Dating and authentication of Rembrandt's etchings with the help of computational intelligence

Jan C.A. van der Lubbe (*), *Eugene P. van Someren* (*), *Marcel J.T. Reinders* (*)

(*) Delft University of Technology, Faculty of Information Technology and Systems
Department of Electrical Engineering, Information and Communication Theory Group
P.O. Box 5031, 2600 GA Delft, The Netherlands
E-mail: J.C.A.vanderLubbe@ITS.TUDelft.NL

ABSTRACT
In this paper it is studied how the analysis of the structure of the paper of Rembrandt's etchings can help determining the authenticity of the etchings and the dates of the prints. Digitised X-ray images of the etchings make the structure of the paper visible and reveal landmarks that originate from its production process. Methods for automatic extraction of these features are developed. On the basis of these features, etchings printed on similar papers can be recognized by means of automatic matching procedures. With this questions concerning authenticity and dating can be answered. Since the aim is to construct a database of all the papers used by Dutch artists, specific attention is paid to how to make this database self-organizing and self-learning. This requires the application of advanced techniques from computational intelligence.

KEYWORDS: Rembrandt's etchings, dating and authenticity, computational intelligence, self-organizing and self-learning databases

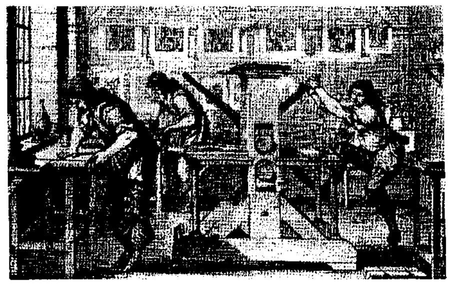

Figure 1: The etching process (by Abraham Bosse, 1642)

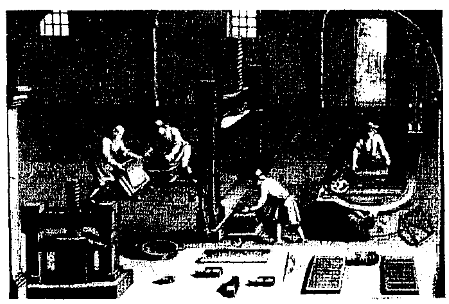

Figure 2: Paper manufacturing (from P.J. Kasteleyn, "De Papiermaaker', 1792)

INTRODUCTION

In art historical research of etchings, dating and authenticity play an important role. A typical question is for instance: "That etching-plate is made by Rembrandt in 1648, but is the corresponding print also from 1648 or is it a later print?" The structure of the paper the print was made on, provides essential information that may lead to the answer of this kind of questions. In the 16th and 17th century paper was produced in paper-mills by means of paper sieves, consisting of a grid of thin copper wires. The pattern of the grid is visible in the paper like a kind of watermark on a banknote. Each sieve produces a characteristic pattern. Now, if there are two etchings, one dated and the other not, both made by the same artist and printed on sheets with identical pattern, then one can conclude that both originate from the same sieve and therefore are probably made in the same period.

X-ray images are employed to make the paper structure visible. Image processing tools are developed for the automatic extraction of the characteristic pattern from those images. Since image processing techniques in themselves are not sufficient, knowledge-based techniques are used in order to control the process of the detection of the so-called chain and laid lines, together forming the pattern. Reasoning processes are used in order to compare sheets of paper and finally to conclude whether they are identical or not and to interpret the results.

At the end the aim is the generation of a database consisting of all 16th and 17th century papers used by Dutch artists (thus not only Rembrandt) and consisting of additional information like date, place, expert information etc. The great challenge here is the development of tools by means of which all kinds of queries can be answered. This implies an adequate reasoning mechanism for combining both image information and textual meta-information. Moreover, the database should be self-organizing and incremental, i.e. new information that becomes available from answering queries should be added to the database. However, the main problem is that new information can bring on an avalanche effect in the database. If one etching is assigned a new date this can effect the dates of the other etchings, influence the validity of theories of art historians, and so on. Reasoning techniques are used to control the updating process and especially for guaranteeing the internal consistency within the database.

