

A Supra-institutional Infrastructure for Image Processing in the Humanities?

Espen S. Ore

Culture in the widest sense, and artifacts, cultural produce, are the objects of study in the humanities. Hence, the ability to process, store and retrieve images and other non-alpha-numerical data types, e.g. sound, are of high importance for humanities studies in general. The same objects may also be studied for different purposes in the different humanities disciplines and specialities. A page from an illuminated medieval MS may for instance be studied for its textual content by linguists, literature scholars and historians. The same page may be studied for its picture content by art historians and historians. The writing is in itself the study object of paleographers, and paleography is a tool for many of the other above mentioned disciplines. Hence, once we start discussing how the object is viewed and treated within the various sectors of the humanities, we can probably no longer consider "the humanities" a unified category. This means that it is difficult to conceive of "image processing in the humanities". It is even difficult to view image processing within the limited area of history studies as a single field of interest (Jaritz 1991).

If we look at computer methods for dealing with humanist material, we will see that the diversity of approaches and technological methods corresponds to the many ways in which the humanist scholars "set" images as objects of study. In short, a complete supra-institutional service centre for computerized image processing would have to possess competence within a wide range of humanist disciplines as well as competence of the appropriate computer methods to fulfil the needs of the humanists.

Under these circumstances, is it possible to imagine a supra-institutional organization on e.g. a national scale that could support computerized work with images in the humanities? If such an organization is deemed useful, which services should it offer? These questions may be turned around: Given the services that can be offered by a national organization, would such an institution be of value to the humanities? Computerized image processing is after all just a special case of humanities computing in general. In Norway there has existed a national institution for humanities computing since 1972. Image processing software and hardware is now on the level where text processing was in 1985/87. Thus, experience from the work both within this institution and at the Norwegian universities gives us a practical basis that may assist in answering these questions.

The Norwegian Computing Centre for the Humanities 1972 — 1992

In 1972 it was decided by the Norwegian Research Council for the Humanities (NRCH) to establish (for a 5 year period) the Norwegian Computing Centre for the Humanities (NCCH) located at the University of Bergen. The aim of NCCH, somewhat simplified, was to stimulate researchers in the humanities into using computers in their work and to help develop computer competence at the local institutions so that they could take over the responsibility for this work from NRCH. In 1976 NRCH decided to prolong NCCH until 1980, and in 1978 it was decided that NCCH should be a permanent institution and its aims were changed accordingly. In 1991 NCCH was transferred from NRCH to the University of Bergen although NRCH will still provide part of the funding, demanding

that NCCH spend a certain amount of its resources on institutions such as museums that do not have the level of computer resources found at the four Norwegian universities.

Looking back over the last 20 years one finds that the original aim of NCCH has more or less been attained. During the first years of its existence a consultant was placed at each of the Norwegian universities while the rest of the NCCH staff worked at the Bergen office which had its own computing facilities and also access to the central computing facilities at the University of Bergen. The consultants at the universities were eventually transferred organisationally (and economically) to the local universities, and eventually three of them provided the basis for local departments for humanities computing.¹ The departments at the universities of Oslo and Bergen give courses that may be included in an academic degree. It thus seems that the universities have the necessary infrastructure to support computing in the humanities. On the other hand NRCH recognizes that there are institutions, particularly museums, that do not have the staff and economic resources to become self supplied in IT².

During its first years NCCH's work was aimed at developing computer literacy within the humanities in Norway. This was done through seminars, workshops, etc. A system of visiting scholars was later established. These scholars stayed at NCCH for periods up to two years, working on specific projects. At the same time, staff from NCCH collaborated with other projects at the local universities. The idea was that computer competence from NCCH would be spread to the local milieus, while NCCH would gain better insight in the computational needs of the various humanities disciplines. Some of the projects where NCCH collaborated with humanities scholars aimed at developing software to solve a given problem or for use in a specific humanities discipline. Mostly, these programs had a brief lifespan, either because the programs at that time tended to be rather hardware dependent and/or because commercial software appeared that answered the same needs as the software developed at NCCH. On the other hand there were some special problems that would probably never be solved by commercial software (e.g. alphabetic sorting of transcribed Tibetan), and here special ad hoc programs certainly have played an important part.

