

Image Restoration of Arbitrarily Warped Documents

Michael S. Brown, *Member, IEEE*, and W. Brent Seales, *Member, IEEE*

Abstract—We present a framework for acquiring and restoring images of warped documents. The purpose of our restoration is to create a planar representation of a once planar document that has undergone an arbitrary and unknown rigid deformation. To accomplish this restoration, our framework acquires and flattens the 3D shape of a warped document to determine a nonlinear image transform that can correct for image distortion caused by the document's shape. Our framework is designed for use in library and museum digitization efforts where very old and badly damaged manuscripts are imaged.

Index Terms—Manuscript restoration, document deskewing, image restoration, cultural heritage digitization.

1 INTRODUCTION

IN this paper, we address the digital restoration of damaged manuscripts. The goal is to move beyond 2D image enhancement techniques, which cannot adequately address distortions that arise in damaged manuscripts. Shape-based distortion, for example, is a common element in damaged manuscripts that cannot be removed via 2D image enhancement algorithms. We achieve results using 3D geometric information together with 2D imaging to formulate a restoration framework for removing shape-based (projective) distortion in the image [9], [8], [7]. Our approach uses a physically-based model of shape and material properties and iteratively computes the transformation that maps the erratic, nonplanar shape of the damaged manuscript representation back to its original shape (a plane). The relationship of the resulting “flattened” manuscript to the original shape defines the piecewise image warp necessary to create a distortion-free image of the document as it would appear if the original had been planar.

While purely 2D image operations can improve the *perceptual quality* of the image, they cannot address shape-based distortions without additional constraints. The pinhole camera is a projective transformation engine and by its nature creates an image that is systematically *projectively* distorted. This *projective distortion* is a function of the distance of a 3D point from the camera: Two points that are far from the camera appear closer in the image than the same two points when they are nearer to the camera. The unique imaging orientation that does not produce projective distortion occurs when all points on the object are the same distance from the camera, as is the case for truly flat manuscripts imaged in an orthogonal orientation to the camera's optical axis.

Consider, for example, bumps on the surface of the manuscript. The variation in distance to the camera over the

bumpy surface causes a distortion of the appearance of letters written on that surface. The bump, which is a 3D shape distortion, and the imaging process, which produces projective distortion, are confounded in the image to yield a complex 2D distortion of the text of a manuscript that was intended originally to be flat. Fig. 1 shows a real example of this effect. The corner of the Medieval manuscript shown on the left includes the distortion due to the nonplanar shape of the page. The line on the paper frame and the text on the vellum manuscript show the distortion because the page is not flat. The recovered 3D view illustrates the degree of distortion.

Even when a document is completely flat, its image can exhibit distortion if the imaging process did not obey orthonormal constraints, i.e., if the optical axis was not orthogonal to the plane defined by the manuscript. This special type of projective distortion is called *skew* and can be systematically removed by estimating the plane on which the document was imaged.

Projective distortion is almost universally ignored in manuscript analysis. In the case of orthonormal imaging, which is satisfied by many scanning setups (e.g., flat-bed scanning), this is justified. When imaging a document that is assumed to be planar but is in reality warped and bumpy, however, projective distortion can be significant. Most scholars use 2D image enhancement algorithms to manipulate the intensity distribution of the image, but systematically ignore projective distortion. As a result, the projective distortion remains in an image after 2D restoration and enhancement. Because it is difficult if not impossible (except under strict theoretical conditions of lighting and surface properties) to estimate surface shape from a single image, there is no reliable way to identify and remove projective distortion via 2D image processing. Image processing algorithms cannot, without a priori information, determine document shape, which is necessary to solve the restoration problem.

In this paper, we present a technique that uses a 2D image of a document together with its 3D shape in order to restore it by *removing* the projective distortion present in the image. We show that this distortion is profound for certain classes of documents that should be flat but have been damaged and now contain bumps, creases, and warps. Our approach improves current restoration techniques applied by scholars who are editing digital collections of manuscripts by

- M.S. Brown is with the Department of Computer Science, Hong Kong University of Science and Technology, Clearwater Bay, Kowloon, Hong Kong, China. E-mail: brown@cs.ust.hk.
- W.B. Seales is with the Laboratory for Advanced Networking, Hardyman Building, 2nd Floor, University of Kentucky, Lexington, KY 40506. E-mail: seales@uky.edu.

Manuscript received 28 Aug. 2003; revised 30 Dec. 2003; accepted 2 Mar. 2004. Recommended for acceptance by M. Pietikainen.

For information on obtaining reprints of this article, please send e-mail to: pami@computer.org, and reference IEEECS Log Number TPAMI-0252-0803.

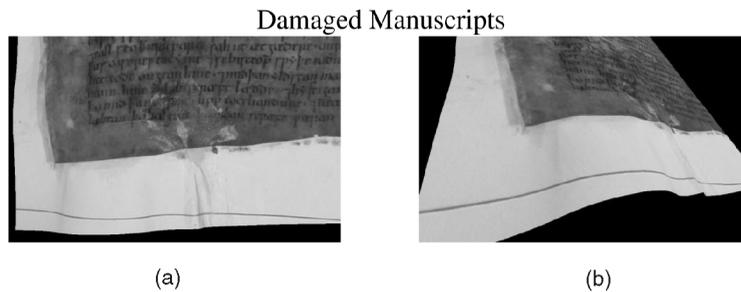


Fig. 1. Damaged manuscript. (a) Bottom corner of original, photographed from directly above. (b) Novel view of the 3D reconstruction of the manuscript page.

addressing restoration limitations and by providing the foundation for new types of manuscript restoration based on digitized geometry. Our new restoration technique presents the following advantages:

- **Perceptual Restoration:** Removing projective distortion can improve the perceptual quality and fidelity of an image, which is especially useful for digitized manuscripts that exhibit text appearing wavy and warped because of shape variation.
- **Document Analysis Preprocessing:** In addition to scholarly study, 2D images are the basic input for an array of *document analysis* applications designed to derive high-level semantics from an image. For example, Optical Character Recognition (OCR) algorithms detect alpha-numeric characters in imagery of a document [19]. Increasingly similar algorithms are emerging to derive features from nontraditional materials such as handwritten documents, music, etc. These algorithms rely on segmentation components that are task-specific, such as staff-line detection in music and identification of rows/columns of text in handwritten materials [17], [27].

The implicit assumption in these document analysis algorithms is that the original image is of a *flat* document. When this is not the case, projective distortion in the image causes adverse results for document analysis algorithms. Projective distortion removal can serve as an important preprocessing step for document analysis and can be decoupled from the particulars of the document domain.

- **Image Scale Normalization:** Our restoration framework can “flatten” damaged manuscripts to recover their appearance as if they had been imaged on a flat surface. By choosing a common planar surface for a set of individual manuscripts, the resulting *restored* images are “scale-normalized.” This can be very useful when imaging thick bound collections, where pages near the front of the collection appear slightly larger than pages near the end because of distance variations during imaging. It can also be important for comparing different collections side-by-side in terms of scale and textual analysis. Note that the scale variation effect is very difficult to remove during the digitization process since the camera would need to be slightly (and correctly) repositioned before imaging each page. Scale normalization is very desirable for consistency in analysis as well as for addressing the follow-on problem of mosaicing when damaged objects are fragmented.

2 BACKGROUND AND RELATED WORK

Digital restoration is a manipulation intended to produce an *improved* digital representation of an object. We define *improvement* as a modification that makes the representation more like a digitization of the undamaged original. Given this definition, three elements are necessary in order to formulate a restoration process: an idealized or “restored” goal state, the process of manipulation, and a metric that measures how close an object is to its original, undamaged state.

In practice, it is very difficult to design a restoration process that objectively defines the goal state, the process, and the metric. Specifically, physical restoration presents the greatest challenge because the goal is subjective: The true original condition can only be guessed. Furthermore, the physical process of restoration can be unpredictable and the results can vary widely depending on factors such as human expertise, materials, environment, and the material nature of the damage.

The restoration process can be made somewhat less subjective when manipulating a digital representation rather than a physical artifact. The digital representation admits metrics and manipulations that are quantifiable and precise. Digital image enhancement and restoration techniques, for example, operate through models of the idealized image together with precise definitions of processes that cause damage and a metric that measures improvement. We have formulated our restoration technique as a quantifiable process that uses a digital representation (3D shape and associated texture), a physics-based model of the allowable manipulation (shape deformation under the rigidity constraint), and the exact goal state (flat original).