PAPER MANUFACTURING AND THE ETCHING PROCESS

With respect to the dating of etchings distinction should be made between 1) the date of the plate and 2) that of its prints. An etching plate can be made by Rembrandt in 1648, but the prints of it can be made in later years. In the case of Rembrandt a number of etching plates have been preserved and also in the 19th century prints are made from it. Of course these prints are considered as less authentic than the prints made by Rembrandt himself. Sometimes the art historian is interested in the date of the etching plate when this is not on the plate, sometimes the date of the etching plate is known and one wants to know the date of the print.

During the last years art historians show an increased interest in the paper structure to support their research. In 1985, serious research on the paper used by Rembrandt was started in Holland by the Rembrandt-expert Theo Laurentius

[1,2]. Since research on paintings had advanced enormously through scientific research of the wooden panel or canvas, he stated that study of the paper would yield equally valuable information about etchings. During the 16^{th} and 17^{th} century paper was made in paper-mills. To make paper, linen rag, rope, sailcloth etc. was pounded, breaking them down in small fibers. By adding water pulp was obtained which was put in large tubs (see figure 2). By means of sieves some pulp was dipped out. The sieve stops the fibers, but let the water through. After its content had been dried up, a paper sheet was the result. It is especially the sieve which determines the paper characteristics in which we are interested (figure 3). It consists of a grid of brass-wires, which leaves an impression in the sheet that can easily be seen when you hold it to the light. The reason for this is that just above a brass-wire of the sieve the pulp layer is smaller, which makes the sheet there become slightly thinner (figure 4). The horizontal lines are called the laid lines and are typically spaced every millimeter. The vertical lines are the chain lines, they are about 0.5 mm in diameter and spaced circa 15 mm. Especially the chain lines are important. All the sheets made in a particular sieve will have the same pattern of lines, and since the sieves were themselves all made by hand, no two are exactly the same. More clearly, by studying the pattern of lines in the paper it is possible to identify and distinguish different sieves. The chain and wire lines are in fact the 'fingerprint of the mould'. If one has two sheets with identical chain lines they must have been made in the same sieve.

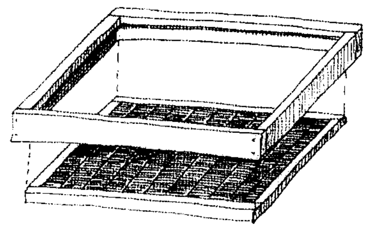

Figure 3: Sieve

DETECTION OF CHAIN LINES

Since the ultimate aim is to construct a large database of all papers used in the Netherlands, automatic extraction of patterns of chain lines is important. In order to make these visible, X-ray images of the etchings are made. By means of this process the paper structure becomes visible, but not the etching itself. X-ray images will show the vertical chain lines, the horizontal laid lines, sometimes a watermark and a number of black or white spots related to remaining pieces of pulp, impenetrable for X-rays. In order to enable automatic processing these X-rays are digitized. Experiments were performed with various X-ray techniques and with various combinations of voltage and exposure time. Nevertheless, also in the optimal case the quality of the images is still poor and the contrast is low. For that very reason advanced image processing techniques are necessary in order to detect the chain lines automatically.

Figure 4: Cross-section of sheet

The identification process consists of two basic steps. First, the chain lines are enhanced in the image. Then the horizontal positions of the chain lines

are determined.
To enhance white vertical lines (i.e. the chain lines) the following image processing steps are done (see figure 5 and 7).

Figure 5: Chain line enhancement

- **Filtering (X^f):** To reduce smudges and fragments of laid lines, the original image (X) is first uniformly filtered. The filtering element is rectangular rather than square, i.e. it is smaller in the horizontal than in the vertical direction. This is to avoid filtering out the chain lines themselves.
- **Shading Estimation (X^s):** An estimation of the shading can be obtained from the filtered image (X^f) by a morphological opening step, followed by a morphological closing step.
- **Shading Correction (X^e):** Shading correction on the filtered image is simply done by subtraction of the shading estimation (X^s) from the filtered image (X^f), resulting in an image with enhanced chain lines.