Computing in the humanities in Norway

Already the very beginning of computing in the humanities in Norway generated discussion on the principles on which it should be based, what possible benefits it could have for the humanities, what threats it might pose etc. This is not surprising since technology and methods from what was then generally considered a spinoff from mathematics was now entering the culture based area of the humanities.

¹ The department for humanities computing at the University of Oslo has later been merged with the departments for linguistics and philosophy due to a general organizational reshuffle at that university.

² In 1983 there were about 475 registered museums in Norway. In an enquete that year 223 museums reported that they had an accumulated workforce of ca. 3000 persons, less than half of whom on an all year basis. (Udjus 1985) In 1991 a survey showed that 520 Norwegian museums employed 1200 person full-time, or 2.3 full-time persons pro museum. (Østby 1991)

From 1973 to 1991 NCCH has published what was originally a newsletter and then expanded into a journal — *Humanistiske Data* (HD). Reading through the volumes one finds in the no. 1, 1974 issue an article by a historian (Johansen 1974) discussing the possible consequences computer related activities might have for the disciplines of the humanities. In his article Johansen gives a sharp warning against building data archives that are not related to a given research problem. If this warning is not heeded, Johansen writes, one will risk that the established data sets may determine what will be used for research rather than that the needs of research define what data should be stored. In the following issue of HD Johansen is answered by another historian (Fonnes 1974) who, although he agrees with Johansen in principle, finds that data in a machine-readable form are a resource and represent an investment in both labour and money that other researchers than the original data collector should be able to use. This, Fonnes writes, means that the computerized data must be as true as possible to the original, i.e. the entire information content of the original data set must be kept in the computerized version. Today one takes a more pragmatic view: Digitized data means available data, so let us digitize the humanities archives now, and worry later about whether having computerized data sets is A Good Thing (see below).

The 1/74 issue of HD also contains a status report on computing in the humanities in Norway for 1973 (HDa 1974), information on available computer programs for computing in the humanities (HDb 1974) and information on available machine-readable data in Norway (HDC 1974). These three articles combined give a picture of humanities computing in Norway which at that time was loosely divided into three main categories:

- Data entry for specific research projects.
- Data entry aimed at building archives for general research purposes.
- Software development.

Software and data archives tended to be field specific but not overly so. History projects used data which were to a certain degree already structured/formalized such as church registers and census papers, while the philologists based their research on free text. Still, the computer tools of choice for both groups (and for musicologists) were those that could be used to generate concordances, word frequency lists and other kinds of lists and indices. Data were stored on punch cards, paper tape and magnetic tape. They were thus not immediately available to the scholars at their own desks, and this may be one of the reasons why many of the archive projects either died completely or went into a long slumber.

At all the universities considerable effort was devoted to programs that could be used in text input and editing. Also listed as research tools were programs for communicating with phototypesetters. In general a lot of work was invested in special computer programs for special hardware that could do a small part of what we expect of any integrated business package today, the generation of concordances excepted. And this last is important: humanists may have needs for computer tools that they at a given time and place consider particular to the humanities. A natural response from people involved in humanities computing if the need were for software would be — and has been — to write special programs. But if or when those humanists' needs coincide with commercial interests, stan-

dard off-the-shelf hardware and software will become available. In image processing, to foreshadow a later argument, we can now buy photo retouching software which makes use of algorithms that a few years ago could only be found in programs developed in special image processing laboratories (Ore E.S. 1992b). We can thank the computer revolution in the graphics industries for this. On the other hand, there has never been a large market for concordance programs in commercial business. Therefore the programs that have been developed within a humanities context represent laudable efforts whether they are sold commercially (e.g. WordCruncher) or are distributed as freeware (e.g. TACT).

