Digital Solutions for Archival Family Records

Rodney Shaw
Hewlett-Packard Research Laboratories
Palo Alto, California, USA

Abstract

The onset of the information era and the general availability of both the global internet and sophisticated digital technologies have ushered in a consumer revolution in attitudes and needs so far as family history and archival records are concerned. Photography and family history constitute two of the world's most popular hobbies, and the latter represents one of the leading uses of the web. Whereas digital photography now offers convergence with other forms of digital records, the historical problem remains of the 'shoe-box-in-the-attic' syndrome, and the hoards of vintage family photographs awaiting future digitizing and restoration. Similarly, important family documents deteriorate in dusty storage archives. There is also the crucial but problematic question of individual ownership versus universal distribution within a given family or any defined group or corresponding circle.

The author has approached the problem on a general basis, first by sampling the general consumer demand and existing practice, next by exploring various approaches based on this personal experience, and finally by bringing together a set of readily-available digital technologies to provide a satisfactory solution. This has involved technologies ranging from image-processing and photo-restoration, to document-scanning and desk-top publishing for complex document types. This present contribution describes things learned during this experience, and discusses the nature of overall solutions and generic formats that to date have proved satisfactory to a majority of consumers sampled within this field.

Introduction

In general the technology community has been slow to acknowledge or address consumer needs and interests within the burgeoning field centered around family archives and photographic collections. While the digital revolution has seen a staggering array of cheap and sophisticated digital gadgetry, now readily within the economic grasp of the majority of consumers, any general back-up support is largely lacking. This support ranges from easily understandable guidelines and usage rules, through to associated software support, and most of all, the almost complete lack of an attractive overall integrated systems solution.

Any rapid survey of the web demonstrates both this need and the lack of any systematic high-technical presence or penetration within the field, perhaps since no single technology-group or commercial enterprise can claim 'ownership' of these broad-based consumer needs. Thus, although the start-up costs for any interested consumer have become relatively trivial, the learning-curve through the formidable technical barriers (both in choice and usage of hardware and software) is prohibitively steep for all but the very persistent, and the low- or no- technology approach therefore prevails. Even personal desktop publishing remains a problem without a general consumer solution, in spite of the vast implicit consumables market involved, and thus directly against the self-interests of the vast desk-top printing industry.

Background

Starting around two years or so, the author informally identified a dozen or so existing and potential users of contemporary digital imaging products (scanners, digital-cameras, printers, and various associated software programs). These consumers ranged from naïve enthusiasts and hobbyists, to committed but largely non-technical professional users who had at least a partial existing interest in digital imaging technologies. As it transpired, not one had any real inclination of the range of capabilities of existing digital products already in their possession or well within their individual budget limits. Thus the lesson was rapidly learned that this lack of awareness, combined with difficulty of operation and lack of any continuing customer explanation, support and back-up interest, resulted in some degree of disillusionment and, of importance in this context, *a severe under-use of installed digital equipment.*

This investigation was also coincident with the author's own interest in various aspects of image processing, photographic restoration and enhancement, and initially those consumers with various interests in digital photography were the main focus. In parallel, as a learning task the author undertook the digitization and optimal image processing of several thousand analog and digital images taken from a diverse range of sources, from high-resolution but degraded historic photos photographs to low-

resolution web-images. Thus, although at first the investigation concentrated on photographic consumers (for whom web-services are still in their infancy, a separate story in itself), based on interest and experience, the focus quickly moved to the greater concern of family archives and genealogical records. During this the author learned on the hand of the vast web activity in this field, ranging from raw genealogical data, archival documents, historic photographs; yet on the other hand, and as discussed above, the almost total lack of technical support, help, or infrastructure, either overall or in detailed and specific areas.

Technical Approach

Based on this experience the author worked in an *ad hoc* manner with a sample of family historians, genealogists and historians, in addition to becoming familiar with the vast web-resources in this field. Over a period of a year the concept evolved of 'the definitive family history', and the bringing together of all existing archives, documents, photographs and genealogical records within a central family resource. Such a history, in the guise of a desktop published document becomes a powerful 'up-datable chain-letter within an even larger chain-letter', in that after initial establishment there is an ever-expanding interest and requirement for additional copies within any group or family. The ultimate product is the professional-quality coffee-table book with its accompanying up-datable CD. The larger part of the chain develops as interest then passes naturally to other families. In this form it becomes of interest to a majority of potential consumers overall, and not merely the large and expanding genealogy population, and as a family history resource is perhaps the modern equivalent of the important role played by the inscribed family bible of a century or more ago.

This presentation will attempt to summarize some of the equipment, software and services required in moving towards such a preferred general solution. The author started as novice in most of these tasks (and remains so), yet none of them present a technical obstacle that could not quickly be integrated within either a general solution or a contributory consumer service. An attempt is made below to show some of these main contributory areas, and, perhaps most important of all, the combined technical skill-set needed for any overall solution.
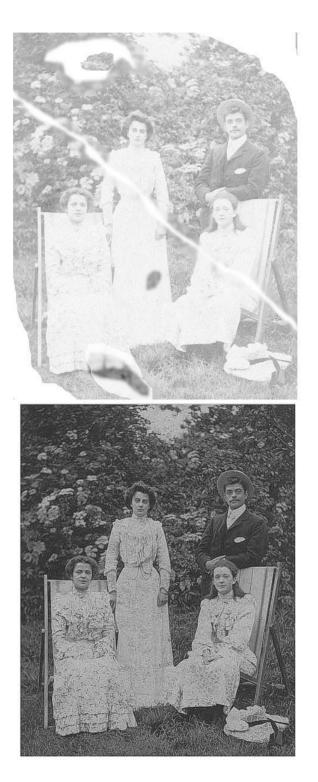
Requirements for Digital Family Records

We leave aside here the separate activities, *sine qua non*, of organizing the central resources, documents, data and photographs, etc. These are typically already in full and enthusiastic swing among growing groups of non-technical enthusiasts. For example, as a class, these enthusiasts are in fact perhaps the most sophisticated of users of web resources and shared results. Here we concentrate on those individual and overall value-added technical components to enable and facilitate these activities, and move them towards an overall solution.

