the time the operator needs to identify the error and the time the operator needs to figure out how to correct it and correct it. Selection and deselections can be done very fast, but regeneralization of a building might take more time.

At this stage it can only be an arbitrary guess, but lets assume that the operator can make 100 corrections in one hour. Having to do 100 km2 in two hours this would dictate that the automatic processes can leave only 200 error in a block. About 6 % of the area in Denmark is build-up where a higher rate of errors pr. km2 are expected, the error rate in the countryside should be correspondingly lower than 200 errors pr. km2.

In the table below expected number of errors in each process is indicated. At the moment these numbers are only guiding numbers used to indicate where in the process more attention/automation should be set in. They have not yet been discussed with neither cartographers or marketing people, who ultimately decides on matters concerning data quality.

Process to be made or mended	**Number of errors to be corrected pr. 100 km2**
Identification of farms (select /deselect process)	40
Manual classification of roads (select /deselect process) (old map as background)	30
Identificaton of road sidelines	10
selection of roads	20
selection of other linear features	20
selection of buildings with special use	0
general selection of buildings	35
agglomeration of areal features	10
Selection of areal features	10
Generalization of buildings (assuming that the problem of generalisation of farms must and will be solved)	10
Other geometric generalization (manual process)	5
Displacement of road sidelines and buildings	10

References:

This is not a reference list, but the suggested procedures draws directly on the programming also presented in the following references:

[1] Højholt, P.: Solving Local and Global Space Conflicts i Map Generalisation Using a Finite Element Method Adapted from Structural Mechanics. SDH98. p 679

[2] Højholt, P.: Solving Space Conflicts in Map Generalisation Using a Finite Element Method. To appear in CaGis.

[3] Højholt, P. and L.G.Jensen: Automatic identification of build-up areas including a comparison between traditional it-then programming and neural-network programming.Euro-Carta XIII, 1995, p.163

Figure 1. TOP10DK 1:10.000 vector map. Here presented in a smaller scale

Figure 2. Our present 1:50.000 paper map.

Figure 3. Automaticly derived 1:50.000 no manual porcesses are involved. Identification and generalization of road sidelines not included.

Figure 4. Example of house generalization from 1:10.000 to 1: 50.000 made by our existing tool for automatic generalization.