Norwegian museum documentation projects

The Association of Norwegian Museums of Art and Social History (ANMASH) organizes a large number of Norwegian museums with material relevant to the humanities. As mentioned above many of these museums are quite small. In order to support its member museums ANMASH has developed documentation standards. The standards have been implemented in database systems for personal computers by NCCH. A first version of the standards and their implementations were made available in 1985 and the standards were closely linked to the physical database implementation. (Østby 1985) One reason for this link between standard and implementation was that ANMASH itself has no power to enforce its standards: by offering an already developed database system to museums that did not have the resources to have their own computing departments, ANMASH ensured that a large number of museums would accept the standard registration form.

In 1991 NCCH was commissioned by ANMASH to enhance the existing standards. This has led to a complete reworking of the field library which among other things now allows multimedia data types. The field library will also be independent of implementation. ANMASH wishes to be able to offer its members ready-to-use implementations under MS-DOS, Windows, and Macintosh, but any implementation of the field library or a subset thereof will generally be acceptable as complying with the standards. For large museums with their own computing expertise this may be an important factor in deciding whether they will accept the new standard or not.

1992: A national documentation project in the humanities

In 1990 a feasibility study was undertaken at the University of Oslo in order to estimate the resources needed to computerize the archives of the various departments in the Faculty of Arts. The study only evaluated the needs of those departments that were considered archive holders, e.g. the Department of Lexicography and the University Museum of National Antiquities which, among other things, is responsible for the Viking ships found in eastern Norway and the artifacts found with them. After the completion of the pilot study it was discovered that quite a few other departments also held archives which they wanted to computerize, such as the slide collection of the Department of Art History or the slide collection of Herculanean papyri at the Department of Classics (Kleve, Ore, Jensen 1987). However, the work needed to computerize the archives of just the first selected departments was estimated at more than 600 man years. (Ore, C.-E. 1991)

On the basis of the feasibility study it was decided to start work on the complete project. In 1991 it outgrew its origins as a project local to the University of Oslo. It is now a national documentation project for the humanities with all four universities participating.

It is funded partly by the government, partly by the universities. Some of the registration work now being done, e.g. in lexicography, is in fact a continuation, or a revival of the archives reported in *Humanistiske Data* in 1974. The big difference between then and now is the arrival of the personal computer and user friendly commercially available software.

So far the documentation project does not aim at establishing a large unified database, although linking of the registered data in interdisciplinary hypermedia systems is a long term aim. The project has a steering group with representatives from the faculties of art at the Norwegian universities and an administrative and technical management situated at the University of Oslo. There are no centralized decisions taken on which data should be input and in what form, although the cooperation between the universities ideally insures that the same data are not entered twice. The main point is to computerize the data in a form that is of use to the departments responsible for the archives and compatible with the computer systems already in use there. Norwegian universities computing is already fairly standardized on Intel-based systems, Macintoshes and Unix computers that are all more or less networked together. When a reasonable amount of data is entered the next step in the documentation project will be to develop presentation and interchange formats. This will include linking the data from the various specialist departments: The terms in the lexicographic database may be linked with the dialectological databases and with GIS (Geographical Information Systems) bases showing where in Norway a specific term is used for e.g. a piece of fishing gear. In other words, this kind of interconnected data ideally in a hypermedia system opens the way for an interdisciplinary approach towards the humanities archives. The further development of specialized tools for the utilisation of the data will still be the concern of the individual departments concerned. For instance the collection of Greek and Roman coins in the Collection of Coins & Medals is now being registered together with digital images of the coins. These images will be generally available, but the tools needed for e.g. a comparison of coins are not to be supplied by the documentation project. Similarly, there are plans for computerizing the Norwegian collections of runic inscriptions both as transcripts and as images of the inscriptions themselves. For epigraphic and paleographic studies it will be of interest to develop computer tools for the analysis of the images but this too is outside of the immediate range of the documentation project.

What have we learned?

After 20 years of experience in humanities computing, some of the lessons we have learned from the work at NCCH and from humanities computing in Norway in general are:

- Development and maintenance of software for general use is costly. If possible, use commercially available software.
- Ad hoc programming for specific project purposes is often worthwhile.
- On a national or international scale, different hardware and software platforms are used (and will be used) in the humanities. Thus, all systems planning should be as platform independent as possible.
- Humanists will use simple hardware and software located on their desktops rather than advanced equipment located somewhere else. (As a corollary: transparent networking and data exchange becomes more and more important.)
- The emerging use of hypermedia systems is an incentive towards interdisciplinary work.³

³ A good example of a hypermedia system with interdisciplinary aspects is the Perseus project at Harvard (Crane G. 1991 and Mylonas E. 1991).