It goes without saying that access to a minimum range of digital products is essential. These include a computer, preferably with CD read/write capabilities, a high-resolution scanner, perhaps a separate hi-resolution transparency scanner, a photo-quality printer, and preferably a digital camera. All these gadgets must be accompanied by experience in use and especially in optimization (an easy phrase, a rare actuality). Likewise all these devices must be controllable and inter-communicable via sophisticated software (again easy in phrase, almost non-existent in practice). Further, additional and very specific software is essential, for example facile photo and document management and desktop publishing software. Anyone with familiarity with the reality in these areas must know the consumer pitfalls and learning curves associated with activities such as these. Most consumers, even sophisticated enthusiasts, have lost patience or interest by this stage.

Based on two years experience with these various problematic areas, the author has however determined that each and every one of these natural obstacles may be reduced to formulaic simplicity: there is a recipe to be learned the hard way. Once learned it could then be translated into practice when the motive exists, for example, from studies of the market potential. But in each case and for each component the author found it necessary to undertake the long learning curve, and then to cast around for a satisfactory overall solution, an even longer and more difficult learning curve in itself.

The examples of this current situation must of course be based on the specific. The digitization and restoration of archival photographs is one such component, and near to the specific interests of the author. However most of the problems here are not based around advanced image processing routines, but around the less glamorous task of optimized-scanning and digital-repair. Yet even in this sphere, a recipe-like solution is possible such that, say one-hour per photograph by a novice can be replaced by two-minutes per photograph by the 'expert', where expert here denotes somebody who has done it a thousand times, as opposed to a photo-software-engineer who understands all the niceties of the underlying photo-software package.

The above before-and-after versions of a digitized print provide an example of the severer type of fading and imperfections found among typical important family archives. While such examples and the efforts involved in their restoration are well-known to this present audience, the non-existence of an easy commercial solution is just as surely well-known.

A General Solution

There may well be as many satisfactory general solutions are there are individual solvers. One evolutionary solution opted with some success by the author is that of the published document, accompanied by an up-datable version on CD, that also includes high-resolution versions of all key documents and photographs for separate high-quality prints, as appropriate. The latter can then be easily circulated among and updated by all group members.

In this context the following two pages of illustration represent a typical extract from such a final version of a family history document. They have been chosen as examples here since they are in a form that is highly satisfactory for enthusiasts in the field. While each has a mundane content and appearance, the collective data was first assembled from a dozen or separate sources, the rather hard part leading up the much harder part, where the current complexities of digital technology must be grappled with.
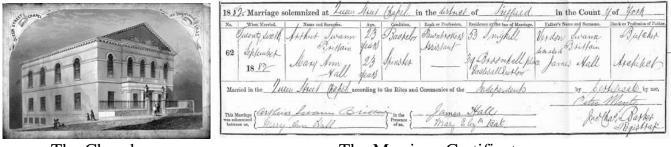
The overall product of course resides most obviously in the final printed document, which cannot be fully represented here, but the author will be happy to demonstrate and discuss this with others having interests in this booming field.
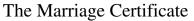
Summary and Conclusions

- The two most popular global hobbies are centered around photography and family history.

- The second largest activity on the world-wide-web is family-history related.

- The web is increasingly image-dominated, yet photographically primitive.

- Almost every family still has a shoe-box approach to vital records, documents and photographs.

- General consumer digital technology solutions are available today

- In general existing consumer digital technology is cumbersome and user-hostile.

- There is an overwhelming desire for any general technical solution to facilitate a digital family history solution.

- The potential market for an integrated consumer-solution is huge, yet largely ignored.

rodshaw@hpl.hp.com

Wedding of Great-Grandparents

The Church

The Marriage Certificate

The Wedding Group

The Family Store

Family Portraits

Census data, Occupations and Residences

Sixty years of family evolution

1841 Census Nether End Farm, Bakewell, Derby, England

William SMYTHE 35	Head	Farmer 24 acres
Mary SMYTH E 32	Wife	Farmers Wife
George SMYTHE 6	Son	Scholar
Martha SMYTHE 4	Daughter	
William SMYTHE 2	Son	
Betsy VICKERS 17		General Servant

1871 Census The Mill, Ashford, Derby, England

George SMYTHE 36	Head	Corn Miller
Harriet SMYTH E 35	Wife	Millers Wife
William SMYTHE 14	Son	Millers Apprentice
Elizabeth SMYTHE 10	Daughter	Scholar
John SMYTHE 8	Son	Scholar
Hannah SMYTHE 3	Daughter	Scholar
Mary SMYTHE 62	Mother	Widow with Annuity
Seth BARTHOLOMEW 19 Paid Helper		Millers Assistant

1901 Census 37 Victoria Promenade, Derby, England

John SMYTHE 38	Head	Bank Manager
Emma SMYTH E 37	Wife	House Wife
George SMYTHE 19	Son	Apprentice Silversmith
Jane SMYTHE 18	Daughter	Trainee Teacher
William SMYTHE 17	Son	Handicapped (deaf)
Catherine SMYTHE 12	Daughter	Scholar
Mary SMYTHE 92	Grandmother	Widow with Annuity
George SMYTHE 66	Father	Retired (infirm)
Molly PEARSON 16		Servant
Liza CUTTS 25		Cook