2.1 Physical Restoration

For certain objects, the most logical solution to the restoration question is direct repair of the actual physical artifact. Libraries and museums have trained conservators and archivists who specialize in precisely this form of restoration and preservation of damaged materials. There is always very serious concern with physical restoration, however. The materials to be restored may be rare and singularly unique. Their condition can be so fragile that physical restoration, especially procedures that must alter shape, can pose a great risk for further and possibly irreversible damage. Given this risk and the subjective nature of the restoration process itself, offering a digitally-based solution becomes very attractive. The risk due to digitization, which is arguably much less than the process of physical restoration, is a one-time thing: Once an object is digitized, there is no further risk of damage as subsequent digital manipulations are explored. If the desired results can be obtained digitally, the need for physical restoration can be postponed until absolutely necessary.

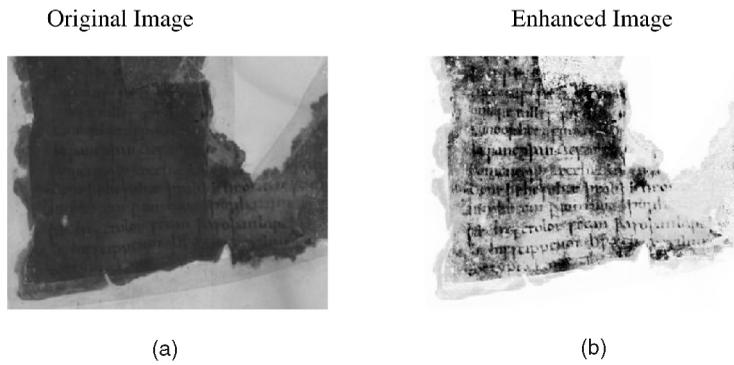


Fig. 2. Example of 2D image enhancement. (a) Original image under white light. (b) The result from simple intensity stretching. The enhanced text is perceptually clearer than the original image.

2.2 Perceptual Image Enhancement

Image enhancement and restoration techniques can be used to help increase the perceptual saliency of features in digitized material, such as characters in a text. The quality of this kind of digital restoration is often quantifiable, and can also be subjectively confirmed without risk to originals (Fig. 2). This is a major advantage that digital representations provide over their physical counterparts: Markings that are difficult or even impossible to detect with the unaided eye can be revealed by manipulating the intensity values of a digital image [4], [2].

Digital image enhancement procedures have been successfully applied in a number of contexts on valuable manuscripts [14], [20]. In one instance, special video imaging techniques radically improved contrast in badly damaged portions of a manuscript [14]. This processing allowed subsequent enhancement and processing algorithms to reveal previously obscured textual content.

2.3 Deskewing Algorithm for Document Analysis

Skew is a document analysis term that refers to image distortion. The process of removing distortion is commonly referred to as “deskewing.” Deskewing is often used as an early preprocessing step to provide distortion-free images for subsequent document analysis routines. Normally, planar deskewing is the only image preprocessing step for document analysis that attempts to align the document’s textual content with the raster representation [18], [12].

Deskewing algorithms almost unilaterally assume a planar document. Under that assumption, the deskewing transformation is formulated as a homography, as shown in Fig. 3. Techniques that address the correction of *nonplanar* materials can be useful in situations where the deformation is

predictable. These techniques focus on binder-curl, where the 3D deformation can be parameterized by a cylindrical model [25], [10], [28]. These approaches correct the curvature distortions present in pages near the binding of a book. While such approaches consider projective distortion, they are limited to a very specific type of curvature distortion.

Our goal is to provide a general deskewing algorithm for *arbitrarily* warped documents. Our assumption is that the documents were *originally* flat, but have become warped over time. The goal is to work from an accurate 3D model of the document. Under these assumptions, we can formulate the three objective components of a restoration process using 3D shape and associated texture as the basic digital representation. The goal state is to produce a planar (flat) shape. The process to manipulate the object involves shape deformations that obey physical laws such as rigidity or material elasticity. The metric is an assessment of the progress of the shape toward the planar goal, which can be objectively measured as an error metric. As we will show, this precise formulation removes subjectivity from the restoration process and produces a result that is closer to the original when it is indeed the case that the damaged object was originally flat.

2.4 Physically-Based Modeling

While the result of our restoration procedure is an undistorted image, the framework begins by manipulating the 3D shape of the manuscript. Our algorithm employs a deformation model similar to *physics-based modeling* approaches in computer graphics. These types of *physics-based* models incorporate classical Newtonian physics into the behavior of 3D models. Most often, these approaches are applied in computer graphics applications to determine changes in the

Typical Planar Deskewing

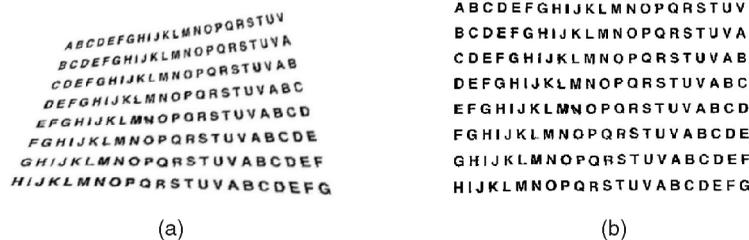


Fig. 3. An example of typical planar deskewing algorithms. (a) An image with content that is not raster aligned. (b) A homography is used to align the content with the raster image.

appearance of materials undergoing dynamic deformations. The goal in computer graphics is to produce a convincing performance by realistically deforming objects according to expected physical laws and allowing them to react and dynamically change their geometric structure based on forces in a simulated environment.

Of specific interest to the restoration problem are approaches that involve the deformation of sheet-like materials. In these approaches [5], [24], [23], [21], an initial flat sheet is deformed and warped under simulation. Typically, a sheet is “dropped” and allowed to collide with various obstacles. Gravity and obstacle collision produce forces that act on the sheet and cause it to deform. These approaches are quite robust for producing realistic effects for materials such as cloth. The techniques are often used to determine how a cloth would appear if draped over an object or how an article of clothing dynamically deforms on an animated character.

Our goal for sheet manipulation can be cast as the *inverse* problem. We start with a sheet that has already been deformed to a complex shape and must reverse the deformations to move the shape back to its original planar state. The unwarping must obey a set of material constraints to simulate the physical process of “flattening.” The physics-based model can provide a framework for approaching the restoration problem in this manner, although it should be noted that this is only a first-order approximation to the actual physical phenomena that take place when a material is manipulated. A much more detailed simulation could incorporate higher-order effects, such as specific material properties of the substance being flattened (animal skin versus paper), interactions of micro-structures in particular materials, and even the amount of moisture present in the material. Clearly, for complex materials like a decaying manuscript, it may be impossible to establish an accurate model that incorporates all of the physical events that have resulted in its current state. Instead, we are interested in manipulating shape in a reasonable and flexible way to obey first order physical laws (rigidity, elasticity) and thereby restore the *perceptual* quality of the digital representations. As will be shown by our controlled experiments, constraints imposed by our first-order physics-based simulation are enough to completely remove distortion caused by rigid deformation.

2.5 Types of Geometric Deformations

Before discussing the distortion removal procedure, we clarify the type of deformation that may occur to the document’s surface. These deformations can be classified into two categories: rigid and nonrigid.

2.5.1 Rigid

We define *rigid* deformation to mean any deformation that preserves surface geodesics. The definition concerns the intrinsic properties of the surface, independent of any embedding. For example, bending and folding a manuscript page is a rigid deformation since points on the manuscript remain the same distance apart when measured on the manuscript. Those points may be closer in 3D since the page can fold, but we define rigidity as a metric on the surface rather than in the space in which the surface is embedded.

Such rigid surface deformations are typical of the damage and buckling caused from decades of handling and mishandling of documents. Years of folding and

unfolding of large materials can leave them warped and creased. Other causes include the natural bending of pages, especially near a book’s binding.