Conclusions

Rather than a supra-institutional infrastructure for image processing in the humanities realized as a separate institution I see a need for an infrastructure in the form of inter-institutional cooperation. The expertise has to be where the scholars are, not the other way round. On the other hand humanities computing is interdisciplinary by nature so all institutions of any size should have an infrastructure for humanities computing that could be organized as a humanities computing centre. Such centres should be part of an international network for humanities computing. This could preferably take place within the scope of organizations such as ACH or ALLC. They should also, nationally, provide assistance to institutions that cannot afford their own humanities computing services. Image processing would be an integral and important part of this.

Today, image databases (and other multimedia databases) depend on descriptive fields for retrieval. In the future it is possible that there will be software that can search for image content using pattern matching algorithms derived from AI-research and military technology. (Radar based homing missiles spring to mind here.) But for the time being, verbal descriptions and classification codes will be the main retrieval aids together with visual browsing of the images. Given the different approaches to the images in the various sectors of the humanities, it is difficult to imagine a completely standardized classification scheme (Jaritz 1991). Still, some standards for classification are better than none at all, so standardization of classification and verbal descriptions should form an important part of inter-institutional collaboration both nationally and internationally.

The computer technology and its use are rapidly changing. This has been true for the last twenty years, and most likely it will be true for the next twenty years. Therefore, little would be gained by trying to establish standardized software and hardware platforms for image processing in the humanities on a supra-institutional level (Roper 1991). More important would be the development of interchange formats and extended use of the international computer networks.

Bibliography

Crane, G. (1991), "Composing Culture: The Authority of an Electronic Text", *Current Anthropology*, Vol. 32, No. 3, 1991, pp. 293-302

Fonnes, I. (1974), "EDB som faktor i kunnskapsproduksjonen" (EDP as a factor in the production of knowledge), *Humanistiske Data*, No 2 1974, pp. 7-8

HDa unsigned article (1974), "EDB-virksomheten i de humanistiske fag ved våre universiteter" (Application of EDP in our universities), *Humanistiske Data*, No. 1 1974, pp. 4-13

HDb unsigned article (1974), "Oversikt over maskinprogram i Norge til bruk i humanistisk forskning" (Overview of computer programs [available] in Norway for use in the humanities), *Humanistiske Data*, No. 1 1974, pp. 14-17

HDc unsigned (1974), "Data i maskinleselig form" (Machine readable data), *Humanistiske Data*, No. 1 1974, pp. 18-24

Jaritz, G. (1991), "Medieval Image Databases: Aspects of Cooperation and Exchange", *Literary and Linguistic Computing*, Vol. 6, No. 1, 1991, pp. 15-19

Johansen, A. B. (1974), "EDB som faktor i kunnskapsproduksjonen" (EDP as a factor in the production of knowledge), *Humanistiske Data*, No. 1 1974, pp. 1-3

Kleve, K., Ore, E.S. and Jensen, R. (1987), "Literalogy: On the Use of Computer Graphics and Photography in Papyrology", *Symbolae Osloensis*, Vol. LXII, 1987, pp. 109-129

Mylonas, E. (1991), "The Perseus Project: Ancient Greece in Texts, Maps and Images", *Proceedings from the Conference: Electronic Books — Multimedia Reference Works*, Bergen 1991, pp. 173-187

Ore, C.-E. (1991), *Dokumentasjonsprosjektet ved Det historisk-filosofiske fakultet, Universitetet i Oslo* (The Documentation Project at the Faculty of Arts, the University of Oslo), Oslo 1991

Ore, E. S. (1992a), "Project Litera: Computer Aids in Restoring Partly Preserved Letters in Papyri", paper given at the ALLC/ACH conference in Siegen June 1990, to be published in *Research in Humanities computing*, 1992

Ore, E. S. (1992b), "Bilder og edb 1991" (Images and Edp 1991), *Proceedings from the Second Norwegian Conference on Photo-preservation 1991, 1992*, in print