The resulting image is now optimized to allow the chain lines to be detected. The procedure to find the actual line positions consists of the following steps (see figure 6 and 7):

- **Vertical Projection (S):** Using the fact that chain lines are vertical, a vertical projection can be performed, by summing the pixel values over the columns, resulting in a one-dimensional representation. This line (S) contains all information of the image about the whereabouts of the chain lines.
- **De-trend (S^{dt}):** To determine the trend of the line a min-max filter is employed. This trend (S^t) is subtracted from the line (S) to remove the trend and to obtain the local maxima.
- **RelativeThreshold (S^{thr}):** Potential chain line locations are found by thresholding the de-trended line (S^{dt}), resulting in a block signal that represents likely chain lines positions. The used threshold is set at 65% of the maximal value of the line.
- **Calculate Positions (S^{pos}):** The actual chain line positions are estimated to be located at the centers of the intervals containing consecutive potential chain line positions.

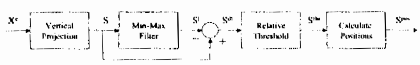

Figure 6: Determination of line positions

The above procedure to find the chain lines is based on the assumption that all chain lines are straight and exactly vertical with respect to the image. Although these assumptions are quite strong, many X-ray images have been processed without problems. However, as the collection of X-ray images will extend, patterns of chain lines are to be expected that cannot be found by the simple approach described above. As a matter of fact we believe that the curves in the chain lines will prove to be excellent features to distinguish between two apparently similar patterns. For that reason we have also investigated other approaches to find chain lines in X-ray images. A promising technique is known as the minimum cost method [3]. This technique employs dynamic programming to calculate a path from top to bottom with the minimum cumulative costs. A polynomial fit through the calculated points can be employed to obtain a compact set of features for comparing different patterns.

THE MATCHING PROCESS

The reason for the chain line extraction is to be able to distinguish patterns that come from one sieve from patterns that come from other sieves. To this end, a set of features must be assigned to each image using the found chain lines. For the current set of X-ray images it suffices to use for each image the average line distance as a single unique feature and to classify images by means of a nearest neighbor classifier. By applying a threshold to the distance, a new sample can be accurately classified, or determined to come from a sieve not earlier observed. When using more complex chain line detection techniques (like in [3]) also more complex classifiers are necessary. In that case the search for the right set of features will be challenging.

Thus it is possible on basis of the paper structure to conclude whether two papers originate from the same sieve. That we can make conclusions with respect to dating and authenticity, only on the basis of similarity of paper structure, rests on some art historical assumptions. If two sheets have the same pattern of chain lines, they are made in the same mould. If these two sheets are also used by the same artist, it is highly probable that they come from the same stack of paper. The wholesaler usually had a limited stock of paper from any one paper-mill – perhaps a few reams (of 480 sheets). Since there were a large number of mills supplying the whole, his stock would contain paper from a several hundred different moulds. And since the retailer in general only bought a small quantity from the wholesaler (Rembrandt bought probably no more than one book of 24 sheets together), the likelihood that two identical sheets come from different stacks is very small. Furthermore, it may be assumed that sheets with the same structure and printed in the same studio will have been used at the same time. Rembrandt for instance always printed a series after each other. So, if one has a print which is dated and another with the same paper structure, it may be concluded that the other print is also from that date, i.e. if the artist is the same in both cases. If one has two prints of different artists this conclusion can not be made, as it is very improbable that two papers from the same sieve and used by different artists are bought and used at the same time.

When using the paper patterns as a means for dating etchings, especially the early states of etchings are important, even more so when they have been dated on the plate. It is reasonable to assume that the first prints are made in the same year as mentioned on the plate. Papers used for the early states can be dated to the year stated in the etchings. But this also holds for other papers with the same structure, since they almost surely come from the same stack.