2.5.2 Nonrigid

Nonrigid surface deformations do not preserve surface geodesics. Such deformations can cause the overall surface area of the document to shrink and/or expand. This type of deformation can certainly be found in damaged materials, especially those suffering from fire and water exposure. For such damage it may be impossible to determine the extent of the nonrigid deformation without a priori knowledge of the material. For example, certain knowledge of the proper appearance of written text can lead a scholar to believe that the document has undergone shrinkage. This type of deformation requires the subjective involvement of an expert because the rigid deformation assumption is violated.

Our restoration framework provides a mechanism for the correction of the damage due to *rigid* surface deformations. We do not address nonrigid effects, although, in cases where nonrigid deformations are present, we believe the two types of deformations can be addressed separately. The idea is to first correct the rigid component of the deformation and then address the remaining effects from the nonrigid component. This holds promise for giving scholars a new approach to the difficult problem of decomposing the rigid and nonrigid components of a badly damaged manuscript.

3 DISTORTION RESTORATION

The primary purpose in developing these digital restoration techniques was to apply them to a portion of a badly damaged collection located in the manuscript collection at the British Library. With that goal, we first developed and tested the restoration methods on control test data in the laboratory. After that development, we scanned a portion of the manuscript collection at the British Library and applied the technique to that data. In order to produce the highest quality results, we carefully engineered an imaging environment that could be reliably transported and deployed in a library setting. The next sections detail the data acquisition process and the details of the restoration method we applied to laboratory control data and to the data we gathered of the manuscript collection.

3.1 2D Images

The British Library provided us with access to its color digital camera (Kontron Elektronik GmbH [15]). We used this camera, mounted on a gantry perpendicular to the work table where the manuscript was positioned, to obtain high-quality color images (see Fig. 4). The particular medieval manuscript we digitized has a long and interesting history and is the subject of substantial scholarly study [13]. More important within the context of the restoration framework is the manuscript’s composition and state of damage. The manuscript leaves, made of vellum, are brittle and badly deteriorated. Making matters worse are the acidic paper frames into which the leaves were set in the nineteenth century, together with other unfortunate early attempts at conservation such as glue and gauze that subsequently became opaque.

The paper frames are bound and hold the vellum leaves (Fig. 4). The manuscript is digitized by positioning the bound volume on the table under the camera gantry. Lighting is

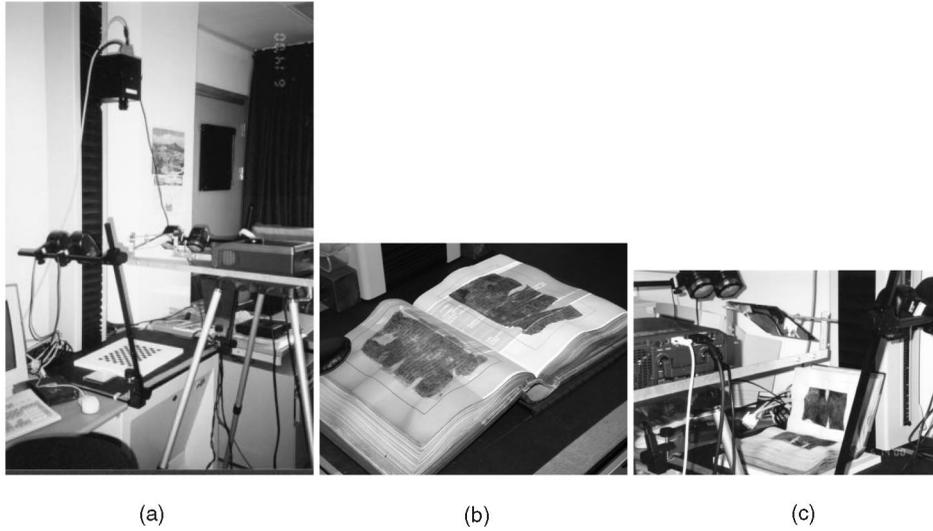


Fig. 4. (a) The high resolution camera is positioned on a gantry above the manuscript. (b) The velum is fastened in paper frames, which are bound together like a book. (c) Lighting around the manuscript is configurable and 3D surface reconstruction makes use of a mirror-deflected LCD projector.

positioned on either side of the manuscript and both white and ultraviolet light are used to illuminate the manuscript. Several images of the manuscript under different lighting conditions are captured without moving the manuscript, which ensures that the multiple images are registered. This image registration allows portions from different images to be composited together to form new images. After capture, most images are enhanced using standard intensity manipulation algorithms such as histogram equalization and contrast stretching.

3.2 3D Data Acquisition

We acquired a 3D reconstruction of each manuscript page using a structured-lighting system composed of a light projector and the British Library’s digital camera. In our implementation, the light source (LCD projector) was directed onto the surface of the manuscript using a front surface mirror (Fig. 4). Both the camera and projector are calibrated to obtain their 3×4 projection matrices, \tilde{P}_{cam} and \tilde{P}_{proj} . In order to recover a 3D point, M , on the surface of the manuscript, a correspondence between the camera coordinates $C(u, v)$ and projector coordinates $P(x, y)$ must be established. Once this correspondence is found, the 3D point can be reconstructed by solving the following system:

$$\begin{aligned} [u \ v \ 1]^T &= \tilde{P}_{cam} M \\ [x \ y \ 1]^T &= \tilde{P}_{proj} M, \end{aligned}$$

where M is the 3D point to be determined, \tilde{P}_{cam} and \tilde{P}_{proj} are the projection matrices of the devices, and $[u \ v \ 1]^T$ and $[x \ y \ 1]^T$ are the 2D device coordinates of the point M for the camera and projector, respectively.

Structured-light gives a robust method to establish device correspondences. Turning on a projector pixel $P(x, y)$ illuminates a point M on the object surface. The camera can observe this illuminated point in its image plane as coordinate $C(u, v)$, thus establishing the correspondence. This process can be accelerated by striping a line instead of a single point. In our implementation, 3D points are reconstructed in millimeters. The accuracy of the reconstructed points is dependent on calibration accuracy and camera positioning. Mean image-plane reprojection error measurements for known pattern

points vary between 0.1 and 0.3 pixels, which translates into 0.1 to 0.5 mm given a typical experimental setup [6].

The 3D points are reconstructed in a coordinate frame established on the table, with the positive z-axis in the direction of the camera. The origin is defined on the table as shown in Fig. 4. Note that, with this setup, an imaged manuscript that is completely flat would have depth value $z = 0$ for each imaged pixel.

3.3 2D/3D Data Relationship

The acquired 2D and 3D data are tightly coupled. Because the camera is used to capture 2D images *and* to triangulate the 3D points, we have a mapping between each image pixel $I(u, v)$ and a corresponding 3D points, M . Thus, each image coordinate contains a tuple of data as:

$$I(u, v) = (r, g, b, x, y, z),$$

where r, g, b are the intensity of the red, green, and blue color channels and x, y, z are the recovered 3D points in millimeters. Because the camera’s projection matrix is known, the corresponding camera coordinate of any 3D point $M = (x, y, z)$ can be computed as $(u, v) = \tilde{P}_{cam}[x \ y \ z \ 1]^T$.

3.4 Virtual Flattening

Restoration begins by flattening the 3D geometric representation of the document’s structure. Under the assumption that the warped shape of the document is due to some rigid surface deformation (Section 2.5), 3D points on the surface of the document are allowed to move. The motion of these points is constrained so that overall surface area is maintained. We accommodate for errors in the recovered 3D points by providing a small element of elasticity in the sheet.

We use a physics-based modeling approach similar to that of Provot [21], who used a mass-spring system to model rigid cloth deformation. Provot’s design has been demonstrated to be highly effective in modeling rigid sheet deformation and contains elastic properties. The overall implementation is straightforward and provides a flexible framework for future modifications.

The approach uses a 3D particle system of point masses. These particles are logically connected by Hookian springs

to form a sheet. *External forces* are applied to the sheet to cause the overall structure to change, while the *internal forces* of the springs hold the sheet together, striving to keep particles equal distance apart while allowing a slight elastic effect. The slight elastic effect is needed to accommodate errors in the 3D reconstruction.