Roper, J. G. (1991) "The New Humanities Workstation", *Literary and Linguistic Computing*, Vol. 6, No. 2, 1991, pp. 131-133

Udjus, I. (1985), "Fakta om de norske museene" (Facts about the Norwegian museums), *Musumsnytt*, Vol. 34, No. 3 4, 1985, pp. 100-102

Østby, J. B. (1985), *Edb-metoder for kunst og kulturhistoriske museer* (Edp-methods for museums of art and social history), Report No. 38 from NCCH, 1985

Østby, J. B. (1991) Working paper presented at a seminar on computerized museum documentation at Gran, Norway, 18 Sept. 1991. Unpublished.

**Halbgraue Reihe
zur Historischen Fachinformatik**

Herausgegeben von
Manfred Thaller
Max-Planck-Institut für Geschichte

Serie A: Historische Quellenkunden

Band 14

Erscheint gleichzeitig als:

MEDIUM AEVUM QUOTIDIANUM

HERAUSGEGEBEN VON GERHARD JARITZ

Manfred Thaller (Ed.)

Images and Manuscripts in Historical Computing

Max-Planck-Institut für Geschichte
In Kommission bei
SCRIPTA MERCATURAE VERLAG



St. Katharinen, 1992

© Max-Planck-Institut für Geschichte, Göttingen 1992
Printed in Germany
Druck: Konrad Pachnicke, Göttingen
Umschlaggestaltung: Basta Werbeagentur, Göttingen
ISBN: 3-928134-53-1

Table of Contents

Introduction	
<i>Manfred Thaller</i>	1
 I. Basic Definitions	
Image Processing and the (Art) Historical Discipline	
<i>Jörgen van den Berg, Hans Brandhorst and Peter van Huisstede</i>	5
 II. Methodological Opinions	
The Processing of Manuscripts	
<i>Manfred Thaller</i>	41
Pictorial Information Systems and the Teaching Imperative	
<i>Frank Colson and Wendy Hall</i>	73
The Open System Approach to Pictorial Information Systems	
<i>Wendy Hall and Frank Colson</i>	87
 III. Projects and Case Studies	
The Digital Processing of Images in Archives and Libraries	
<i>Pedro González</i>	97
High Resolution Images	
<i>Anthony Hamber</i>	123
A Supra-institutional Infrastructure for Image Processing in the Humanities?	
<i>Espen S. Ore</i>	135
Describing the Indescribable	
<i>Gerhard Jaritz and Barbara Schuh</i>	143
Full Text / Image DBMSs	
<i>Robert Rowland</i>	155

Introduction

Manfred Thaller

This book is the product of a workshop held at the International University Institute in Firenze on November 15th, 1991. The intention of that workshop has been to bring together people from as many different approaches to "image processing" as possible. The reason for this "collecting" approach to the subject was a feeling, that while image processing in many ways has been the "hottest" topic in Humanities computing in recent years, it may be the least well defined. It seems also much harder to say in this area, what is specifically important to historians, than to other people. In that situation it was felt, that a forum would be helpful, which could sort out what of the various approaches can be useful in historical research.

To solve this task, the present volume has been produced: in many ways, it reflects the discussions which actually have been going on less, than the two companion volumes on the workshops at Glasgow and Tromsø do. This is intentional. On the one hand, the participants at the workshop in Firenze did strongly feel the need to have projects represented in the volume, which were not actually present at the workshop. On the other, the discussions for quite some time were engaged in clarifying what the *methodological* issues were. That is: what actually are the topics for scholarly discussion beyond the description of individual projects, when it comes to the processing of images in historical research?

The situation in the area is made difficult, because some of the underlying assumptions are connected with vigorous research groups, who use fora of scholarly debate, which are only slightly overlapping; so, what is tacitly assumed to hold true in one group of research projects may be considered so obviously wrong in another one, that it scarcely *deserves* explicit refutation.