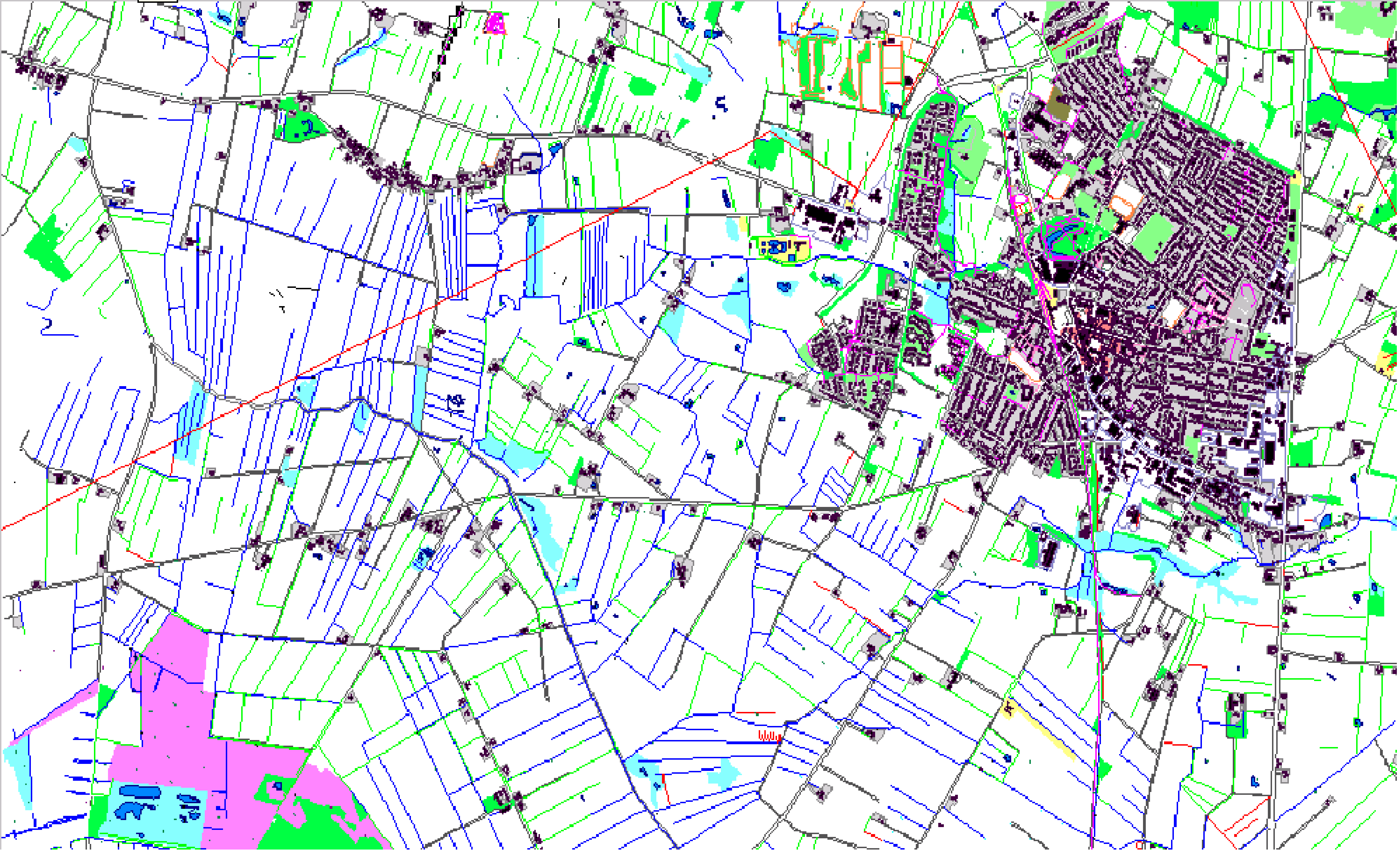

Figure 1. TOP10DK 1:10.000 vector map. Here presented in a smaller scale

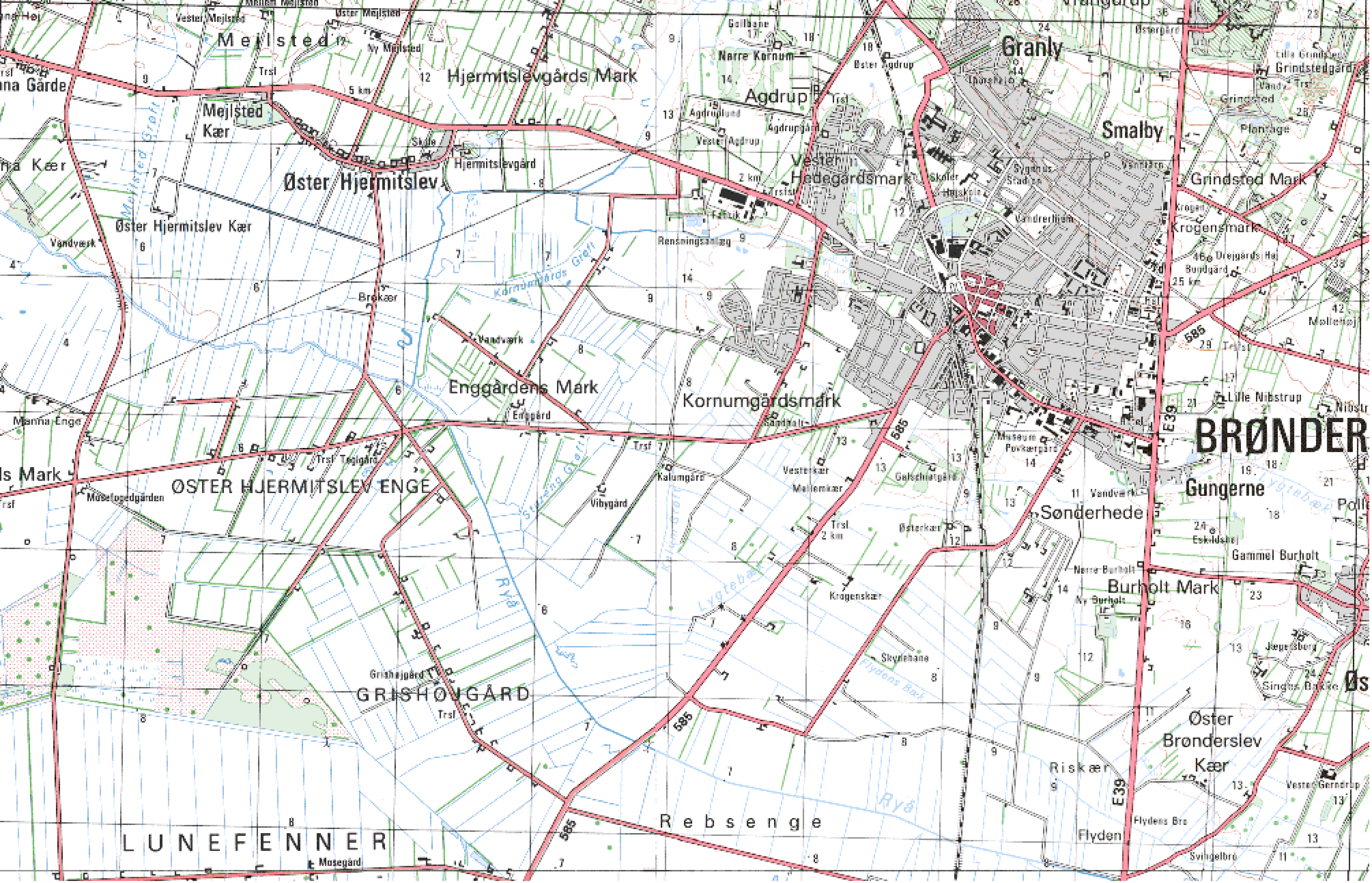

Figure 2. Our present 1:50.000 paper map.