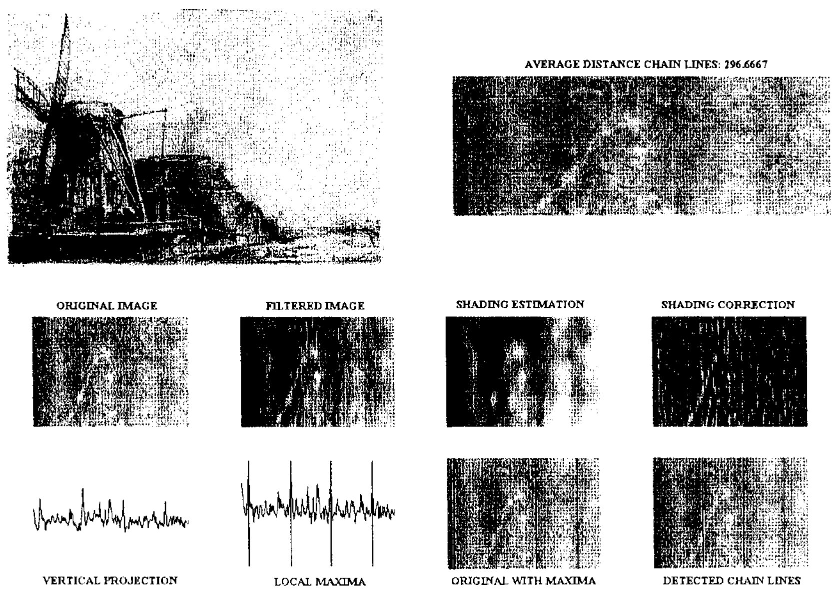

Figure 7: The wind mill by Rembrandt, 1641: detection of chain lines

SELF-ORGANIZING AND SELF-LEARNING DATABASE

The ultimate goal is the generation of a large database of all 16^{th} and 17^{th} century Dutch etchings, the corresponding X-rays of the paper with detected chain lines and meta-information. The meta-information includes: name of the artist, description of the picture, origin of the etching, date of the etching plate, date of the print, origin of the paper, features of the paper structure, description of watermarks if present, related etchings printed on the same papers, art historical comments etcetera. This database has two aims:

1) It can function as a knowledge base on the basis of which (new) etchings or papers can be identified.
2) It is self-organizing and self-learning in that sense that after each query the knowledge in the database is updated with the new results and as such can generate new information.

In the case of e.g. dating a new sheet containing an etching, the x-ray image and available meta-information are input to the system. In the first instance the system looks for sheets with the same paper structure as the sheet in question by using the methods described above. Due to the large data volume search space reduction is preferable. This can be done on the basis of hypotheses by the user of the system or on the basis of meta-information related to the undated sheet. If for instance the artist is known, the search can be limited (in the first instance) to that part of the database which is related to that artist. If such a sheet is found and the similarity is above some threshold, both sheets are considered to be originating from the same sieve. If such a sheet is not found a more detailed research is performed to be on the safe side. Depending on the size of the etching, more than one print was made from one sheet of paper. A sheet was divided in smaller pieces of paper, just large enough for one print. It can happen that two pieces of paper are from the same sieve, but from different positions in such a way that they only have a small overlap, i.e. only in a small part of the sheet the structure of the papers is similar. It is difficult to detect this on the basis of the parameters used in the first search, describing the structure of the paper in general properties, e.g. the average distances between the chain lines. Therefore, in that case search is made on a lower level, e.g. comparing the 2-dimensional representation of the chain lines and their mutual positions. This surely is needed when one tries to identify a paper that in retrospect forms the missing link between two papers which were considered until then to originate from two different sieves.