We hope, that we have been successful in bringing some of these hidden differences in opinion out into the open. We consider this extremely important, because only that clarification allows for a fair evaluation of projects which may have started from different sets of assumption. So important, indeed, that we would like to catalogue here some of the basic differences of opinion which exist between image processing projects. The reader will rediscover them in many of the contributions; as editor I think however, that summarizing them at the beginning may make the contributions — which, of course, have been striving for impartiality — more easily recognizable as parts of one coherent debate.

Three basic differences in opinion seem to exist today:

(1) Is image processing a genuine and independent field of computer based research in the Humanities, or is it an auxiliary tool? Many projects assume tacitly — and some do so quite outspokenly — that images on the computer act as illustrations to more conventional applications. To retrieval systems, as illustrations in catalogues and the like. Projects of this type tend to point out, that with currently easily available equipment and currently clearly understood data processing technologies, the analysis of images, which can quite easily be handled as illustrations today, is still costly and of uncertain promise. Which is the reason why they assume, that such analytical approaches, if at all, should be undertaken

as side effects of projects only, which focus upon the relatively simple administration of images. Their opponents think, in a nutshell, that while experiments may be needed, their overall outcome is so promising, that even the more simple techniques of today should be implemented only, if they can later be made useful for the advanced techniques now only partially feasible.

(2) Connected to this is another conflict, which might be the most constant one in Humanities data processing during the last decades, is particularly decisive, however, when it comes to image processing. Shall we concentrate on levels of sophistication, which are available for many on today's equipment or shall we try to make use of the most sophisticated tools today, trusting that they will become available to an increasingly large number of projects in the future? This specific battle has been fought since the earliest years of Humanities computing, and this editor has found himself on both sides at different stages. A "right" answer does not exist: the debate in image processing is probably one of the best occasions to understand mutually, that both positions are full of merit. It is pointless to take permanently restrictions into consideration, which obviously will cease to exist a few years from now. It discredits all of us, if computing in history always promises results only on next years equipment and does not deliver here and now. Maybe, that is indeed one of the more important tasks of the *Association for History and Computing*: to provide a link between both worlds, lending vision to those of us burdened down by the next funding deadline and disciplining the loftier projects by the question of when something will be affordable for all of us.

(3) The third major underlying difference is inherently connected to the previous ones. An image as such is beautiful, but not very useful, before it is connected to a description. Shall such descriptions be arbitrary, formulated in the traditionally clouded language of a historian, perfectly unsuitable for any sophisticated technique of retrieval, maybe not even unambiguously understandable to a fellow historian? Or shall they follow a predefined catalogue of narrow criteria, using a carefully controlled vocabulary, for both of which it is somewhat unclear how they will remain relevant for future research questions which have not been asked so far? — All the contributors to this volume have been much to polite to phrase their opinions in this way: scarcely any of them does not have a strong one with regard to this problem.

More questions than answers. "Image processing", whether applied to images proper or to digitalized manuscripts, seems indeed to be an area, where many methodological questions remain open. Besides that, interestingly, it seems to be one of the most consequential ones: a project like the digitalization of the *Archivo General de Indias* will continue to influence the conditions of historical work for decades in the next century. There are not only many open questions, it is worthwhile and necessary to discuss them.

While everybody seems to have encountered image processing in one form or the other already, precise knowledge about it seems to be relatively scarce. The volume starts, therefore, with a general introduction into the field by J. v.d. Berg, H. Brandhorst and P. v. Huisstede. While most of the following contributions have been written to be as self supporting as possible, this introduction attempts to give all readers, particularly those

with only a vague notion of the techniques concerned, a common ground upon which the more specialized discussions may build.

The contributions that follow have been written to introduce specific areas, where handling of images is useful and can be integrated into a larger context. All authors have been asked in this part to clearly state their own opinion, to produce clearcut statements about their methodological position in the discussions described above. Originally, four contributions were planned: the first one, discussing whether the more advanced techniques of image processing can change the way in which images are analysed and handled by art historians, could unfortunately not be included in this volume due to printing deadlines: we hope to present it as part of follow up volumes or in one of the next issues of *History and Computing*.