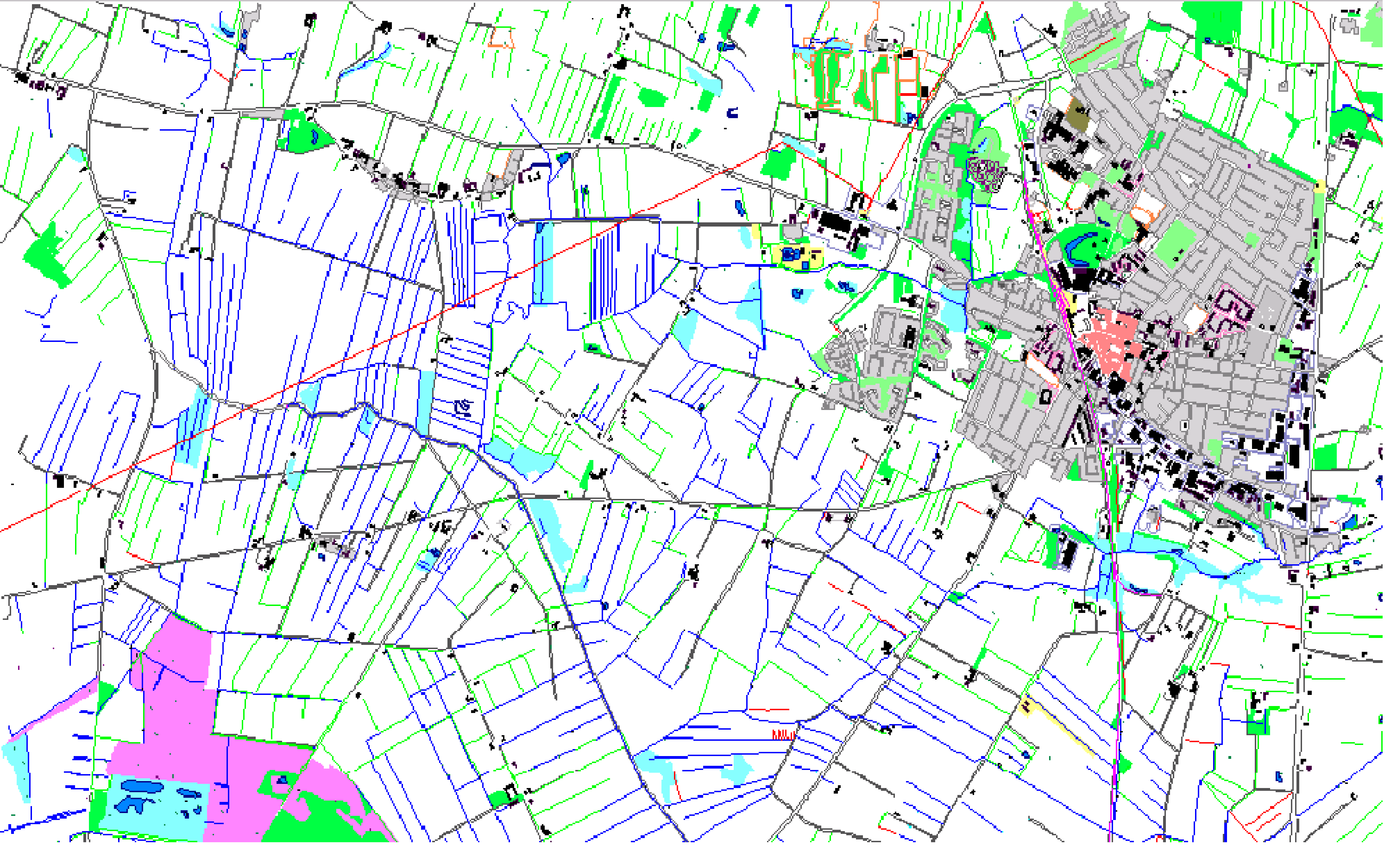

Figure 3. Automaticly derived 1:50.000 no manual porcesses are involved. Identification and generalization of road sidelines not included.

Figure 4. Example of house generalization from 1:10.000 to 1: 50.000 made by our existing tool for automatic generalization.