Anyway, if the matching procedure determines a sheet to be identical with another, this is always with a certain degree of uncertainty. This uncertainty influences the interpretation process. Until now it is the expert who interprets the observed similarity and tries to answer the relevant questions like dating, authenticity etc. In the present situation this interpretation process is performed by the system itself. For that very reason, knowledge and reasoning patterns of the expert as mentioned in the foregoing section are formalized in knowledge rules. Examples are:

a) *If an etching is made in year x and x is on the plate and the print concerns the first state of the etching, then that sheet of paper was possessed by the artist in year x.*
b) *If a type of paper is used for proofs in year x and that type does not appears in years before x, then the artist has probably bought the paper in year x.*
c) *If a type of paper is bought in year x, then all etchings and prints on sheets with the same paper structure date from year x or later.*
d) *If a paper is dated as year x, then probably all prints of the same etching and on a paper with an identical structure, date from year x.*

These kinds of rules determine whether the observed similarity is sufficient to associate the unknown sheet with information concerning the known sheet. E.g. that the undated sheet should be dated as the same year as the known sheet. Uncertainty now propagates. The date of the known sheet has some uncertainty, the similarity between the paper structures is burdened with uncertainty, and therefore the date of the unknown sheet has some (more) uncertainty. Computation of propagated uncertainties is based on methods developed in fuzzy logic.

After a successful query session the image of the etching, X-ray image of the new print and meta-information are added to the database. Each query is related to things one wants to know, but also to things one already knows. The latter is important, since this information can be used in order to update the database and as such to generate new information and knowledge. When one has on the one hand an undated print made by a known artist, whereas on the other hand one has a dated print from a unknown artist, then in the case of identical structure of the papers information can be exchanged. The possible artist of the print with the known date is found as well as the date of the print of the known artist. Just as in the case of the interpretation of matching results, rules for updating are formulated which control what data may be exchanged and on what conditions. Although the incremental process of updating leads to knowledge enrichment and an increasingly powerful database, exchange of information between the various records in the database can cause an avalanche-effect. Each adjustment in one part of the database can imply adjustments in other parts of the database and so on. The problem of truth maintenance and conflict management asks for specific attention. Updating should also imply uncertainty. Generated interpretations can become false later due to new information or insights. Therefore, the conclusions that are made are added to the meta-information file together with the reasoning path that has lead to these conclusions. This information is important since this gives the art historian the possibility of tracing and in the case of falsified interpretation the possibility of repair.

CONCLUSIONS

From an art historical point of view the research of the papers has led to a new and exact dating of various etchings of Rembrandt, to new insights into the various editions and their extent. The problem of automatic feature extraction of X-ray images of etchings and the problem of comparing and matching procedures can be tackled adequately by the proposed methods as shown in this paper. However, the ultimate goal is the development of a large database of all papers used by Dutch artists. The great challenge there is making this database self-organizing with the property of incremental learning. Although the first steps are made towards this direction the research on this is still in progress. From a technical point of view the application is interesting since it requires all kinds of mechanisms of computational intelligence related to the various reasoning types, reasoning with uncertainty and truth maintenance among others.

ACKNOWLEDGEMENTS

The authors would like to thank Theo Laurentius, art historian and Rembrandt-expert, for his contribution. This research has been performed within the framework of ToKeN2000, a interdisciplinary research program on information and knowledge systems, and is supported by the Dutch Ministry of Economic Affairs, the Ministry of Education, Culture and Science and the Netherlands Organization for Scientific Research NWO.

REFERENCES

1. Th. Laurentius: Adriaen van Ostade: The Watermarks. In *Everyday life in Holland's Golden Age; The complete etchings of Adriaen van Ostade*, P. van der Coelen et al., Rembrandt Information Center, Amsterdam, 1998, pp. 61-68

2. Th. Laurentius et al.: The Amsterdam research on Rembrandt's paper: radiography of the watermarks in the etchings of Rembrandt (in Dutch), In *Bulletin Rijksmuseum*, Amsterdam, no 4, 1992, pp. 353-382

3. B. Nap: *Rembrandt, fact or artifact*, MSc thesis, Delft University of Technology, 1998