3.5 Mass-Spring Particle System

A particle system is governed by the classic second order Newtonian equation, $f = ma$, where f is a force, m is the mass of a particle, and a is its acceleration. This is also commonly written as: $\ddot{x} = f/m$. A particle modeled by this equation can be described with six variables,

$$[x_1, x_2, x_3, v_1, v_2, v_3],$$

where x_i represents the particle's 3-space position and v_i represents its velocity. This position/velocity product space is referred to as the *phase space*. The phase space is derivative with respect to time and the subsequent motion equation is $[v_1, v_2, v_3, f_1/m, f_2/m, f_3/m]$. This system describes a particle's mass, position, and velocity at a given instant in time. Dynamic forces can be exerted on these particles over time and new positions are calculated at advancing time intervals.

In a basic particle system, individual particles respond only to *external* forces and have no influence on other particles. This basic system can be extended to incorporate forces between particles. One common extension, referred to as a mass-spring particle system, arises by logically connecting particles together via springs. The resulting forces on such a system can be categorized into two types: *internal*, or forces between particles, and *external* forces. The slightly modified equation expressing this is: $F_{int} + F_{ext} = ma$. In the following sections, we explain the particulars of the mass-spring particle system used in our experiments.

3.6 Internal Forces

3.6.1 Ideal Hookian Spring

The internal forces in our system are from springs used to connect particles. Using Hooke's law [22], the equations for forces between two particles, a and b , are:

$$f_a = \left[k_s (|\vec{x}_a - \vec{x}_b| - r) + k_d \frac{\vec{v}_a - \vec{v}_b \cdot \vec{x}_a - \vec{x}_b}{|\vec{x}_a - \vec{x}_b|} \right] \frac{\vec{x}_a - \vec{x}_b}{|\vec{x}_a - \vec{x}_b|}, \quad (1)$$

$$f_b = -f_a, \quad (2)$$

where f_a and f_b are the forces on a and b , r is the rest length of the spring, \vec{x} and \vec{v} are the position and velocity vectors from a particle's current phase space state, k_s and k_d are the spring stiffness and damping constants, respectively, and $|\cdot|$ is the l_2 norm. It is apparent from the equations that, as long as two particles maintain their initial starting distance, r (the resting length of the spring), the spring exerts no force. When external forces begin to move the individual particles from their resting distance, internal spring forces f_a and f_b are induced.

3.6.2 Finite Elements

Fig. 5 shows how the springs are logically connected, forming a finite-element of the mass-spring model. The basic finite-element is a quadrangle. For each quadrangle, springs are attached. Each element is composed of *structural*

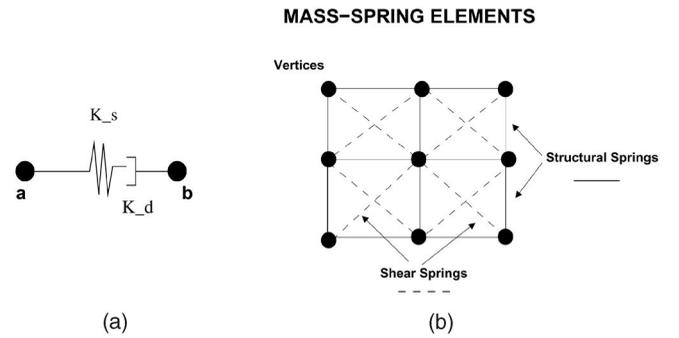


Fig. 5. (a) The ideal Hookian spring with damper acts on two particles. K_s is the stiffness coefficient of the spring and K_d is the dampening coefficient. (b) The finite element structure of the particles consists of structural and shear springs.

springs, which form a quadrangles' hull, and two *shear* springs, which connect diagonally. This particular structure has been shown to be effective and robust for modeling rigidly flexible sheet materials [21], [16], [26]. More springs may be used to create additional rigidity if required [21]. Fig. 6 shows an example of the finite-element meshes initialized with a document's acquired geometry.

3.7 External Forces

3.7.1 Global Downward Force and Damping

A global downward force is exerted on all particles, forcing them toward the $Z = 0$ plane.¹ This force is typically modeled as a constant force, such as gravity, and can be represented as $f = mg$. In addition to a global downward force, a *viscous drag* is applied in the form $f = -\mu_d v$, where μ_d is the drag coefficient and v is the velocity state of a particle. This force encourages a particle to come to rest and improves the numerical stability of the system.

3.7.2 Plane Collision

The goal of "flattening" is to push the 3D shape to the $Z = 0$ plane. When the x_3 phase space component of a particle is equal to 0, it has collided with the $Z = 0$ plane (or when $x_3 < \epsilon$, where ϵ is a threshold close to zero). The response of a particle to this collision changes the velocity of the particle. This new velocity is obtained by separating the velocity at the time of collision into two orthogonal components: the *normal* component and *tangential* component defined by

$$v_n = (N \cdot v)N \quad (3)$$

$$v_t = v - v_n, \quad (4)$$

where N is the collision plane's normal, v is the colliding particle's velocity, and v_n and v_t are the respective components. In our case, these components are trivial to calculate because the plane's normal vector is $[0 \ 0 \ 1]^T$, resulting in the following solutions: $v_n = [0 \ 0 \ v_3]^T$ and $v_t = [v_1 \ v_2 \ 0]^T$. If the collision is purely elastic, a particle's resulting velocity will travel in the same tangential direction and the opposite normal direction such that $v_{new} = v_t - v_n$.

1. In the reconstructed manuscript's world coordinate frame, the $Z = 0$ plane defines the plane the reconstructed manuscript was placed onto during 3D reconstruction. Thus, if the manuscript had been flat, its reconstructed 3D points would all have $Z = 0$ coordinates.

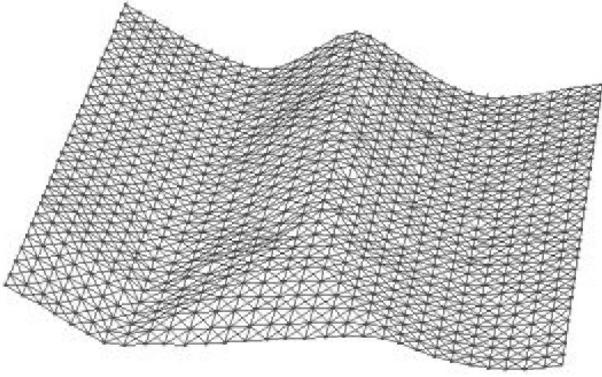


Fig. 6. Examples of mass-spring finite-element particle system mesh initialized with the geometry of a 3D document.

However, the normal direction can be dampened with a coefficient of restitution, k_r , such that $v_{new} = v_t - k_r v_{nr}$, where k_r is a value from $[0 - 1]$. At $k_r = 0$, a particle's normal component will be ignored completely and the particle will essentially "stick" to the plane; alternatively, $k_r = 1$ will cause the particle to reflect elastically. For our experiments, we found $k_r = 0.1$ to be appropriate.

We note that it is possible to set the system's advancing time-step, δt , too high such that $x_3 < 0$ for a particle. This implies that the particle has *penetrated* the collision plane, which is an undesirable effect. In this case, we restore the system to the previous time, $t - 1$, and calculate a new time step such that $\delta t_{new} = \delta t_{prev}/2$, until the penetration condition is avoided.

3.8 Numerical Approach

Incorporating the internal and external forces, Baraff and Witkin [1] describe the system with the following general ordinary differential equation:

$$\ddot{x} = M^{-1} \left(-\frac{\partial E}{\partial x} + F \right), \quad (5)$$

where E is a function of the internal springs, F is a function of external forces, M^{-1} is a diagonal matrix of particles $1/m$, and x is the positional state of the particles.

Our objective is to solve the system of equations over time until all the particles have a phase space with position $x_3 = 0$ and the internal energy (spring forces) reaches a steady state such that the value of $\sum |F_{int}|$ is minimized.

Our implementation solves this objective function implicitly in its implementation. For each particle, we have a structure that encodes the following information: the position in phase space $[x_1, x_2, x_3, v_1, v_2, v_3]$, a force accumulator $[f_x, f_y, f_z]$, and the particle mass m , which is set to 1. The particle's structure also contains a list of all of the springs connected to this particle. At each advancing time step in the simulation, the internal, external, and collision forces on all of the particles are calculated. Once the forces are calculated, they are integrated to give the current velocity of each particle (v_1, v_2, v_3) . A second integration determines the new position (x_1, x_2, x_3) of each particle. The accuracy of the simulation is tied to the numerical approach used to integrate the system over time. We tested this with both a simple explicit *Euler* method and a fourth order

Runge-Kutta integrator. To avoid numerical instability, the Euler method must take small time-steps. Better integrators exist and can be incorporated [1]. For our purposes, the Euler method proved to be sufficiently fast and numerical instability was not encountered.