The paper of M. Thaller argues that scanning and presenting corpora of manuscripts on a work station can (a) save the originals, (b) introduce new methods for palaeographic training into university teaching, (c) provide tools for reading damaged manuscripts, the comparison of hand writing and general palaeographic studies. He further proposes to build upon that a new understanding of editorial work. A fairly long technical discussion of the mechanisms needed to link images and transcriptions of manuscripts in a wider context follows.

F. Colson and W. Hall discuss the role of images in teaching systems in university education. They do so by a detailed description of the mechanism by which images are integrated into Microcosm / HiDES teaching packages. Their considerations include the treatment of moving images; furthermore they enquire about relationships between image and text in typical stages in the dialogue between a teaching package and a user.

W. Hall and F. Colson argue in the final contribution to this part the general case of open systems, exemplifying their argument with a discussion of the various degrees in which control about the choices a user has is ascertained in the ways in which navigation is supported in a hyper-text oriented system containing images. In a nutshell the difference between "open" and closed systems can be understood as the following: in an "open system" the user can dynamically develop further the behaviour of an image-based or image-related system. On the contrary in static "editions" the editor has absolute control, the user none.

Following these general description of approaches, in the third part, several international projects are presented, which describe in detail the decisions taken in implementing "real" image processing based applications, some of them of almost frightening magnitude. The contributors of this part were asked to provide a different kind of introduction to the subject than those to the previous two: all of them should discuss a relatively small topic, which, however, should be discussed with much greater detail than the relatively broad overviews of the first two parts.

All the contributions growing out of the workshop came from projects, which had among their aims the immediate applicability of the tools developed within the next 12 - 24 months. As a result they are focusing on corpora not much beyond 20.000 (color) and 100.000 (b/w) images, which are supposed to be stored in resolutions manageable within ≤ 5 MB / image (color) and ≤ 0.5 MB / image (b/w). The participants of the workshop felt strongly, that this view should be augmented by a description of the rationale behind

the creation of a large scale project for the systematic conversion of a complete archive. The resulting paper, by P. González, describes the considerations which lead to the design of the *Archivo General de Indias* project and the experiences gained during the completed stages. That description is enhanced by a discussion of the strategies selected to make the raw bitmaps accessible via suitable descriptions / transcriptions / keywords. A critical appraisal, which decisions would be made differently after the developments in hardware technology in recent years, augments the value of the description.

The participants of the workshop felt furthermore strongly, that their view described above should be augmented by a description of the techniques used for the handling of images in extremely high resolution. A. Hamber's contribution, dealing with the *Vasari project*, gives a very thorough introduction into the technical problems encountered in handling images of extremely high quality and also explains the economic rationale behind an approach to start on purpose with the highest quality of images available today on prototypical hardware.

As these huge projects both were related to institutions which traditionally collect source material for historical studies, it seemed wise to include also a view on the role images would play in the data archives which traditionally have been of much importance in the considerations of the AHC. E.S. Ore discusses what implications this type of machine readable material should have for the infrastructure of institutions specifically dedicated to Humanities computing.

Image systems which deal with the archiving of pictorial material and manuscript systems have so far generally fairly "shallow" descriptions. At least in art history, moreover, they rely quite frequently on pre-defined terminologies. G. Jaritz and B. Schuh describe how far and why historical research needs a different approach to grasp as much of the internal structure and the content of an image as possible.

Last not least R. Rowland, who acted as host of the workshop at Firenze, describes the considerations which currently prepare the creation of another largescale archival database, to contain large amounts of material from the archives of the inquisition in Portugal. His contribution tries to explore the way in which the more recent developments of image processing can be embedded in the general services required for an archival system.

This series of workshop reports shall attempt to provide a broader basis for thorough discussions of current methodological questions. Their main virtue shall be, that it is produced sufficiently quick to become available, before developments in this field of extremely quick development make them obsolete. We hope we have reached that goal: the editor has to apologize, however, that due to the necessity to bring this volume out in time, proofreading has by necessity be not as intensive as it should have been. To which another shortcoming is added: none of the persons engaged in the final production of this volume is a native speaker of English; so while we hope to have kept to the standards of what might be described as "International" or "Continental" English, the native speakers among the readers can only be asked for their tolerance.

Göttingen, August 1992