A document model is flattened by initializing the particle system with the described finite element structure and initial conditions. A global force is then applied to drive the model to the plane. The process is complete when the mass-spring system's internal energy (spring forces) is minimized after collision with the $Z = 0$ plane. Upon completion, the new position state of each particle is saved.

The mass-spring system has no inherent mechanism to keep a surface from folding over on itself during the flattening process. In practice, this problem can be avoided when surfaces to be flattened are height maps, i.e., one-to-one functions from the ground plane to 3D. If 3D points in a height map are projected directly to the $Z = 0$ collision plane, no folding or overlap of the quads would occur. Without spring internal forces, the external gravity force vector, which is perpendicular to the collision, would produce the same effect as projecting these points to the collision plane. The additional internal spring forces ensure preservation of equal distances between points as flattening occurs. Since boundary edges of the surface are not constrained with any force, they are free to spread and ensure that a surface that begins as a height map under a gravity force will not form a fold.

3.9 Flattening Example

Fig. 7 shows before and after views of a document's 3D shape unwarped by the mass-spring particle system. These images were created by applying the mapping, determined by the mass-spring system, from the initial 3D shape of the document to its resulting 2D "flattened" shape.

3.10 Image-Space Restoration

The information from the geometric structure of the flattened document can be used to determine the appropriate image-space warp necessary to produce a distortion-free image. The following sections describe how this image-space warp can be implemented.

3.10.1 Camera Reprojection

From the 3D acquisition phase described in Section 3.2, we obtained a projection matrix of the camera, \tilde{P} , used to capture the original image of the manuscript being restored. The pin-hole camera and virtual flattened manuscript are in the same coordinate frame. This allows us to create a new image of the flattened manuscript as it would have appeared when imaged by the same camera that captured the image of the deformed manuscript. The following describes how to calculate the new *restored* image.

Each vertex in the quad mesh of the mass-spring element sheet stores the coordinates of its initial 3D position, $M_i = [X, Y, Z]_i^T$ ($[x_1, x_2, x_3]$ in phase space notation) in the world coordinate frame. The camera image's (u, v) coordinates can be determined by projecting the 3D point into the camera's frame by $m_i = [us \ vs \ s]_i^T = \tilde{P}M_i$, where \tilde{P} is the camera's projection matrix. The matrix \tilde{P} is obtained during the 3D acquisition step described in the previous chapter. After the flattening process, each vertex is associated with a new 3D point, $\hat{M}_i = [X', Y', Z']_i^T$. These new points are still in the world coordinate frame of the calibration pattern; their new position can be computed by projecting them into the camera frame as before such that $\hat{m}_i = [us' \ vs' \ s']_i^T = \tilde{P}\hat{M}_i$. The

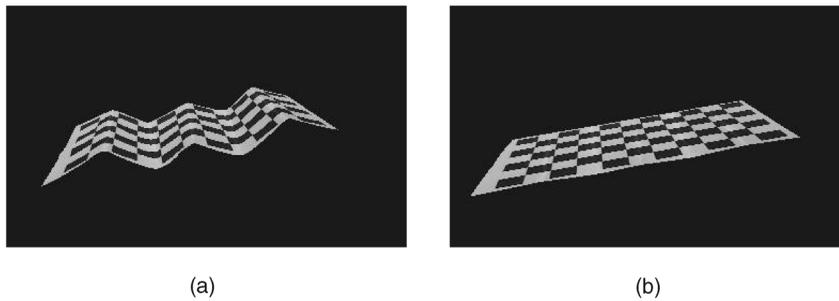


Fig. 7. (a) The *before* image shows the recovered 3D shape. (b) The *after* image shows the final position of the mass-spring system after convergence, textured with the original 2D image.

restored image is created by texturing each finite element (quad) from its initial m_i position to the new \hat{m}_i image coordinate position. We do not alter the pixel intensities in the original image and, therefore, do not compensate for shading effects that may have been captured due to the deformed shape of the document. Fig. 8 shows an example of the restoration process. Fig. 9 shows an example of the 2D restoration.

3.10.2 Orthographic Reprojection

Because we have a full 3D description of the document's shape, we can create a restored image from any desired view under any viewing model. Using the same camera view as the original image is a natural choice. Restoring the image from the original camera position minimizes the resampling that needs to be done to the original image, resulting in a flattened image with minimal blurring and aliasing degradation.

For the purpose of visualization, it may be desirable to create a restored image from a new point of view. For example, a virtual orthographic camera (with center of projection at infinity) can be specified such that the camera axis is normal to the flattened document and the center of the image is the center of the flattened document. For our approach, the normal to the flattened document is the $[0, 0, 1]$ vector since the 3D points have been flattened to the XY

plane. In the context of the physical setup, this corresponds to an orthographic view of the calibration pattern.

Using this orthographic projection, only the XY components, $[X', Y']_i^T$, of the flattened 3D points, \hat{M}_i , in the quad mesh are needed. To create a restored orthographic image, the original 2D points m_i are warped to correspond to their 2D location $[X', Y']_i^T$. Recall that the $[X', Y']_i^T$ points will be in the metric coordinate frame of the calibration pattern. Their values are therefore real metric distances. The user can specify how they wish to scale these distances into pixel values, thus specifying the desired "dpi" in the X and Y direction. The $[X', Y']$ coordinates will then be normalized and scaled to produce an image with accurate dpi such that m_i map to $[s_x X', s_y Y']_i^T$, where s_x and s_y are the dpi scale. The orthographic image can encode accurate scale that can be stored as metadata within the collection itself. As with the previous example, the restored image is created by texturing each quad mesh from its initial m_i position to the new $[s_x X', s_y Y']_i^T$ image coordinate position.

Restoration in this manner creates images of uniform scale. This can be useful if the camera that created the original image was positioned off-axis to the flattening plane. In this case, the camera itself will produce an image that appears slightly skewed (in a planar sense) from the off-axis projection. This uniform scale can be useful for making metric measurements and can be applied to the problem of mosaicing physically fragmented documents into an accurate, coherent whole.

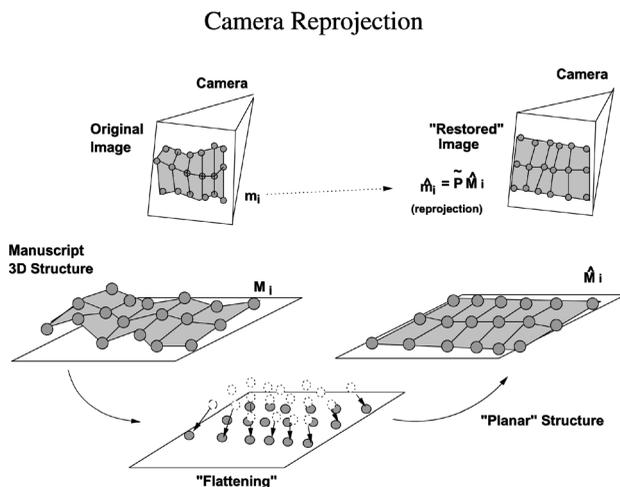


Fig. 8. Image restoration in the camera's image plane. Image restoration in the camera's image plane. Three dimensional points on the surface of the deformed document, M_i , correspond to their position in the original image at m_i . The flattening process moves these points to new locations, \hat{M}_i . These new points are reprojected back into the camera's image plane using the projection matrix, P . This determines how to warp the original distorted imagery by moving 2D points m_i to new image coordinates \hat{m}_i .

4 EXPERIMENTAL RESULTS

We present three experiments that demonstrate the effectiveness of the restoration approach. For the first two experiments, models were acquired using a Sony TRV1000 digital video camera (640×480 pixel resolution) and an Epson Powerlight ($1,024 \times 768$ pixel resolution) projector. The third experiment uses an Epson Powerlight projector

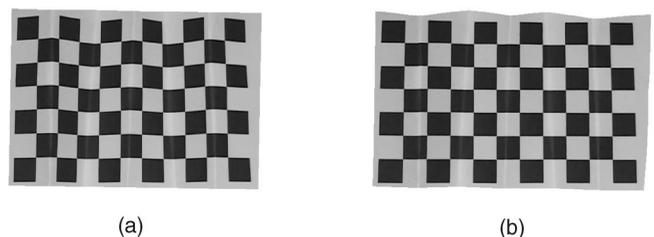


Fig. 9. Example of image-space distortion removal. This figure shows the effects from warping 2D points, m_i , to their new locations, \hat{m}_i , defined by the flattening procedure.

and the Kontron digital camera at the British library. The reconstruction accuracy of our structured-light setup is estimated at $0.25mm$. The accuracy is estimated by reconstructing several planes at different orientations and measuring the mean error from the best-fit plane [9].

The first and second experiments are performed with controlled test cases, where printed materials have been rigidly crumpled by hand and restored. By using a control we are able to compare the original (before damage) and the results from our restoration framework. This allows a quantitative analysis of the restoration framework. The first experiment provides metric and image-based measurements between the experiment’s control document and the resulting restored image. The second experiment provides a qualitative comparison between warped and unwarped text to show the improvement in text readability without any assumptions about document content. The third experiment applies our technique to damaged medieval manuscripts housed at the British library. This experiment shows how surface markings (such as lines drawn on the manuscript) can be restored.

In all the experiments, a mass-spring finite-element mesh of quadrilaterals is constructed from 45×45 samples of points on the surface of recovered documents. Initial spring lengths are assigned based on 3D points distances apart and vary according to the surface shape (i.e., the springs do not have uniform length). The physics-based simulation is then performed to flatten 3D mesh to the $Z = 0$ plane, as described in Section 3.4, and the resulting restored image is computed. The flattening procedure takes approximately one minute to converge for each experiment running on a Dell Pentium 800Mhz (512 MB RAM).

4.1 Experiment I: Pixel Distance

The first experiment quantifies the ability of the mass-spring system to restore a rigidly deformed 3D sheet to its original 2D planar shape. Fig. 10 shows a 2D image of a flat piece of paper with a checkerboard pattern printed on it. The flat document is imaged on the $Z = 0$ plane and removed. The image of the document before it is crumpled serves as the experimental control. The document is then crumpled by hand. The shape of this “damaged” document is acquired using our structured-light system. The image is restored using the framework outlined in the previous section. For this experiment, we use *camera reprojection* to create a restored image in the camera’s coordinate frame.

After the restored image has been constructed, we compare it with the original control image. The shadows and other shading/illumination cues on the 3D surface of the crumpled sheet are still present. Hence, we make a comparison between the control image and the restored image by comparing *features* in the two images (instead of pixels). In this case, we use the corners of the checkerboard pattern. Because the orientations of the two images will differ slightly, the two point sets are brought into alignment by calculating a least-squares homography between the two point sets.

Fig. 10 shows five trials with their results. experiments. The 2D points are back-projected onto the $Z = 0$ plane to provide metric measurements in the world coordinate frame.

Results from this experiment show that the restored image can be registered to the control image within half an image pixel and within $0.25mm$ in the world coordinate frame.

4.2 Experiment II: Text

The second experiment uses typeset text as a way to compare warped versus restored letterforms. A document of letters is

imaged while flat, as a control image, and afterward is crumpled. The four experiments shown in Fig. 11 illustrate typical results. The readability of the text improves substantially under the flattening transformation. Note that rows of text become straight lines and individual letters are warped back into their proper proportion. Closeups show a side-by-side comparison of selected letters (noted in the Row I originals by a transparent patch) from the original and restored images. The shape of the letters appears corrected. For example, the ascender of the letter “L” is lengthened to its proper height.

It is important to observe that no textual or content-based information is necessary in order to perform this restoration. While, in certain problem domains, it may be important to use constraints such as “lines of text are straight,” our method does not require any a priori knowledge about the content. With respect to a general restoration process for handwritten, damaged manuscripts, decoupling the restoration from the content is a decided advantage.

4.3 Experiment III: Manuscripts

In the third experiment, illustrated by Figs. 12 and 13, we apply the flattening restoration to manuscript data from the British Library. We acquired and restored the folios of the Otho B. X manuscript, which consists of the front and the back of approximately 67 badly damaged folios, each set in a paper frame and bound together as a book. Fig. 12 shows two selected examples from this extensive application (134 restored pages, each at a resolution of $2,000 \times 1,500$ pixels). The first example (Fig. 12, Row I) shows the flattened portion of Otho B. X 18v, which is shown before restoration in Fig. 1. It is clear that the line on the bottom of the paper frame, which does not appear as a straight line in the original imagery because of shape distortion, becomes linear. Fig. 12 Row II shows the shape of a portion of the top of Otho B. X 22v. The pronounced bubble makes the text appear to ride across the page along a curve. After restoration, the lines of text become nearly linear.

Fig. 13 shows a second manuscript example that makes clear the large changes that occur when the restoration removes shape-based distortion. The two lines of text shown side-by-side show a pronounced effect in terms of the shape of the manuscript letterforms.

4.4 Results Summary

Our experiments demonstrate that we can successfully remove distortion present in images of rigidly deformed documents. The framework manipulates the 3D geometry of the document with a physically-based model which uses a mass-spring particle system to simulate the process of flattening the manuscript. The mass-spring system iteratively calculates the results of a force pushing down on the manuscript’s geometric (nonplanar) structure, forcing that surface back to a plane. The mass-spring simulation computes the trajectory that points on the surface should take in order to deform into a plane. The result is a realistic representation of how the actual text would appear if it were flattened. The relationship of the final position of the points from the simulation to their starting positions determines an image-space warp that maps the original, distorted image onto the restored, flattened image.

5 LIMITATIONS

Our restoration algorithm starts with the original image and works directly to recover the warp that will remove distortion. There is no correction for image cues such as shading or shadows that may be present in the original

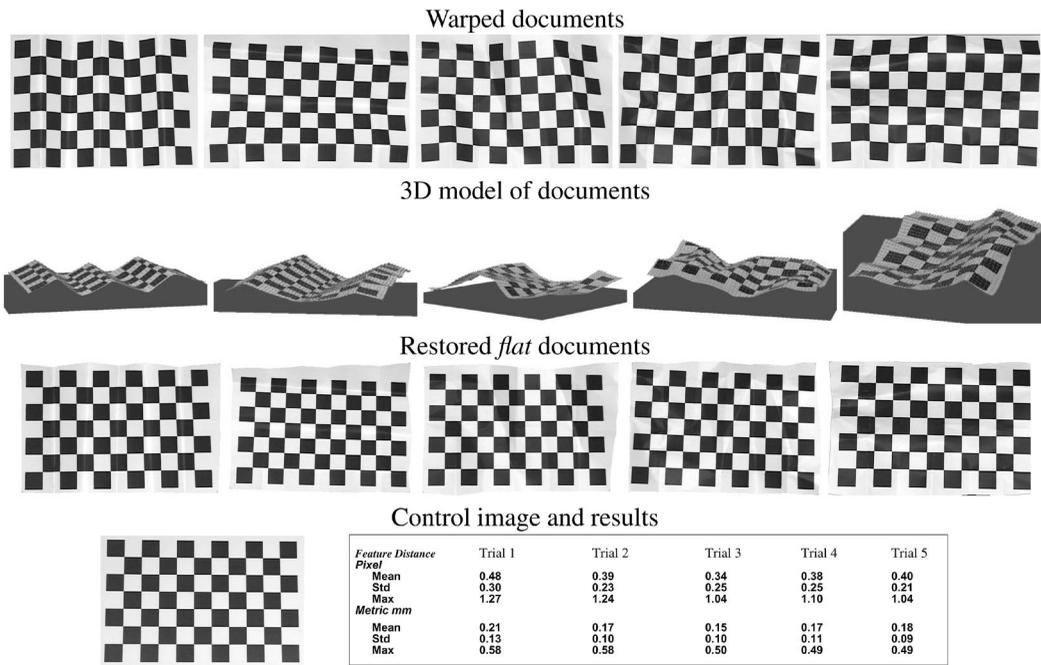


Fig. 10. Experiment I: Row I shows the 2D images of a crumpled document. Row II shows views of the recovered 3D models. Row III presents the restored (flattened) images. Row IV shows the experiment's control image and the pixel and metric distances between corners in the flattened images and corners in the control image.

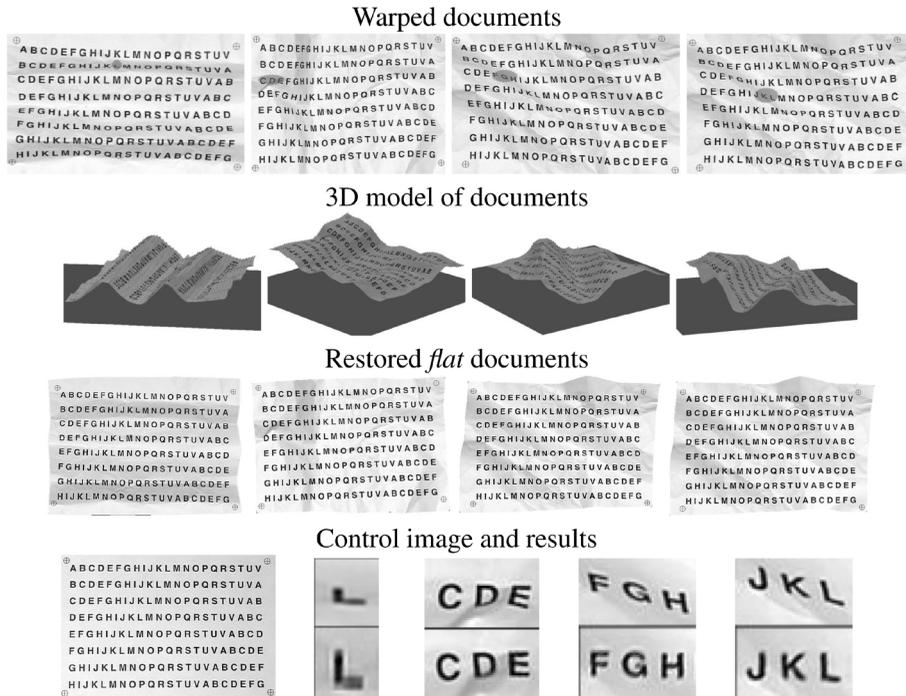


Fig. 11. Experiment II: Row I shows the 2D images of a crumpled and warped lettered document. Row II shows views of the recovered 3D models. Row III presents the restored (flattened) image. Row IV shows the control image and individual letters from the crumpled document (highlighted letters in Row I) compared with the same letters from the restored image.

image. Shading cues give the viewer a strong sense of nonplanar shape and, with these cues still present in the undistorted image, at first glance some restored images still appear to be distorted. On closer inspection, however, distortion removal is apparent. When control images are available for comparison, the differences are readily quantifiable. This shading effect can be lessened with image processing or almost completely removed at image acquisi-

tion by imaging under a range of lighting conditions and positions. Fig. 14 shows an example that uses 2D image processing to lessen the shading effect and highlight the corrections made by removing the projective distortion.

While the algorithm removes projective distortion, it cannot fill in missing intensity information that was lost due to projection. Therefore, unwarped regions of high deformation may appear blurry due to the lack of intensity

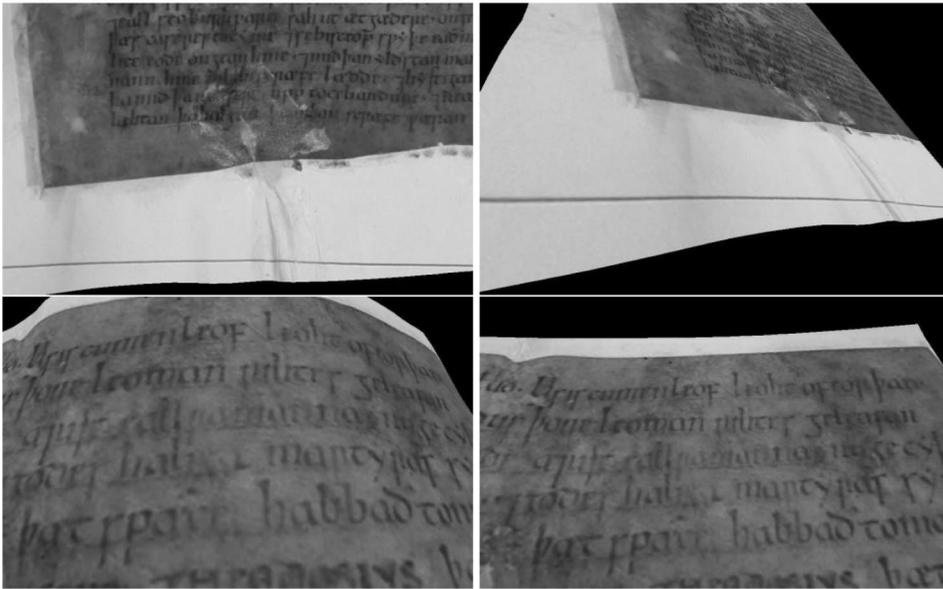


Fig. 12. Experiment III: Results using damaged manuscripts. Row I shows the flattened portion of the damaged manuscript from Fig. 1. Row II shows the profound effect that the flattening restoration can create for manuscript pages that are very badly curved.

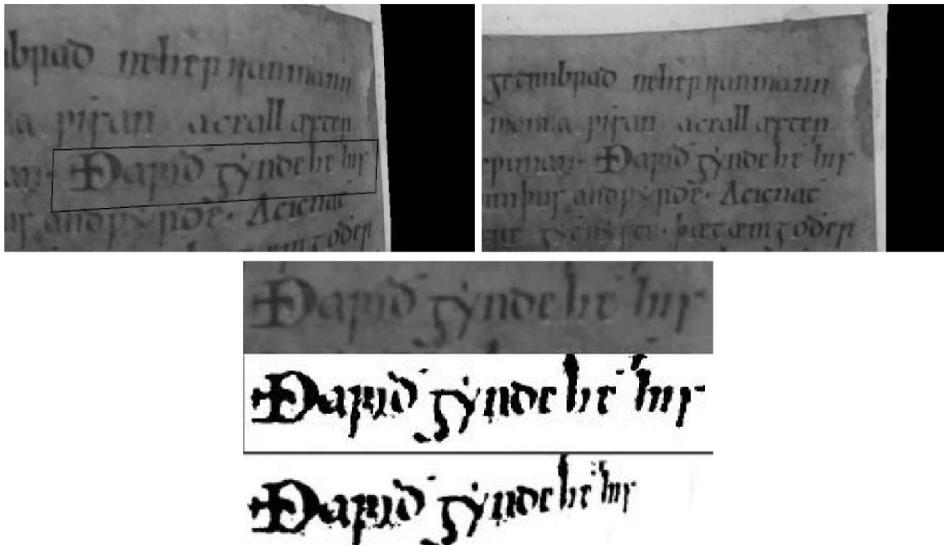


Fig. 13. Experiment III: damaged manuscript closeup: Row I shows a side-by-side view of a manuscript page and its restoration. Row II shows the effect flattening has on letterforms. The line of flattened, enhanced text differs substantially from the shape and size of the original. The difference is purely from removing variation due to shape distortion.

information (Fig. 15). Image processing routines, such as sharpening, can help reduce this effect. However, a more appropriate solution would involve the capture of multiple images of the original warped document to help fill in missing intensity information. Research on super-resolution imaging, where multiple calibrated images are combined to form a composite higher-resolution texture, can be applied in this situation [3], [11]. In particular, we have done preliminary experiments with five low-resolution cameras and a super-resolution algorithm for combining their images into a single texture. This method can potentially remove the need for the additional “texture” camera and has the advantage of supplying data from a number of camera locations. The most challenging problem at acquisition time is accurate calibration of the imaging devices so that their fusion is accurate.

6 SUMMARY

We have presented a new framework for restoring manuscript images by removing the shape distortion that results from images of damaged (warped and deformed) documents. Our framework uses the 3D geometry of a document, which can be captured using a number of techniques, to impose first-order physical constraints such as rigidity. We use a physically-based mass-spring particle system to guide the iterative unwarping process. The restoration algorithm produces a new image that realistically represents how the manuscript should appear if it were once again flat. The underlying assumption is that the manuscript was originally flat and that the present (damaged) shape resulted from a rigid deformation process.

The controlled experiments we presented using the restoration algorithm demonstrate that the framework can



Fig. 14. (a) Shadows in the original image are still present after restoration. (b) The appearance of these shadows can be lessened by 2D image processing routines.

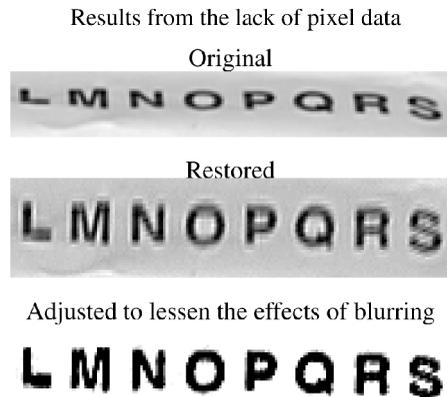


Fig. 15. The restored image is created by flattening the original image according to the image warp found through application of the mass-spring system. Deformations in the image can cause blurring in regions with little pixel coverage. Image processing can help to lessen these effects.

restore documents to within 0.25mm of their original shape and within one pixel in image-space. The real examples from collections at libraries show the results we can obtain "in the field" for real digital editions.

REFERENCES

- [1] D. Baraff and A. Witkin, "Large Steps in Cloth Simulation," *Computer Graphics (Proc. SIGGRAPH)*, pp. 43-52, Aug. 1998.
- [2] J. Becos, "Image Processing Algorithms for Readability Enhancement of Old Manuscripts," *Proc. Int'l Electronic Imaging Exposition and Conf.*, pp. 392-397, 1989.
- [3] S. Borman and R.L. Stevenson, "Super-Resolution from Image Sequences—A Review," *Proc. 1998 Midwest Symp. Circuits and Systems*, 1998.
- [4] G. Braudaway, "Restoration of Faded Photographic Transparencies by Digital Image Processing," *Proc. IS and T's 46 Ann. Conf.*, pp. 287-289, 1993.
- [5] D.E. Breen, D.H. House, and M.J. Wozny, "Predicting the Drape of Woven Cloth Using Interacting Particles," *Computer Graphics (Proc. SIGGRAPH)*, vol. 28, pp. 365-372, 1994.
- [6] M.S. Brown, "New Techniques and Algorithms for Acquiring, Restoring, and Displaying Digital Collections," PhD thesis ukycocs-2001-d-003, Univ. of Kentucky, Lexington, 2001.
- [7] M.S. Brown and W.B. Seales, "3D Imaging and Processing of Damaged Texts," *Proc. Assoc. for Computing in the Humanities and Assoc. of Linguistic and Literacy Computing Joint Conf.*, June 2001.
- [8] M.S. Brown and W.B. Seales, "Digital Atheneum: New Approaches for Preserving, Restoring and Analyzing Damaged Manuscripts," *Proc. First ACM-IEEE Joint Conf. Digital Library*, June 2001.
- [9] M.S. Brown and W.B. Seales, "Document Restoration Using 3D Shape: A General Deskewing Algorithm for Arbitrarily Warped Documents," *Proc. Int'l Conf. Computer Vision (ICCV '01)*, July 2001.
- [10] H. Cao, X. Ding, and C. Liu, "Rectifying the Bound Document Image Captured by the Camera: A Model Based Approach," *Proc. Int'l Conf. Document Analysis and Recognition (ICDAR '03)*, pp. 71-75, 2003.

- [11] P. Cheeseman, B. Kanefsky, R. Kraft, J. Stutz, and R. Hanson, "Super-Resolved Surface Reconstruction from Multiple Images," G.R. Heidbreder, ed., *Maximum Entropy and Bayesian Methods*, pp. 293-308, Kluwer Academic, 1996.
- [12] B. Gatos, N. Papamarkos, and C. Chamzas, "Skew Detection in Text Line Position Determination in Digitized Documents," *Pattern Recognition*, vol. 30, no. 9, pp. 1505-1519, 1997.
- [13] K. Kiernan, B. Seales, and J. Griffioen, "The Reappearances of St. Basil the Great in British Library MS Cotton Otho B.X.," *Computers and the Humanities*, vol. 36, pp. 7-26, 2002.
- [14] K.S. Kiernan, "Digital Image Processing and the Beowulf Manuscript," *Literary and Linguistic Computing: Special Issue on Computers and Medieval Studies*, vol. 6, pp. 20-27, 1991.
- [15] Kontron Elektronik GmbH, Kontron Embedded Computers ag, <http://www.kontron.com>, year?
- [16] J. Lander, "Devil in Blue Faceted Dress: Real-Time Cloth Animation," *Game Developer*, May 1999.
- [17] E. Lecolinet, L. Likforman-Sulem, L. Robert, F. Role, and J-L. Lebrave, "An Integrated Reading and Editing Environment for Scholarly Research on Literary Works and Their Handwritten Sources," *Proc. Third ACM Conf. Digital Libraries*, pp. 144-151, 1998.
- [18] U. Pal and B. Chaudhuri, "An Improved Document Skew Angle Estimation Technique," *Pattern Recognition Letters*, vol. 8, no. 17, pp. 899-904, July 1996.
- [19] T. Pavlidis and S. Mori, "Optical Character Recognition," *Proc. IEEE*, vol. 80, no. 7, pp. 1027-1209, July 1992.
- [20] A. Prescott, "The Electronic Beowulf and Digital Restoration," *Literary and Linguistic Computing*, vol. 12, no. 197, pp. 185-95, 1997.
- [21] X. Provot, "Deformation Constraints in a Mass-Spring Model to Describe Rigid Cloth Behavior," *Graphics Interface*, pp. 174-155, 1995.
- [22] R. Serway, "The Law of Motion," *Physics for Scientists and Engineers*, fourth ed., pp. 95-128, Saunders College Publishing, 1996.
- [23] D. Terzopoulos and K. Fleischer, "Modeling Inelastic Deformations: Viscoelasticity, Plasticity, Fracture," *Computer Graphics (Proc. SIGGRAPH)*, vol. 22, pp. 269-278, Aug. 1988.
- [24] D. Terzopoulos, J.C. Platt, and A.H. Barr, "Elastically Deformable Models," *Computer Graphics (Proc. SIGGRAPH)*, vol. 21, pp. 205-214, 1987.
- [25] T. Wada, H. Ukida, and T. Matsuyama, "Shape from Shading with Interreflections under Proximal Light Source: 3D Shape Reconstruction of Unfolded Book Surface from a Scanner Image," *Int'l Conf. Computer Vision (ICCV '95)*, pp. 66-71, 1995.
- [26] Y. Wu and D. Thalmann, "Deformable Surfaces Using Physically-Based Particle Systems," *Computer Graphics Int'l*, pp. 205-216, year?
- [27] C.J. Yuan and W.B. Seales, "Guided Linking: Efficient Making Image-to-Transcript Correspondence," *Proc. First ACM-IEEE Joint Conf. Digital Libraries*, June 2001.
- [28] Z. Zhang and C.L. Tan, "Correcting Document Image Warping Based on Regression of Curved Text Lines," *Proc. Int'l Conf. Document Analysis and Recognition (ICDAR '03)*, pp. 589-593, 2003.



and analysis, and computer graphics.

Michael S. Brown received the BEng and PhD degrees, both in computer science, from the University of Kentucky, Lexington, in 1995 and 2001, respectively. He was a visiting PhD student at the University of North Carolina at Chapel Hill from 1998 to 2001. In 2001, he joined the Computer Science Faculty at the Hong Kong University of Science and Technology as an assistant professor. His research interests include image processing, document restoration



W. Brent Seales received the BS degree in computer science from the University of Southwestern Louisiana and the MS and PhD degrees in computer science from the University of Wisconsin-Madison. In 1991, he joined the Computer Science Faculty of the University of Kentucky and now holds the post of associate professor. His central research interest is in computer vision and image processing, with applications in digital libraries, medical visualization, and multimedia.